TECHNISCHE UNIVERSITÄT DARMSTADT

Macroeconomic Forecast Evaluation Under Asymmetric Loss

Dissertation

Approved as fulfilling the requirements for the degree of

doctor rerum politicarum (Dr. rer. pol.)

by the Department of Law and Economics of the Darmstadt University of Technology

by

Diplom-Mathematiker Julian Hoss
(Place of birth: Frankfurt am Main)

First supervisor:	**Prof. Dr. Jens Krüger**
Second supervisor:	**Prof. Dr. Michael Neugart**
Filing date:	**28.01.2016**
Date of defense:	**23.03.2016**

Place of publication:	**Darmstadt**
Year of publication:	**2016**

D 17

Acknowledgments

I am deeply grateful to my supervisor, Prof. Dr. Jens Krüger. He is an extraordinary mentor who has always been open to discussing my work and has provided me with invaluable support and advice. I would also like to thank Prof. Dr. Michael Neugart for his support as my second supervisor, Prof. Dr. Volker Caspari for chairing the examination committee as well as the committee members Prof. Dr. Alexander Kock and Prof. Dr. Dirk Schiereck for their questions and insights during the defense.

Moreover, I am grateful to Sebastian Ruths Sion and Till Weigt for helpful comments and discussion on early drafts of some chapters and to Thomas Büttner for his thought-provoking questions from a non-economic perspective. My colleagues at Darmstadt University of Technology provided me with unwavering moral support and the opportunity to engage in scientific debate. I would like to extend particular thanks to my colleagues in the department of Empirical Economics Vanessa Belew, Dr. Benjamin Hampf, Dr. Stephan Hitzschke, Mathias Rhiel and Maike Schmitt.

Furthermore, I want to thank the researchers who participated in the 35th International Symposium on Forecasting in Riverside California, the 16th IWH-CIREQ Macroeconometric Workshop in Halle (Saale) and the Workshops of the Center for Statistics at the TU Darmstadt for their valuable comments on earlier drafts of individual chapters of this thesis.

Finally, I am thankful for the continuous support of my friends and my family. I am especially grateful to my fiancée, Alexandra Katherine LeCrone, not only for supporting me in many ways throughout the past years but also for her contribution to the style and language of this dissertation.

Contents

List of Tables

List of Figures

1 Introduction

1.1 Economic Forecasting

A proverb often attributed to the famous physicist Niels Bohr says, "it is very difficult to predict - especially the future" (Mencher, 1971, p. 37). While it is hard to disagree with this proverb, it leads us to a series of questions concerning predictions or forecasts. First of all, it is necessary to determine what a forecast actually is and what types of forecasts exist, before asking what makes a forecast a "good" forecast and how the quality of a forecast or series of forecasts could be measured.

Making a prediction or forecast essentially means making a statement about an event in the future. The task certainly becomes more difficult if this event is to happen in the more distant future and the forecast's horizon is therefore larger. A forecast can be based on anything from a simple guess, to an expert's judgment, to a model-based prediction. While professional forecasts certainly should not be founded on guesses, judgmental forecasts have shown to improve forecasts if they are combined with a model forecast (see e.g. Fildes and Stekler (2002)). When making a forecast, a forecaster can either predict the value of a target variable (a so-called point forecast), or provide an interval in which this variable falls with a certain possibility, or even give an estimate of the target variable's entire density.

A forecast's accuracy and precision are the two factors that first come to mind when distinguishing a good forecast from a poor one. Predicting that the economy will fall into a recession during the next century is very accurate, as this very likely will happen, but not precise at all. In contrast, it is highly precise to forecast a recession starting on September 25th, 2025, but most likely not very accurate. Hence, a good, or useful, forecast has to be accurate as well as precise.

As the title suggests, this dissertation is concerned with the evaluation of macroeconomic forecasts. Forecasting macroeconomic variables, such as GDP growth, inflation or (un)employment, has a long tradition, as these forecasts are essential to politicians, central bankers and decision-makers in business as well as in finance. Hence, a large body of literature on macroeconomic forecasting has developed. Particularly thorough overviews of the most important areas in this field are to be found in the handbooks published by Clements and Hendry (2002, 2011), Elliott, Granger and Timmermann (2006) and Elliott and Timmermann (2013).

Predicting macroeconomic developments is closely linked to understanding business cycles, which are highly complex and evolve over time. Furthermore, Blanchard and Watson (1986) state that "business cycles are not at all alike" (p. 125) and point out that economies are exposed to small and frequent shocks. Thus business cycles are very difficult to model adequately and thus hard to predict. On this subject, Clements and Hendry (2002, p. 540) have written the following: "[s]ince an imperfect tool is being used to forecast a complicated and changing process, it is perhaps hardly surprising that forecasts sometimes go badly awry." The tools these authors refer to are mainly time series models and econometric systems, but also leading indicators as well as consumer and business surveys (see Clements and Hendry (1998)). The models referred to are generally distinguished by their nature of being structural or nonstructural models, i.e.

based on a specific theory, or exploiting the information available in past observations using time series tools (see Diebold (1998)).

To better understand why forecasting models often fail it is useful to have a closer look at the components such models usually contain. According to Clemens and Hendry (2002), these components are deterministic variables, observed stochastic variables and unobserved error terms, and the main characteristic in which they can be distinguished is how much the forecaster knows about the future value of each component. While a deterministic variable (e.g. an intercept or a trend) is supposed to be known, the future of a stochastic variable (an explanatory variable or the target variable) is unknown, although its past values can be observed. Unobserved errors are unknown at all times. In a model that contains these three components, each one could be measured inaccurately or change over time, the respective parameters could be falsely estimated, or the entire model could be misspecified. Surprisingly, Clemens and Hendry (2002) find that unexpected changes in the deterministic terms are the main source of forecasting failure. These structural breaks can be caused by a lot of factors. Some examples for sources of change in the economic environment that may cause structural breaks are policy revisions or disruptions to the political climate such as wars, riots, changing legislations, regulations or deregulations of markets (e.g. financial markets), privatizations and the establishment or dissolution of trade, monetary, or political unions. Moreover, global factors such as rising or falling oil prices or technical innovations can affect the economy.[1]

Summing up the challenges a forecaster faces when making macroeconomic predictions, we can establish that:

- macroeconomic systems are highly complex structures with multi-causal connections and complicated feedback relations between variables
- economies are evolving over time and so are the relationships between the variables
- these changes can be very abrupt and economies may suffer shocks
- individuals may adapt their behavior in reaction to a forecast and thus forecasts can become self-fulfilling prophecies
- many macroeconomic time series are observed at a rather low frequency and published with a considerable lag
- variables are often revised substantially between their first and their final publication

These challenges can lead to imperfect models that suffer from parameter instabilities and structural breaks. The data quality adds to the difficulties as low frequencies, publication lags and data revisions increase the uncertainty about the current state of the economy (see e.g. Öller and Teterukovsky (2007) for a study on the quality of macroeconomic data). In any case, as long as macroeconomic forecasts are being made, they need to be evaluated in order to determine which models are the most promising, or which forecaster appears to do better for a certain variable

[1] See Clements and Hendry (2008) for further discussion.

(see e.g. Ashley (1988) and Stekler (1991)). Granger and Newbold (1973) differentiate between a subjective and an objective approach to evaluating forecasts or forecast errors. On the one hand, researchers can try to find out more about the specific reasons for past forecast failures and learn from these past errors. This subjective approach is also supported by Stekler (2007). On the other hand, researchers can use accuracy measures in order to objectively quantify series of forecast errors, for example, by comparing different series of forecasts across variables, forecasters, horizons or time periods.

Considering the apparent difficulties in macroeconomic forecasting, one could ask whether it is worth taking the effort of producing these forecasts and if there is any additional benefit at all. Especially with regard to the experiences from the last recession and the Lehman Brothers' bankruptcy in 2008, which most forecasters failed to predict, the guild of professional forecasters' reputation sank considerably. See Ng and Wright (2013) and Giacomini and Rossi (2013) for studies that discuss the lessons for macroeconomic forecasting to be learned from this crisis.

The following statement clearly expresses one important reason why it is necessary to strive for better macroeconomic forecasts and forecasting methods: "Quite simply, good forecasts lead to good decisions," (Diebold and Lopez, 1996, p. 242). In addition to the goal of predicting a future event as accurately as possible, forecasters may also consider the preferences of or the certain restrictions on the decision-makers who use their forecasts. Thus, the forecaster or forecast-producer and the decision-maker or forecast-user may be the same person, or at least have a similar preference structure (e.g. a bank's in-house forecast that is used for the bank's own commercial operations). An economic research institute, for example, may produce forecasts for policy without having to make its own economic decisions based on them. In this case, the preferences of both parties are not necessarily identical, as the forecasting institute may be concerned about its reputation as well as the accuracy of the forecast, whereas politicians are certainly more interested in the costs of a potential forecast error when acting on a given forecast. The close connection between forecasts and the decisions that are made based on these forecasts is emphasized in Granger and Pesaran (2000), Granger and Machina (2006) and Pesaran and Skouras (2002), to name a few.

The focus on decision making is strongly connected to the forecaster's optimization problem when making a prediction. Thus, an optimal forecast depends on the objective function, either a utility function or a loss function, that the forecaster seeks to either maximize or minimize. Gaining insight on the - potentially asymmetric - shape of the forecaster's loss function is the focal point of this dissertation. A loss function helps quantify the costs related to certain forecast errors by assigning a function value to each error. In this regard, the concept of a loss function is very useful for analyzing and comparing forecast errors of different signs and magnitudes, even when the exact monetary costs of an error are unknown. If the forecaster's or decision-maker's preferences are asymmetric for some reason, this can be modeled using an asymmetric loss function. Potential reasons for these asymmetric preferences may be the different costs associated with positive versus negative forecast errors, the forecaster's incentive for building a certain reputation for being particularly optimistic or pessimistic, herding or anti-herding behavior in individual forecasters with respect to a consensus forecast, or intentionally smoothing a forecast when revising it in order to conceal the gravity of a past mistake.

When analyzing a series of forecast errors, the forecaster's preferences and thus the form of the loss function used when the forecast was made are usually unknown to the researcher. Elliott, Komunjer and Timmermann (2005, 2008), henceforth referred to as EKT, introduce a seminal approach that enables the researcher to gain insight into the forecaster's preference structure. The central idea of their test reverses the traditional way of analyzing an observed series of forecast errors, which (implicitly) assumes a specific loss function (mainly symmetric quadratic loss) and then tests for the optimality or rationality of the forecasts. This traditional procedure is usually based on an approach suggested by Mincer and Zarnowitz (1969). EKT generalize this approach, proposing a flexible family of loss functions and estimating the most plausible degree of asymmetry in the forecaster's preferences for a given series of forecast errors. Herein, the rationality of the series of forecasts is formulated as an implicit hypothesis that is tested within their framework.

1.2 Research Objectives

Using EKT's method as the central approach, this dissertation contributes to the question of whether macroeconomic forecasters produce forecasts and forecast errors that indicate that they have asymmetric preferences. If this is the case, we are also interested in the direction of the asymmetry in order to learn whether forecasters' predictions tend to be optimistic or pessimistic, or whether they are indifferent to the direction of their forecast errors. As mentioned above, the insight into a certain forecaster's loss function can improve a decision-maker's ability to act on this forecaster's prediction. Hence, the decision-maker may act more optimistically or more conservatively than the forecast suggests, given the information on the forecaster's loss function.

Along with the asymmetry in the loss function, we are particularly interested in finding out more about the forecaster's rationality and how efficiently information is used when producing the forecast. Pointing out situations in which a series of forecasts appears to be irrational and inefficient may be understood as an attempt to discredit a forecaster's abilities. In contrast to this line of thinking, our intention here is to try to find indications for unused information in economic variables that could be used to help further improve future forecasts.

When applying the EKT approach to data, we are particularly interested in potential differences in the results with respect to real-time and revised realization data of the target variable. At first glance, it seems likely that forecasters aim to predict a target variable's final realization, regarding earlier vintages only as rough estimates of the true value. For various variables, data revisions may nevertheless take several years or may be caused by ex post redefinitions of certain statistics that cannot be foreseen by the forecaster. Thus, forecast errors need to be evaluated with respect to both types of realizations. For a more elaborate discussion of the use of real-time data see Croushore (2006, 2011).

Apart from applying Elliott, Komunjer and Timmermann's approach to a wide range of novel data and forecasting situations, this dissertation provides an extensive Monte-Carlo analysis of the procedure's statistical properties. Herein, we are particularly interested in the test's ability to detect the correct degree of asymmetry in the loss function as well as the size and power properties of the corresponding test for forecast efficiency. We start by investigating the properties of the EKT test for a wide range of scenarios, considering different error term distributions, such as normally distributed error terms with different variances, autocorrelated error terms and fat-tailed error term distributions. We also analyze a scenario with an induced single outlier in order to find out how the properties of the EKT test change in the presence of a major crisis similar to the recent recession. Disengaging ourselves from the arbitrary choice of a data generating process characteristic of most Monte-Carlo analyses, we simulate a more realistic scenario reflecting the US economy.

In addition to the test's statistical properties with respect to different scenarios, we are mainly concerned with two more questions. First, we are interested in how the results change for different information sets that are available to the forecaster. In particular, we want to find out what happens to the results when some of the variables in the data generating process are omitted from the model the forecaster uses to make the predictions, or when irrelevant variables are included. Second, we are interested in the robustness of the results generated under the different covariance matrices used for the GMM optimization.

Finally this dissertation addresses the question of whether higher moments of the forecast error play a role in macroeconomic forecasting, and if so, how the inclusion of these moments changes the loss function's form. Especially in regards to the last crisis, learning more about possible changes in the forecasters' loss functions in times of increased uncertainty may help to better understand the forecasting process. In financial forecasting, for example, up to four moments are also regarded in the context of a profit-oriented utility maximization (see e.g. Jondeau and Rockinger (2006)).

While various utility functions exist that allow higher moments to be included in financial applications, we face the task of determining a suitable loss function for conducting an analysis similar to the EKT approach. In our macroeconomic data context, the appropriate calculation of the moments is also of interest, as the low frequency of the available time series makes calculating the moments using past observations unappealing.

1.3 Outline of the Dissertation

The following gives a brief overview of this dissertation's structure. Chapter 2 revolves around the general methods needed to measure economic loss. It thus starts by introducing the concept of a loss function, presents a variety of symmetric accuracy measures as well as methods for testing forecast rationality, and, in addition, gives an overview of the relevant literature. The chapter provides several arguments in favor of the possibility of asymmetry in loss functions and discusses the most relevant examples for asymmetric loss functions along with some studies that employ these functions. After introducing the EKT approach as the central method of this dissertation, the chapter concludes with a brief overview of the approach's applications.

Chapters 3 to 5 pursue the research questions briefly outlined in this introduction. Chapter 3 applies the EKT approach to German employment forecasts. The chapter starts with a short discussion of the relevance of forecasting employment separately from GDP, before presenting the forecasting institutions, their forecasts and the realization data. Apart from the potential asymmetry in the institutions' forecasts, we are primarily interested in investigating information efficiency and consider multiple instrumental variables that are arranged in various sets of variables. As a first step, we discuss the results of the EKT test with respect to these instrument sets. At the end of the chapter, we concentrate the information contained in the individual instrumental variables using factor methods and conduct the EKT tests using the first three factors as instrumental variables. At this point, we repeat the test with and without a pre-selecting algorithm before extracting the factors. To our knowledge, the EKT approach has not been combined with factor methods so far. Thus, we are particularly interested in learning more about the form of the loss function when this bundled information is used and whether the information available in these factors is used efficiently.

Chapter 4 provides a Monte-Carlo analysis of the EKT approach's statistical properties. Once again, we start by discussing contexts which necessitate such an analysis and review the rather scarce literature addressing this topic. The ensuing sections illustrate the experiment's design and provide a discussion of the results regarding our baseline simulations. The next step of our analysis introduces several variations into our setting in order to cover a series of data situations that each reflect a single aspect that potentially could impact the results. Moreover, we simulate a scenario that mimics the US economy to gain insight into how the statistical properties might change in an environment with a more realistic parameter setting.

Chapter 5 asks what role the forecast error's higher moments play when they are included in the loss function. After discussing why this may be relevant to macroeconomics, we concentrate on how to alter the EKT approach in order to include higher moments. Focusing on Linex-based[2] loss functions, we analyze how higher moments might be introduced and how this changes the optimality condition for the EKT test. Having laid the methodological groundwork, we apply the new procedure to the ECB's Survey of Professional Forecasters' predictions for growth, inflation and unemployment in the euro area. A special feature of this dataset is its (unbalanced) panel structure which allows to compare the different asymmetry preferences across individuals and which we exploit to calculate the moments in one setup. We start by introducing the survey and the data and discussing previous studies that have been conducted on these data. We then use different approaches to calculate the higher moments and analyze the forecasters' asymmetry preferences with and without higher moments in their loss functions.

Chapter 6 puts our results in perspective by highlighting the main insights we have gained from our analyses and by showing how the results in chapters 3, 4, and 5 relate to each other. The study concludes by outlining possible extensions to our analyses and discussing several related areas that deserve further attention.

[2] The Linex loss function is a loss function with a linear and an exponential part.

2 Measuring Economic Loss

This chapter addresses the theoretical concepts that are needed to measure economic loss. In order to do so, the next section introduces the notation that is used throughout this dissertation as well as the concept of the loss function. We also discuss various forecast accuracy measures that are frequently applied in the literature and present the classic Mincer-Zarnowitz approach for unbiasedness, as well as efficiency tests based on their approach. Section 2.2 argues in favor of potential asymmetries in the forecasters' preferences. These asymmetries are generally modeled using the concept of a loss function. This concept is scrutinized, and the axioms needed to define a loss function are presented. Section 2.3 introduces the approach suggested by Elliott, Komunjer and Timmermann (2005, 2008). This approach is the core method in this dissertation and will be applied, analyzed and extended in the subsequent chapters. The last section of this chapter provides an overview of the literature in which their approach is applied.

2.1 Fundamental Concepts

This section introduces the fundamental concepts used in forecast evaluation. First, we establish the notation and then present the concept of the loss function along with its basic properties, before discussing a variety of accuracy measures as well as tests for unbiasedness and rationality. The section concludes with a review of empirical applications of these tests.

2.1.1 Loss Functions

This dissertation uses the following notation: y_{t+h} and $f_{t+h,t}$ denote the realization of the target variable y and the forecast for this variable f at time $t+h$, respectively. The forecast horizon is h periods. We assume the forecast to be based on the information that is available in the previous period t. The available information is summarized in the information set Ω_t. In the remainder of this dissertation, the subscript t is often suppressed in the forecasts for notational convenience, writing simply f_{t+h}. The difference between the realizations and the forecasts $e_{t+h} = y_{t+h} - f_{t+h}$ is the forecast error e. Thus, a positive forecast error ($e_{t+h} > 0$) is the outcome of an underprediction of the target variable ($f_{t+h} < y_{t+h}$), while a negative forecast error ($e_{t+h} < 0$) is the outcome of an overprediction ($f_{t+h} > y_{t+h}$). Of course, a correct forecast implies that $e_{t+h} = 0$. Throughout this dissertation, we consider series of forecast errors e_{t+h} of length T, i.e. $t = 0, \ldots, T-1$.

As a continuing series of perfect forecasts is highly unlikely to occur and forecast errors tend to differ from zero in general, concepts for the quantification of forecast errors are needed. In the words of Granger and Newbold (1986), "[a]n intellectually satisfying way to proceed is to introduce the idea of a *cost function*" (p. 121). Granger (1969) first introduced this formal concept for quantifying the cost or the economic loss related to a forecast error. Today, the more common term for this concept is a *loss function*. From a mathematical point of view, the loss function $L(.)$ maps realizations and forecasts on the non-negative real numbers, given the information set available at time t, $L\left(y_{t+h}, f_{t+h} \middle| \Omega_t\right) \in \mathbb{R}_+$. Frequently, the loss function is specified as a function only of the forecast error, i.e. $L\left(e_{t+h} \middle| \Omega_t\right) \in \mathbb{R}_+$.

To be considered as a loss function, the following properties first stated in Granger (1969) and refined in Granger (1999) need to be fulfilled. A loss function is usually a continuous and differentiable function that takes its unique minimum at zero for perfect forecasts $f_{t+h} = y_{t+h}$. This implies $e_{t+h} = 0$ and $L(e_{t+h}) = 0$ if and only if $e_{t+h} = 0$. $L(e_{t+h})$ needs to be positive if $f_{t+h} \neq y_{t+h}$ and therefore $e_{t+h} \neq 0$. Furthermore, the loss function is required to be monotonic non-decreasing for errors increasing in absolute value, $L(e_1) \geq L(e_2)$ for $|e_1| \geq |e_2| \geq 0$. Moreover, the conditional expectation of the loss function is assumed to exist. According to Elliott and Timmermann (2008), the latter property depends on the functional design of the loss function itself, as well as on the conditional distribution of the outcome variable.

2.1.2 Accuracy Measures

Having introduced the concept of the loss function, we will now discuss a variety of forecast accuracy measures which are useful for purposes of comparison, before further broadening the concept of the loss function in section 2.2. We will see that these accuracy measures each represent a specific loss function. A simple measure of forecast accuracy is the mean error ($ME = \frac{1}{T}\sum_{t=0}^{T-1} e_{t+h}$), which provides a first and often very useful impression of whether a series of forecasts systematically diverges from the target variable in one direction.

As pointed out by Hyndman and Koehler (2006), the choice of a reasonable measure depends on multiple factors. For one, it is important whether forecasts for the same or for different target variables are compared, as different variables generally require accuracy measures that are independent of the scale of the data. If the forecasts for a single target variable are compared for different forecasters or different time periods (e.g. decades or business cycles), the scaling is less important. Furthermore, the nature of the target variable matters for the choice of an appropriate accuracy measure, as some measures are infinite, or not defined, when either the target variable equals zero at any point, or when subsequent realizations of the target variable are identical.

Addressing these concerns, Hyndman and Koehler (2006) classify measures of forecast accuracy in the following four categories: scale-dependent measures, measures based on percentage errors, measures based on relative errors and relative measures. Commonly used representatives of the first category are the mean and median absolute error ($MAE = \frac{1}{T}\sum_{t=0}^{T-1} |e_{t+h}|$; $MdAE = \textit{median}\,(|e_{t+h}|)$), the mean square error ($MSE = \frac{1}{T}\sum_{t=0}^{T-1} e_{t+h}^2$) and the root mean square error ($RMSE = \sqrt{\frac{1}{T}\sum_{t=0}^{T-1} e_{t+h}^2}$). While these measures are well suited for comparing forecasts of the same target variable, i.e. produced by different forecasters, there is no point in using these measures when comparing the forecast accuracy of various variables that are scaled differently. Traditionally, the MSE and $RMSE$ are the most frequently used measures. According to Hyndman and Koehler (2006), the latter is often preferred because it is on the same scale as the data. Because of the squared errors, however, the influence of single large forecast errors is disproportionately high and hence the $(R)MSE$ exhibits a distinct sensitivity to outliers. Thus, Armstrong (2001) recommends using the more robust MAE and $MdAE$ instead. Nevertheless, the sensitivity of the squared error measures can be useful, as high values of the error statistics signalize the presence of large errors. Throughout this dissertation, we focus on the scale-dependent measures MAE and $RMSE$ for the preliminary descriptive analysis of the forecast errors (see chapters 3 and 5 in particular), as we do not explicitly compare the forecast accuracy of different variables.

The accuracy measures in the other categories all aim to solve the problem of scale-dependence. Herein, two of the four categories contain measures directly based on the forecast errors in the form of either percentage errors ($p_{t+h} = 100e_{t+h}/y_{t+h}$) or relative errors ($r_{t+h} = e_{t+h}/e^{\star}_{t+h}$), where $e^{\star}_{t+h}$ is the error of a benchmark forecast, which is often simply the naïve forecast (i.e. $f_{t+h} = y_t$). These percentage and relative errors (p_{t+h} and r_{t+h}) are then used to compute the scale-independent analogons for the measures in the first group. A drawback to these measures results from having to divide by either the target variable or an naïve forecast errors. If the denominator is zero or close to zero, the accuracy measure statistic is either undefined or very large. Hence, the percentage error-based measures are not suited if the target variable crosses or even draws near the zero line, while the relative error-based measures need subsequent observations to differ sufficiently.

Regarding the macroeconomic data considered in this dissertation, there is no single scale-independent measure which is suited to measure the forecast accuracy of variables such as GDP growth, inflation and the unemployment rate. While growth and inflation may equal zero at some points, the unemployment rate is likely to remain unchanged in two subsequent periods. An alternative measure, proposed by Hyndman and Koehler (2006), relates the actual out-of-sample error to the in-sample MAE of the naïve forecast. The authors argue that this measure has the advantages of being scale-independent and avoiding the dependency of single values for the target variable. It thus reduces the probability of undefined or infinite values considerably.

Regarding the relative measures, the fourth category discussed by Hyndman and Koehler (2006), one of the scale-dependent forecast measures is related to the same measure calculated with a benchmark forecast. The most prominent representative of the group is Theil's U statistic (Theil, 1966), where the $RMSE$ is divided by the $RMSE$ of the benchmark forecasts, which are the naïve forecasts.

Armstrong and Collopy (1992) provide an empirical analysis of percentage error-based and relative error-based accuracy measures. They evaluate forecast errors for nearly 200 economic time series and find the median relative absolute error ($MdRAE$) and the median absolute percentage error ($MdAPE$) to be the most favorable accuracy measures.

Probably the most comprehensive sources for applications of forecast accuracy measures are the forecasting competitions organized by a group of researchers working with Spyros Makridakis. These competitions, known as the M-Competitions, seek to evaluate and compare the accuracy of different forecasting methods. The results of three competitions have been published, all indicating simple models often outperform more sophisticated models that in practice (see Makridakis et al. (1982), Makridakis et al. (1993) and Makridakis and Hibon (2000)).

Apart from comparing the predictive ability using the measures introduced above, statistical tests for comparing two forecast models also exist. Diebold and Mariano (1995) propose a test statistic that has been widely applied to test the hypothesis of equal predictive accuracy of two series of forecasts. For the Diebold-Mariano (DM) statistic, the average loss differential (i.e. $\bar{d} = \frac{1}{T}\sum_{t=0}^{T-1} g(e_{1,t+h}) - g(e_{2,t+h})$, where $g(.)$ is a loss function) is related to a consistent estimate of its standard deviation. The authors show that this test statistic is asymptotically standard normal under the null hypothesis of equal predictive accuracy. West (1996) has further formalized this test and Giacomini and White (2006) have extended it in two ways. On the one hand, the authors consider forecasts based on limited memory estimators, which means that past observations are discarded at some point. On the other hand, they suggest conducting forecast evaluation conditional on a forecaster's past performances. In his recent study, Diebold (2015) points out that the Diebold-Mariano approach is a useful tool for comparing two series of forecasts, but advises against using it to analyze out-of-sample forecasts when comparing forecasting models.

Apart from the question of the forecast accuracy, forecast evaluators are generally interested in whether or not forecasts are unbiased. A simple examination of the mean error provides a first indication for a bias, as unbiased forecasts should have a ME equal to zero. Observing a bias in a series of forecasts implies that the forecaster who produced these forecasts is, for some reason, either incapable of adapting or unwilling to adapt the knowledge gained from past forecast errors to future forecasts. Although this may appear unreasonable at first glance, various studies suggest that it can be optimal for forecasters to strategically bias their forecasts (see e.g. Scharfstein and Stein (1990), Ehrbeck and Waldmann (1996) and Lim (2001)). Batchelor's (2007) study specifically focuses on the bias in macroeconomic forecasts and finds evidence for bias toward optimism in the GDP forecasts for four G7 countries (Japan, Italy, Germany and France). The author argues that there are only three reasons why a forecaster would produce biased forecasts. Along with the possibility of a strategic and rational bias, or the lack of skill, Batchelor argues that the forecaster may have insufficient data and thus may not be able to formulate adequate predictions concerning the developments of the target variable.

2.1.3 Unbiasedness and Rationality Tests

The unbiasedness of a series of forecasts can be tested by performing simple linear regressions. According to Mincer and Zarnowitz (1969), an optimal (or rational) forecast is called unbiased if the forecast error is zero on average. This can be checked by estimating the following linear regression

$$y_{t+h} = \beta_0 + \beta_1 f_{t+h} + u_{t+h} \tag{1}$$

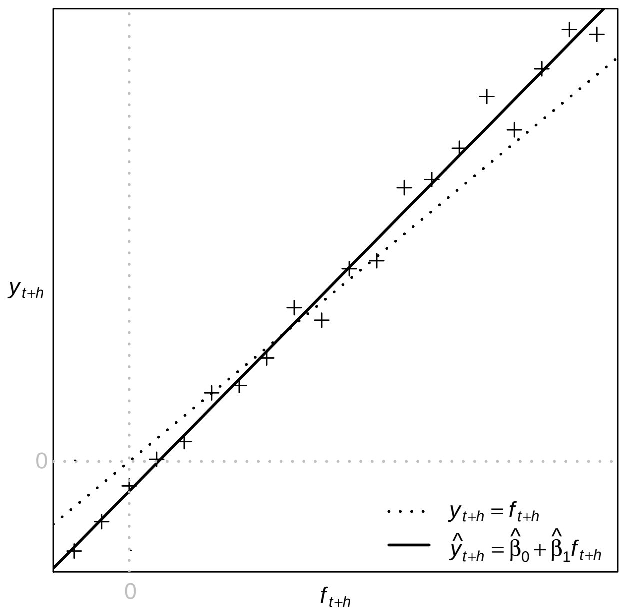
and testing the null hypothesis $H_0 : \beta_0 = 0, \beta_1 = 1$. The idea of this test is quite straightforward and can be illustrated nicely in a graph. Figure 2.1 shows a scatter plot with fictitious forecasts f_{t+h} on the x-axis plotted against the realizations of the target variable y_{t+h} on the y-axis, along with the regression line of equation (1) and the 45° line. If the null hypothesis in the Mincer-Zarnowitz regression cannot be rejected, the regression line is close to the 45° line, suggesting that there is no systematic difference between the forecasts and their realizations.

As an extension of the Mincer-Zarnowitz approach, the forecast efficiency can be tested by expanding equation (1) by adding variables $\boldsymbol{w}_t$ selected from the information set Ω_t. This leads to the regression

$$y_{t+h} = \beta_0 + \beta_1 f_{t+h} + \boldsymbol{\beta}_2' \boldsymbol{w}_t + u_{t+h} \tag{2}$$

where efficiency is tested as the null hypothesis $H_0 : \boldsymbol{\beta}_2 = \mathbf{0}$. If this null hypothesis is rejected, variables from the information set have explanatory power and hence contain information that is not adequately used by the forecasters and could be further exploited to reduce future forecast error. An alternative way to formulate the Mincer-Zarnowitz test in equations (1) and (2) is to enforce the restriction $\beta_1 = 1$ and to subtract the forecasts f_{t+h} on both sides of the equation. This results in $e_{t+h} = \beta_0 + u_{t+h}$ and $e_{t+h} = \beta_0 + \boldsymbol{\beta}_2' \boldsymbol{w}_t + u_{t+h}$ with the focus solely on the forecast errors e_{t+h}.

Figure 2.1: Mincer-Zarnowitz test for unbiasedness



Nordhaus (1987) distinguishes between concepts of weak efficiency and strong efficiency. He defines a forecast as being weakly efficient if there is no information left in the past forecasts that could be used to further improve the forecast. A forecast is strongly efficient if the forecaster uses all relevant information at his or her disposal. As Nordhaus argues, strong efficiency is "equivalent to the strong form of rational expectations" (p. 667) and is difficult to test in practice because the entire information set of the forecaster is hardly ever known. Nevertheless, he points out that "weak efficiency is a necessary condition for strong efficiency, [...] but it is clearly not sufficient," (p. 673). Given the lack of knowledge considering the forecaster's (full) information structure, it has become standard practice to test weak efficiency as well as subsets of the forecaster's information set $\boldsymbol{w}_t \subset \Omega_t$, as in the Mincer-Zarnowitz based regression above.

The Mincer-Zarnowitz test and several extensions based on their approach have become the standard procedure for evaluating the unbiasedness and efficiency of forecasts, as argued by Holden and Peel (1990), who refer to various studies on this approach. Nevertheless, the approach is limited by its rather restrictive underlying assumption of loss symmetry. This will be discussed in depth in the next section. Beforehand, we briefly provide an overview of the forecast rationality testing literature focusing on Mincer-Zarnowitz based methods.

2.1.4 Review of Empirical Applications

Considering the inflation expectations of individual forecasters in the Livingston data, Figlewski and Wachtel (1981) find evidence against forecast rationality. Testing several Mincer-Zarnowitz based models, they conclude that a single linear equation is rather insufficient for modeling the complex expectation formation process of the survey respondents.

Testing the unbiasedness and efficiency in survey forecasts of real GNP growth, inflation, unemployment and other macroeconomic variables, Zarnowitz (1985) mainly finds evidence for bias in inflation forecasts, while there is less evidence for bias in the other forecasts. He analyzes efficiency by testing for serial correlation in the forecast errors. He finds evidence for serial correlation in inflation as well as in unemployment forecasts.

Keane and Runkle (1990) analyze a panel of price forecasts of the individual forecasters in the survey of the American Statistical Association (ASA) and the National Bureau of Economic Research (NBER), the predecessor of the Survey of Professional Forecasters (SPF), the oldest quarterly survey of macroeconomic forecasts in the United States, and find these forecasts to be unbiased and efficient. This is a rather surprising result, as "almost the entire existing literature has rejected the rationality of price forecasts, even from professional forecasters," (Keane and Runkle, p. 730). In their study, Keane and Runkle also emphasize the importance of four factors when testing forecast rationality for a panel of forecasters. First, they argue in favor of using the individual instead of the aggregated forecasts, as, on the one hand, each individual forecast is based on a different information set and, on the other hand, biases of individual forecasters toward different directions may be concealed by the aggregation. Second, Keane and Runkle emphasize the necessity of paying attention to data revisions, as revised data can vary considerably from the real-time data in the forecaster's information set at the time when the forecast was made. Moreover, they point out that professional forecasters have stronger incentives for providing accurate forecasts than non-professional forecasters, as the former aim to sell their forecasts on the market. Last, they argue that the forecast errors may be influenced by effects that are common to all individual forecasters, such as shocks to the economy, and therefore the errors may not be independent across forecasters.

While Bonham and Cohen‘s (1995) response to Keane and Runkle does not argue with any of the factors that need to be considered when testing forecast rationality. They point out that Keane and Runkle do not consider the unit roots in some of their time series. Amending for this, Bonham and Cohen obtain results that lead to a rejection of the forecast efficiency in the same forecasts, which is more in line with other results.

Bonham and Cohen (2001) also analyze the SPF and focus on the conditions under which unbiasedness tests using pooled individual forecasts and consensus forecasts are consistent. They find that microhomogeneity (i.e. equal parameters $\beta_0 = 0$ and $\beta_1 = 1$ in each individual's efficiency regression) has to hold in the presence of unit roots and cointegration in forecasts and target variable to obtain consistent pooled estimates. If this is not the case, they recommend testing unbiasedness individually for each forecaster.

A nonparametric alternative to the Mincer and Zarnowitz regression is proposed by Campbell and Ghysels (1995), who analyze the unbiasedness and efficiency of US federal budget projections using sign and sign-rank tests that are robust in non-normal and heteroskedastic data situations. Employing parametric and non-parametric tests, they find evidence for bias and observe stronger evidence for inefficiencies in the forecasts when using non-parametric methods.

Looking at the GNP growth and CPI forecasts of the Blue Chip Economic Indicator, a survey of US economists, Swidler and Ketcher (1990) analyze how the forecasts improve after being revised. A key feature of these monthly forecasts is that forecasters repeatedly make a forecast for the same target variable at the same point in time, with a shorter forecast horizon each time. For inflation, they find that new information is incorporated well in the forecasts and hence forecasts become more accurate. Interestingly, they do not observe the same effect for GNP growth forecasts and argue that, unlike CPI, GNP realizations are revised frequently and therefore the changes in the forecaster's information set are more severe for this variable.

Batchelor and Dua (1991), who also study the Blue Chip forecasts, but focus more on the individual forecast, observe that inflation forecasts appear to be less rational than real GNP growth and interest rate forecasts. Moreover, they find that information contained in past forecasts made by the same individual forecaster for other variables is often not used well and thus is the main reason for inefficiency. Their finding that forecasts that are partly judgmental and not only based on an econometric model appear to be rational more often is another interesting result.

Romer and Romer (2000) analyze the Federal Reserve's inflation forecasts and compare these to the commercial forecasts of Blue Chip, Data Resources Inc. and the SPF. Their central finding considering inflation is that the Federal Reserve possesses better information than the commercial forecasters and hence is able to produce more efficient forecasts.

While the focus has been predominantly on forecasts concerning the US, in the studies mentioned so far, Öller and Barot (2000) are interested in the accuracy, unbiasedness and weak efficiency of the OECD's and national institutes' growth and inflation forecasts for European countries. They find inflation forecasts to be significantly better than GDP growth forecasts in terms of accuracy and find the OECD's and the national institutes' forecasts to be similarly accurate. Moreover, the authors observe only a few rejections of the hypotheses of unbiasedness and weak efficiency.

In a more recent paper, Ager et al. (2009) conduct bias and efficiency tests for the pooled forecasts of the Consensus Economics survey. They analyze GDP growth and price forecasts for 12 countries and find growth forecasts to be biased for Italy and Germany (although the German bias is not statistically significant) and inflation forecasts to be biased for Sweden and Switzerland. Testing for weak efficiency, they discover inefficiencies in GDP forecasts for all 12 countries. For the inflation forecasts, efficiency can be rejected in all countries except Belgium and the United Kingdom.

Dovern and Weisser (2011) analyze a three-dimensional panel of annual forecasts for GDP growth, the inflation rate and the growth rates of private consumption and industrial production for the G7 countries, which are also reported in the Consensus Economics survey. Their main findings are that forecasters differ considerably in their forecast accuracy and that the quality of the forecasts differs across countries and variables. For example, they find inflation forecasts to be rather unbiased compared to the other variables. In addition, biases are found to be correlated across individual forecasters. Moreover, weak efficiency is rejected frequently for all G7 countries except Japan.

2.2 Asymmetric Loss

2.2.1 Reasons for Loss Asymmetry

So far, all considerations relating to the measurement of forecast errors and the evaluation of the forecast rationality have been made under the underlying assumption that the forecaster is indifferent to over- and underestimating the target variable and thus has a symmetric loss function. Recalling two of the error measures introduced in the previous section, we can see that the MAE and the MSE are two specific examples of symmetric loss functions

$$L(e_{t+h}) = a \cdot |e_{t+h}| \tag{3}$$

$$L(e_{t+h}) = a \cdot e_{t+h}^2, \tag{4}$$

with $a > 0$. Here, the main difference is that larger forecast errors are weighted proportionately under linear loss (equation (3)), while larger errors are weighted disproportionately high under quadratic loss (equation (4)). Considering that the regressions based on Mincer-Zarnowitz are all estimated using least squares, we also identify the implicit quadratic loss function.

The first to argue that this assumption may not always be justified is Granger (1969), who states that the costs implied by a forecast error of a certain magnitude may be different for positive and negative forecast errors. To illustrate his point, he gives the following example: a bank that is planning to buy a new computer for a certain task can either spend too much money on a computer that is more powerful than needed for the purpose, or buy one that is not capable of fulfilling the task. In both cases, the costs that come along with the forecast error increase with the magnitude of the error. Nevertheless, there is no reason why the costs should be the same in both cases and therefore symmetric.

Another rather simple example is the different costs implied by arriving five minutes too early or five minutes too late at the gate of the airport to catch a flight. Here it is unlikely that the opportunity costs of waiting at the airport are the same as the costs of purchasing a new plane ticket and flying hours later. The same principle applies to the costs caused by macroeconomic forecast errors. However, there may be many consequences implied by the forecast errors that can be quantified as direct costs only with difficulty. Thus, the exact costs, or the economic loss of a certain error, are hard or nearly impossible to quantify in monetary units at all in many cases.

Regarding forecasts of GDP growth, for example, an overly optimistic forecast of a forecaster with high credibility may, on the one hand, serve as a self-fulfilling prophecy and further increase economic activity, while a pessimistic forecast may lead to the opposite behavior (see e.g. Morgenstern (1928)). Underpredicting GDP growth, on the other hand, can cause decision-makers to budget less income (e.g. taxes) and thus face needless constrains (see Krüger and Hoss, 2012).

Overpredicting growth could lead to excessive government spending, or overly optimistic budgeting at the very least. Similar arguments can be used for (un)employment forecasts for which an overly optimistic outlook on future employment may prompt employers to create more jobs, while an overly negative prediction may cause premature cutbacks.

Due to the inflation targeting pursued by many central banks (see e.g. Bernanke et al. (2001)), the costs of over- and underestimating inflation may even vary over time, following the general inflation regime at a certain period. In times in which inflation is generally higher than the central banks' inflation target it may be preferable to overestimate rather than to underestimate inflation since an underestimation means that the actual inflation is even further away from the target variable. In recent years, inflation has been rather low in the US and Europe and the aversion against high inflation is probably less pronounced than the fear of deflation. On the other hand, many macroeconomists fear the consequences of deflation more than the consequences of (mild) inflation. Hence, these different costs depend on the state of the economy. Capistrán (2008) argues that "[i]f inflation above the target is more costly for the central bank than inflation below the target [...] the fear of having inflation above the target will induce the central bank to maintain inflation below it," (p. 1425).

The common feature of all of the examples above is that even if the sign of the more costly forecast error is not immediately obvious, as this may vary over time and for different variables, there is no good reason why the assumption of symmetric cost or loss should be valid. Moreover, if a forecaster has strict preferences toward negative instead of positive forecast errors, it may even be rational for this forecaster to produce biased forecasts. This economical point of view is strengthened by the well-established psychological notion that humans tend to be incapable of evaluating the outcome of a future event independently of the probability of the event's outcome (see Weber, 1994). Thus, being averse to a certain outcome biases human perception of the probability of that outcome.

Another string of argumentation emphasizes the connection between forecasting and decision making. To use the words of Granger and Pesaran (2000, p. 537), "[i]n the real, non-academic world forecasts are made for a purpose and the relevant purpose in economics is to help decision-makers improve their decisions." These authors argue in favor of a closer connection between forecast evaluation and the decision problem that is related to the forecasts. As an example they discuss the Kuipers score that is used in meteorology and is based on the difference in the percentages of correctly forecast bad events and incorrectly forecast good events. The survey articles by Granger and Machina (2006) and Pesaran and Skouras (2002) discuss methods that allow for a decision-based forecast evaluation. These methods require the flexibility to evaluate the costs related to an over- or an underestimation of the target variable differently.

The concept of asymmetric preferences also is in line with the theory of herding and anti-herding behavior in forecasters. In this context anti-herding means that an individual forecaster induces an intentional bias away from the consensus forecast in addition to his or her usual objective of producing an accurate forecast. This could be rational for the forecaster in order to place his or her own forecast more prominently on the market or to establish a reputation as being an exceptionally optimistic or conservative forecaster. Herding behavior can be interpreted the other way around, as a forecaster that herds may fear losing his or her good reputation when a forecast that deviates from the consensus fails, whereas failing with the consensus might go unnoticed or at least attract less attention. Both herding as well as anti-herding can explain a forecaster's willingness to intentionally bias his or her forecast and the literature provides empirical evidence for both phenomena. While Trueman (1994) observes herding behavior for analysts, Pierdzioch and Stadtmann (2011) argue that anti-herding can be observed for yen-dollar and dollar-euro exchange rate forecasters. More recently, Clements (2015) has found evidence for anti-herding in the GDP growth and inflation forecasts submitted to the US Survey of Professional Forecasters.

2.2.2 Loss Functions and Optimal Forecasts

After establishing that the forecaster's or decision-maker's preferences for over- and underestimating a certain target variable can potentially be asymmetric, the methodical implications are quite substantial. As a consequence, all forecast measures and tests for unbiasedness and efficiency that have been introduced in the previous section, such as the Mincer-Zarnowitz test, need to be modified. This can be accomplished by using more flexible loss functions than those discussed in the previous section. According to Elliott and Timmermann (2008) the choice of a loss function mainly depends on the costs of over- and underestimation and whether these are symmetric or not. Generally, loss functions are specified with the forecast error as their only argument. However, Patton and Timmermann's (2007b) study argues in favor of loss functions that depend on the level of the target variable as well as the forecast. To underline their argument, these authors give the perspicuous example of meteorological forecasts, which may lead to rather symmetric cost while weather is normal, whereas cost related to a forecast error in extreme weather conditions like a flood or a hurricane are probably highly asymmetric.

Here, we focus on the first case, in which L only depends on the forecast error. As mentioned in section 2.1, a loss function has its unique minimum at zero, a function value of zero in this point and is non-decreasing when the forecast error departs from zero. In addition to these properties, there are three more properties that a loss function may possess, although these are not necessary (Granger, 1999). These are symmetry of the loss function ($L(e) = L(-e)$), homogeneity ($L(\alpha e) = g(\alpha) \cdot L(e)$, for some positive function $g(\alpha)$) and differentiability to order k (i.e. $L^{(m)}(e)$ exists for all $m \leq k$). Hence, a loss function is called symmetric if all positive and negative forecast errors of equal absolute magnitude imply the same function value. Otherwise it is called asymmetric.

Having fully established the concept of the loss function, an optimal forecast can be defined as the forecast $f^{\star}_{t+h}$ that minimizes the expected value of the loss function L, given the forecaster's information set Ω_t available at time t,

$$f^{\star}_{t+h} = \underset{f_{t+h}}{\operatorname{argmin}} \, \mathrm{E}(L(y_{t+h}, f_{t+h}) \,|\, \Omega_t), \tag{5}$$

(see e.g. Patton and Timmermann (2007a, 2007b)). Hence, with respect to the definition of forecast efficiency in section 2.1, we can now say that an optimal forecast has to be efficient, but not necessarily unbiased, if the underlying loss function is asymmetric. In the remainder of this dissertation, the expressions optimal, rational and efficient are often used interchangeably. Nevertheless, we keep in mind that optimality rather accentuates the value of the forecast given a certain loss function, while efficiency and rationality emphasize the forecaster's use of information.

As argued above, it is often difficult to justify the assumption of indifference in symmetric loss functions like the absolute loss or the squared loss function (equations (3) and (4)). Nevertheless, the quadratic loss function is without any doubt the most widespread approach. The simple explanation for this fact is that if forecasts are made using a least squares predictor, a quadratic loss function is automatically implied and the optimal forecast is the conditional mean of the target variable (see e.g. Granger and Newbold, 1986). Such forecasts are unbiased and are not autocorrelated, as shown by Diebold and Lopez (1996). However, Patton and Timmermann (2007a) show that these properties of optimal forecasts are specific to a quadratic loss function and do not hold in general. Hoque et al. (1988) and Magnus and Pesaran (1989, 1991) discuss violations of the standard properties of optimal forecasts caused by estimation error rather than by a choice of a loss function different from MSE.

A natural way to generalize the symmetric loss functions in equations (3) and (4) and to allow for asymmetry in the loss caused by positive and negative errors of the same magnitude can be found in Granger (1969). The piecewise linear (Lin-Lin) and piecewise quadratic (Quad-Quad) loss function permit the slope of the loss function to differ on both sides of the axis of ordinates,

$$L(e_{t+h}) = \begin{cases} a \cdot |e_{t+h}|, & e_{t+h} \geq 0, \\ b \cdot |e_{t+h}|, & e_{t+h} < 0, \end{cases} \tag{6}$$

$$L(e_{t+h}) = \begin{cases} a \cdot e_{t+h}^2, & e_{t+h} \geq 0, \\ b \cdot e_{t+h}^2, & e_{t+h} < 0, \end{cases} \tag{7}$$

where $a > 0$ and $b > 0$. For Lin-Lin loss it can be shown straightforwardly that the optimal forecast is a quantile of the cumulative distribution function G of the target variable y_{t+h} conditional on the information set Ω_t, i.e. $f_{t+h}^{\star} = G^{-1}\left(\frac{a}{a+b} \,|\, \Omega_t\right)$ (see e.g. Christoffersen and Diebold (1997) and Gneiting (2011)). If $a = b$ and the loss function is symmetric, the optimal forecast is the conditional median. Hence, forecasts are assumed to be made using quantile regression methods if the loss function is piecewise linear (Koenker and Bassett, 1978). For Quad-Quad loss, the optimal forecast can be computed using expectile methods (Newey and Powell, 1987), although a closed analytic solution for the optimal forecast only exists under very special conditions, as argued by Christoffersen and Diebold (1996).

The *Linex* loss function proposed by Varian (1975) is another frequently used asymmetric loss function. It has been used by Zellner (1986a), who introduced the reparametrization of the original function that is known today,

$$L(e_{t+h}; a, b) = b \cdot (\exp(a \cdot e_{t+h}) - a \cdot e_{t+h} - 1), \tag{8}$$

with parameters $a \in \mathbb{R} \setminus \{0\}$, $b > 0$, where b scales the function and a determines the degree of asymmetry. For positive a, the right side of the function is approximately exponential and the left side is approximately linear, while for negative a the opposite holds. In addition, Zellner states that for small absolute values of a, the Linex is similar to a symmetric squared error loss function,[3] whereas the function's asymmetry increases for a further away from zero. In contrast to the Lin-Lin and Quad-Quad loss functions for which the loss ratio of positive and negative errors of the same magnitude is constant, the loss ratio for the Linex loss function depends on the absolute value of the forecast error. Assuming normality of the target variable, Christoffersen and Diebold (1997) show that under Linex loss the optimal h-step-ahead forecast is, $f^{\star}_{t+h} = \mu_{t+h} + \frac{a}{2}\sigma^2_{t+h}$, where μ_{t+h} is the expected value of y_{t+h} and σ^2_{t+h} is the expected error variance, and both are conditional on the information set Ω_t. Hence, the bias of the optimal forecast is time-varying under this setting, as it depends on the changes in the forecast error's variance over time as well as a.

Keeping these examples for symmetric as well as asymmetric loss functions in mind, it is easy to construct further loss functions. According to Granger (1999), given two loss functions $L_1(e)$ and $L_2(e)$, the following functions are loss function as well:

- $\varphi_1(e) = a \cdot L_1(e) + b \cdot L_2(e)$, $a \geq 0$, $b \geq 0$
- $\varphi_2(e) = [L_1(e)]^a \cdot [L_2(e)]^b$, $a > 0$, $b > 0$
- $\varphi_3(e) = I(e < 0) \cdot L_1(e) + [1 - I(e < 0)] \cdot L_2(e)$
- $\varphi_4(e) = \psi(L_1(e)) - \psi(0)$,

where in φ_3 the indicator function $I(.)$ is equal to unity for negative forecast errors and equal to zero elsewhere, and in φ_4 there is a positive monotonic, non-decreasing function ψ defined over $\mathbb{R} \setminus \{0\}$ and with $|\psi(0)| < \infty$. As an example, Granger suggests the Double Linex loss function, $D(e_{t+h}) = L(e_{t+h};\, a_1) + L(e_{t+h};\, -a_2) = \exp(a_1 \cdot e_{t+h}) + \exp(-a_2 \cdot e_{t+h}) - (a_1 - a_2) \cdot e_{t+h} - 2$, as a composition of two Linex loss functions (see equation (8)), where b is set to unity and $a_1 > 0$ and $a_2 > 0$. The Double Linex loss function is exponential on both sides and is differentiable at any order. Furthermore, while the Linex loss function does not nest a symmetric special case, the Double Linex loss function is symmetric for $a_1 = a_2$, but asymmetric for $a_1 \neq a_2$.

2.2.3 Review of Empirical Studies

Before introducing the forecast evaluation approach that has been suggested by Elliott, Komunjer and Timmermann (2005) in the next section, the remainder of this section provides a brief overview of the applied as well as the theoretical literature centered around asymmetric loss. Herein, the focus lies on the work preceding EKT. Some of the studies that have extended EKT's approach are presented at the end of the next section.

Zellner (1986b) shows that it is rational to produce biased forecasts if preferences are asymmetric and he employs a Linex loss function to demonstrate this. Batchelor and Peel (1998) find that when a forecaster has an underlying asymmetric loss function, conventional rationality tests, like

[3] Zellner uses the truncated power series representation of the exponential function $exp(ae) \approx 1 + ae + \frac{1}{2}a^2e^2$ for his argumentation and implies that higher powers of a are small for a close to zero.

one of Mincer and Zarnowitz' (1969), may lead to false conclusions about the rationality of the forecasts. They propose an augmented version of the Mincer-Zarnowitz test that is useful for forecasts made under Linex loss if the time-varying bias is controlled for as an ARCH-in-Mean effect. Ruge-Murcia (2003) also employs a Linex loss function to study the preferences of the central banks in Canada, Sweden and the UK regarding deviations from the inflation target. He finds these preferences to be asymmetric and better reflected by a Linex loss function than by a quadratic loss function. Nobay and Peel (2003) come to the same conclusion regarding the preferences of central banks in general and also show that these preferences lead to an "inflation premium" that is driven by the variance of inflation.

Analyzing the government deficit forecasts provided by the IMF, OECD and EC for the G7 countries between 1975 and 1995 under the assumption of a Quad-Quad as well as Linex loss function, Artis and Marcellino (2001) do not find general affirmation for asymmetric preferences of the three institutions. Weiss (1996) focuses on Quad-Quad loss functions and allows for heteroskedasticity in the target variable. Herein, he considers approximations to the optimal forecast in situations for which the conditional distribution of y_{t+h} is unknown. Ulu (2007) predicts returns for three exchange rates and five stock market indices both under symmetric quadratic and under Lin-Lin loss and finds the latter superior to the former if agents have asymmetric preferences.

Focusing on how several forecasts should be optimally combined when the different underlying loss functions are potentially asymmetric, Elliott and Timmermann (2004) argue that both the loss functions and the forecast error distributions have to be considered. Using the Lin-Lin, Quad-Quad and Linex as examples, they show how the optimal combination weights differ from those under MSE loss.

Patton and Timmermann (2007b) argue that loss may depend on the levels of the target variable y_{t+h} and the forecast f_{t+h} instead of their mere difference, the forecast error, and suggest a flexible loss function based on linear and quadratic splines. They apply their proposed test to the Fed's quarterly Greenbook forecasts of GDP growth between 1968 and 1999 and find that forecast optimality is rejected when the loss function only depends on the forecast error. They find this not to be the case when both target variable and forecasts are accounted for separately.

The mean absolute percentage error measure ($MAPE$), discussed in the previous section, treats realizations of the target variable above and below a given prediction asymmetrically. McKenzie (2011) analyzes the $MAPE$ and finds that under this loss function, the optimal forecast is lower than the mean when the realizations of the target variable are positive.

Ulu (2013) tests for the joint rationality of US inflation and output and obtains different results when using a GARCH-M model versus univariate Lin-Lin and Linex loss functions, as there is stronger evidence for rationality when using the multivariate approach. Nevertheless, she assumes the loss of both variables to be additively separable for her multivariate approach.

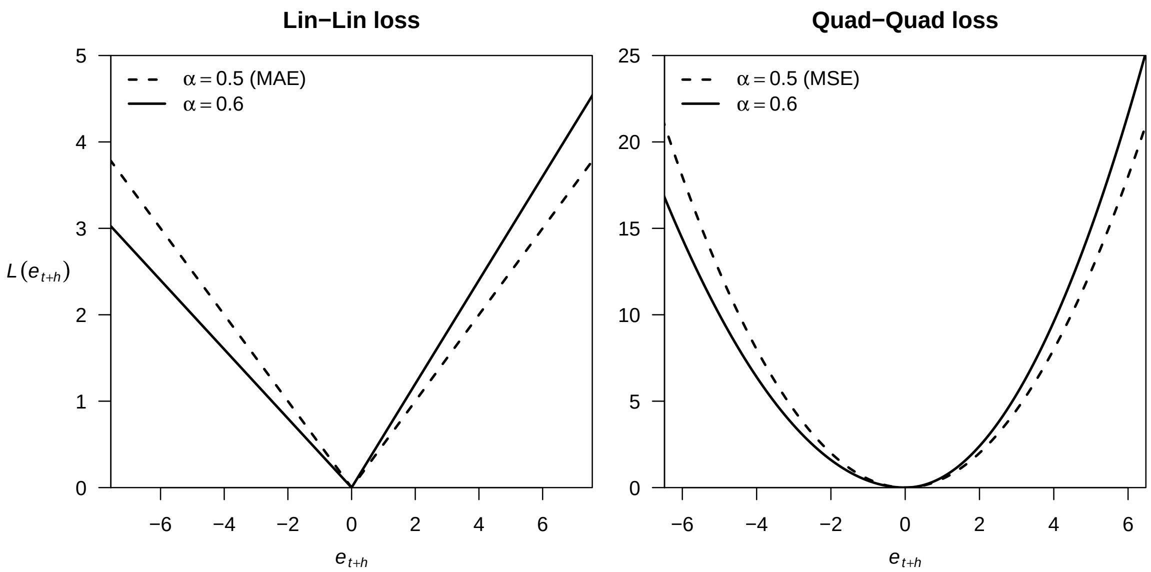
2.3 The EKT approach

This section introduces the approach proposed by Elliott, Komunjer and Timmermann (2005, 2008). It centers on the idea of gaining insights about the forecaster's preferences concerning the asymmetry of the underlying loss function from a given series of forecasts errors. In contrast to the rationality tests of unbiasedness and efficiency based on Mincer-Zarnowitz (1969) and centered on the assumption of symmetric and quadratic loss, the EKT approach is much more general. Therefore, EKT propose the following family of loss functions

$$L(e_{t+h}; \alpha, p) = [\alpha + (1 - 2\alpha) I(e_{t+h} < 0)] \cdot |e_{t+h}|^p \tag{9}$$

which, apart from e_{t+h}, depends on the parameters α and p. Herein, $I(.)$ is the indicator function equal to unity where the condition $e_{t+h} < 0$ holds and equal to zero elsewhere. The parameter $\alpha \in (0, 1)$ represents the degree of asymmetry, while the parameter $p > 0$ determines the curvature of the loss function. This family of loss functions also nests the symmetric case as a special case for $\alpha = 0.5$. Fixing the curvature parameter at $p = 1$ and $p = 2$, the loss function corresponds to MAE and MSE, in the symmetric case, respectively. Allowing for $\alpha \neq 0.5$, the loss function turns into a piecewise linear (Lin-Lin) or a piecewise quadratic (Quad-Quad) function, respectively. In the latter case, a value of $\alpha > 0.5$ indicates that underpredicting the target variable (a positive forecast error) is weighted more heavily than an overprediction (a negative forecast error), whereas for $\alpha < 0.5$ the opposite holds. Figure 2.2 shows the Lin-Lin as well as the Quad-Quad loss function with an asymmetry parameter of $\alpha = 0.6$, along with their symmetric counterparts. The choice of the parameter $\alpha = 0.6$ illustrates the higher weights of positive forecast errors.

Figure 2.2: Loss functions



Note that even small deviations of α from the value of 0.5, which is associated with symmetric loss, imply rather large loss differences. Take, for example $\alpha = 0.45$, which implies a loss ratio of positive and negative forecast errors of $\alpha/(1-\alpha) = 0.45/0.55 \approx 0.82$ meaning a loss difference of about 18 percent. Even this large loss difference would hardly be detected by testing the null hypothesis that $\alpha = 0.5$ on usual significance levels. Therefore, results in the following sections have to be regarded under the aspect of their statistical significance, as well as their economical significance. Hence, if the α estimates fail to differ from the symmetric case on a conventional significance level, even a small deviation from 0.5 might nevertheless be an indication of considerable asymmetry in the forecaster's preferences. Elliott and Timmermann (2008) raise another related issue concerning the estimates of α by stating that an estimate close to zero or unity, e.g. $\alpha = 0.1$, indicates an unrealistically high loss ratio. They argue that if such a high loss ratio is observed in a series of forecast errors, the bias in the forecasts may be driven by other reasons than the forecaster's asymmetric preferences, e.g. irrationality.

The fundamental idea in EKT's approach is to estimate α while either fixing p, or estimating both parameters jointly, using the following first order condition

$$\mathrm{E}(L'(e_{t+h};\alpha,p)\,|\,\Omega_t) = 0, \tag{10}$$

where $L'(.)$ is the derivative of the loss function with respect to e_{t+h} and Ω_t again denotes the information set of time t. Optimality is achieved if the forecast errors fulfill condition (10) above. Under optimality, there is no further information in Ω_t to reduce the forecast errors. To test for optimality, we use the orthogonality condition

$$\mathrm{E}(\boldsymbol{w}_t \cdot [\alpha - I(e_{t+h} < 0)] \cdot \left|e_{t+h}\right|^{p-1}) = \mathbf{0}, \tag{11}$$

where $\boldsymbol{w}_t$ is a $d \times 1$-dimensional subset of instrumental variables from the information set Ω_t that are required to be available when the forecast is established. The test is performed by applying GMM estimation (Hansen, 1982) of the parameter α (while holding p fixed), or the joint estimation of both parameters.[4]

Apart from the information gained on the shape parameters, EKT's procedure offers the possibility of assessing whether the forecaster used the available information efficiently. This can be illustrated nicely by deriving the moment conditions needed for the GMM estimation. An efficient use of information implies that there is no information left in the set of instrumental variables at hand, $\boldsymbol{w}_t$, and hence $H_0 : \boldsymbol{\beta} = \mathbf{0}$ in $L'(e_{t+h};\alpha,p) = \boldsymbol{\beta}'\boldsymbol{w}_t + u_{t+h}$ cannot be rejected. Therefore, the moment conditions $\mathrm{E}(\boldsymbol{w_t} \cdot u_{t+h}) = 0$ can be written as $\mathrm{E}(\boldsymbol{w_t} \cdot (L'(e_{t+h};\alpha,p) - \boldsymbol{\beta}'\boldsymbol{w}_t)) = 0$ and reduce to equation (11) under the null of rationality. In this sense, EKT exactly reverse the traditional procedure of assuming a certain loss function and then testing the rationality of the forecasts. In their approach, estimating the shape parameter α (and p) leads to the form of the loss function that is most compatible with the implicit assumption that the forecasts are rational and the available information has been used efficiently.

[4] In most applications, p is fixed at $p = 1$ or $p = 2$ and the estimation of p is the exception (see Krüger and Hoss (2012) for such an exception using the Nelder–Mead method for numerical optimization). The reason for this may be the frequent convergence problems in the numerical optimization when p is estimated. For the computation, we rely on the R package "gmm" described in Chaussé (2010).

To justify the assumption of information efficiency, the optimal value of the target function of the GMM estimation should have a function value of zero. This can be tested by performing the well-known J-test for overidentifying restrictions. To this end, EKT (2005) suggest the following statistic

$$J = T\left\{\frac{1}{T}\sum_{t=0}^{T-1}\boldsymbol{w}_t[\hat{\alpha} - I(e_{t+h} < 0)]\left|e_{t+h}\right|^{p-1}\right\}' \times \hat{\boldsymbol{S}}^{\mathbf{-1}} \\ \times\left\{\frac{1}{T}\sum_{t=0}^{T-1}\boldsymbol{w}_t[\hat{\alpha} - I(e_{t+h} < 0)]\left|e_{t+h}\right|^{p-1}\right\}, \qquad (12)$$

where, for p fixed, the test is asymptotically distributed as a χ^2_{d-1} random variable for d moment conditions and a consistent estimator $\hat{\boldsymbol{S}}$ of $\boldsymbol{S} = \mathrm{E}(\boldsymbol{w}_t\boldsymbol{w}_t'[I(e_{t+h} < 0) - \alpha]^2\left|e_{t+h}\right|^{2p-2})$ as well as a linear instrumental variable (IV) estimator $\hat{\alpha}$ of α:

$$\hat{\alpha} = \frac{\left[\sum_{t=0}^{T-1}\boldsymbol{w}_t\left|e_{t+h}\right|^{p-1}\right]'\hat{\boldsymbol{S}}^{\mathbf{-1}}\left[\sum_{t=0}^{T-1}\boldsymbol{w}_t\left|e_{t+h}\right|^{p-1}I(e_{t+h} < 0)\right]}{\left[\sum_{t=0}^{T-1}\boldsymbol{w}_t\left|e_{t+h}\right|^{p-1}\right]'\hat{\boldsymbol{S}}^{\mathbf{-1}}\left[\sum_{t=0}^{T-1}\boldsymbol{w}_t\left|e_{t+h}\right|^{p-1}\right]}. \qquad (13)$$

Accordingly, the entire procedure is a joint estimation of the shape parameters of the loss function with a test of forecast optimality. More precisely, the procedure is designed to estimate values of α and p that are consistent with forecast optimality, given the instrumental variables $\boldsymbol{w}_t$. If such values of α and p cannot be found because the estimates differ considerably for some of the moment conditions, then the J-test rejects the null hypothesis of forecast optimality. EKT (2008) point out that the J-test is a consistent test even if the loss function depends on further unknown parameters. If there is only one instrumental variable, i.e. $d = 1$, and overidentification therefore cannot be tested, a unique closed form solution for $\hat{\alpha}$ exists (see equation (13) with scalar $\boldsymbol{w}_t$). According to EKT (2005), the estimate can then be interpreted as the estimate that shapes the loss function in such a way that bias in the forecasts is justified or rationalized.

2.4 Applications of EKT

Since EKT (2005, 2008) themselves have applied their test to data on IMF and OECD budget deficit forecasts for the G7 countries and to data on individual forecasts from the Survey of Professional Forecasters on US growth and inflation, their approach has been adopted by various other authors. EKT find evidence for asymmetric loss in both studies and also observe a considerable decline in the number of rationality rejections once asymmetry is allowed for. In this section, we review other applications of their approach. Herein, we start with studies that evaluate business cycle forecasts, before turning to applications of the approach that are concerned with other variables, such as energy prices as well as financial and fiscal variables.

Like most forecast evaluation exercises, applications of the EKT approach mainly focus on business cycle forecasts (output growth and inflation). Capistrán (2008) and Capistrán and Timmermann (2009) are two early studies that use the EKT approach to analyze US inflation forecasts. The former study finds structural breaks in the bias of the Federal Reserve's inflation forecasts dating back to 1968 that coincide with the appointments of different Fed chairmen. While in the era before Volcker, inflation was systematically underpredicted, there was no observable bias during the time he was in office and a tendency to overpredict inflation afterwards. Hence, Capistrán (2008) estimates the Fed's loss function before and since Volcker and finds it to have changed drastically. Capistrán and Timmermann (2009), on the other hand, find that individual forecasters of the US Survey of Professional Forecasters vary considerably in their preferences towards asymmetry. Nevertheless, they argue that asymmetry in the loss functions alone is unable to explain the change from underpredicting inflation to overpredicting it observed in many of the individual forecasters around 1982. Christodoulakis and Mamatzakis (2008) analyze forecasts of the EU Commission for 12 member states between 1969 and 2004 and find the asymmetry preferences for GDP growth forecasts to be different among the member states.

Business cycle forecast evaluations using the EKT approach for the German economy are Döpke et al. (2010) and Krüger and Hoss (2012). Döpke et al. (2010) analyze business cycle forecasts of 17 forecasting institutes and hardly find any evidence for the rejection of rationality of growth forecasts and merely can reject the rationality of inflation rate forecasts for some of the institutes. Similar results are reported in Krüger and Hoss (2012) who, in addition, find certain financial variables to be promising for reducing future forecast errors in the inflation forecasts. In both articles, business cycle forecasts appear to be approximately symmetric, while the underlying loss function for inflation forecasts is found to be asymmetric.

In more recent years, Pierdzioch et al. (2012, 2015) have analyzed the underlying loss function of inflation as well as growth forecasts for the Bank of Canada, the central banks of Argentina, Brazil, Chile, and Mexico, respectively using the EKT approach. Their general findings are that these central banks tend to have asymmetric preferences considering the two variables at hand and that there is less evidence for irrationality of the forecasts when a flexible loss function is used. Furthermore, they find an increasing degree of of asymmetry for larger forecast horizons. Wang and Lee (2014) analyze Greenbook and SPF forecasts on US growth and inflation, and confirm previous findings in general, but focus additionally on a rolling window strategy for detecting changes in the asymmetry parameters over time.

There is also an increasing number of studies focusing chiefly on other forecasts such as those for energy prices and financial and fiscal variables. Auffhammer (2007) analyzes forecasts for oil, coal and electricity prices but also the natural gas consumption, the electricity sales, GDP and the energy intensity of the United States Energy Information Administration. He predominantly finds evidence for these forecasts having been made under an asymmetric loss function and argues that a lack of knowledge of this loss function may lead forecast users to make suboptimal decisions. Pierdzioch et al. (2013) find that the oil price forecasts in the ECB's Survey of Professional Forecasters on the one hand appear to be rather asymmetric, as an overestimation seems to cause a larger loss than an underestimation. On the other hand, they note that the forecasts are not necessarily rational even if the loss function is allowed to be asymmetric.

Aretz et al. (2011) combine the approaches of EKT and Patton and Timmermann (2007b) and find S&P 500 return expectations from the Livingston Surveys to be rational in many cases, once they allow for heterogeneity in the individual forecasters' asymmetry preferences. Fritsche et al. (2014) look at monthly euro-dollar exchange rate forecasts on a microeconomic level. They use survey data from Consensus Forecasts Inc. and find asymmetry preferences to vary over the individuals. Mamatzakis (2014) finds that West Texas Intermediate's oil 1-month futures tend to be optimistic for the years from 1995 to 2012 and observes that this is even more pronounced if only the years after the Lehman Brothers' bankruptcy are taken into account.

Christodoulakis and Mamatzakis (2009) conclude that government balance forecasts tend to be optimistic for most of the EU-12 states between 1970 and 2004 and argue that this is rather unsurprising, as it allows governments more flexibility in their budget planning. On the basis of Californian data, Krol (2013) finds that the common underprediction of revenue in budget forecasts can be explained by the lower costs of an underprediction compared to an overprediction of the target variable, and hence argues in favor of an asymmetric loss function.

Generalizing the approach of EKT, Komunjer and Owyang (2012) argue that one forecaster's forecast errors of two or more different variables, e.g. output and inflation, often cannot be regarded independently of each other. Hence, they propose a multivariate framework in which a vector of forecast errors is analyzed jointly. Applying their approach to monthly forecast data from the Blue Chip Economic Indicators between 1976:08 and 2004:12, they find a reduced degree of asymmetry when jointly evaluating output, inflation and short-term interest rate forecasts.

Further applications of Komunjer and Owyang's multivariate approach are Caunedo et al. (2013) and Krüger (2014). The former jointly test the rationality of the Federal Reserve's Greenbook forecasts of inflation, unemployment, and output growth from 1966 to 2005 and find a considerable degree of asymmetry in output and unemployment forecasts and less asymmetry in inflation forecasts. Krüger (2014) analyzes German output growth and inflation forecasts of the German Council of Economic Experts and finds a moderate degree of asymmetry in these forecasts.

3 Information Efficiency in German Employment Forecasts

3.1 Motivation

This chapter focuses on information efficiency in the employment growth forecasts published by two of the most important institutions for macroeconomic forecasts in Germany, the Council of Economic Experts and the Joint Forecasts of the leading economic research institutes. First, the efficiency is tested under the assumption of symmetric and quadratic loss, using the test based on Mincer-Zarnowitz presented in section 2.1. Second, as the assumption of an incorrect loss function may lead to false inferences considering forecast efficiency, we use the EKT approach that has been discussed in section 2.3 to test the information efficiency. This approach also allows us to estimate the shape of the forecasters' underlying loss function and hence make inferences about their preferences regarding an underestimation versus an overestimation of the target variable.[5]

In the forecast evaluation literature, the primary focus of interest lies on output growth along with inflation forecasts rather than employment forecasts. As news reports of GDP growth expectations or their upward or downward revisions appear almost every day in the media, output forecasts dominate public awareness. In spite of their evidently high relevance on a macroeconomic level as much as for each individual person of working age, one reason for the rather low public and academic interest in employment growth forecasts could be the close connection between the business cycle and the labor market. This may lead to the conclusion that separately analyzing employment growth forecasts is not important and rather redundant.

However, the economic laws that describe this relationship appear to be unstable over time and across countries. In this context Klinger and Weber (2014) analyze Verdoorn's law (1949), which states that output growth induces productivity growth and can be transformed into a linear relationship between output growth and employment growth, as first argued by Kaldor (1966). For Germany, Klinger and Weber find that the so-called Verdoorn coefficient has changed over time due to phenomena such as labor hoarding and jobless recovery and find "effects of a further autonomous component of employment growth" (p. 25) besides the one induced by GDP growth. In a more recent study, Ball et al. (2015) examine Okun's law (1962), i.e. the negative correlation between output and the unemployment rate and confirm Okun's law in general in its version that considers output growth and changes in the unemployment rate. While their main interest is to explore whether forecasters incorporate Okun in their forecasts, Ball et al. find a variation of the Okun coefficients across the G7 countries plus Australia and New Zealand for the years 1989-2012. Although the coefficients for all countries show the expected negative sign for the realizations as well as the forecasts, none of the German coefficients differs significantly from zero, indicating at best a weak presence of Okun's law. In their comment on Ball et al.'s article, Guisinger and Sinclair (2015) point out that, when using real-time data, only four of the nine countries show Okun coefficients that are negative and statistically significant. Thus, Guisinger and Sinclair argue in favor of a weaker existence of Okun's law in real-time. All of these findings

[5] This chapter is based on Hoss (2014).

point out the importance of analyzing employment forecasts independently of the predominant business cycle forecasts.

In the next section, we introduce the employment forecast and the real-time and revised realizations of the target variable and analyze the resulting forecast errors descriptively. In sections 3.3 and 3.4 the instrument sets under consideration and the respective time series they contain are presented. The efficiency tests with respect to the information contained in each instrument set are discussed in section 3.5, with each subsection concentrating on another set and the last subsection summarizing the results. In sections 3.6 and 3.7, the information in the individual variables is combined, using factor methods. In the latter section, a subset of the variables is pre-selected before the factors are extracted. Section 3.8 summarizes the results.

3.2 Employment Forecasts and Realizations

The data for the employment forecasts have been assembled from the German Council of Economic Experts' (CEE) annual reports.[6] These forecasts for the following year have been published each November since 1969. The years in which the forecasts are published and the subsequent year to which they refer are labeled as period t and period $t+1$, respectively. Likewise, the ifo Institute publishes the Joint Forecast (JF) from the forecasts of the leading German economic research institutes each spring and autumn.[7] To allow for a better comparison to the CEE forecasts, only the autumn forecasts are used here. The forecasts of both institutions are basically judgmental forecasts. In CEE's case a group of five experts makes the forecasts. This group is supported by a staff of assistants who also prepare the forecasts using statistical methods. The Joint Forecast, on the other hand, which is based on the individual institutes' forecasts, uses an iterative-analytic method that iteratively makes use of forecasts for several subareas and combines them consistently. An exception to the rather nontransparent process of forecast production is Heilemann (2002), who argues in favor of more transparency and discusses a the RWI-business cycle model in detail.

Throughout the analysis, forecast errors are considered with respect to both revised and real-time realizations. Revised realizations of employment have been taken from the German Central Bank's time series database.[8] Real-time realizations have been taken from the CEE's reports for the year following the forecast. Henceforth, the corresponding forecast errors simply will be called revised forecast errors and real-time forecast errors.

It is striking that there is a tendency towards a rather systematic upward revision of the data for over three decades of the observed time period. This upward revision is clearly depicted in figure 3.1. This underlines the importance of analyzing the results for both variants of realizations, as they are expected to differ considerably. Especially if one is interested in the potential asymmetry in the forecasting institutions' loss functions, it is necessary to look at the real-time as well as

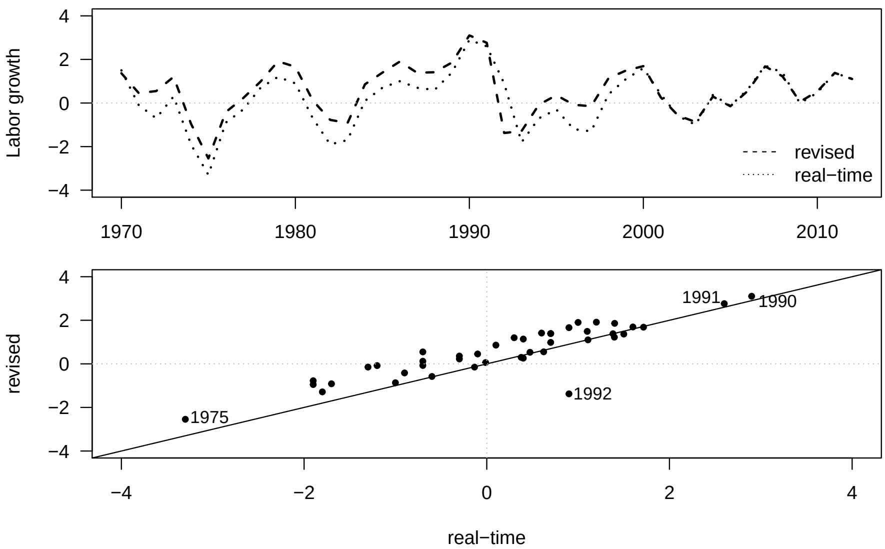
[6] See http://www.sachverstaendigenrat-wirtschaft.de/gutachten.html?&L=1.

[7] See http://www.cesifo-group.de/ifoHome/facts/Forecasts/Gemeinschaftsdiagnose/Archiv.html.

[8] The time series used are those for the German labor force BBK01.JJ5007 (before reunification) and BBK01.JJ5009 (after reunification). Both series are expressed in yearly averages per million people according to the domestic concept. See http://www.bundesbank.de/Navigation/EN/Statistics/Time_series_-databases/time_series _databases.html.

the revised realizations, as it is not clear which of the two is targeted by the forecasters. Another argument in favor of the analysis of both variants of realizations is the higher variance of the real-time data, which is persistent over the entire observed period. However, the standard deviation of the real-time realizations exceeds that of the revised realizations more distinctly before 1990 (1.43 versus 1.29) than after (1.12 versus 1.04).

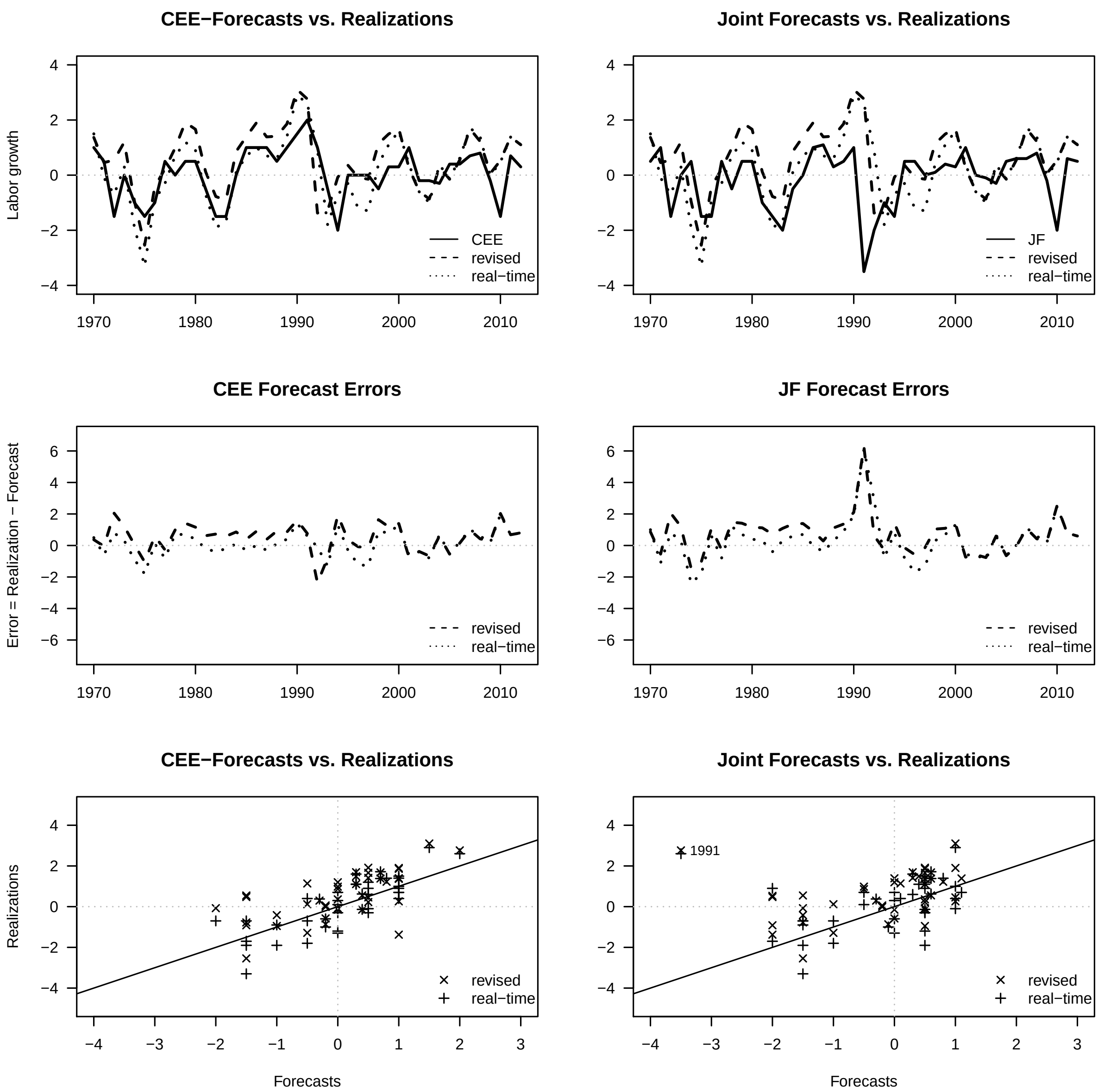
Figure 3.1: Realizations of labor force growth



Unfortunately, many of the time series used as instrumental variables in section 3.3 are unavailable in real-time for Germany over the entire analyzed period. Consequently, in what follows the real-time as well as the revised data have been analyzed only for the realizations of employment growth, while the latest realizations available have been used for all variables employed as instrumental variables.

Figure 3.2 reports the CEE's forecasts and realizations in the left column and the JF's corresponding forecasts in the right column. In the first row, we find the time series plots of the forecast (the solid line) along with the realizations (shown by the dashed line for revised and the dotted line for real-time). The forecast errors subject to the two variants of realizations are quite similar, as shown in the second row. Nevertheless, the revised forecast errors between 1970 and 1990 lay constantly above the real-time errors, which indicates an upward revision in these years. No such general tendency in data revisions can be detected for the decade between 1990 and 2000, although years with upward revisions also dominate this period. From 2000 on, hardly any difference between the two types of realizations exists, which might be an indication of more accurate real-time data. It is also possible that the most recent data have not yet fully been revised. These visual impressions considering the tendency in the revision process are enforced by the median and mean of the two variants of realizations. While the real-time realizations' mean and median are 0.02 and 0.30 whereas the revised analogues are 0.70 and 0.98 before 1990, both types of realizations differ less after 1990 with a mean of 0.32 for real-time and 0.45 for revised realizations and median 0.40 and 0.33, respectively. The final diagrams in each column compare forecasts and realizations in a scatter-plot in which pluses (+) indicate real-time and crosses (×) denote revised realizations. While the majority of the points is scattered around the 45° line, one severe underprediction appears in the JF forecast in 1991. All diagrams depicted in figure 3.2 indicate a general tendency to underestimate employment growth, i.e. forecast errors tend to be positive.

Figure 3.2: Forecasts of labor force growth and forecast errors



This possible bias in the forecasts is analyzed in more detail in section 3.5, while table 3.1 provides a more detailed descriptive overview on the forecasts errors of both institutions. Along with the mean forecast errors (ME), the table contains the mean absolute errors (MAE) and root mean squared errors ($RMSE$) that where presented in section 2.1, as well as an error measure based on the asymmetric Linex loss function (equation (8)). Here, the Linex loss is measured as $Linex = \frac{1}{T}\sum_{t=t_1}^{t_2} b\cdot(\exp(a\cdot e_{t+1}) - a\cdot e_{t+1} - 1)$, where $T = t_2 - t_1$. Following the parametrization of Christoffersen and Diebold (1997), we set the parameters $a = -1$ and $b = 1$. Negative forecast errors (overprediction of the target variable) are weighted approximately exponentially, whereas positive forecast errors (underprediction of the target variable) are weighted approximately linearly. This setting is denoted as *Linex 1*. To show the effect of exponentially weighted positive and linearly weighted negative forecast errors, the analogous *Linex 2* loss has been calculated as well, using the parameter $a = 1$. As long as the forecast errors take moderate absolute values, both Linex function values are alike. Only if the forecasters considerably overpredict or underpredict employment growth will the respective function values of the *Linex 1* and *Linex 2* be large.

Table 3.1: Forecast error measures

		Real-Time Forecast Errors					Revised Forecast Errors				
		ME	*MAE*	*RMSE*	*Linex 1*	*Linex 2*	*ME*	*MAE*	*RMSE*	*Linex 1*	*Linex 2*
CEE	1971-1980	-0.110	0.710	0.832	0.522	0.287	0.608	0.880	1.070	0.415	0.989
	1981-1990	0.060	0.340	0.506	0.096	0.195	0.785	0.785	0.851	0.264	0.547
	1991-2000	0.070	0.910	1.005	0.576	0.564	0.388	1.066	1.294	1.114	1.136
	2001-2010	0.234	0.701	0.853	0.277	0.661	0.211	0.678	0.846	0.269	0.624
	1971-2007	-0.006	0.645	0.781	0.363	0.327	0.465	0.848	1.017	0.523	0.743
	1971-2012	0.096	0.669	0.816	0.361	0.412	0.510	0.847	1.020	0.501	0.802
JF	1971-1980	-0.210	1.010	1.184	1.341	0.536	0.508	1.170	1.262	0.677	1.260
	1981-1990	0.420	0.580	0.766	0.201	0.517	1.145	1.145	1.231	0.496	1.349
	1991-2000	0.750	1.670	2.332	1.311	45.714	1.068	1.299	2.152	0.782	52.315
	2001-2010	0.234	0.781	1.007	0.361	1.062	0.211	0.768	1.004	0.350	1.103
	1971-2007	0.235	1.004	1.451	0.822	12.694	0.706	1.097	1.478	0.577	14.894
	1971-2012	0.317	0.995	1.426	0.774	11.403	0.731	1.076	1.451	0.558	13.354

Note: *Linex 1* refers to exponential weights for negative forecast errors and *Linex 2* to positive forecast errors.

With the exception of real-time forecast errors in the 1970s, the error measures in the upper panel of table 3.1 reveal a positive mean error (meaning an underprediction of employment growth) for the CEE forecasts for all decades. The accuracy measures MAE and $RMSE$ are largest from 1991 to 2000, reflecting the effects of German reunification, and both versions of the Linex loss function show slightly higher losses during this decade with no considerable difference between the two specifications of the Linex. This overall assessment is rather similar for revised and real-time forecast errors, except for the mean errors that are larger for revised realizations in most decades.

In the case of the JF forecasts reported in the lower panel of the table, positive mean errors can be observed with one exception. Compared to the CEE forecasts, mean absolute and root mean squared errors show an even stronger increase during the decade after German reunification. Thus, the CEE seems to have been better at predicting the labor market turbulences following reunification. Especially in this decade, the Linex loss function clearly differs depending on which sign of the forecast errors is weighted exponentially. When positive forecast errors are weighted exponentially, the loss function's value increases eminently, penalizing the JF's underpredictions after unification. Again, we get a similar picture from both revised and real-time forecast errors.

Regarding the last two rows of each panel, which depict the entire time period since 1971 both up to and including the recent crisis, the CEE appears to produce more accurate employment forecasts than the JF. Although all mean errors increase when the post-crisis years are included, indicating a stronger tendency towards underprediction, the accuracy measures do not differ systematically.

The appendix shows the corresponding figures as well as a table with the forecast error measures discussed above for the GDP growth forecasts of the same institutions. This puts the findings above into perspective since GDP growth forecasts have been examined more frequently in the literature. With the exception of the 1980s, the results here show negative mean errors, revealing a tendency to overpredict output growth. The largest mean absolute errors and root mean squared errors can be observed for the 1970s (the decade in which the oil price shocks of 1973/74 and 1979/80 occurred) and during 2001-10 (the decade that witnessed the breakdown of Lehman brothers and the subsequent recession). The Linex loss underlines the extent of overprediction during these periods. Overall, the figures reveal strong similarities in both institutes' predictions and forecast errors.

3.3 Instrument Sets

This section presents the sets of instrumental variables that are used to test how the forecasting institutions under consideration here incorporate the information available in these instruments into their forecasts and whether there is information left that could be used to improve the forecasts. Below we present the instrument sets and their structure before giving a detailed introduction of the data employed in each instrument set in section 3.4 and testing the information efficiency in section 3.5.

Along with the weak efficiency, i.e. the information contained in the lagged forecast error, seven further instrument sets are analyzed:

A	weak efficiency	E	price indices
B	labor market	F	financial variables
C	aggregate demand	G	foreign trade
D	leading indicators	H	business climate

Apart of set A, which is obviously shorter, each instrument set contains three instrumental variables that are used to form the following twelve instrumental subsets. For subsets 1-3, the single lags of the instruments are used. The next three subsets are formed using the first and second lags of the variables, and subsets 7-9 combine the first lags with their squares. All possible combinations of interaction terms formed by two of the instrumental variables and their associated first lags are used at the end of each instrumental set. All of the twelve instrumental subsets contain an intercept. Hence, the typical structure of any of the instrument sets B to H, given three instrumental variables w_1, w_2 and w_3, is as follows:

1	$\boldsymbol{w}_t = (1,\, w_{1,t})'$	7	$\boldsymbol{w}_t = (1,\, w_{1,t},\, w_{1,t}^2)'$
2	$\boldsymbol{w}_t = (1,\, w_{2,t})'$	8	$\boldsymbol{w}_t = (1,\, w_{2,t},\, w_{2,t}^2)'$
3	$\boldsymbol{w}_t = (1,\, w_{3,t})'$	9	$\boldsymbol{w}_t = (1,\, w_{3,t},\, w_{3,t}^2)'$
4	$\boldsymbol{w}_t = (1,\, w_{1,t},\, w_{1,t-1})'$	10	$\boldsymbol{w}_t = (1,\, w_{1,t},\, w_{2,t},\, w_{1,t} \cdot w_{2,t})'$
5	$\boldsymbol{w}_t = (1,\, w_{2,t},\, w_{2,t-1})'$	11	$\boldsymbol{w}_t = (1,\, w_{1,t},\, w_{3,t},\, w_{1,t} \cdot w_{3,t})'$
6	$\boldsymbol{w}_t = (1,\, w_{3,t},\, w_{3,t-1})'$	12	$\boldsymbol{w}_t = (1,\, w_{2,t},\, w_{3,t},\, w_{2,t} \cdot w_{3,t})'$

To cover the most recent information that should be at a forecasting institutions' disposal while working on their forecasts, and keeping in mind that the JF's fall forecast is published in October and CEE's in November of each year, the following definition of lagged data is used throughout the remainder of chapter 3. The first lag of a variable always refers to the most recent information available and therefore covers the first six months of a year, or the growth measured from the first half of the preceding year to the first half of a year if growth rates are used. Second lags are defined in the exact same manner using the second half of the preceding year. For growth rates, this means that the second lag for year t is defined as the growth measured from the second half of year $t-2$ to the second half of year $t-1$. For some variables, the quarterly or monthly data needed to construct the half-year data are unavailable for the time period considered. In these cases, the standard first two lags of the annual data are used.

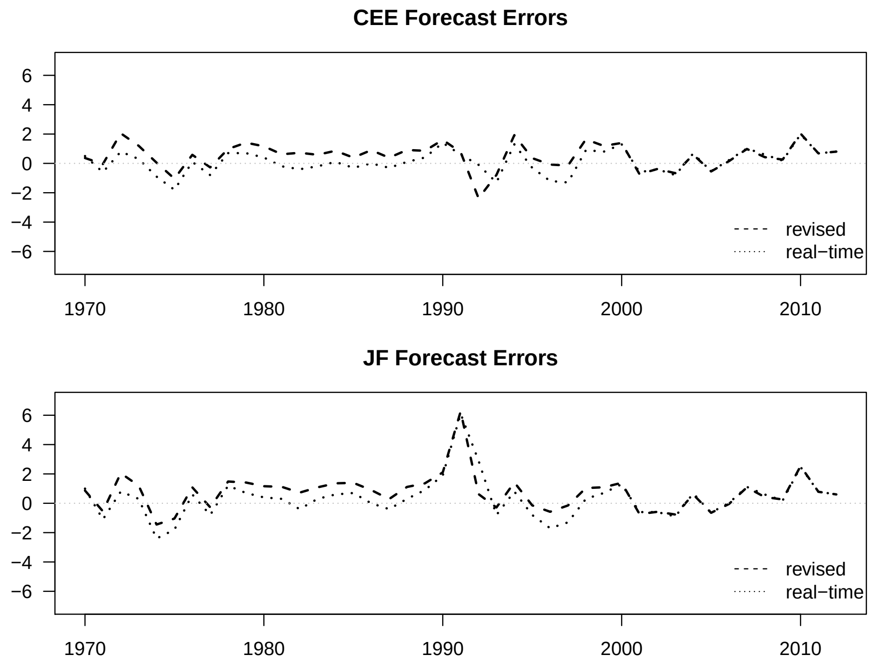
Besides the institutions' reports, the main data sources are the German Central Bank's time series database and some additional data directly requested from the German Central Bank. The oil price data has been taken from the UNCTAD database and the ifo business climate series have been received from the ifo institute. Raw data have been transformed into differences of the natural logarithm in most cases; exceptions will be discussed later on. Stationarity has been tested using the unit root test proposed by Elliot, Rothenberg and Stock (1996), and we have found a rejection of the unit root null for all growth rate series except the nominal labor cost growth. This has led us use second differences of this variable in the subsequent analysis in order allow for an interpretation of this variable as the acceleration of nominal labor cost growth.

3.4 Data

The following section presents all eight instrumental sets in detail. Our discussion of the data begins with set A, which tests for weak efficiency. Due to the lack of other variables apart of the forecast errors, set A contains less subsets than the other sets. Hence, it only consists of the lagged forecast errors, the forecast errors' first and second lags and the forecast errors' first lags and their squares. Moreover, the structure of the lags differs from the half-year definition introduced above. This is inevitable because only annual forecasts are analyzed here and therefore realizations and the resulting forecast errors must be annual as well. In spite of these differences with respect to the instrumental variables, which will be discussed later on, it is still important to analyze how well the institutions used the information they might have gained from their past errors.

Thus the forecast errors of the CEE are depicted in the upper panel of figure 3.3, while the lower panel shows the JF forecast errors. The errors with respect to either realization generally do not appear to differ a lot. However, we observe a tendency toward a systematic upward revision for most of the first three decades, as discussed in section 3.2. Besides this observation, the forecast errors seem to evolve in a rather stable manner over time, with a period of higher volatility between 1989 and 2000.

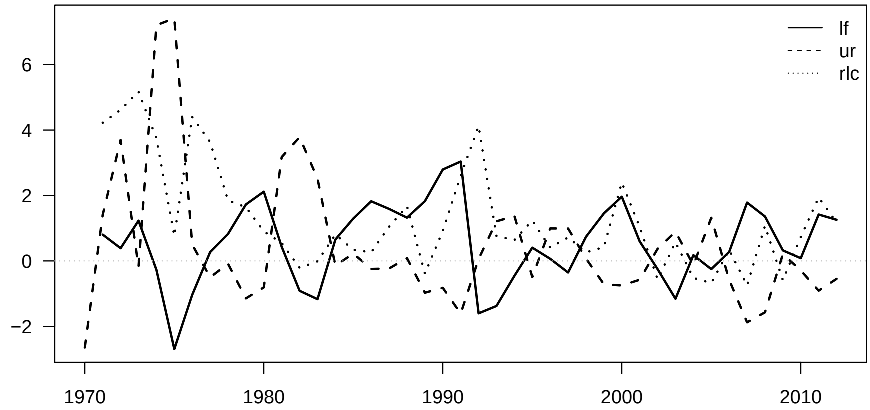
Figure 3.3: Weak efficiency (A)



Instrument set B is probably the most self-evident set in this section, as it contains three labor market variables in order to test whether forecasting institutions make adequate use of the information they contain. Specifically, these three variables are the German labor force (lf), the unemployment rate (ur) and the real labor costs per employee (rlc).[9] Figure 3.4 presents the growth rates of the three series. For graphical convenience, the growth rate of the unemployment rate has been divided by 10. One reason for the remarkable peaks in the growth of the unemployment rate might be the low unemployment rate up until the early eighties. Hence, absolute changes of a certain amount have had a stronger impact in this period compared to later decades. All three series seem to have been influenced by German reunification, as they show strong movement in the early 1990s. Nevertheless, a reduction in volatility can be observed for more recent years.

[9] With the exception of the GDP deflator, all data have been taken from the German Central Bank's time series database. The GDP deflator was used to calculate the real labor costs and was thus additionally requested from the German Central Bank's statistical department. The following series are used: JB5007 (employment in western Germany prior to 1992 in millions, quarterly data, seasonally adjusted), UABA14 (German employment starting 1991 in thousands, quarterly data, seasonally adjusted), US02CC (unemployment rate as a percentage of the total civilian labor force in western Germany prior to 1993, monthly data, seasonally adjusted), USCC02 (unemployment rate as a percentage of the total civilian labor force in Germany starting 1992, monthly data, seasonally adjusted), JB5008 (nominal labor costs per employee, index series with 2005 = 100) and the GDP deflator in the form of index numbers with 2005 = 100. For all series the semiannual means are calculated and the resulting data are converted to growth rates by differences of logarithms and lagged as illustrated above.

Figure 3.4: Labor market (B)

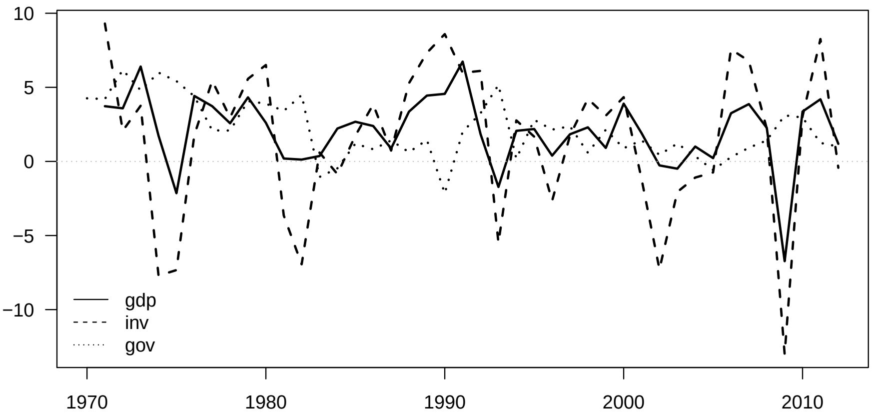


Set C tests whether the forecasters have considered information on aggregate demand. Therefore, the following three components of the national accounts are analyzed. The first instrument is the real gross domestic product (*gdp*), as it is used the most for expressing movement in business cycles, and according to Ball et al. (2013), largely coincides with the employment, or even leads it. As a second instrument, real investment (*inv*) is examined. Though highly correlated with GDP, it shows a higher cyclical volatility that GDP. The last variable in set C is government consumption (*gov*). It covers governmental behavior, such as stimulus programs, that does not necessarily follow the business cycle, but is still capable of influencing employment growth (see Ramey (2012)). This counter-cyclical characterization is indicated by the low correlation between government consumption and GDP as well as investments.[10] See table 3.2 on page 40 for the correlation coefficients. Private consumption is not used as an instrument here even though it accounts for the largest share of the GDP. In contrast to real investments, it is less volatile than GDP, but unlike government consumption, is positively correlated with GDP with a correlation coefficient of 0.56.

Figure 3.5 depicts the data for the instruments in growth rates. On the one hand we have a nice representation of how real investment growth tends to move in accordance with GDP growth, but in more extreme amplitudes, while on the other hand, growth in government consumption tends to behave counter-cyclically; see, for example, the years 1974 and 1975, 1990, 1993 and, more recently, 2009.

[10] The following time series have been taken from the German Central Bank's time series database: JB5000 (gross domestic product, quarterly data, seasonally and price adjusted), JB5004 (gross fixed capital formation, quarterly data, seasonally and price adjusted) and JJ5003 (government consumption, annual data, price adjusted). While semiannual means have been calculated from the original data for GDP and investment, annual data has been used for government consumption, as national budgets are determined per fiscal year. All series are organized in the database in the form of index numbers, with 2005 = 100, and have been converted to growth rates by differences of logarithms and lagged as previously described.

Figure 3.5: Aggregate demand (C)



Set D captures the information included in certain leading indicators using the following instruments. The first instrument set is the index series of industrial orders received (ord), followed by the total number of building permissions (bp) along with the term spread (rs).[11] The term spread has been calculated as the difference in the monthly average yields on outstanding debt securities issued by residents with a mean residual maturity of more than 9 and up to 10 years and the monthly averages of the money market rates for three-month funds reported by Frankfurt banks until 1999 and by EURIBOR starting in 1999. Because the data available for the series on long-term interest rates only dates back to 1973, the term spread series starts accordingly.

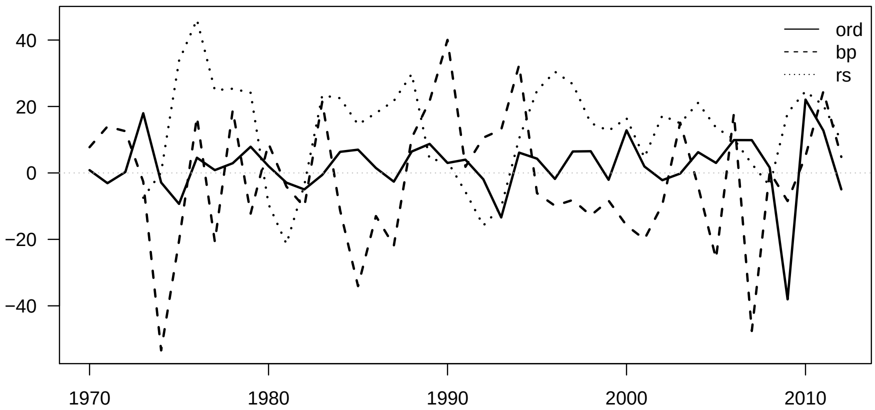
Following the argumentation of Jones and Tuzel (2013), the common use of new orders as a leading indicator for macroeconomic activity is justified, as new orders "measure investment at the time that the purchase decision is made [...] rather than the time that the goods are delivered or installed" (p. 116). In his recent paper, Strauss (2013) argues that an increase and a decline in building permits are driving employment growth in the United States, while Álvarez and Cabrero (2010) argue that building permissions are leading the business cycle in Germany and other European countries. As Estrella and Mishkin (1998), Estrella and Trubin (2006) and Adrian and Estrella (2008) argue, a flattening yield curve indicates reduced real output growth in the short term, as a tightening monetary policy leads to rising short-term interest rates, while its influence on long-term rates tends to be minor.

Figure 3.6 shows the growth rates of the first two instruments along with the level of the term spread. For graphical convenience, the term spread series was multiplied by 10. While the growth rates of industrial orders and the term spread stay above zero for most of the observed

[11] One of the time series used in this set has been taken from the German Central Bank's time series database, i.e. BBDE1.M.DE.Y.AEA1.A2P300000.F.C.I10.L (industrial orders received in the form of index numbers with 2010 = 100, quarterly data, seasonally and price adjusted). The data on monthly building permissions was requested from and provided by the German Central Bank's statistical department. The term spread is constructed as the difference of long-term and short-term interest rates, both monthly data and presented below in the description of instrument set F. While semiannual means have been calculated for the first and last series in the exact same manner as for the other sets, the semiannual sums have been calculated for the building permissions, as the series contains the absolute number of building permissions. The industrial orders and the building permissions have been converted to growth rates by differences of their logarithms, while the term spread has been left without further transformation. All three series have been lagged accordingly.

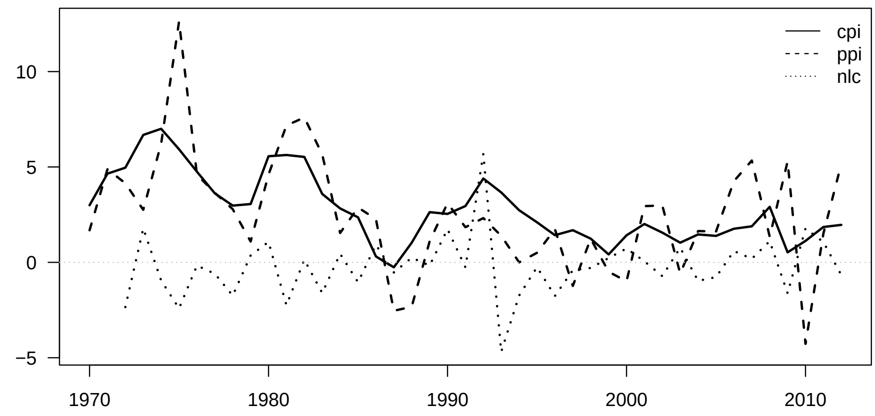
year and only take negative values occasionally, the growth rate series of building permissions shows a high volatility during the observed time period and changes between positive and negative growth frequently.

Figure 3.6: Leading indicators (D)



Instrument set E focuses on the development of several price indices over the years. Specifically, these indices are the consumer price index (*cpi*), which covers the individual costs of living, the producer price index for industrial products (*ppi*), which covers the costs of production and the nominal labor costs employee (*nlc*), which covers the costs of labor. All three may affect a firm's or an individual's decisions to offer or seek employment and hence indicate whether the labor force grows or shrinks.[12]

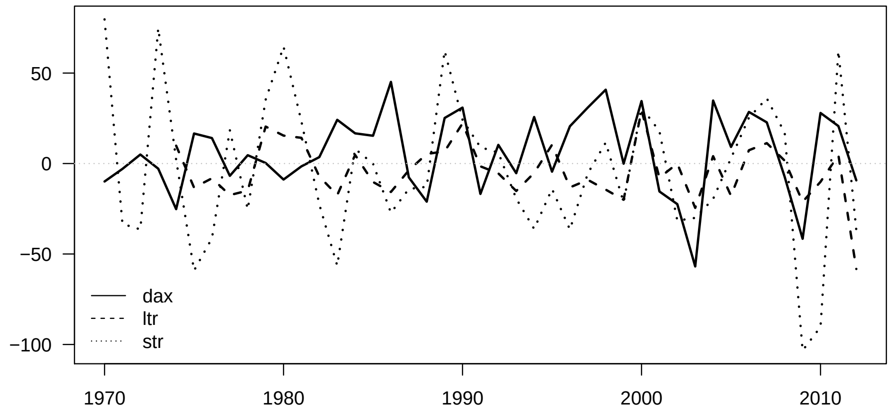
Figure 3.7: Prices (E)



[12] The time series used in instrument set E have been taken from the German Central Bank's time series database: BBDP1.M.DE.Y.VPI.C.A00000.I10.L (CPI, seasonally adjusted, monthly data), BBDP1.A.DE.N.EPG.G.GP09SA000000.I10.L (producer prices for industrial product, annual data) and JB5008 (nominal labor costs per employee). All three time series are organized in the database in the form of index numbers with 2010 = 100 in the case of CPI and PPI and 2005 = 100 for nominal labor costs. As monthly data on PPI before 1976 are not available, annual data have been used here in order to avoid the loss of observations. For the monthly data, i.e. CPI and nominal labor costs, semiannual means have been calculated as described before and all series have been converted to growth rates by differences of logarithms, or acceleration rates by second differences in the case of nominal labor costs, and lagged as described earlier.

In figure 3.7, the growth rates of CPI and the PPI are depicted along with the acceleration of the nominal labor costs. For CPI and PPI, growth rates seem to decline gradually over the course of the years. With the exception of 2008, CPI constantly conforms to the inflation goal of a maximum inflation at 2 percent from the mid-1990s. PPI shows a higher volatility for the same period. The acceleration of nominal labor costs hovers around the zero line with a tendency to stay below zero until 1990, indicating a reduction of growth dynamics. It shows two peaks in opposite directions in 1992 and 1993, right after German reunification, before returning to a path around the zero line.

Figure 3.8: Financial variables (F)



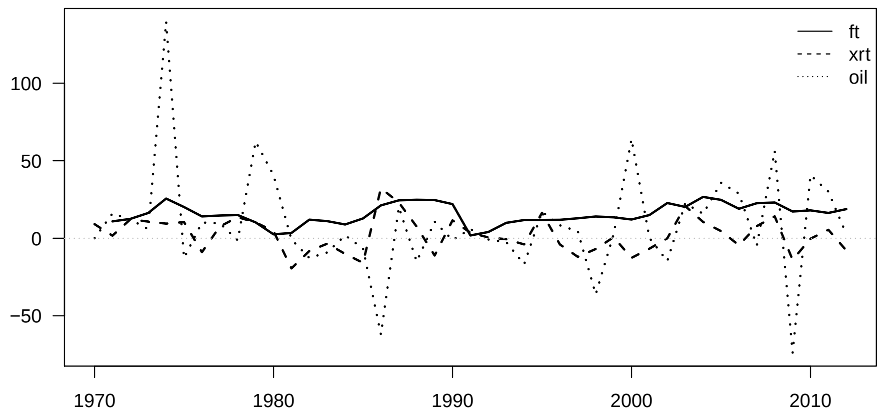
Instrument set F considers the information contained in financial variables. As argued by Sinai (1992) and, more recently, by Ng and Wright (2013) financial variables are of importance for business cycle forecasts, as they can be the origin of recessions and economic crisis. Sinai states: "The overwhelming verdict of history is that both real and financial factors matter for growth, employment, inflation, interest rates and other financial market prices, financial flows, debt, and sectoral balance sheet behavior. Both can provide a useful framework for analysis of what actually goes on in the real and financial world" (p. 47). Estrella and Mishkin (1998) point out that stock prices contain different information than the interest spreads, which are analyzed in set D, as stock prices reflect future expectations on dividend streams and are therefore strongly connected to future economic activity. Moreover, they argue that these two variables (stock prices and interest spreads) provide financial indicators that are simple and the most useful. Hence the three variables used in instrument set F to represent financial information are the DAX performance index (*dax*) along with the two components of the interest spread, *ltr* and *str*, discussed separately in set D.[13] Figure 3.8 shows the growth rates of all series. While the short-term interest rate is more volatile from the 1970s until the 1990s and after 2008, the other two series exhibit a roughly constant volatility during the observed period. As mentioned above,

[13] All time series, used for the financial variable set F, have been taken from the German Central Bank's time series database: WU3141 (DAX performance index, 1987 = 1000, end of month), WU8608 (yields on debt securities outstanding issued by residents with a mean residual maturity of more than 9 and up to 10 years, monthly average), SU0107 and SU0316 (three-month fund money market rate reported respectively by Frankfurt banks and EURIBOR, monthly averages). All series in the database have been converted to growth rates by differences of logarithms and lagged as described above.

the series on long-term interest only dates back to 1973. Thus, after transforming into differences of the natural logarithm the first observation is in 1974.

Instrument set G focuses on variables that represent the German foreign trade activity in order to account for information related to global events such as oil price shocks and measures Germany's international competitiveness. Here, the three variables selected are the foreign trade balance (ft), calculated as the natural logarithm of the ratio of exports and imports, the euro-US dollar exchange rate (xrt) and the index of the world market price for crude petroleum (oil).[14] While it is self-evident that goods produced for exportation tend to increase employment in the producing country, imported goods can be understood as a form of outsourcing production and thus employment. The trade balance is used to account for the interaction of these two effects. Concerning the two other instrumental variables in this set, studies by Klein et al. (2003) and Moser et al. (2010) find an effect of exchange rate shocks on employment for the United States and Germany, respectively. Jiménez-Rodríguez and Sánchez (2005), among others, find rising oil prices to negatively impact economic activity in all oil-importing OECD countries except Japan. However, Lutz and Meyer (2009) argue that this impact has diminished over the last decades due to the reduced energy intensity of the major economies.

Figure 3.9: Foreign trade (G)



In figure 3.9 the foreign trade balance is plotted along with the growth rates of the other two time series. The behavior of the three time series differs considerably. On the one hand, the solid line representing foreign trade is strictly positive and rather stable over time with the exception of a

[14] The time series of foreign trade and the euro-dollar exchange rate have been taken from the German Central Bank's database and can be found there under the following denominations: EU2001 (total exports in million euros, monthly data), EU3001 (total imports in million euros, monthly data), WU5009 (exchange rates on Frankfurt exchange, USD 1 = DM ..., monthly data), BBEX3.M.USD.EUR.BB.AC.A02 (Euro reference exchange rate of the ECB / EUR 1 = USD ... monthly data). The oil price date has been taken from the UNCTAD data base and can be found under the following url: http://unctadstat.unctad.org. It presents the monthly world market price for crude petroleum as an equally weighted average of UK Brent (light), Dubai (medium) and Texas (heavy) (USD/barrel), and is available in the database in the form of index numbers with 2000 = 100. To obtain a single time series for exchange rates, the Frankfurt exchange rates have been converted to Euro using the irrevocable Euro conversion rate (EUR 1 = DM 1.95583). While the foreign trade balance has been calculated as the natural logarithm of the ratio of exports and imports, exchange rates and oil prices have been converted to growth rates by differences of logarithms and all series have been lagged as described above.

smaller decline around 1980 and a stronger one in the early 1990s, after German reunification. The growth rates, on the other hand, show a higher volatility in the time period observed. The growth rate of the oil price, in particular, has some extreme peaks and valleys. Whereas the highest values can be observed in 1974, 1979, 2000 and 2008, the lowest values occur in 1986, 1998 and 2009.

The final instrument set H contains the ifo business climate index and its components, the business situation and expectation indices.[15] According to Henzel and Rast (2013) the ifo business climate index is based on approximately 7000 monthly survey responses from firms in manufacturing, construction, wholesaling and retailing in Germany. The firms are asked to state their current business situation (*bs*) as "good," "satisfactory," or "poor" and their business expectations (*be*) for the next six month as "more favorable," "unchanged," or "more unfavorable." After weighting the responses according to the industries' importance and aggregating the data, the business climate (*bc*) is calculated as the geometric mean of its two components. The main advantage of survey data like this is its early availability compared to official macroeconomic data, see e.g. Nardo (2003). The growth rates of all three indices are presented in figure 3.10. The figure shows a high volatility in the 1970s, followed by a comparatively calm phase, which, with the exception of the years right after German reunification, lasts until the mid-2000s and ends with the recent crisis, as the corresponding peaks and valleys in the business climate indicate.

Figure 3.10: ifo business climate (H)

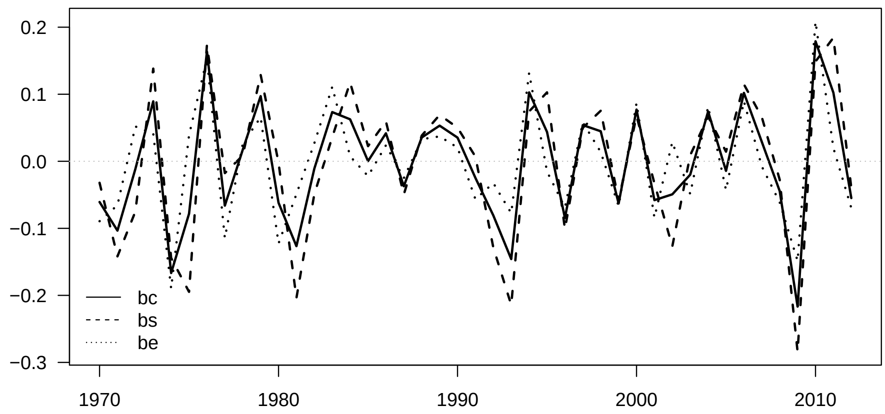


Table 3.2 presents the correlation coefficients between both institutions' forecasts, the real-time and revised realization data and all instrumental variables introduced above. Focusing on the top left corner of the table, the expected positive correlation between forecasts and realizations can be found, although the coefficients are positive but below 0.4 for the JF forecasts. Although the correlation between real-time and revised realizations takes an expected high value of 0.89, it still indicates substantial differences between both types of realizations in some of the years. Following

[15] The time series used in this set have been taken from the ifo database for all years from 1991 to present. It can be found under the following url: http://www.cesifo-group.de/ifoHome/facts/Time-series-and-Diagrams/Zeitreihen/Reihen-Geschaeftsklima-Deutschland.html. All data before 1991 have been requested directly at the ifo institute. After having converted the monthly data to semiannual averages, growth rates have been calculated by differences of logarithms and the series have been suitably lagged.

the first four rows of the table through the columns containing the instrumental variables, we can observe a positive correlation between labor force growth, GDP growth, real investment growth, short-term interest rate growth, and the acceleration of nominal labor costs across both forecasting institutions and types of realizations. Nevertheless, the JF forecasts and GDP growth do not appear to be correlated. Both institutions and both types of realizations are negatively correlated with the growth rates of the unemployment rate, the government consumption, the consumer price index and the producer price index.

Turning to the correlation coefficients between the individual instrumental variables, we find an unsurprisingly high negative correlation of -0.64 between labor force and unemployment rate growth. Furthermore, labor force growth shares the highest positive correlation coefficients with the GDP, real investment and short-term interest growth rates. Accordingly, the growth of the unemployment rate negatively correlates with these three variables, the acceleration of nominal labor cost and the growth of the ifo business sentiment index. Moreover, the growth of the unemployment rate is positively correlated with the growth rates of government consumption as well as CPI and PPI. Here, the correlation between unemployment growth and government consumption might be interpreted as an indication for a counter-cyclical fiscal policy. The last variable in instrument set B, the growth of real labor costs per employee, positively correlates with GDP, government consumption and CPI growth rates.

Rows 8 to 10 show the correlation coefficients of instrument set C (aggregate demand) for all instrumental variables apart from those in set B. Apart from the unsurprisingly high correlation between GDP and real investment growth, it is interesting to note that neither of these two variables seems to be correlated with the growth of government consumption. According to these results, the correlation structure of GDP and investment growth is very similar, showing positive correlations with the growth rates of industrial orders, long-term and short-term interest rates as well as the acceleration of nominal labor costs and the growth of the ifo business climate and sentiment indices. The growth of government consumption positively correlates with CPI and PPI growth and negatively correlates with the growth rates of the DAX performance index, the ifo business climate and sentiment indices.

Continuing with the variables in instrument set D (leading indicators), the growth of industrial orders shows a correlation structure similar to GDP and investment growth, although the negative correlation with PPI growth and the positive correlations with all three ifo business indices are slightly higher here. The correlation coefficients with the CPI and the short-term interest growth rates are negative for the term spread, and there are positive correlations with all three ifo business indices and the foreign trade balance. However, there is no correlation between any other instruments with a coefficient absolute higher than 0.25. The same holds for the growth of building permissions with maximal correlation coefficients of 0.36 with real investment and of 0.4 with the ifo business expectation index.

Apart from the correlations of the instruments in set E (price indices), which already have been mentioned above, there is only one more positive correlation coefficient above 0.5 for variables in this set, which is the correlation between CPI and PPI growth (0.71). Furthermore, there is a weak negative correlation between the ifo business sentiment and CPI (-0.31) as well as PPI

Table 3.2: Correlation table of forecasts, realizations and instruments

		CEE	JF	real-time	revised	*lf*	*ur*	*rlc*	*gdp*	*inv*	*gov*	*ord*	*bp*	*rs*	*cpi*	*ppi*	*nlc*	*dax*	*ltr*	*str*	*ft*	*xrt*	*oil*	*bc*	*bs*	*be*
1	CEE	1	0.46	0.77	0.66	0.69	-0.71	-0.02	0.47	0.57	-0.30	0.27	-0.07	-0.20	-0.41	-0.39	0.49	-0.07	0.23	0.54	0.02	0.16	-0.02	-0.07	0.16	-0.35
2	JF		1	0.30	0.36	0.39	-0.31	-0.19	0.06	0.22	-0.21	0.22	-0.28	0.17	-0.40	-0.29	0.15	0.02	0.22	0.35	0.50	0.24	0.20	-0.04	0.15	-0.28
3	real-time			1	0.89	0.86	-0.81	0.00	0.68	0.76	-0.41	0.54	0.13	-0.22	-0.40	-0.44	0.52	0.15	0.44	0.60	-0.01	0.13	-0.05	0.27	0.43	-0.02
4	revised				1	0.98	-0.70	-0.03	0.72	0.69	-0.34	0.57	0.08	-0.09	-0.36	-0.42	0.28	0.18	0.45	0.55	0.00	0.13	-0.05	0.33	0.50	0.01
5	*lf*					1	-0.64	-0.01	0.69	0.63	-0.30	0.52	-0.03	-0.17	-0.30	-0.37	0.25	0.10	0.50	0.59	0.00	0.13	0.06	0.20	0.38	-0.11
6	*ur*						1	0.12	-0.52	-0.68	0.44	-0.52	-0.20	0.01	0.56	0.62	-0.42	-0.09	-0.19	-0.52	0.09	-0.05	0.23	-0.38	-0.54	-0.06
7	*rlc*							1	0.51	0.24	0.57	0.08	0.11	-0.05	0.54	0.07	0.25	-0.26	0.10	0.07	-0.28	0.13	0.31	-0.07	-0.04	-0.10
8	*gdp*								1	0.76	-0.04	0.72	0.15	-0.06	0.08	-0.20	0.42	0.13	0.47	0.56	-0.12	0.02	0.12	0.47	0.59	0.17
9	*inv*									1	-0.28	0.57	0.36	-0.04	-0.21	-0.32	0.52	0.29	0.30	0.46	-0.14	0.08	-0.09	0.41	0.51	0.16
10	*gov*										1	-0.36	-0.14	-0.21	0.70	0.39	-0.20	-0.35	-0.02	0.02	-0.29	0.07	0.30	-0.42	-0.43	-0.29
11	*ord*											1	0.00	0.10	-0.20	-0.36	0.41	0.32	0.45	0.52	0.12	-0.12	0.08	0.75	0.84	0.43
12	*bp*												1	-0.04	0.04	-0.15	0.07	0.06	0.00	-0.09	-0.22	-0.02	-0.29	0.29	0.15	0.40
13	*rs*													1	-0.39	-0.17	-0.21	0.18	-0.25	-0.53	0.34	0.16	-0.09	0.44	0.37	0.43
14	*cpi*														1	0.71	-0.07	-0.19	0.17	0.12	-0.35	-0.15	0.29	-0.27	-0.31	-0.15
15	*ppi*															1	-0.20	0.06	0.00	-0.14	-0.10	-0.11	0.01	-0.28	-0.38	-0.06
16	*nlc*																1	0.03	0.26	0.44	-0.05	0.12	0.03	0.24	0.36	0.01
17	*dax*																	1	0.12	-0.02	0.04	-0.25	-0.32	0.50	0.36	0.55
18	*ltr*																		1	0.63	-0.04	-0.12	0.44	0.20	0.22	0.11
19	*str*																			1	-0.08	-0.11	0.25	0.06	0.28	-0.26
20	*ft*																				1	0.41	0.08	0.14	0.14	0.12
21	*xrt*																					1	0.01	-0.06	0.03	-0.15
22	*oil*																						1	-0.18	-0.04	-0.32
23	*bc*																							1	0.92	0.84
24	*bs*																								1	0.56
25	*be*																									1

(-0.38) and positive correlations between the acceleration of nominal labor costs, the short-term interest growth rate (0.44) and the ifo business sentiment index.

Rows 17 to 19 depict the remaining correlation results of the financial set F. The highest coefficients are found between long-term and short-term interest rate growth (0.63), the DAX performance index and the three ifo indices as well as long-term interest rate growth and the oil price growth (0.44). Disregarding this last correlation coefficient and the foreign trade balance that is positively correlated with exchange rate growth (0.41), none of the instrumental variables in set G (foreign trade) appears to be correlated with any other instrument with an absolute correlation coefficient above 0.4 considered here. See columns 20 to 22. In addition to the correlations discussed above, the instruments in H (business climate) are highly correlated among each other.

The next section presents the empirical results of tests for information efficiency in the forecasts with respect to the instrumental sets discussed above. Herein, symmetric and asymmetric loss have been accounted for.

3.5 Efficiency Tests and Results

Subsections 3.5.1 to 3.5.8 all have an identical structure, while subsection 3.5.9 summarizes the main results. First, the instrumental variables used in the set are shortly introduced, along with the twelve instrumental subsets they form (four in the case of weak efficiency in set A). We then test for the efficient use of the information contained in these subsets under the assumption of symmetric and quadratic loss by applying a procedure based on Mincer-Zarnowitz (1969). See equations (1) and (2) in section 2.1. Herein, the unbiasedness and efficiency hypotheses are tested by a standard F-test based on a heteroscedasticity and autocorrelation consistent (HAC) covariance matrix estimator.[16] The next step relaxes the assumption of a symmetric loss function and discusses the results using the GMM approach of Elliott, Komunjer and Timmermann (2005, 2008) presented in section 2.3. In order to keep the tables compact, the GMM results for each forecasting institution are depicted in separate tables. The GMM estimation is implemented using the continuously updating estimator discussed in Hansen et al. (1996) with a quadratic spectral kernel and bandwidth choice according to Andrews (1991). For minimization, the derivative-free Nelder–Mead algorithm (Nelder and Mead, 1965) is used, as it is a robust method and has advantages over quasi-Newton methods if the objective function is non-differentiable. This is true for the loss function in the case at hand. We faced convergence problems in the numerical optimization when estimating α and p simultaneously. Thus, we fixed p at $p = 1$, $p = 1.5$ and $p = 2$.

The results discussed in sections 3.5.1 to 3.5.8 have been obtained by constraining the data to the pre-crisis period, i.e. the years 1970 to 2007. This choice has been made to assure that the results are not influenced by the irregular behavior of some of the time series presented above during the crisis. . Tables A3 to A27 in the appendix present analogous results including the crisis and expanding the observed time period to 2012, and illustrate that the general results still persist.

[16] This is actually implemented using the R package “sandwich” explained in Zeileis (2004).

3.5.1 Instrument Set A: Weak Efficiency

To analyze the weak efficiency of the CEE and JF forecasts, the following combinations of the lagged forecast errors are taken into account.

$$\begin{array}{llll} \text{A0} & \boldsymbol{w}_t = 1' & \text{A2} & \boldsymbol{w}_t = (1,\, e_t,\, e_{t-1})' \\ \text{A1} & \boldsymbol{w}_t = (1,\, e_t)' & \text{A3} & \boldsymbol{w}_t = (1,\, e_t,\, e_t^2)' \end{array}$$

Table 3.3 shows the results of the Mincer-Zarnowitz (1969) test for unbiasedness for set A0 in which only a constant is used as an instrument. It also shows the results of the efficiency tests for sets A1-A3. These tests imply the assumption of quadratic and symmetric loss. While unbiasedness cannot be rejected for both CEE and JF forecasts regarding real-time realizations, it is rejected for both institutions with respect to revised realizations. This similarity between the institutions does not hold for the efficiency results, in which there is evidence against CEE's forecast efficiency for all sets and for revised realizations (once only on a 10 percent level), but not for real-time realizations. Regardless of the type of realization, weak efficiency is rejected for the JF forecasts for all sets except set A1 and revised realizations. Again, efficiency can only be rejected on a 10 percent level in one case (set A2 and revised data).

Table 3.3: Regression tests of efficiency (A)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
A0	0.514 (0.602)	1.765 (0.186)	4.691 (0.015)	5.234 (0.010)
A1	0.043 (0.837)	5.243 (0.028)	4.283 (0.046)	0.962 (0.334)
A2	0.418 (0.662)	6.102 (0.006)	3.951 (0.029)	2.883 (0.071)
A3	1.062 (0.357)	4.812 (0.015)	2.819 (0.074)	3.953 (0.029)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Turning to the results for CEE's loss function shown in table 3.4, the difference in the results again depends on the type of realized data used. Whereas the α estimates are below 0.5 in all but three cases, only the parameters estimated using the revised realizations significantly differ from 0.5. Given these estimates, the CEE's loss function appears to be asymmetric, with higher weights on negative forecast errors (overpredicting employment growth) when revised realizations are considered. In three cases, all of which are on a 10 percent level, a high J-statistic leads to a rejection of forecast rationality. These rejections all appear in the results using revised realizations, with two rejections in set A1 (for curvature parameters $p = 1$ and $p = 1.5$) and one in A3 (for $p = 1$). The real-time estimates of α take values around 0.46, with an average standard deviation of 0.09. Hence, they do not differ significantly from 0.5 on any common level. Furthermore, the rationality hypothesis is not rejected in any instrument set. Thus, one could assume that the forecasters' loss function is close to symmetric with respect to real-time realizations.

Table 3.4: GMM estimates of the loss function - CEE (A)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
A0	1	0.474 (0.079)	0	0.289 (0.079)	0
	1.5	0.490 (0.090)	0	0.243 (0.082)	0
	2	0.494 (0.103)	0	0.223 (0.092)	0
A1	1	0.480 (0.082)	0.484 (0.487)	0.214 (0.078)	3.714 (0.054)
	1.5	0.503 (0.095)	0.001 (0.972)	0.146 (0.071)	3.032 (0.082)
	2	0.506 (0.109)	0.213 (0.644)	0.117 (0.082)	2.438 (0.118)
A2	1	0.431 (0.081)	1.895 (0.388)	0.174 (0.077)	3.907 (0.142)
	1.5	0.441 (0.094)	1.228 (0.541)	0.143 (0.068)	2.910 (0.233)
	2	0.431 (0.107)	1.257 (0.533)	0.119 (0.058)	2.215 (0.330)
A3	1	0.486 (0.083)	2.329 (0.312)	0.218 (0.077)	5.143 (0.076)
	1.5	0.493 (0.096)	2.513 (0.285)	0.163 (0.075)	3.892 (0.143)
	2	0.504 (0.110)	3.034 (0.219)	0.111 (0.080)	3.154 (0.207)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

Similar observations can be made for the results on JF's presented in table 3.5. Again, the tendency for α estimates to differ significantly from 0.5 is a lot stronger for revised data, although now three estimates only differ significantly on a 10 percent level and two do not differ at all. For real-time data, there are even two estimates above 0.5 (sets A1 and A3 and curvature parameter $p = 2$) although they do not differ significantly. The other α estimates take values around 0.38. With the exception of set A2 and $p = 1$, where α differs on a 10 percent level, they fail to differ significantly. Furthermore, two of the J-statistics in set A1 indicate a rejection of the rationality hypothesis for revised data. The rejections occur for shape parameters $p = 1$ and $p = 1.5$, with the former holding on a 5 percent level and the latter only on a 10 percent level.

Table 3.5: GMM estimates of the loss function - JF (A)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
A0	1	0.368 (0.079)	0	0.342 (0.087)	0
	1.5	0.394 (0.096)	0	0.245 (0.080)	0
	2	0.373 (0.139)	0	0.175 (0.075)	0
A1	1	0.377 (0.082)	0.147 (0.701)	0.418 (0.090)	3.929 (0.047)
	1.5	0.434 (0.099)	1.093 (0.296)	0.337 (0.087)	3.578 (0.059)
	2	0.540 (0.115)	1.376 (0.241)	0.279 (0.085)	2.600 (0.107)
A2	1	0.347 (0.084)	1.034 (0.596)	0.354 (0.092)	4.338 (0.114)
	1.5	0.391 (0.105)	1.830 (0.401)	0.104 (0.058)	4.597 (0.100)
	2	0.431 (0.138)	2.698 (0.259)	0.188 (0.074)	4.307 (0.116)
A3	1	0.378 (0.082)	0.287 (0.866)	0.337 (0.085)	3.616 (0.164)
	1.5	0.432 (0.093)	1.064 (0.587)	0.365 (0.090)	3.978 (0.137)
	2	0.513 (0.129)	1.790 (0.409)	0.350 (0.074)	3.570 (0.168)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

To sum up, the revised results show a stronger indication for asymmetry in both institutions' loss function than the real-time results do. In the cases, for which we find evidence for asymmetric loss, the direction of the asymmetry suggests a preference toward underestimating employment growth and higher costs associated with an overprediction. For the JF real-time results, the estimates of α tend to be below 0.5. Although not statistically significant, the degree of asymmetry is just big enough to prevent the rationality hypothesis from being rejected, as observed under symmetric loss. Ziliak and McCloskey (2004) argue in favor of the interpretation of results, that fail to be statistically significant, as these results may nevertheless bear an economical value. Elliott and Granger (2004) reinforce their point, although they warn against completely abandoning the concept of statistical significance, as it has proved to be a useful tool in many fields.

3.5.2 Instrument Set B: Labor Market

Instrument set B is employed to analyze how the forecasting institutions use the available information on the labor market, such as the growth rates of the labor force (lf), the unemployment rate (upr) and the real labor cost per employee (rlc). To this end, the three variables form the following twelve instrument sets introduced in section 3.3:

B1	$\boldsymbol{w}_t = (1, lf_t)'$	B7	$\boldsymbol{w}_t = (1, lf_t, lf_t^2)'$
B2	$\boldsymbol{w}_t = (1, upr_t)'$	B8	$\boldsymbol{w}_t = (1, upr_t, upr_t^2)'$
B3	$\boldsymbol{w}_t = (1, rlc_t)'$	B9	$\boldsymbol{w}_t = (1, rlc_t, rlc_t^2)'$
B4	$\boldsymbol{w}_t = (1, lf_t, lf_{t-1})'$	B10	$\boldsymbol{w}_t = (1, lf_t, upr_t, lf_t \cdot upr_t)'$
B5	$\boldsymbol{w}_t = (1, upr_t, upr_{t-1})'$	B11	$\boldsymbol{w}_t = (1, lf_t, rlc_t, lf_t \cdot rlc_t)'$
B6	$\boldsymbol{w}_t = (1, rlc_t, rlc_{t-1})'$	B12	$\boldsymbol{w}_t = (1, upr_t, rlc_t, upr_t \cdot rlc_t)'$

Table 3.6: Regression tests of efficiency (B)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
B1	11.392 (0.002)	91.583 (0.000)	707.17 (0.000)	742.74 (0.000)
B2	29.365 (0.000)	31.110 (0.000)	17.398 (0.000)	26.808 (0.000)
B3	0.027 (0.871)	0.131 (0.720)	0.010 (0.920)	0.171 (0.682)
B4	11.385 (0.000)	86.394 (0.000)	168.86 (0.000)	660.86 (0.000)
B5	14.020 (0.000)	18.288 (0.000)	8.053 (0.001)	15.731 (0.000)
B6	9.780 (0.000)	2.575 (0.092)	9.182 (0.001)	3.997 (0.028)
B7	6.126 (0.005)	52.968 (0.000)	489.22 (0.000)	388.19 (0.000)
B8	16.224 (0.000)	40.749 (0.000)	8.422 (0.001)	17.905 (0.000)
B9	0.159 (0.853)	0.230 (0.796)	0.039 (0.962)	0.481 (0.622)
B10	31.895 (0.000)	311.25 (0.000)	405.86 (0.000)	1007.5 (0.000)
B11	15.996 (0.000)	47.578 (0.000)	265.36 (0.000)	438.61 (0.000)
B12	13.384 (0.000)	18.079 (0.000)	8.063 (0.000)	11.940 (0.000)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

As shown in table 3.6, forecast efficiency is rejected under symmetric loss for all instrumental sets for both institutions regardless of the type of realization data used, with three exceptions.

Efficiency cannot be rejected at any common level in sets B3 and B9, and the rejection only holds on a 10 percent level for real-time realizations and JF forecast errors in set B6. All three sets contain information on the growth of real labor costs that appears to be used efficiently under the assumption of symmetric loss.

Table 3.7: GMM estimates of the loss function - CEE (B)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
B1	1	0.690 (0.075)	15.790 (0.000)	0.113 (0.066)	8.276 (0.004)
	1.5	0.749 (0.070)	10.771 (0.001)	0.077 (0.047)	5.442 (0.020)
	2	0.261 (0.092)	7.415 (0.006)	0.488 (0.092)	19.124 (0.000)
B2	1	0.395 (0.067)	7.588 (0.006)	0.275 (0.080)	1.824 (0.177)
	1.5	0.393 (0.077)	4.961 (0.026)	0.215 (0.081)	1.503 (0.220)
	2	0.376 (0.087)	3.125 (0.077)	0.200 (0.092)	1.121 (0.290)
B3	1	0.484 (0.083)	0.207 (0.649)	0.299 (0.082)	0.011 (0.918)
	1.5	0.499 (0.095)	0.109 (0.741)	0.246 (0.084)	0.030 (0.863)
	2	0.503 (0.108)	0.025 (0.874)	0.205 (0.081)	0.177 (0.674)
B4	1	0.220 (0.076)	15.007 (0.001)	0.077 (0.063)	7.937 (0.019)
	1.5	0.182 (0.079)	13.950 (0.001)	0.071 (0.047)	5.824 (0.054)
	2	0.185 (0.087)	8.979 (0.011)	0.056 (0.037)	4.534 (0.104)
B5	1	0.452 (0.070)	11.378 (0.003)	0.333 (0.076)	7.861 (0.020)
	1.5	0.388 (0.079)	9.824 (0.007)	0.389 (0.069)	23.526 (0.000)
	2	0.330 (0.092)	8.725 (0.013)	0.513 (0.089)	26.433 (0.000)
B6	1	0.428 (0.084)	5.263 (0.072)	0.149 (0.068)	5.077 (0.079)
	1.5	0.429 (0.099)	4.949 (0.084)	0.116 (0.058)	4.660 (0.097)
	2	0.412 (0.113)	4.703 (0.095)	0.084 (0.053)	4.222 (0.121)
B7	1	0.518 (0.084)	16.703 (0.000)	0.334 (0.072)	16.794 (0.000)
	1.5	0.572 (0.087)	12.800 (0.002)	0.033 (0.025)	8.045 (0.018)
	2	0.586 (0.100)	9.990 (0.007)	0.030 (0.021)	4.566 (0.102)
B8	1	0.359 (0.062)	17.338 (0.000)	0.291 (0.076)	4.295 (0.117)
	1.5	0.709 (0.049)	15.307 (0.000)	0.246 (0.075)	3.695 (0.158)
	2	0.682 (0.056)	11.711 (0.003)	0.223 (0.084)	2.248 (0.325)
B9	1	0.483 (0.084)	0.224 (0.894)	0.298 (0.082)	0.114 (0.944)
	1.5	0.496 (0.095)	0.171 (0.918)	0.245 (0.082)	0.055 (0.973)
	2	0.495 (0.106)	0.136 (0.934)	0.205 (0.078)	0.167 (0.920)
B10	1	0.518 (0.071)	17.002 (0.001)	0.095 (0.056)	9.125 (0.028)
	1.5	0.659 (0.074)	10.757 (0.013)	0.082 (0.041)	6.239 (0.101)
	2	0.676 (0.090)	9.349 (0.025)	0.060 (0.032)	4.066 (0.254)
B11	1	0.250 (0.079)	14.089 (0.003)	0.093 (0.065)	9.423 (0.024)
	1.5	0.270 (0.088)	10.784 (0.013)	0.079 (0.045)	6.290 (0.098)
	2	0.260 (0.095)	7.006 (0.072)	0.048 (0.031)	5.306 (0.151)
B12	1	0.466 (0.052)	8.988 (0.029)	0.281 (0.082)	3.340 (0.342)
	1.5	0.352 (0.069)	9.272 (0.026)	0.202 (0.076)	4.366 (0.225)
	2	0.812 (0.097)	12.472 (0.006)	0.146 (0.065)	4.281 (0.233)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

Allowing for an asymmetric loss function and turning to the CEE results in table 3.7, the rejections of efficiency strongly depend on the choice of realization data taken into account. For real-time realizations, sets B3 and B9 show J-statistics that do not lead to a rejection of forecast rationality. These sets contain information on the growth of real labor costs and are the same two instrument sets that did not lead to a rejection of efficiency under symmetric loss. Because the α estimates for these two sets are only slightly below 0.5 and do not differ significantly, a rather symmetric underlying loss function can be assumed. Many of the α estimates in the other instrument sets, for which the hypothesis of forecast rationality can be rejected, differ significantly from 0.5, indicating asymmetric loss. In some cases, the values of $\hat{\alpha}$ are above 0.5. See, for example, set B8.

Turning to the results for the revised realizations, all but three estimates of α are significantly below 0.5. This holds on a 5 percent level regardless of the accompanying J-statistic. Rationality is rejected less often compared to the real-time results. Instrument sets B1 and B5 are the only ones with rationality rejections for all curvature parameters p, while B4, B6 to B7 and B11 show two rejections each. That leaves set B10 with only one rejection on a 5 percent level and B2, B3, B8, B9 and B12 without any rejection of rationality. While sets B3 and B9 contain information on the growth of real labor costs, sets B2 and B8 are built around the growth of the unemployment rate, and set B12 contains the first lags of both instruments and their interaction term. Surprisingly, all sets that include information about past labor force growth also show a strong tendency to reject of forecast rationality. This indicates that fully incorporating the information comprised in this instrumental variable might be useful for improving future forecasts.

Focusing on table 3.8, which shows the results for JF's forecast errors with respect to real-time realizations, there are less rejections of forecast rationality than in the CEE's forecasts. Now, the indication for asymmetry in the loss function is somewhat stronger in sets that do not reject rationality. The sets without any rejections are B3, B6 and B9. Interestingly, many of the α estimates among those in the sets that reject forecast rationality take values above 0.5, although most of these values do not differ significantly.

Regarding the revised realizations, the results for each instrument set are quite robust over the different curvature parameters p, with rejections of rationality in sets B1, B4, B6 to B8, B10 and B11 (mostly containing labor force growth, with the exception of sets B6 and B8). Similar to the real-time results, the sets for which rationality is not rejected, i.e. B2, B3, B5, B9 and B12, contain the growth rates of the unemployment rate and the real labor costs, either individually or interacting. However, these sets now strongly indicate an asymmetric underlying loss function because the associated α estimates are all significantly below 0.5.

Table 3.8: GMM estimates of the loss function - JF (B)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
B1	1	0.515 (0.081)	12.120 (0.000)	0.542 (0.091)	10.261 (0.001)
	1.5	0.621 (0.090)	5.896 (0.015)	0.440 (0.092)	8.927 (0.003)
	2	0.708 (0.106)	3.020 (0.082)	0.351 (0.092)	5.188 (0.023)
B2	1	0.318 (0.078)	3.831 (0.050)	0.285 (0.096)	1.967 (0.161)
	1.5	0.297 (0.092)	3.179 (0.075)	0.181 (0.077)	1.882 (0.170)
	2	0.240 (0.112)	2.561 (0.110)	0.111 (0.060)	1.768 (0.184)
B3	1	0.375 (0.081)	0.157 (0.692)	0.336 (0.075)	0.082 (0.774)
	1.5	0.402 (0.102)	0.420 (0.517)	0.238 (0.071)	0.086 (0.770)
	2	0.443 (0.137)	0.537 (0.464)	0.171 (0.075)	0.119 (0.730)
B4	1	0.566 (0.082)	14.684 (0.001)	0.532 (0.091)	12.755 (0.002)
	1.5	0.586 (0.076)	7.441 (0.024)	0.047 (0.043)	6.498 (0.039)
	2	0.598 (0.087)	4.870 (0.088)	0.032 (0.028)	4.872 (0.088)
B5	1	0.307 (0.073)	5.770 (0.056)	0.262 (0.085)	3.975 (0.137)
	1.5	0.289 (0.076)	3.382 (0.184)	0.133 (0.058)	3.880 (0.144)
	2	0.244 (0.101)	2.477 (0.290)	0.083 (0.045)	2.906 (0.234)
B6	1	0.298 (0.080)	4.104 (0.128)	0.270 (0.081)	3.651 (0.161)
	1.5	0.398 (0.108)	4.543 (0.103)	0.094 (0.051)	6.042 (0.049)
	2	0.480 (0.138)	3.449 (0.178)	0.213 (0.076)	5.974 (0.050)
B7	1	0.482 (0.081)	11.806 (0.003)	0.524 (0.093)	10.876 (0.004)
	1.5	0.536 (0.074)	7.947 (0.019)	0.299 (0.078)	9.948 (0.007)
	2	0.517 (0.091)	4.964 (0.084)	0.173 (0.059)	6.482 (0.039)
B8	1	0.477 (0.071)	13.799 (0.001)	0.294 (0.088)	4.862 (0.088)
	1.5	0.593 (0.080)	9.771 (0.008)	0.245 (0.079)	6.349 (0.042)
	2	0.681 (0.087)	5.400 (0.067)	0.227 (0.079)	4.977 (0.083)
B9	1	0.370 (0.080)	0.269 (0.874)	0.328 (0.071)	0.209 (0.901)
	1.5	0.392 (0.087)	0.426 (0.808)	0.241 (0.056)	0.085 (0.958)
	2	0.444 (0.093)	0.527 (0.768)	0.197 (0.051)	0.271 (0.873)
B10	1	0.560 (0.075)	13.650 (0.003)	0.602 (0.090)	11.996 (0.007)
	1.5	0.529 (0.083)	8.707 (0.033)	0.521 (0.093)	15.238 (0.002)
	2	0.534 (0.105)	4.637 (0.200)	0.445 (0.086)	15.819 (0.001)
B11	1	0.575 (0.081)	16.112 (0.001)	0.492 (0.080)	11.412 (0.010)
	1.5	0.684 (0.085)	9.219 (0.027)	0.374 (0.072)	10.981 (0.012)
	2	0.788 (0.073)	8.384 (0.039)	0.286 (0.078)	9.545 (0.023)
B12	1	0.262 (0.074)	6.720 (0.081)	0.290 (0.070)	1.731 (0.630)
	1.5	0.646 (0.104)	9.249 (0.026)	0.159 (0.055)	2.257 (0.521)
	2	0.169 (0.103)	5.518 (0.138)	0.079 (0.040)	2.775 (0.428)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

3.5.3 Instrument Set C: Aggregate Demand

To cover the information available on aggregate demand, set C contains twelve instrument sets, shown below, that are built around the growth rates of the gross domestic product (*gdp*), real investment (*inv*) and government consumption (*gov*).

C1	$\boldsymbol{w}_t = (1,\, gdp_t)'$	C7	$\boldsymbol{w}_t = (1,\, gdp_t,\, gdp_t^2)'$
C2	$\boldsymbol{w}_t = (1,\, inv_t)'$	C8	$\boldsymbol{w}_t = (1,\, inv_t,\, inv_t^2)'$
C3	$\boldsymbol{w}_t = (1,\, gov_t)'$	C9	$\boldsymbol{w}_t = (1,\, gov_t,\, gov_t^2)'$
C4	$\boldsymbol{w}_t = (1,\, gdp_t,\, gdp_{t-1})'$	C10	$\boldsymbol{w}_t = (1,\, gdp_t,\, inv_t,\, gdp_t \cdot inv_t)'$
C5	$\boldsymbol{w}_t = (1,\, inv_t,\, inv_{t-1})'$	C11	$\boldsymbol{w}_t = (1,\, gdp_t,\, gov_t,\, gdp_t \cdot gov_t)'$
C6	$\boldsymbol{w}_t = (1,\, gov_t,\, gov_{t-1})'$	C12	$\boldsymbol{w}_t = (1,\, inv_t,\, gov_t,\, inv_t \cdot gov_t)'$

Starting with table 3.9, which exhibits the efficiency results under symmetric loss, most of the calculated F-statistics once again indicate an inefficient use of the information available. Only three sets (C3, C6 and C9) do not reject the efficiency assumption under symmetric loss in all cases, i.e. across both types of realizations and both institutions. While for CEE's forecast errors in set C9, efficiency is rejected with respect to real-time realizations only on a 10 percent level, there are rejections for JF's forecast errors in set C6 for both types of realizations (once only on a 10 percent level) and in set C3 for real-time realizations (again only on a 10 percent level). Hence, the results suggest that the institutions do not seem to use the information contained in GDP and investment growth efficiently if the underlying loss function is symmetric. Under this assumption, information on government consumption growth appears to be used rather efficiently.

Table 3.9: Regression tests of efficiency (C)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
C1	21.345 (0.000)	19.108 (0.000)	31.918 (0.000)	26.570 (0.000)
C2	19.119 (0.000)	24.059 (0.000)	20.188 (0.000)	24.163 (0.000)
C3	2.528 (0.121)	3.508 (0.069)	0.712 (0.404)	1.777 (0.191)
C4	14.724 (0.000)	15.309 (0.000)	14.502 (0.000)	14.195 (0.000)
C5	18.601 (0.000)	18.494 (0.000)	16.025 (0.000)	19.923 (0.000)
C6	1.330 (0.278)	5.170 (0.011)	0.548 (0.583)	3.210 (0.053)
C7	56.272 (0.000)	21.379 (0.000)	109.52 (0.000)	33.119 (0.000)
C8	11.113 (0.000)	14.640 (0.000)	12.032 (0.000)	19.558 (0.000)
C9	2.678 (0.083)	2.398 (0.106)	1.220 (0.308)	0.957 (0.394)
C10	13.868 (0.000)	9.935 (0.000)	19.878 (0.000)	42.326 (0.000)
C11	18.985 (0.000)	71.019 (0.000)	16.258 (0.000)	36.576 (0.000)
C12	12.510 (0.000)	14.059 (0.000)	10.727 (0.000)	9.909 (0.000)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Table 3.10 presents the GMM estimates of the loss function for CEE forecast errors with respect to both real-time and revised realization data. Regardless of the type of realization under consideration, the three instrument sets that do not allow rejecting efficiency under symmetric loss (i.e. C3, C6 and C9) still do not suggest a rejection of the rationality hypothesis. For real-time results, none of the estimated shape parameters for α differs significantly from 0.5 in these sets, although all but two $\hat{\alpha}$ take values below 0.5. The revised results indicate asymmetry in the loss function with α estimates significantly below 0.5. Although the majority of the other α estimates is significantly below 0.5 for real-time as well as for revised results, rationality is rejected in all sets that include information on GDP or investment growth for real-time data and in the sets C1, C4, and C7 (also containing GDP growth) for revised data. In addition, there are two rejections in sets C2 and C10 for revised data and one rejection in set C11.

Table 3.10: GMM estimates of the loss function - CEE (C)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
C1	1	0.745 (0.077)	18.938 (0.000)	0.155 (0.068)	5.894 (0.015)
	1.5	0.252 (0.084)	11.738 (0.001)	0.072 (0.051)	5.879 (0.015)
	2	0.249 (0.093)	7.549 (0.006)	0.465 (0.092)	17.627 (0.000)
C2	1	0.320 (0.078)	9.765 (0.002)	0.235 (0.077)	2.490 (0.115)
	1.5	0.299 (0.087)	8.837 (0.003)	0.150 (0.074)	3.720 (0.054)
	2	0.274 (0.096)	6.617 (0.010)	0.176 (0.094)	2.965 (0.085)
C3	1	0.464 (0.080)	1.075 (0.300)	0.291 (0.079)	0.362 (0.547)
	1.5	0.483 (0.091)	1.811 (0.178)	0.228 (0.080)	0.530 (0.467)
	2	0.464 (0.107)	2.629 (0.105)	0.183 (0.078)	0.736 (0.391)
C4	1	0.786 (0.078)	23.249 (0.000)	0.088 (0.056)	7.025 (0.030)
	1.5	0.198 (0.081)	12.041 (0.002)	0.060 (0.039)	5.896 (0.052)
	2	0.210 (0.094)	7.620 (0.022)	0.041 (0.030)	5.172 (0.075)
C5	1	0.243 (0.076)	12.025 (0.002)	0.151 (0.069)	4.555 (0.103)
	1.5	0.231 (0.086)	10.197 (0.006)	0.133 (0.070)	4.154 (0.125)
	2	0.221 (0.096)	7.362 (0.025)	0.164 (0.088)	3.331 (0.189)
C6	1	0.466 (0.081)	1.338 (0.512)	0.294 (0.080)	0.705 (0.703)
	1.5	0.474 (0.092)	2.005 (0.367)	0.223 (0.077)	0.556 (0.757)
	2	0.443 (0.107)	2.615 (0.270)	0.159 (0.067)	0.980 (0.613)
C7	1	0.257 (0.079)	20.609 (0.000)	0.169 (0.066)	6.396 (0.041)
	1.5	0.414 (0.091)	17.134 (0.000)	0.048 (0.035)	7.182 (0.028)
	2	0.459 (0.102)	13.082 (0.001)	0.008 (0.021)	5.641 (0.060)
C8	1	0.297 (0.073)	13.880 (0.001)	0.226 (0.075)	3.139 (0.208)
	1.5	0.260 (0.080)	14.970 (0.001)	0.141 (0.071)	3.733 (0.155)
	2	0.243 (0.089)	9.462 (0.009)	0.142 (0.079)	3.216 (0.200)
C9	1	0.473 (0.083)	1.225 (0.542)	0.276 (0.078)	1.380 (0.502)
	1.5	0.505 (0.092)	2.080 (0.353)	0.223 (0.079)	1.204 (0.548)
	2	0.536 (0.104)	2.851 (0.240)	0.151 (0.070)	1.297 (0.523)
C10	1	0.219 (0.074)	18.848 (0.000)	0.149 (0.066)	6.721 (0.081)
	1.5	0.202 (0.084)	16.403 (0.001)	0.059 (0.040)	6.282 (0.099)
	2	0.155 (0.088)	11.409 (0.010)	0.019 (0.025)	6.239 (0.101)
C11	1	0.639 (0.072)	20.474 (0.000)	0.147 (0.064)	7.026 (0.071)
	1.5	0.658 (0.093)	17.105 (0.001)	0.062 (0.045)	6.191 (0.103)
	2	0.324 (0.101)	12.342 (0.006)	0.030 (0.032)	5.837 (0.120)
C12	1	0.391 (0.074)	15.089 (0.002)	0.256 (0.073)	5.925 (0.115)
	1.5	0.349 (0.079)	14.413 (0.002)	0.106 (0.055)	5.281 (0.152)
	2	0.316 (0.089)	12.752 (0.005)	0.195 (0.076)	4.939 (0.176)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

Turning to the JF results presented in table 3.11, we find a similar pattern. Rationality is once again rejected in sets C1, C4, and C7. In sets C3, C6 and C9, the α estimates are below 0.5, but only significantly so for revised realizations. The total number of rejections of forecast rationality is still higher for the results on real-time data and sets built around investment growth, i.e. C2, C5 and C8. The sets under consideration with the interaction term of the variables, C10 to C12, mostly show high J-statistics that lead to a rejection of the rationality hypothesis. For revised realization data we find rationality rejections in the sets around GDP growth (C1, C4 and C7) as well as in sets C2 (although twice only on a 10 percent level), C10 and C12 (only two rejections each) and C6 and C11 (only one rejection each, both on a 10 percent level).

Table 3.11: GMM estimates of the loss function - JF (C)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
C1	1	0.200 (0.073)	10.453 (0.001)	0.200 (0.074)	5.551 (0.018)
	1.5	0.639 (0.094)	7.118 (0.008)	0.090 (0.044)	5.757 (0.016)
	2	0.690 (0.097)	2.761 (0.097)	0.322 (0.091)	5.122 (0.024)
C2	1	0.231 (0.074)	7.782 (0.005)	0.271 (0.086)	2.793 (0.095)
	1.5	0.666 (0.073)	7.713 (0.005)	0.126 (0.058)	3.721 (0.054)
	2	0.762 (0.068)	3.494 (0.062)	0.404 (0.107)	7.597 (0.006)
C3	1	0.356 (0.080)	1.843 (0.175)	0.351 (0.086)	0.568 (0.451)
	1.5	0.374 (0.096)	2.053 (0.152)	0.240 (0.080)	0.866 (0.352)
	2	0.271 (0.124)	2.098 (0.148)	0.151 (0.071)	0.941 (0.332)
C4	1	0.605 (0.083)	30.614 (0.000)	0.143 (0.072)	6.310 (0.043)
	1.5	0.084 (0.069)	10.133 (0.006)	0.044 (0.032)	6.254 (0.044)
	2	0.384 (0.100)	8.152 (0.017)	0.016 (0.016)	5.448 (0.066)
C5	1	0.152 (0.066)	10.578 (0.005)	0.198 (0.085)	4.404 (0.111)
	1.5	0.655 (0.069)	10.245 (0.006)	0.078 (0.046)	4.663 (0.097)
	2	0.704 (0.075)	5.075 (0.079)	0.039 (0.026)	3.809 (0.149)
C6	1	0.357 (0.081)	1.731 (0.421)	0.347 (0.085)	0.587 (0.745)
	1.5	0.393 (0.094)	2.366 (0.306)	0.243 (0.079)	0.889 (0.641)
	2	0.429 (0.108)	2.976 (0.226)	0.179 (0.071)	1.298 (0.523)
C7	1	0.249 (0.073)	12.159 (0.002)	0.222 (0.075)	6.252 (0.044)
	1.5	0.355 (0.096)	10.658 (0.005)	0.097 (0.046)	7.310 (0.026)
	2	0.337 (0.116)	7.409 (0.025)	0.101 (0.044)	6.852 (0.033)
C8	1	0.222 (0.074)	7.782 (0.020)	0.244 (0.087)	3.679 (0.159)
	1.5	0.203 (0.091)	7.981 (0.018)	0.109 (0.057)	4.036 (0.133)
	2	0.502 (0.099)	6.833 (0.033)	0.052 (0.036)	4.063 (0.131)
C9	1	0.360 (0.081)	2.119 (0.347)	0.352 (0.087)	0.615 (0.735)
	1.5	0.409 (0.098)	2.284 (0.319)	0.249 (0.080)	0.921 (0.631)
	2	0.368 (0.139)	2.970 (0.227)	0.160 (0.071)	0.986 (0.611)
C10	1	0.181 (0.073)	11.113 (0.011)	0.159 (0.070)	6.656 (0.084)
	1.5	0.344 (0.097)	11.427 (0.010)	0.081 (0.044)	6.465 (0.091)
	2	0.431 (0.115)	7.480 (0.058)	0.043 (0.029)	6.242 (0.100)
C11	1	0.157 (0.067)	14.357 (0.002)	0.210 (0.072)	5.659 (0.129)
	1.5	0.509 (0.096)	9.025 (0.029)	0.077 (0.042)	6.787 (0.079)
	2	0.546 (0.106)	4.665 (0.198)	0.246 (0.069)	6.007 (0.111)
C12	1	0.317 (0.076)	14.393 (0.002)	0.273 (0.086)	3.188 (0.364)
	1.5	0.553 (0.096)	8.189 (0.042)	0.164 (0.058)	6.522 (0.089)
	2	0.614 (0.085)	5.103 (0.164)	0.137 (0.045)	7.236 (0.065)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

3.5.4 Instrument Set D: Leading Indicators

In the present set, the growth rates of industrial orders received (ord) and building permissions (bp), along with the term spread (rs) are used as instruments to investigate whether the forecasting institutions use the information contained in these leading indicators efficiently. For this purpose, the same twelve instrument sets have been formed with these three variables.

D1	$\boldsymbol{w}_t = (1, ord_t)'$	D7	$\boldsymbol{w}_t = (1, ord_t, ord_t^2)'$
D2	$\boldsymbol{w}_t = (1, bp_t)'$	D8	$\boldsymbol{w}_t = (1, bp_t, bp_t^2)'$
D3	$\boldsymbol{w}_t = (1, rs_t)'$	D9	$\boldsymbol{w}_t = (1, rs_t, rs_t^2)'$
D4	$\boldsymbol{w}_t = (1, ord_t, ord_{t-1})'$	D10	$\boldsymbol{w}_t = (1, ord_t, bp_t, ord_t \cdot bp_t)'$
D5	$\boldsymbol{w}_t = (1, bp_t, bp_{t-1})'$	D11	$\boldsymbol{w}_t = (1, ord_t, rs_t, ord_t \cdot rs_t)'$
D6	$\boldsymbol{w}_t = (1, rs_t, rs_{t-1})'$	D12	$\boldsymbol{w}_t = (1, bp_t, rs_t, bp_t \cdot rs_t)'$

Table 3.12 shows the results of the efficiency tests under quadratic loss. The most apparent pattern is that instrument sets D1, D4 and D7, all of which contain the growth of industrial orders, lead to a rejection of efficiency for both institutions and both types of realizations. Sets D10 and D11, which contain the growth of industrial orders together with any of the two other instruments and their interaction, exhibit the same robustness concerning the rejections across institutions and types of realizations. Apart from these results, efficiency is rejected regardless of the type of realization under consideration for JF's forecast errors and in instrument set D6, which contains the first and second lags of the term spread. For the CEE results, there is evidence against efficiency in sets D5 (first and second lag of building permission growth) and D12 (first lags and interactions of growth building permissions and term spread) for both realizations, while there are two more rejections that only occur in the real-time results, i.e. sets D2 and D8. All rejections hold on a 5 percent level, except for two rejections on a 10 percent level, i.e. set D2 for real-time data and D12 for revised data (both for CEE's forecast errors).

Table 3.12: Regression tests of efficiency (D)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
D1	21.689 (0.000)	13.150 (0.001)	20.252 (0.000)	16.512 (0.000)
D2	3.772 (0.060)	1.221 (0.277)	1.135 (0.294)	0.975 (0.330)
D3	0.462 (0.502)	2.420 (0.130)	0.094 (0.761)	1.058 (0.311)
D4	9.804 (0.000)	7.456 (0.002)	11.442 (0.000)	9.147 (0.001)
D5	6.981 (0.003)	1.868 (0.170)	3.358 (0.047)	1.848 (0.173)
D6	0.707 (0.501)	24.613 (0.000)	0.384 (0.685)	16.219 (0.000)
D7	22.921 (0.000)	8.857 (0.001)	18.737 (0.000)	10.321 (0.000)
D8	4.357 (0.021)	0.896 (0.418)	1.544 (0.228)	0.610 (0.549)
D9	0.746 (0.483)	1.351 (0.274)	0.876 (0.426)	1.067 (0.356)
D10	10.810 (0.000)	7.944 (0.000)	8.501 (0.000)	7.636 (0.001)
D11	12.555 (0.000)	7.945 (0.000)	7.564 (0.001)	5.930 (0.003)
D12	5.524 (0.004)	1.231 (0.316)	2.883 (0.052)	0.537 (0.660)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Table 3.13: GMM estimates of the loss function - CEE (D)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
D1	1	0.245 (0.073)	15.102 (0.000)	0.146 (0.065)	6.687 (0.010)
	1.5	0.276 (0.084)	10.544 (0.001)	0.098 (0.052)	4.541 (0.033)
	2	0.283 (0.095)	6.817 (0.009)	0.058 (0.043)	3.488 (0.062)
D2	1	0.467 (0.082)	3.873 (0.049)	0.288 (0.079)	0.358 (0.550)
	1.5	0.500 (0.091)	3.276 (0.070)	0.248 (0.083)	0.603 (0.437)
	2	0.501 (0.102)	2.528 (0.112)	0.237 (0.092)	0.360 (0.549)
D3	1	0.480 (0.084)	0.112 (0.738)	0.289 (0.085)	0.104 (0.747)
	1.5	0.500 (0.099)	0.342 (0.559)	0.241 (0.080)	0.100 (0.752)
	2	0.475 (0.111)	0.821 (0.365)	0.199 (0.078)	0.388 (0.533)
D4	1	0.249 (0.073)	15.063 (0.001)	0.149 (0.066)	7.195 (0.027)
	1.5	0.282 (0.085)	10.523 (0.005)	0.083 (0.049)	6.126 (0.047)
	2	0.298 (0.097)	6.975 (0.031)	0.031 (0.038)	6.201 (0.045)
D5	1	0.424 (0.083)	7.434 (0.024)	0.228 (0.074)	4.111 (0.128)
	1.5	0.374 (0.090)	7.316 (0.026)	0.153 (0.069)	3.872 (0.144)
	2	0.350 (0.101)	4.937 (0.085)	0.110 (0.066)	3.891 (0.143)
D6	1	0.442 (0.085)	2.376 (0.305)	0.296 (0.088)	0.157 (0.924)
	1.5	0.463 (0.101)	1.556 (0.459)	0.231 (0.078)	0.432 (0.806)
	2	0.440 (0.111)	1.458 (0.482)	0.166 (0.069)	1.178 (0.555)
D7	1	0.402 (0.082)	19.962 (0.000)	0.065 (0.051)	10.855 (0.004)
	1.5	0.498 (0.090)	14.676 (0.001)	0.035 (0.026)	7.557 (0.023)
	2	0.485 (0.099)	10.457 (0.005)	0.025 (0.018)	4.278 (0.118)
D8	1	0.492 (0.081)	3.963 (0.138)	0.295 (0.075)	3.861 (0.145)
	1.5	0.524 (0.094)	3.531 (0.171)	0.264 (0.079)	2.368 (0.306)
	2	0.544 (0.104)	2.754 (0.252)	0.241 (0.084)	1.777 (0.411)
D9	1	0.482 (0.085)	0.560 (0.756)	0.297 (0.086)	0.551 (0.759)
	1.5	0.507 (0.099)	3.168 (0.205)	0.243 (0.081)	0.754 (0.686)
	2	0.356 (0.103)	3.804 (0.149)	0.169 (0.073)	1.260 (0.533)
D10	1	0.222 (0.072)	20.350 (0.000)	0.123 (0.062)	8.061 (0.045)
	1.5	0.327 (0.081)	12.823 (0.005)	0.092 (0.051)	6.507 (0.089)
	2	0.310 (0.086)	7.277 (0.064)	0.091 (0.044)	6.369 (0.095)
D11	1	0.212 (0.074)	16.494 (0.001)	0.085 (0.066)	7.697 (0.053)
	1.5	0.223 (0.085)	11.595 (0.009)	0.108 (0.053)	5.714 (0.126)
	2	0.209 (0.085)	7.996 (0.046)	0.079 (0.044)	4.278 (0.233)
D12	1	0.525 (0.085)	4.328 (0.228)	0.258 (0.084)	3.345 (0.341)
	1.5	0.462 (0.098)	4.559 (0.207)	0.206 (0.075)	3.459 (0.326)
	2	0.416 (0.109)	3.861 (0.277)	0.128 (0.062)	3.487 (0.322)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

Table 3.13 shows the CEE results under a flexible loss function. Once we allow for asymmetry in the loss function, forecast rationality can be rejected in all instrument sets containing growth of industrial orders. These rejections are analogous to the efficiency rejections under symmetry. This indicates that even under a more flexible loss function there is information left in this instrumental variable that could be used to further improve the forecasts. Especially for revised results, the rejections are not robust over the different curvature parameters p. In all sets but set C4 there are either less than three rejections, or some rejections only on a 10 percent level. For revised realization results, the sets that do not contain industrial order growth do not indicate further rationality rejections. For real-time results, rejections can be found in sets D2 and D5, both of which contain growth of building permissions. Another pronounced difference between the two types of realizations lies in the estimates of the asymmetry parameter. The estimates of α only differ significantly from 0.5 for the revised results. In some cases, the estimates even take values above 0.5 for real-time realizations.

Table 3.14: GMM estimates of the loss function - JF (D)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
D1	1	0.177 (0.070)	11.726 (0.001)	0.264 (0.083)	4.366 (0.037)
	1.5	0.191 (0.083)	9.326 (0.002)	0.115 (0.056)	5.386 (0.020)
	2	0.158 (0.100)	5.569 (0.018)	0.058 (0.038)	4.194 (0.041)
D2	1	0.311 (0.075)	4.547 (0.033)	0.316 (0.084)	1.482 (0.223)
	1.5	0.319 (0.090)	3.394 (0.065)	0.187 (0.069)	2.190 (0.139)
	2	0.601 (0.149)	2.934 (0.087)	0.109 (0.053)	2.062 (0.151)
D3	1	0.366 (0.083)	0.912 (0.339)	0.330 (0.097)	1.259 (0.262)
	1.5	0.466 (0.110)	1.857 (0.173)	0.289 (0.094)	1.678 (0.195)
	2	0.577 (0.116)	1.480 (0.224)	0.254 (0.092)	1.492 (0.222)
D4	1	0.174 (0.071)	13.172 (0.001)	0.268 (0.083)	4.812 (0.090)
	1.5	0.192 (0.085)	9.064 (0.011)	0.115 (0.056)	5.266 (0.072)
	2	0.101 (0.092)	7.620 (0.022)	0.058 (0.038)	4.500 (0.105)
D5	1	0.304 (0.078)	4.900 (0.086)	0.299 (0.083)	1.794 (0.408)
	1.5	0.306 (0.091)	4.309 (0.116)	0.147 (0.060)	3.178 (0.204)
	2	0.558 (0.102)	3.060 (0.217)	0.056 (0.034)	4.132 (0.127)
D6	1	0.336 (0.085)	2.758 (0.252)	0.310 (0.102)	1.824 (0.402)
	1.5	0.403 (0.113)	3.438 (0.179)	0.165 (0.085)	3.024 (0.220)
	2	0.525 (0.126)	3.196 (0.202)	0.201 (0.080)	3.172 (0.205)
D7	1	0.109 (0.055)	15.460 (0.000)	0.266 (0.079)	4.932 (0.085)
	1.5	0.119 (0.068)	11.597 (0.003)	0.087 (0.038)	6.527 (0.038)
	2	0.124 (0.075)	6.080 (0.048)	0.036 (0.018)	5.149 (0.076)
D8	1	0.317 (0.074)	5.904 (0.052)	0.313 (0.082)	1.766 (0.414)
	1.5	0.420 (0.095)	5.175 (0.075)	0.199 (0.066)	3.041 (0.219)
	2	0.620 (0.154)	3.403 (0.182)	0.118 (0.051)	3.421 (0.181)
D9	1	0.363 (0.083)	1.182 (0.554)	0.336 (0.093)	1.678 (0.432)
	1.5	0.458 (0.108)	1.859 (0.395)	0.270 (0.087)	1.877 (0.391)
	2	0.570 (0.121)	1.506 (0.471)	0.230 (0.083)	1.786 (0.409)
D10	1	0.176 (0.073)	13.722 (0.003)	0.239 (0.086)	4.171 (0.244)
	1.5	0.220 (0.090)	11.333 (0.010)	0.120 (0.059)	5.984 (0.112)
	2	0.178 (0.101)	9.604 (0.022)	0.072 (0.040)	8.375 (0.039)
D11	1	0.104 (0.063)	13.350 (0.004)	0.191 (0.082)	5.660 (0.129)
	1.5	0.101 (0.066)	10.921 (0.012)	0.095 (0.054)	6.415 (0.093)
	2	0.081 (0.064)	7.807 (0.050)	0.082 (0.045)	6.281 (0.099)
D12	1	0.267 (0.078)	6.307 (0.098)	0.278 (0.093)	2.805 (0.423)
	1.5	0.230 (0.091)	5.100 (0.165)	0.152 (0.075)	3.435 (0.329)
	2	0.556 (0.122)	3.906 (0.272)	0.077 (0.053)	3.463 (0.326)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

The analog JF results are presented in table 3.14. They follow a similar pattern, as rejections of forecast rationality are most robust for sets that include growth of industrial orders. In many of these sets there are less than three rejections over the different curvature parameters as well as rejections that only hold on a 10 percent level. For the results using the real-time realizations, there are further rejections for sets D2 and D5 and also D8. These sets are all constructed around the growth rate of building permissions. As in the CEE results in table 3.13, the values of the estimates of the asymmetry parameter $\hat{\alpha}$ differ substantially between revised and real-time results. For revised results, the estimates take values considerably below, and hence significantly different from, 0.5. In some cases they are even close to zero. Regarding the real-time results, seven of the α estimates are even higher than 0.5, although they do not differ significantly. Except for instrument sets that lead to a rejection of forecast rationality and for which the interpretation of $\hat{\alpha}$ is thus questionable, only about one third of the α estimates on real-time realizations is significantly smaller than 0.5.

3.5.5 Instrument Set E: Price Indices

In this set, the growth rates of the consumer price index (*cpi*) and the producer price index for industrial products (*ppi*), as well as the acceleration of the nominal labor costs per employee (*nlc*) are used as instruments. Sets E1-E3 contain the first lags, E4-E6 the first and second lags, E7-E9 the first lags and squares and E10-E12 contain the first lags and an interaction term. As before, all instruments were tested with an intercept.

E1	$\boldsymbol{w}_t = (1, cpi_t)'$	E7	$\boldsymbol{w}_t = (1, cpi_t, cpi_t^2)'$
E2	$\boldsymbol{w}_t = (1, ppi_t)'$	E8	$\boldsymbol{w}_t = (1, ppi_t, ppi_t^2)'$
E3	$\boldsymbol{w}_t = (1, nlc_t)'$	E9	$\boldsymbol{w}_t = (1, nlc_t, nlc_t^2)'$
E4	$\boldsymbol{w}_t = (1, cpi_t, cpi_{t-1})'$	E10	$\boldsymbol{w}_t = (1, cpi_t, ppi_t, cpi_t \cdot ppi_t)'$
E5	$\boldsymbol{w}_t = (1, ppi_t, ppi_{t-1})'$	E11	$\boldsymbol{w}_t = (1, cpi_t, nlc_t, cpi_t \cdot nlc_t)'$
E6	$\boldsymbol{w}_t = (1, nlc_t, nlc_{t-1})'$	E12	$\boldsymbol{w}_t = (1, ppi_t, nlc_t, ppi_t \cdot nlc_t)'$

Table 3.15: Regression tests of efficiency (E)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
E1	0.548 (0.464)	3.540 (0.068)	0.369 (0.548)	1.773 (0.192)
E2	1.259 (0.270)	3.769 (0.060)	2.241 (0.143)	3.885 (0.057)
E3	2.380 (0.132)	13.410 (0.001)	0.100 (0.753)	1.379 (0.249)
E4	0.267 (0.767)	5.409 (0.009)	0.187 (0.831)	3.781 (0.033)
E5	0.705 (0.501)	2.895 (0.069)	1.179 (0.320)	2.287 (0.117)
E6	8.252 (0.001)	5.988 (0.006)	5.944 (0.007)	0.727 (0.491)
E7	1.668 (0.204)	2.514 (0.096)	0.193 (0.826)	1.196 (0.315)
E8	11.431 (0.000)	10.430 (0.000)	5.322 (0.010)	8.884 (0.001)
E9	7.446 (0.002)	7.835 (0.002)	21.009 (0.000)	2.829 (0.074)
E10	6.482 (0.001)	6.697 (0.001)	1.263 (0.303)	4.099 (0.014)
E11	1.152 (0.344)	5.942 (0.003)	0.109 (0.954)	1.524 (0.228)
E12	2.213 (0.106)	9.607 (0.000)	2.105 (0.120)	2.520 (0.076)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

The results of the efficiency tests under quadratic loss, shown in table 3.15, are fairly robust across both types of realization as far as CEE results are concerned. Efficiency is rejected in sets E6, E8 and E9 for both realization types and in set E10 for real-time results. Sets E6 and E9 both contain the acceleration of nominal costs.

Table 3.16: GMM estimates of the loss function - CEE (E)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
E1	1	0.466 (0.080)	0.629 (0.428)	0.287 (0.078)	0.033 (0.855)
	1.5	0.488 (0.093)	0.648 (0.421)	0.240 (0.082)	0.069 (0.793)
	2	0.465 (0.100)	0.975 (0.324)	0.198 (0.083)	0.391 (0.532)
E2	1	0.449 (0.080)	1.731 (0.188)	0.292 (0.078)	0.139 (0.709)
	1.5	0.462 (0.093)	1.279 (0.258)	0.237 (0.082)	0.418 (0.518)
	2	0.454 (0.105)	1.039 (0.308)	0.210 (0.093)	0.624 (0.429)
E3	1	0.458 (0.085)	2.195 (0.138)	0.280 (0.085)	0.261 (0.609)
	1.5	0.427 (0.098)	4.884 (0.027)	0.232 (0.080)	0.095 (0.758)
	2	0.365 (0.111)	4.256 (0.039)	0.183 (0.074)	0.425 (0.514)
E4	1	0.467 (0.082)	0.858 (0.651)	0.286 (0.079)	0.057 (0.972)
	1.5	0.480 (0.093)	1.076 (0.584)	0.240 (0.083)	0.068 (0.967)
	2	0.463 (0.101)	0.987 (0.610)	0.195 (0.082)	0.392 (0.822)
E5	1	0.452 (0.081)	1.717 (0.424)	0.287 (0.079)	0.399 (0.819)
	1.5	0.479 (0.093)	2.061 (0.357)	0.230 (0.081)	0.575 (0.750)
	2	0.473 (0.106)	2.285 (0.319)	0.199 (0.090)	0.819 (0.664)
E6	1	0.417 (0.084)	6.439 (0.040)	0.249 (0.082)	2.241 (0.326)
	1.5	0.334 (0.098)	8.471 (0.014)	0.185 (0.067)	2.359 (0.307)
	2	0.314 (0.114)	5.888 (0.053)	0.134 (0.059)	2.645 (0.267)
E7	1	0.463 (0.082)	0.942 (0.624)	0.289 (0.078)	0.937 (0.626)
	1.5	0.456 (0.092)	1.446 (0.485)	0.235 (0.082)	0.398 (0.819)
	2	0.421 (0.088)	1.547 (0.461)	0.185 (0.079)	0.533 (0.766)
E8	1	0.436 (0.083)	2.571 (0.276)	0.283 (0.078)	0.372 (0.830)
	1.5	0.448 (0.091)	1.603 (0.449)	0.230 (0.075)	0.497 (0.780)
	2	0.440 (0.100)	1.244 (0.537)	0.208 (0.084)	0.598 (0.741)
E9	1	0.327 (0.072)	9.554 (0.008)	0.222 (0.085)	2.429 (0.297)
	1.5	0.614 (0.093)	6.417 (0.040)	0.184 (0.075)	1.948 (0.378)
	2	0.440 (0.107)	3.578 (0.167)	0.133 (0.062)	1.734 (0.420)
E10	1	0.424 (0.082)	3.300 (0.348)	0.287 (0.080)	0.447 (0.930)
	1.5	0.420 (0.082)	2.177 (0.537)	0.239 (0.081)	0.460 (0.928)
	2	0.415 (0.091)	1.642 (0.650)	0.202 (0.080)	0.711 (0.871)
E11	1	0.462 (0.086)	2.514 (0.473)	0.250 (0.085)	1.248 (0.742)
	1.5	0.606 (0.095)	5.994 (0.112)	0.234 (0.080)	0.150 (0.985)
	2	0.598 (0.092)	5.897 (0.117)	0.171 (0.071)	0.577 (0.902)
E12	1	0.439 (0.087)	2.760 (0.430)	0.273 (0.086)	0.650 (0.885)
	1.5	0.424 (0.100)	4.081 (0.253)	0.205 (0.078)	0.946 (0.814)
	2	0.381 (0.109)	3.535 (0.316)	0.142 (0.066)	1.511 (0.680)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

The efficiency rejections here either indicate that the CEE did not adequately use the information contained in this variable, or that their underlying loss function is not quadratic. However, it is hard to find a similar interpretation for the JF results, as they differ considerably across both types of realizations. While there are rejections for revised data in sets E4, E8 and E10 on a 5 percent level as well as in sets E2, E9 and E12 on a 10 percent level, efficiency is rejected in all sets for real-time results. Again, four of the rejections (E1, E2, E5 and E7) only hold on a 10 percent level.

Table 3.17: GMM estimates of the loss function - JF (E)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
E1	1	0.368 (0.080)	0.002 (0.965)	0.341 (0.087)	0.003 (0.958)
	1.5	0.393 (0.098)	0.080 (0.777)	0.244 (0.081)	0.077 (0.781)
	2	0.347 (0.134)	0.286 (0.593)	0.163 (0.073)	0.263 (0.608)
E2	1	0.365 (0.080)	0.242 (0.623)	0.340 (0.089)	0.391 (0.532)
	1.5	0.375 (0.096)	0.598 (0.439)	0.223 (0.078)	0.697 (0.404)
	2	0.324 (0.130)	0.750 (0.386)	0.138 (0.066)	0.900 (0.343)
E3	1	0.355 (0.083)	3.022 (0.082)	0.324 (0.096)	2.576 (0.109)
	1.5	0.467 (0.105)	2.419 (0.120)	0.220 (0.087)	3.082 (0.079)
	2	0.578 (0.123)	1.891 (0.169)	0.123 (0.072)	2.704 (0.100)
E4	1	0.371 (0.080)	0.423 (0.810)	0.341 (0.086)	0.003 (0.998)
	1.5	0.375 (0.097)	1.549 (0.461)	0.241 (0.080)	0.099 (0.952)
	2	0.261 (0.120)	1.994 (0.369)	0.159 (0.072)	0.317 (0.854)
E5	1	0.363 (0.080)	0.249 (0.883)	0.323 (0.086)	1.496 (0.473)
	1.5	0.374 (0.093)	0.585 (0.746)	0.218 (0.075)	1.799 (0.407)
	2	0.315 (0.119)	0.890 (0.641)	0.148 (0.062)	2.534 (0.282)
E6	1	0.337 (0.081)	4.002 (0.135)	0.288 (0.094)	3.527 (0.171)
	1.5	0.380 (0.097)	3.672 (0.159)	0.176 (0.074)	3.762 (0.152)
	2	0.405 (0.102)	3.335 (0.189)	0.110 (0.058)	2.973 (0.226)
E7	1	0.334 (0.079)	1.846 (0.397)	0.338 (0.085)	0.030 (0.985)
	1.5	0.301 (0.092)	3.549 (0.170)	0.210 (0.071)	1.040 (0.594)
	2	0.575 (0.133)	3.880 (0.144)	0.108 (0.050)	2.278 (0.320)
E8	1	0.337 (0.081)	1.882 (0.390)	0.330 (0.086)	0.483 (0.785)
	1.5	0.347 (0.094)	1.761 (0.415)	0.216 (0.073)	0.751 (0.687)
	2	0.269 (0.111)	2.267 (0.322)	0.136 (0.060)	0.930 (0.628)
E9	1	0.383 (0.083)	3.158 (0.206)	0.345 (0.090)	2.678 (0.262)
	1.5	0.414 (0.106)	2.632 (0.268)	0.277 (0.085)	3.397 (0.183)
	2	0.640 (0.128)	3.030 (0.220)	0.074 (0.044)	3.210 (0.201)
E10	1	0.314 (0.079)	3.380 (0.337)	0.255 (0.074)	2.103 (0.551)
	1.5	0.230 (0.080)	4.681 (0.197)	0.133 (0.047)	2.565 (0.464)
	2	0.585 (0.146)	4.179 (0.243)	0.297 (0.083)	7.004 (0.072)
E11	1	0.361 (0.085)	3.403 (0.334)	0.338 (0.094)	3.555 (0.314)
	1.5	0.476 (0.102)	2.297 (0.513)	0.287 (0.082)	4.720 (0.193)
	2	0.556 (0.123)	2.383 (0.497)	0.280 (0.078)	4.560 (0.207)
E12	1	0.350 (0.085)	2.871 (0.412)	0.307 (0.094)	2.606 (0.456)
	1.5	0.430 (0.103)	2.504 (0.475)	0.219 (0.077)	3.448 (0.328)
	2	0.511 (0.129)	2.421 (0.490)	0.194 (0.066)	3.277 (0.351)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

As depicted in table 3.16, the results for the estimation of the CEE's loss function using the EKT approach reflect the results under symmetric loss. For real-time realizations, rejections of forecast rationality can once again only be found in sets that contain nominal labor cost acceleration.

There are no rejections of the forecast rationality for revised results at all. As before, the α estimates differ remarkably with respect to the type of realization data. For revised data, all values of $\hat{\alpha}$ are significantly below 0.5, taking values around 0.23 whereas for real-time results, the average value of $\hat{\alpha}$ is 0.45 and thus none of the estimates differs significantly from 0.5. The α estimates with an accompanying J-statistic that indicates a rejection of rationality are the exceptions here. In these cases, a high J-statistic suggests that there is no estimate of α that is consistent with the rationality hypothesis. Therefore, these estimates are not interpreted here.

Table 3.17 shows analogous results for the JF forecast errors. Similar to the CEE results discussed above, all α estimates are significantly below 0.5, as far as the revised realization results are concerned. However this now holds for about one third of the results on real-time realizations as well. Similar to the results on CEE's forecasts, which only indicate a few rejections of rationality, there are only three rejections for JF's forecasts. Two of these rejections can be observed in sets E3 and E10 and for curvature parameters $p = 1.5$ and $p = 2$ for revised realizations. Rationality can be rejected once in set E3 and $p = 1$ for real-time realizations. All of these rejections only hold on a 10 percent level. Hence, the total number of rejections dramatically decreases once we allow for asymmetry in the loss function.

3.5.6 Instrument Set F: Financial Variables

To test if and how information from financial markets is incorporated in the institutions' forecasts, the growth rate of the DAX performance index (dax) along with the growth rates of a long-term and a short-term interest rate (ltr and str) are used as instruments. The structure of the twelve instrument sets under consideration remains unchanged.

F1	$\boldsymbol{w}_t = (1,\, dax_t)'$	F7	$\boldsymbol{w}_t = (1,\, dax_t,\, dax_t^2)'$
F2	$\boldsymbol{w}_t = (1,\, ltr_t)'$	F8	$\boldsymbol{w}_t = (1,\, ltr_t,\, ltr_t^2)'$
F3	$\boldsymbol{w}_t = (1,\, str_t)'$	F9	$\boldsymbol{w}_t = (1,\, str_t,\, str_t^2)'$
F4	$\boldsymbol{w}_t = (1,\, dax_t,\, dax_{t-1})'$	F10	$\boldsymbol{w}_t = (1,\, dax_t,\, ltr_t,\, dax_t \cdot ltr_t)'$
F5	$\boldsymbol{w}_t = (1,\, ltr_t,\, ltr_{t-1})'$	F11	$\boldsymbol{w}_t = (1,\, dax_t,\, str_t,\, dax_t \cdot str_t)'$
F6	$\boldsymbol{w}_t = (1,\, str_t,\, str_{t-1})'$	F12	$\boldsymbol{w}_t = (1,\, ltr_t,\, str_t,\, ltr_t \cdot str_t)'$

We start with the efficiency test results under symmetric loss, presented in table 3.18. Apart from sets F1 and F7 in the JF forecasts, for which efficiency cannot be rejected for either type of realization, all other tests of efficiency across institutions and types of realizations lead to a rejection of the hypothesis. In other words, there is strong evidence that the information contained in the financial variables chosen here is not used efficiently. However, efficiency is not rejected for sets F1 and F7 (both containing the DAX growth rate). It thus seems that the JF uses the information provided by this variable efficiently given the assumption of a symmetric loss function.

Table 3.18: Regression tests of efficiency (F)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
F1	6.737 (0.014)	1.014 (0.321)	7.692 (0.009)	1.728 (0.197)
F2	10.128 (0.003)	6.706 (0.015)	9.805 (0.004)	6.312 (0.017)
F3	9.838 (0.003)	12.981 (0.001)	8.408 (0.006)	10.220 (0.003)
F4	5.955 (0.006)	3.966 (0.029)	10.628 (0.000)	5.350 (0.010)
F5	10.443 (0.000)	6.103 (0.006)	8.287 (0.001)	5.299 (0.011)
F6	4.461 (0.019)	5.984 (0.006)	4.821 (0.014)	6.127 (0.005)
F7	5.021 (0.012)	0.910 (0.412)	3.824 (0.032)	1.486 (0.241)
F8	19.632 (0.000)	4.688 (0.017)	7.246 (0.003)	3.856 (0.032)
F9	6.292 (0.005)	10.044 (0.000)	4.140 (0.025)	7.845 (0.002)
F10	11.616 (0.000)	3.837 (0.020)	8.298 (0.000)	6.575 (0.002)
F11	6.145 (0.002)	7.210 (0.001)	5.662 (0.003)	5.408 (0.004)
F12	5.413 (0.004)	3.792 (0.021)	5.884 (0.003)	4.008 (0.017)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Table 3.19 shows the GMM results for the CEE's forecast errors. Although most of the $\hat{\alpha}$ are significantly below 0.5 for revised realization results, the rationality hypothesis still can be rejected for more than half of the results. Particularly the instrument sets including the growth of the long-term or short-term interest rate show high J-statistics and therefore suggest a rejection of the rationality hypothesis, (which, in some cases, only holds on a 10 percent level). The sets with only one or zero rejections of forecast rationality (i.e. sets F1, F4, F7 and F11) all contain the DAX instrument.

Turning to the first two columns of table 3.19 and the results using real-time realizations, all sets but F7, F8 and F12 lead to a rejection of rationality on a 5 percent level for all three curvature parameters. In these three sets, there are either less than three rejections, or rejections that only hold on a 10 percent level. Despite the high J-statistics, it is noteworthy that many of the α estimates take values above 0.5, in some cases even significantly. Once we allow for asymmetry in the loss function, the results nevertheless indicate that CEE uses information on DAX growth more efficiently than suggested by the results under symmetric loss.

A similar pattern can be seen in the JF results presented in table 3.20. Throughout the real-time realizations, most of the J-statistics again lead to a rejection of the forecasts' optimality in connection with the financial instruments used here. The results of the forecast errors related to the revised realizations imply a rejection for about half the instrument sets. Regardless of the type of realization, information on DAX growth appears to be used rather efficiently, as rationality is hardly ever rejected in sets F1, F4 and F7. Moreover, when the revised and the real-time realization results are compared, rationality cannot be rejected in sets F11 and F12 in the former and rejections also hold only on a 10 percent level more often here. The asymmetry parameters $\hat{\alpha}$ are considerably smaller for the revised data (they are significantly below 0.5), while the real-time estimates do not differ significantly from 0.5 in most cases and often take values even above 0.5.

Table 3.19: GMM estimates of the loss function - CEE (F)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
F1	1	0.438 (0.080)	4.366 (0.037)	0.244 (0.070)	1.048 (0.306)
	1.5	0.476 (0.090)	4.057 (0.044)	0.174 (0.062)	1.627 (0.202)
	2	0.499 (0.102)	4.178 (0.041)	0.159 (0.085)	2.144 (0.143)
F2	1	0.583 (0.084)	5.377 (0.020)	0.471 (0.086)	16.077 (0.000)
	1.5	0.696 (0.097)	6.259 (0.012)	0.441 (0.092)	11.819 (0.001)
	2	0.722 (0.106)	5.551 (0.018)	0.414 (0.101)	8.517 (0.004)
F3	1	0.566 (0.080)	6.736 (0.009)	0.250 (0.076)	4.358 (0.037)
	1.5	0.607 (0.093)	7.985 (0.005)	0.221 (0.080)	3.291 (0.070)
	2	0.297 (0.094)	6.153 (0.013)	0.216 (0.092)	2.154 (0.142)
F4	1	0.373 (0.080)	9.078 (0.011)	0.141 (0.062)	6.488 (0.039)
	1.5	0.376 (0.093)	7.510 (0.023)	0.100 (0.048)	4.476 (0.107)
	2	0.342 (0.102)	7.237 (0.027)	0.085 (0.072)	3.839 (0.147)
F5	1	0.632 (0.085)	12.004 (0.002)	0.462 (0.090)	16.384 (0.000)
	1.5	0.745 (0.096)	10.914 (0.004)	0.415 (0.091)	14.498 (0.001)
	2	0.784 (0.102)	8.807 (0.012)	0.422 (0.103)	10.809 (0.004)
F6	1	0.463 (0.080)	12.115 (0.002)	0.176 (0.071)	6.219 (0.045)
	1.5	0.279 (0.085)	10.711 (0.005)	0.087 (0.056)	5.973 (0.050)
	2	0.271 (0.093)	6.965 (0.031)	0.047 (0.056)	4.988 (0.083)
F7	1	0.491 (0.084)	4.985 (0.083)	0.252 (0.070)	1.447 (0.485)
	1.5	0.518 (0.091)	4.263 (0.119)	0.165 (0.066)	2.975 (0.226)
	2	0.544 (0.102)	4.286 (0.117)	0.144 (0.071)	3.910 (0.142)
F8	1	0.636 (0.080)	5.805 (0.055)	0.452 (0.087)	15.537 (0.000)
	1.5	0.702 (0.095)	6.708 (0.035)	0.417 (0.091)	11.235 (0.004)
	2	0.711 (0.108)	5.851 (0.054)	0.404 (0.102)	8.087 (0.018)
F9	1	0.563 (0.083)	6.468 (0.039)	0.329 (0.077)	5.688 (0.058)
	1.5	0.598 (0.092)	7.828 (0.020)	0.314 (0.082)	5.309 (0.070)
	2	0.629 (0.105)	7.759 (0.021)	0.304 (0.090)	4.763 (0.092)
F10	1	0.574 (0.090)	10.888 (0.012)	0.012 (0.027)	11.835 (0.008)
	1.5	0.655 (0.098)	9.278 (0.026)	0.262 (0.068)	16.957 (0.001)
	2	0.672 (0.106)	8.356 (0.039)	-0.030 (0.040)	6.839 (0.077)
F11	1	0.516 (0.070)	19.829 (0.000)	0.099 (0.052)	8.504 (0.037)
	1.5	0.672 (0.080)	18.428 (0.000)	0.107 (0.058)	5.156 (0.161)
	2	0.200 (0.087)	9.953 (0.019)	0.123 (0.085)	4.177 (0.243)
F12	1	0.649 (0.075)	6.045 (0.109)	0.416 (0.076)	18.508 (0.000)
	1.5	0.731 (0.078)	7.731 (0.052)	0.371 (0.076)	14.040 (0.003)
	2	0.737 (0.100)	7.295 (0.063)	0.374 (0.085)	9.797 (0.020)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

Table 3.20: GMM estimates of the loss function - JF (F)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
F1	1	0.289 (0.074)	4.717 (0.030)	0.286 (0.085)	2.215 (0.137)
	1.5	0.346 (0.095)	2.934 (0.087)	0.166 (0.064)	2.643 (0.104)
	2	0.317 (0.129)	1.235 (0.266)	0.096 (0.049)	2.390 (0.122)
F2	1	0.417 (0.086)	5.526 (0.019)	0.428 (0.104)	5.278 (0.022)
	1.5	0.636 (0.113)	6.514 (0.011)	0.323 (0.103)	5.241 (0.022)
	2	0.788 (0.181)	4.564 (0.033)	0.114 (0.080)	4.036 (0.045)
F3	1	0.420 (0.079)	7.098 (0.008)	0.316 (0.084)	3.504 (0.061)
	1.5	0.573 (0.102)	7.970 (0.005)	0.101 (0.051)	5.583 (0.018)
	2	0.815 (0.124)	4.843 (0.028)	0.360 (0.096)	6.569 (0.010)
F4	1	0.264 (0.073)	5.682 (0.058)	0.244 (0.084)	4.211 (0.122)
	1.5	0.287 (0.092)	4.304 (0.116)	0.125 (0.054)	4.050 (0.132)
	2	0.227 (0.107)	2.841 (0.242)	0.819 (0.138)	32.398 (0.000)
F5	1	0.156 (0.075)	12.405 (0.002)	0.143 (0.089)	7.595 (0.022)
	1.5	0.554 (0.100)	13.989 (0.001)	0.035 (0.043)	7.222 (0.027)
	2	0.599 (0.096)	15.024 (0.001)	0.331 (0.078)	13.537 (0.001)
F6	1	0.186 (0.069)	11.105 (0.004)	0.279 (0.083)	4.210 (0.122)
	1.5	0.165 (0.073)	8.649 (0.013)	0.103 (0.051)	5.425 (0.066)
	2	0.573 (0.115)	13.992 (0.001)	0.042 (0.031)	4.951 (0.084)
F7	1	0.300 (0.076)	6.352 (0.042)	0.300 (0.085)	2.562 (0.278)
	1.5	0.355 (0.094)	2.995 (0.224)	0.172 (0.065)	3.673 (0.159)
	2	0.311 (0.125)	1.301 (0.522)	0.100 (0.050)	3.055 (0.217)
F8	1	0.492 (0.090)	5.832 (0.054)	0.425 (0.105)	5.152 (0.076)
	1.5	0.629 (0.111)	6.443 (0.040)	0.343 (0.101)	5.309 (0.070)
	2	0.730 (0.121)	4.629 (0.099)	0.288 (0.103)	4.288 (0.117)
F9	1	0.454 (0.080)	7.626 (0.022)	0.377 (0.084)	5.132 (0.077)
	1.5	0.547 (0.094)	8.310 (0.016)	0.301 (0.077)	6.719 (0.035)
	2	0.707 (0.059)	6.403 (0.041)	0.304 (0.081)	7.214 (0.027)
F10	1	0.387 (0.085)	13.593 (0.004)	0.335 (0.092)	7.791 (0.051)
	1.5	0.547 (0.107)	10.561 (0.014)	0.006 (0.022)	9.370 (0.025)
	2	0.576 (0.115)	8.680 (0.034)	0.001 (0.009)	7.637 (0.054)
F11	1	0.096 (0.049)	13.779 (0.003)	0.223 (0.080)	4.569 (0.206)
	1.5	0.107 (0.062)	9.770 (0.021)	0.086 (0.044)	5.470 (0.140)
	2	0.614 (0.119)	16.578 (0.001)	0.042 (0.025)	4.783 (0.188)
F12	1	0.519 (0.090)	6.305 (0.098)	0.456 (0.106)	5.282 (0.152)
	1.5	0.637 (0.109)	6.176 (0.103)	0.403 (0.109)	5.355 (0.148)
	2	0.723 (0.084)	5.199 (0.158)	0.381 (0.115)	4.748 (0.191)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

3.5.7 Instrument Set G: Foreign Trade

The following set of instrumental variables consists of the foreign trade balance (ft), the growth rates of the euro-dollar exchange rate (xrt) and the oil price of crude petroleum (oil). The set is used to analyze how information on foreign trade is incorporated into the institutions' forecasts.

G1	$\boldsymbol{w}_t = (1,\, ft_t)'$	G7	$\boldsymbol{w}_t = (1,\, ft_t,\, ft_t^2)'$
G2	$\boldsymbol{w}_t = (1,\, xrt_t)'$	G8	$\boldsymbol{w}_t = (1,\, xrt_t,\, xrt_t^2)'$
G3	$\boldsymbol{w}_t = (1,\, oil_t)'$	G9	$\boldsymbol{w}_t = (1,\, oil_t,\, oil_t^2)'$
G4	$\boldsymbol{w}_t = (1,\, ft_t,\, ft_{t-1})'$	G10	$\boldsymbol{w}_t = (1,\, ft_t,\, xrt_t,\, ft_t \cdot xrt_t)'$
G5	$\boldsymbol{w}_t = (1,\, xrt_t,\, xrt_{t-1})'$	G11	$\boldsymbol{w}_t = (1,\, ft_t,\, oil_t,\, ft_t \cdot oil_t)'$
G6	$\boldsymbol{w}_t = (1,\, oil_t,\, oil_{t-1})'$	G12	$\boldsymbol{w}_t = (1,\, xrt_t,\, oil_t,\, xrt_t \cdot oil_t)'$

As presented in table 3.21, there is hardly any indication for the rejection of the efficiency hypothesis under quadratic loss across institutions and types of realization. Efficiency cannot be rejected at all in the CEE results, with the exception of a single rejection on a 10 percent level in set G11 and real-time data. For the JF forecast errors, there is one rejection each in set G4 regarding the revised realizations and in set G9 with respect to real-time realizations. The latter only holds on a 10 percent level. There is also a rejection for both types of realizations in set G11. These results do not give a clear picture about information that is possibly neglected in the JF's forecasts. Instrument set G11 contains the interaction term of the foreign trade balance and the growth rate of the oil price along with both individual variables. None of the instrument sets containing one of the two individual variables (with the exception of G4 for revised data and G9 for real-time data) strictly indicates an inefficient use of the information under symmetric loss.

Table 3.21: Regression tests of efficiency (G)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
G1	0.151 (0.700)	1.006 (0.323)	0.007 (0.935)	1.410 (0.243)
G2	0.003 (0.956)	0.177 (0.677)	0.056 (0.815)	0.108 (0.745)
G3	0.107 (0.745)	0.406 (0.528)	0.113 (0.739)	0.702 (0.408)
G4	0.460 (0.635)	2.156 (0.132)	0.992 (0.382)	4.452 (0.020)
G5	0.488 (0.618)	0.091 (0.913)	0.033 (0.968)	0.230 (0.796)
G6	0.237 (0.790)	0.227 (0.798)	0.062 (0.940)	0.374 (0.691)
G7	0.155 (0.857)	1.546 (0.228)	0.276 (0.761)	1.689 (0.200)
G8	0.568 (0.572)	0.091 (0.913)	0.551 (0.581)	0.945 (0.399)
G9	0.275 (0.761)	2.503 (0.097)	0.133 (0.876)	1.246 (0.300)
G10	0.947 (0.429)	0.785 (0.511)	0.098 (0.961)	0.872 (0.466)
G11	2.722 (0.061)	17.943 (0.000)	1.785 (0.170)	19.121 (0.000)
G12	0.201 (0.895)	0.960 (0.423)	0.584 (0.630)	1.617 (0.204)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

The results under symmetric loss show that efficiency is not rejected at all for the CEE forecasts. Accordingly, table 3.22 only shows two rejections of forecast rationality for a flexible loss function. These are to be found in set G9 for revised data, and both hold only on a 10 percent level. Surprisingly, all α estimates for the revised results are significantly below 0.5, despite the indication of symmetry shown above. This points toward a higher loss associated with negative forecast errors. The real-time results are more in line with the earlier results under symmetry, as none of the parameters differs significantly from 0.5 and one third of the estimates even takes values slightly above 0.5.

Table 3.22: GMM estimates of the loss function - CEE (G)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
G1	1	0.486 (0.083)	0.072 (0.788)	0.297 (0.083)	0.000 (0.997)
	1.5	0.502 (0.095)	0.055 (0.815)	0.247 (0.083)	0.000 (0.996)
	2	0.501 (0.108)	0.106 (0.745)	0.210 (0.081)	0.106 (0.745)
G2	1	0.468 (0.080)	0.445 (0.505)	0.287 (0.078)	0.056 (0.812)
	1.5	0.490 (0.091)	0.076 (0.783)	0.242 (0.083)	0.002 (0.962)
	2	0.494 (0.104)	0.026 (0.871)	0.224 (0.093)	0.005 (0.943)
G3	1	0.470 (0.080)	0.312 (0.577)	0.283 (0.077)	0.215 (0.643)
	1.5	0.487 (0.091)	0.104 (0.747)	0.239 (0.081)	0.059 (0.808)
	2	0.491 (0.106)	0.109 (0.741)	0.219 (0.091)	0.114 (0.735)
G4	1	0.463 (0.086)	2.290 (0.318)	0.230 (0.079)	1.608 (0.447)
	1.5	0.518 (0.098)	2.014 (0.365)	0.196 (0.078)	1.541 (0.463)
	2	0.527 (0.107)	1.012 (0.603)	0.188 (0.081)	1.145 (0.564)
G5	1	0.467 (0.081)	0.443 (0.801)	0.285 (0.079)	0.232 (0.891)
	1.5	0.487 (0.093)	0.211 (0.900)	0.220 (0.078)	0.587 (0.746)
	2	0.490 (0.107)	0.485 (0.785)	0.169 (0.079)	1.194 (0.550)
G6	1	0.467 (0.083)	0.623 (0.732)	0.286 (0.077)	0.716 (0.699)
	1.5	0.487 (0.091)	0.127 (0.939)	0.242 (0.080)	0.292 (0.864)
	2	0.473 (0.102)	0.479 (0.787)	0.223 (0.087)	0.143 (0.931)
G7	1	0.485 (0.084)	0.201 (0.904)	0.293 (0.082)	0.971 (0.616)
	1.5	0.505 (0.094)	0.524 (0.770)	0.246 (0.081)	0.020 (0.990)
	2	0.505 (0.110)	0.969 (0.616)	0.193 (0.074)	0.722 (0.697)
G8	1	0.470 (0.078)	0.648 (0.723)	0.334 (0.076)	2.177 (0.337)
	1.5	0.493 (0.091)	1.493 (0.474)	0.244 (0.080)	2.484 (0.289)
	2	0.499 (0.104)	0.642 (0.725)	0.115 (0.063)	2.616 (0.270)
G9	1	0.473 (0.084)	0.289 (0.865)	0.326 (0.077)	3.475 (0.176)
	1.5	0.481 (0.091)	0.364 (0.833)	0.303 (0.078)	5.854 (0.054)
	2	0.482 (0.103)	0.303 (0.859)	0.262 (0.083)	4.645 (0.098)
G10	1	0.502 (0.082)	1.936 (0.586)	0.302 (0.084)	0.649 (0.885)
	1.5	0.512 (0.097)	3.614 (0.306)	0.246 (0.083)	0.207 (0.976)
	2	0.504 (0.110)	2.998 (0.392)	0.204 (0.080)	0.189 (0.979)
G11	1	0.465 (0.085)	5.647 (0.130)	0.294 (0.082)	0.353 (0.950)
	1.5	0.504 (0.098)	3.555 (0.314)	0.267 (0.083)	1.669 (0.644)
	2	0.501 (0.111)	2.982 (0.394)	0.206 (0.078)	3.115 (0.374)
G12	1	0.488 (0.082)	3.560 (0.313)	0.279 (0.077)	0.733 (0.865)
	1.5	0.490 (0.094)	2.727 (0.436)	0.234 (0.079)	0.841 (0.840)
	2	0.502 (0.108)	1.991 (0.574)	0.222 (0.090)	0.803 (0.849)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

Turning to the JF results in table 3.23, the forecast rationality hypothesis can still hardly be rejected. There are three rejections on a 10 percent level for revised realization data only, one in set G1 and curvature parameter $p = 2$ and two in set G4 and $p = 1.5$ and $p = 2$. Both sets contain the foreign trade balance. Apart from this finding, the estimates of the asymmetry parameters α with respect to the revised data again take values significantly below 0.5 for all sets and curvature parameters with only three exceptions. In the real-time results, about one fourth of the α estimates are significantly below 0.5 (on a 10 percent level).

Table 3.23: GMM estimates of the loss function - JF (G)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
G1	1	0.374 (0.082)	1.697 (0.193)	0.349 (0.092)	1.231 (0.267)
	1.5	0.460 (0.102)	2.204 (0.138)	0.181 (0.070)	2.658 (0.103)
	2	0.580 (0.126)	1.733 (0.188)	0.078 (0.043)	3.156 (0.076)
G2	1	0.368 (0.079)	0.801 (0.371)	0.343 (0.087)	0.080 (0.778)
	1.5	0.378 (0.096)	0.953 (0.329)	0.228 (0.076)	0.624 (0.430)
	2	0.337 (0.131)	0.561 (0.454)	0.126 (0.055)	1.123 (0.289)
G3	1	0.350 (0.078)	1.162 (0.281)	0.320 (0.087)	2.085 (0.149)
	1.5	0.364 (0.094)	0.972 (0.324)	0.208 (0.075)	1.567 (0.211)
	2	0.305 (0.124)	1.132 (0.287)	0.131 (0.060)	1.300 (0.254)
G4	1	0.358 (0.083)	4.300 (0.117)	0.300 (0.096)	4.456 (0.108)
	1.5	0.508 (0.091)	2.906 (0.234)	0.339 (0.090)	5.238 (0.073)
	2	0.543 (0.086)	1.930 (0.381)	0.244 (0.076)	4.797 (0.091)
G5	1	0.364 (0.080)	0.996 (0.608)	0.344 (0.088)	0.428 (0.807)
	1.5	0.369 (0.095)	1.087 (0.581)	0.230 (0.077)	0.779 (0.677)
	2	0.323 (0.112)	0.582 (0.748)	0.131 (0.057)	1.032 (0.597)
G6	1	0.350 (0.079)	1.140 (0.566)	0.297 (0.085)	2.926 (0.232)
	1.5	0.359 (0.091)	0.960 (0.619)	0.199 (0.071)	1.704 (0.426)
	2	0.320 (0.115)	1.154 (0.562)	0.128 (0.056)	1.281 (0.527)
G7	1	0.414 (0.083)	2.890 (0.236)	0.423 (0.096)	3.346 (0.188)
	1.5	0.512 (0.082)	2.453 (0.293)	0.296 (0.082)	3.601 (0.165)
	2	0.542 (0.100)	1.758 (0.415)	0.179 (0.062)	3.840 (0.147)
G8	1	0.373 (0.080)	1.272 (0.529)	0.403 (0.089)	1.904 (0.386)
	1.5	0.377 (0.097)	1.078 (0.583)	0.269 (0.080)	2.262 (0.323)
	2	0.440 (0.119)	1.661 (0.436)	0.137 (0.057)	1.866 (0.393)
G9	1	0.362 (0.081)	1.197 (0.550)	0.333 (0.080)	2.387 (0.303)
	1.5	0.370 (0.088)	0.797 (0.671)	0.224 (0.065)	1.686 (0.430)
	2	0.321 (0.103)	1.285 (0.526)	0.146 (0.054)	1.236 (0.539)
G10	1	0.372 (0.083)	2.061 (0.560)	0.395 (0.094)	2.975 (0.395)
	1.5	0.376 (0.102)	2.904 (0.407)	0.262 (0.080)	3.544 (0.315)
	2	0.461 (0.124)	2.652 (0.448)	0.184 (0.066)	4.523 (0.210)
G11	1	0.342 (0.083)	2.795 (0.424)	0.301 (0.091)	2.590 (0.459)
	1.5	0.393 (0.099)	2.598 (0.458)	0.184 (0.070)	2.681 (0.443)
	2	0.434 (0.122)	2.857 (0.414)	0.128 (0.056)	3.334 (0.343)
G12	1	0.349 (0.079)	3.633 (0.304)	0.314 (0.088)	2.605 (0.457)
	1.5	0.359 (0.096)	2.566 (0.463)	0.182 (0.068)	2.290 (0.514)
	2	0.309 (0.125)	2.936 (0.402)	0.102 (0.048)	2.051 (0.562)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

On the one hand, the findings discussed above might indicate that the information provided by the foreign trade instruments has been used efficiently by the forecasting institutions. On the other hand, these results could simply indicate that these foreign trade instruments do not contain information that is helpful for further improving the institutions' employment forecasts.

3.5.8 Instrument Set H: Business Climate

The final set of instrumental variables contains three time series that have been derived from the ifo business survey. Specifically, the series are the growth rates of the business climate index (bc), the business sentiment index (bs) and the business expectation index (be).

H1	$\boldsymbol{w}_t = (1, bc_t)'$	H7	$\boldsymbol{w}_t = (1, bc_t, bc_t^2)'$
H2	$\boldsymbol{w}_t = (1, bs_t)'$	H8	$\boldsymbol{w}_t = (1, bs_t, bs_t^2)'$
H3	$\boldsymbol{w}_t = (1, be_t)'$	H9	$\boldsymbol{w}_t = (1, be_t, be_t^2)'$
H4	$\boldsymbol{w}_t = (1, bc_t, bc_{t-1})'$	H10	$\boldsymbol{w}_t = (1, bc_t, bs_t, bc_t \cdot bs_t)'$
H5	$\boldsymbol{w}_t = (1, bs_t, bs_{t-1})'$	H11	$\boldsymbol{w}_t = (1, bc_t, be_t, bc_t \cdot be_t)'$
H6	$\boldsymbol{w}_t = (1, be_t, be_{t-1})'$	H12	$\boldsymbol{w}_t = (1, bs_t, be_t, bs_t \cdot be_t)'$

Table 3.24 shows the efficiency results under symmetric loss. While efficiency is rejected across all instrument sets for both types of realizations for the CEE forecasts, there is one set (H3) for the JF forecasts without rejections of efficiency, regardless of the type of realization. In set H6 and for real-time realizations, efficiency is only rejected on a 10 percent level. All three sets without any or only with weak rejections are built around the growth of the business expectation index.

Table 3.24: Regression tests of efficiency (H)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
H1	28.292 (0.000)	4.666 (0.038)	15.200 (0.000)	8.425 (0.006)
H2	26.536 (0.000)	11.886 (0.001)	12.313 (0.001)	12.160 (0.001)
H3	13.307 (0.001)	0.083 (0.775)	9.336 (0.004)	0.470 (0.497)
H4	16.853 (0.000)	5.222 (0.011)	7.148 (0.003)	5.956 (0.006)
H5	14.714 (0.000)	5.463 (0.009)	6.959 (0.003)	5.626 (0.008)
H6	8.251 (0.001)	2.515 (0.096)	4.866 (0.014)	3.432 (0.044)
H7	12.793 (0.000)	8.052 (0.001)	7.798 (0.002)	9.554 (0.001)
H8	15.718 (0.000)	5.227 (0.010)	7.493 (0.002)	6.655 (0.004)
H9	7.858 (0.002)	3.312 (0.049)	4.835 (0.014)	3.944 (0.029)
H10	9.818 (0.000)	6.570 (0.001)	6.111 (0.002)	4.612 (0.008)
H11	9.645 (0.000)	5.079 (0.005)	5.955 (0.002)	6.169 (0.002)
H12	10.371 (0.000)	3.721 (0.021)	5.851 (0.003)	4.485 (0.010)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Under asymmetric loss the high share of rationality rejections still persists. Starting with the real-time realization results and CEE forecasts in table 3.25, all instrument sets again indicate an inefficient use of the information contained in the ifo data. Apart from one rejection in set H9 that only holds on a 10 percent level, all other rejections hold on a 5 percent level. Turning to the revised realizations, the results point in a similar direction, but are less conclusive. In five instrument sets there is not a single rejection of rationality for curvature parameter $p = 2$, and about half of the other rejections only hold on a 10 percent level. In spite of the fact that most J-statistics indicate a rejection of the rationality hypotheses, it is interesting to note that the majority of the α estimates takes values significantly below 0.5 for real-time as well as revised data.

Table 3.25: GMM estimates of the loss function - CEE (H)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
H1	1	0.261 (0.076)	14.736 (0.000)	0.101 (0.057)	8.954 (0.003)
	1.5	0.270 (0.085)	11.301 (0.001)	0.078 (0.043)	5.280 (0.022)
	2	0.252 (0.094)	8.270 (0.004)	0.055 (0.033)	3.489 (0.062)
H2	1	0.260 (0.077)	13.520 (0.000)	0.125 (0.060)	7.369 (0.007)
	1.5	0.282 (0.084)	9.863 (0.002)	0.096 (0.051)	4.435 (0.035)
	2	0.282 (0.093)	6.487 (0.011)	0.068 (0.042)	3.089 (0.079)
H3	1	0.442 (0.080)	6.842 (0.009)	0.243 (0.073)	4.754 (0.029)
	1.5	0.437 (0.090)	6.495 (0.011)	0.113 (0.059)	5.033 (0.025)
	2	0.495 (0.105)	6.107 (0.013)	0.064 (0.053)	3.673 (0.055)
H4	1	0.243 (0.081)	15.342 (0.000)	0.103 (0.058)	8.654 (0.013)
	1.5	0.253 (0.086)	11.380 (0.003)	0.077 (0.043)	5.892 (0.053)
	2	0.251 (0.093)	8.266 (0.016)	0.040 (0.028)	5.728 (0.057)
H5	1	0.263 (0.074)	13.551 (0.001)	0.124 (0.062)	7.549 (0.023)
	1.5	0.282 (0.087)	9.631 (0.008)	0.098 (0.052)	4.955 (0.084)
	2	0.286 (0.095)	6.445 (0.040)	0.070 (0.043)	4.047 (0.132)
H6	1	0.295 (0.078)	13.857 (0.001)	0.137 (0.062)	7.157 (0.028)
	1.5	0.297 (0.087)	10.336 (0.006)	0.089 (0.049)	5.305 (0.070)
	2	0.306 (0.100)	8.262 (0.016)	0.075 (0.043)	3.737 (0.154)
H7	1	0.256 (0.072)	18.927 (0.000)	0.051 (0.042)	11.502 (0.003)
	1.5	0.193 (0.081)	16.935 (0.000)	0.027 (0.020)	7.215 (0.027)
	2	0.171 (0.092)	12.080 (0.002)	0.284 (0.050)	32.516 (0.000)
H8	1	0.208 (0.073)	19.021 (0.000)	0.050 (0.042)	11.119 (0.004)
	1.5	0.174 (0.073)	16.537 (0.000)	0.036 (0.022)	6.997 (0.030)
	2	0.668 (0.102)	17.075 (0.000)	0.396 (0.071)	31.623 (0.000)
H9	1	0.455 (0.081)	7.209 (0.027)	0.108 (0.049)	8.900 (0.012)
	1.5	0.472 (0.092)	6.610 (0.037)	0.080 (0.036)	5.316 (0.070)
	2	0.475 (0.104)	5.694 (0.058)	0.063 (0.031)	3.576 (0.167)
H10	1	0.198 (0.075)	20.227 (0.000)	0.050 (0.041)	11.196 (0.011)
	1.5	0.156 (0.070)	17.142 (0.001)	0.024 (0.021)	7.270 (0.064)
	2	0.590 (0.105)	16.459 (0.001)	0.361 (0.065)	36.218 (0.000)
H11	1	0.251 (0.072)	15.685 (0.001)	0.068 (0.046)	10.864 (0.012)
	1.5	0.274 (0.083)	12.976 (0.005)	0.036 (0.021)	6.581 (0.087)
	2	0.321 (0.098)	9.690 (0.021)	0.012 (0.008)	5.423 (0.143)
H12	1	0.247 (0.072)	16.583 (0.001)	0.086 (0.052)	9.527 (0.023)
	1.5	0.259 (0.082)	12.737 (0.005)	0.032 (0.027)	7.238 (0.065)
	2	0.288 (0.096)	9.671 (0.022)	0.008 (0.013)	6.166 (0.104)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

The JF results shown in table 3.26 give a somewhat inconclusive picture. Once again, a large part of the rationality hypotheses can be rejected for both types of realizations. The pattern of the instrument sets, for which rationality cannot be rejected, does not indicate that any of the variables is used efficiently. However, a curvature parameter of $p = 2$ leads to J-statistic values for both real-time and revised results that are considerably smaller and hence only indicate a rejection of rationality on a 10 percent level or no rejection at all. As for the CEE results discussed in table 3.25, it is again noteworthy that all α estimates take values significantly below 0.5 and might thus indicate an asymmetric loss function, regardless of the type of realization under consideration. The α estimates are, however, predominantly accompanied by a J-statistic that suggests a rejection of forecast rationality.

Table 3.26: GMM estimates of the loss function - JF (H)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
H1	1	0.204 (0.071)	10.559 (0.001)	0.249 (0.081)	5.644 (0.018)
	1.5	0.232 (0.080)	7.127 (0.008)	0.109 (0.054)	5.781 (0.016)
	2	0.208 (0.101)	3.476 (0.062)	0.064 (0.039)	3.988 (0.046)
H2	1	0.223 (0.071)	8.498 (0.004)	0.263 (0.082)	4.320 (0.038)
	1.5	0.242 (0.083)	6.321 (0.012)	0.123 (0.058)	4.895 (0.027)
	2	0.207 (0.104)	3.711 (0.054)	0.069 (0.044)	3.628 (0.057)
H3	1	0.265 (0.072)	7.307 (0.007)	0.275 (0.082)	5.489 (0.019)
	1.5	0.285 (0.086)	5.738 (0.017)	0.111 (0.055)	6.413 (0.011)
	2	0.218 (0.100)	2.938 (0.087)	0.068 (0.041)	4.011 (0.045)
H4	1	0.192 (0.070)	10.790 (0.005)	0.241 (0.085)	5.566 (0.062)
	1.5	0.232 (0.085)	6.898 (0.032)	0.110 (0.056)	5.642 (0.060)
	2	0.217 (0.103)	4.039 (0.133)	0.064 (0.039)	3.929 (0.140)
H5	1	0.214 (0.071)	9.014 (0.011)	0.263 (0.086)	4.339 (0.114)
	1.5	0.242 (0.087)	6.270 (0.043)	0.122 (0.061)	4.773 (0.092)
	2	0.276 (0.107)	6.986 (0.030)	0.045 (0.042)	4.530 (0.104)
H6	1	0.248 (0.073)	7.942 (0.019)	0.236 (0.083)	6.648 (0.036)
	1.5	0.270 (0.089)	6.076 (0.048)	0.111 (0.056)	6.898 (0.032)
	2	0.216 (0.102)	2.943 (0.230)	0.071 (0.042)	4.539 (0.103)
H7	1	0.118 (0.058)	15.225 (0.000)	0.246 (0.078)	5.543 (0.063)
	1.5	0.232 (0.075)	10.405 (0.006)	0.067 (0.033)	7.649 (0.022)
	2	0.195 (0.084)	4.840 (0.089)	0.016 (0.013)	7.247 (0.027)
H8	1	0.158 (0.070)	12.120 (0.002)	0.260 (0.080)	4.198 (0.123)
	1.5	0.204 (0.074)	9.068 (0.011)	0.090 (0.040)	6.065 (0.048)
	2	0.198 (0.091)	4.265 (0.119)	0.029 (0.019)	5.590 (0.061)
H9	1	0.291 (0.071)	9.497 (0.009)	0.276 (0.081)	5.251 (0.072)
	1.5	0.306 (0.083)	6.378 (0.041)	0.134 (0.053)	8.910 (0.012)
	2	0.231 (0.098)	4.015 (0.134)	0.087 (0.042)	8.284 (0.016)
H10	1	0.129 (0.064)	13.888 (0.003)	0.252 (0.080)	5.814 (0.121)
	1.5	0.222 (0.073)	9.300 (0.026)	0.079 (0.036)	7.119 (0.068)
	2	0.220 (0.090)	4.433 (0.218)	0.022 (0.015)	5.858 (0.119)
H11	1	0.169 (0.058)	14.253 (0.003)	0.253 (0.079)	5.718 (0.126)
	1.5	0.244 (0.077)	9.572 (0.023)	0.063 (0.034)	9.359 (0.025)
	2	0.191 (0.089)	5.849 (0.119)	0.001 (0.011)	9.636 (0.022)
H12	1	0.166 (0.060)	13.062 (0.005)	0.255 (0.080)	5.825 (0.120)
	1.5	0.235 (0.079)	7.796 (0.050)	0.085 (0.039)	7.083 (0.069)
	2	0.217 (0.096)	4.012 (0.260)	0.021 (0.014)	5.769 (0.123)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

3.5.9 Summarized Results of the Instrument Sets

To give a brief overview of the results presented in this section, table 3.27 shows the absolute number of efficiency rejections for both institutions and types of realization data under symmetric and asymmetric loss. The instrument sets that include interactions between the variables (i.e. sets 10 to 12), which have been discussed in each of the preceding subsections, are left out here in order to prevent the results regarding one instrumental variable from being obscured or covered by those of another. An example is the rate of rejections in the CEE forecasts and the growth rate of government consumption in set C. The total rate of efficiency rejections would rise to 40 percent if the results of sets C11 and C12 were to be included. Efficiency can be rejected in these sets due to a strong indication for inefficient use of the growth rates of real GDP and real investment in these forecasts. For government consumption alone, efficiency is hardly ever rejected. Excluding the interaction terms, for each variable efficiency is thus tested in three different settings under symmetric loss and in nine settings under asymmetric loss because of the variants for the curvature parameters. Unbiasedness has been tested only once under symmetric loss.

Table 3.27: Absolute numbers of efficiency rejections

		CEE				JF			
set	variable	symmetric		asymmetric		symmetric		asymmetric	
		real-time	revised	real-time	revised	real-time	revised	real-time	revised
A	unbiasedness	0 (0)	1 (1)	-	-	0 (0)	1 (1)	-	-
	weak efficiency	0 (0)	2 (3)	0 (3)	0 (0)	3 (3)	1 (2)	1 (2)	0 (0)
B	*lf*	3 (3)	3 (3)	6 (7)	9 (9)	3 (3)	3 (3)	8 (9)	6 (9)
	ur	3 (3)	3 (3)	3 (4)	8 (9)	3 (3)	3 (3)	1 (3)	2 (6)
	rlc	1 (1)	1 (1)	0 (2)	0 (3)	0 (1)	1 (1)	1 (2)	0 (0)
C	*gdp*	3 (3)	3 (3)	6 (9)	9 (9)	3 (3)	3 (3)	8 (9)	8 (9)
	inv	3 (3)	3 (3)	0 (2)	9 (9)	3 (3)	3 (3)	1 (4)	7 (9)
	gov	0 (1)	0 (0)	0 (0)	0 (0)	1 (2)	0 (1)	0 (0)	0 (0)
D	*ord*	3 (3)	3 (3)	7 (8)	9 (9)	3 (3)	3 (3)	4 (8)	9 (9)
	bp	2 (3)	1 (1)	0 (0)	3 (5)	0 (0)	0 (0)	0 (0)	1 (6)
	rs	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)	1 (1)	0 (0)	0 (0)
E	*cpi*	0 (0)	0 (0)	0 (0)	0 (0)	1 (3)	1 (1)	0 (0)	0 (0)
	ppi	1 (1)	1 (1)	0 (0)	0 (0)	1 (3)	1 (2)	0 (0)	0 (0)
	nlc	2 (2)	2 (2)	0 (0)	6 (7)	3 (3)	0 (1)	0 (1)	0 (1)
F	*dax*	3 (3)	3 (3)	1 (1)	6 (7)	1 (1)	1 (1)	1 (1)	2 (4)
	ltr	3 (3)	3 (3)	9 (9)	7 (9)	3 (3)	3 (3)	6 (8)	8 (9)
	str	3 (3)	3 (3)	2 (8)	9 (9)	3 (3)	3 (3)	4 (8)	9 (9)
G	*ft*	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)	0 (3)	0 (0)
	xrt	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
	oil	0 (0)	0 (0)	0 (2)	0 (0)	0 (1)	0 (0)	0 (0)	0 (0)
H	*bc*	3 (3)	3 (3)	5 (9)	9 (9)	3 (3)	3 (3)	5 (8)	6 (8)
	bs	3 (3)	3 (3)	5 (7)	9 (9)	3 (3)	3 (3)	3 (6)	7 (8)
	be	3 (3)	3 (3)	4 (7)	7 (9)	1 (2)	2 (2)	7 (8)	6 (7)

Note: Number of rejections of the efficiency hypothesis on a 5 (10) percent level due to a high F-statistic (J-statistic) under symmetric (asymmetric) loss with 3 (9) tests for each variable. Unbiasedness was tested only once under symmetric loss.

Across all variables and institutions and in comparison to the results under symmetric loss, the share of rejected efficiency hypotheses is much smaller once we allow for asymmetry in the loss function. See table 3.27. Particularly in the case of the results using revised realization data about half of the inefficiencies under symmetric loss can be explained by a possible asymmetry in the loss function.

It seems that both institutions are very much alike in terms of the instrumental variables that bear pronounced results by showing either exceptionally high or low rejection rates. The growth rates of real labor cost, government consumption and the building permissions as well as the term spread and the growth rates of the consumer and the producer price indices, along with all variables in set G representing foreign trade (i.e. the foreign trade balance and growth rates of the euro-dollar exchange rate and the oil price) show only a few or no rejections of efficiency, regardless of the form of the loss function. The efficiency hypothesis is rejected in most tests under symmetric loss and less often under asymmetric loss for most of the other instrument variables discussed in subsections 3.5.1 to 3.5.8.

Instrumental variables with a high share of efficiency rejections for both institutions include the growth rates of labor force, the unemployment rate, GDP, investments, industrial orders, the long-term and the short-term interest rate, along with the growth rates of the ifo business climate index and its components, the business sentiment and expectations. As mentioned above, the share of rejections of efficiency drops for all of these variables once asymmetry is allowed for.

On the one hand, these common tendencies point out which variables appear not to contain any further information that could be exploited to improve future forecasts, either because the forecasting institutions already made use of all valuable information or because the variable simply does not contain any information that is relevant for predicting future employment growth. On the other hand, we find that both institutions potentially could improve their forecasts by trying to extract the additional information from the instruments for which efficiency is frequently rejected, even under asymmetric loss.

Although table A27 in the appendix, which shows analogous summarized results for the data including the post-crisis years until 2012, does not differ in general, there are two aspects in which the results vary that are worth mentioning. Using the growth rates of long-term and short-term interest rates, the number of efficiency rejections considerably declines when the time period is extended. This might indicate that the forecasting institutions paid more attention to these financial instruments in the wake of the financial crisis. Leaving the interest rates aside, the rate of efficiency rejections tends to increase in the results that include the post-crisis data. The higher volatility of some of the time series, along with more noise in the efficiency regressions and hence a potentially reduced test power all indicate potential increases in forecast uncertainty during the crisis. We have suspected this to be true from the beginning and this is the reason why we restricted the data to the pre-crisis period in the first place.

In the next section we use the entire set of variables discussed in this section to aggregate the information they contain using factor methods. These factors are then used as instruments in the same setting to analyze the forecast efficiency that we used in this section. Section 3.7 introduces the LASSO method and then applies it to pre-select a subset of instrumental variables, before extracting factors and conducting the forecast efficiency analysis with these factors.

3.6 Aggregate Information using Factor Analysis

Regarding the variety of different macroeconomic variables used as instruments in the preceding section, the wish to combine the information they contain in only a few variables and the intention of analyzing how well this combined information is used by the forecasting institutions comes naturally. The present section discusses the use of factor methods and their applicability in forecast evaluation. Factor methods offer a widely-applicable solution whenever one is trying to achieve a concentration of information and a substantial reduction of the number of variables dealt with. Subsequently, the first three factors obtained are used to analyze whether the forecasting institutions efficiently use the information these factors contain. Analogous to the approach used in section 3.5, these efficiency tests are conducted under a symmetric as well as an asymmetric loss function.

Using factors to bundle the information contained in a large number of macroeconomic time series has gained popularity in recent years. One reason for this might be the increasing amount of data available, along with the claim, especially from central banks, of being able to monitor as much of the available information as possible and take it into account when making a decision or producing a forecast. Bernanke and Boivin (2003) argue that another motivation for the use of factor methods can be that factors extracted from a large number of relevant variables might be interpreted as latent variables in a large system that possibly influence economic activity. For this reason, factors can be an appealing concept for central banks like the Fed, as they target the simulation of economic activity along with monetary stability.

The classical factor analysis focuses on a rather small number of variables $|N|$ and emphasizes estimating indices of the variables' covariation. Stock and Watson (2002a, 2002b) set the methodical groundwork for large $|N|$ and use factors to improve forecasts. The survey articles by Breitung and Eickmeier (2006), as well as Stock and Watson (2006, 2011), give a broad overview of the further developments in factor analysis for forecasting, whereby one focus lies on the construction of models that use dynamic factors. Recent studies by Kapetanios and Marcellino (2010) and Bai and Ng (2010) analyze the applicability of factors used as instruments and employ them in a GMM framework. Both show that using factors as instrumental variables instead of a large set of individual instruments leads to gains in efficiency.

The main focus here is on factors employed as instruments in the EKT forecast evaluation setting introduced in section 2.3. Herein, the choice of the instrumental variables that are used to extract the factors is mainly influenced by the studies of Boivin and Ng (2006) and Bai and Ng (2008). The former find that using larger datasets in terms of more time series does not necessarily lead to better factors and thus recommend pre-selecting the variables. The latter support these findings and argue in favor of an additional use for the squared instruments when constructing the factors to allow for potential non-linearities.

Hence, the dataset used to extract the factors contains 18 of the 21 variables introduced in sections 3.3 and 3.4 along with their squares. The three series excluded are the growth rate of the long-term interest rate (ltr), the term structure (rs) and the acceleration of the nominal labor costs per employee (nlc). Since no data are available before 1974 (ltr), 1973 (rs) and 1972 (nlc), they have been excluded to prevent the available, albeit short, time period from being shortened further. This leaves us a dataset containing 36 variables for the years 1971 to 2007, again excluding the years after 2007. Analogous results for the period including the post-crisis years 2008 to 2012 can be found in the appendix.

Although the number of variables in this study is rather small compared to others, results of Monte-Carlo studies, such as Kapetanios and Marcellino (2010), encourage the use of factor methods here. The authors find that using instruments derived as factors is to be preferred to standard instrumental variable estimation, even for a relatively small number of variables and observations over time (in their case $N = 30$ and $T = 50$). Moreover, the variables discussed in section 3.4 have been selected to cover a broad range of possible predictors in an effort to avoid redundancies. This can be understood as a pre-selection in the spirit of Boivin and Ng (2006) who argue as follows "[b]ecause the theory is developed for large N and T, there is a natural tendency for researchers to use as much data as are available. But in simulations and the empirical examples considered, the factors extracted from as few as 40 series seem to do no worse, and in many cases, better than the ones extracted from 147 series" (p. 189).

The factors were estimated using the principal components approach, which provides consistent estimates of the factors as proved by Stock and Watson (2002a, 2002b). This approach also offers a robust method in empirical applications and leads to similar or even better results than the more complex dynamic factor models (see Boivin and Ng, 2005). Relying on these results and because of the concern to further shorten the period of available observations by including various lags of the factors when applying a dynamic factor model, the present analysis restricts itself to the use of so-called static factor models.[17]

Applying the principle component approach, we find the first three factors to explain 48 percent of the total variance. Herein, the share of the variance is calculated as the sum of the first $|i|$ largest eigenvalues of the variables' covariance matrix divided by the sum of all eigenvalues. Accordingly, the individual factors explain 25 ($f1$), 13 ($f2$) and 10 ($f3$) percent of the variance. Our 48 percent can be regarded as quite a lot compared to Ludvigson and Ng (2009), who analyze bond risk premia using factors extracted from a macroeconomic dataset with 132 indicators of economic activity from 1964 to 2003 and find five factors to explain around 40 percent of the variation in their series.

The composition of the first three factors in terms of the highest factor loadings is as follows: for the first factor, the highest positive loadings come from ifo business climate and business sentiment indices, along with industrial order, real investment and the labor force growth rates. The unemployment rate and the squared unemployment rate as well as the squared real investment growth, the growth of government consumption and the CPI and PPI load negatively on the first factor. There is no instrument that loads negatively on the second factor in a meaningful way, whereas the loadings of GDP, industrial orders and real investment growth, the growth of real labor costs and the short-term interest rate along with their squares as well as the CPI are

[17] Fewer observations would be undesirable especially for the GMM approach used in the following.

positive. Considering the third factor, the highest positive loadings are the squares of GDP, industrial orders, short-term interest and labor force growth and the growth rates of the building permissions and the ifo business expectations. The highest negative loadings are found for the oil price growth rate and its square, along with the squares of the growth of building permissions and the ifo business expectation index.

According to the structure of instrument sets B to H discussed in section 3.5, set I contains the following instrumental variable subsets. The first three sets (I1 to I3) contain a constant together with factors one to three. The next three sets (I4 to I6) include each factor and its first lag in their respective sets. In sets I7 to I9, the lagged factors have been replaced by the squares of the factors, and sets I10 to I12 each contain two factors along with their interaction term.

I1	$\boldsymbol{w}_t = (1,\, f1_t)'$	I7	$\boldsymbol{w}_t = (1,\, f1_t,\, f1_t^2)'$
I2	$\boldsymbol{w}_t = (1,\, f2_t)'$	I8	$\boldsymbol{w}_t = (1,\, f2_t,\, f2_t^2)'$
I3	$\boldsymbol{w}_t = (1,\, f3_t)'$	I9	$\boldsymbol{w}_t = (1,\, f3_t,\, f3_t^2)'$
I4	$\boldsymbol{w}_t = (1,\, f1_t,\, f1_{t-1})'$	I10	$\boldsymbol{w}_t = (1,\, f1_t,\, f2_t,\, f1_t \cdot f2_t)'$
I5	$\boldsymbol{w}_t = (1,\, f2_t,\, f2_{t-1})'$	I11	$\boldsymbol{w}_t = (1,\, f1_t,\, f3_t,\, f1_t \cdot f3_t)'$
I6	$\boldsymbol{w}_t = (1,\, f3_t,\, f3_{t-1})'$	I12	$\boldsymbol{w}_t = (1,\, f2_t,\, f3_t,\, f2_t \cdot f3_t)'$

Table 3.28: Regression tests of efficiency (I)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
I1	16.328 (0.000)	25.439 (0.000)	10.156 (0.003)	22.374 (0.000)
I2	4.589 (0.039)	1.622 (0.211)	6.097 (0.019)	2.936 (0.096)
I3	0.004 (0.948)	0.001 (0.982)	0.093 (0.762)	0.001 (0.980)
I4	9.058 (0.001)	22.785 (0.000)	8.960 (0.001)	11.705 (0.000)
I5	7.442 (0.002)	7.572 (0.002)	7.098 (0.003)	7.879 (0.002)
I6	2.653 (0.086)	0.204 (0.816)	2.242 (0.123)	0.519 (0.600)
I7	12.495 (0.000)	19.598 (0.000)	9.009 (0.001)	23.500 (0.000)
I8	7.416 (0.002)	5.200 (0.011)	4.482 (0.019)	6.046 (0.006)
I9	7.168 (0.003)	38.321 (0.000)	5.459 (0.009)	28.887 (0.000)
I10	16.027 (0.000)	17.306 (0.000)	10.259 (0.000)	11.825 (0.000)
I11	10.083 (0.000)	23.804 (0.000)	10.806 (0.000)	17.754 (0.000)
I12	2.645 (0.066)	7.989 (0.000)	4.017 (0.016)	8.094 (0.000)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Table 3.28 shows the results of the efficiency tests under symmetric loss using the three factors as instrumental variables as described above. The null hypothesis of efficiency can be rejected on a 5 percent level regardless of the variant of realization data or the forecasting institution considered whenever the first factor is included in the instrumental variable set. The results become less evident for the second factor. For CEE's forecast errors, efficiency can be rejected for both variants of realizations for all sets that contain $f2$. For JF's forecasts, the efficiency hypothesis cannot be rejected in the set with the single second factor along with a constant (I2) for real-time data and is only weakly rejected (on a 10 percent level) for revised realizations. In sets I5, I8, I10 and I12, which include the lag or the quadratic term of the second factor and the other factors and their interaction terms with the second factor, a rejection of the efficiency hypothesis nevertheless is indicated for both institutions and types of realizations. The results are even more ambiguous for the third factor than for the second. On the one hand, apart from one weak rejection in set I6 for CEE's forecast errors subject to revised data efficiency cannot be rejected in sets I3 and I6 at all. This suggests that the forecasters have used all information contained in the third factor efficiently. On the other hand, efficiency is rejected across institutions and realizations in set I9, which contains the factor itself along with its square and a constant, and in sets I11 and I12, both of which contain the interaction terms with the other two factors. There is one exception: efficiency is rejected only on a 10 percent level for CEE's forecast errors and real-time data.

Under a more flexible and possibly asymmetric loss function, the efficiency results for the CEE's forecast change considerably for revised realizations, while the results for real-time data resemble those under symmetric loss, as shown in table 3.29. In the case of real-time realizations, most rejections of forecast rationality occur in sets that include the first and second factor, whereas no rejections of the rationality hypothesis occur in sets that only contain the third factor. None of the estimates of the asymmetry parameter α significantly differs from 0.5 when the accompanying J-statistic does not lead to a rejection of rationality. However, it is interesting to note that in cases that include the first factor, such as I1, I4 and I11, the average value of $\hat{\alpha}$ is 0.38, which is below the average $\hat{\alpha}$ value of 0.44 in sets I3, I6, I9 and I11, all of which contain the third factor. This indicates that some of the inefficiencies detected under symmetric loss can be overcome by allowing for a more general loss function. This finding also suggests that the CEE puts a higher weight on negative forecast errors due to the higher costs associated with overpredicting employment growth.

The results for revised realizations support this hypothesis. In this case, the degree of asymmetry in the loss function is even stronger, with α average estimates equal to 0.18. At the same time, rationality is rejected far less often, as shown in the last column of table 3.29. This striking difference may be explained by the tendency for higher forecast errors regarding revised realizations due to the systematic upward revision in the observed period, as discussed in section 3.2. A logical consequence is the difference in the significance of the estimates' deviations from 0.5 with respect to the type of realization. While all estimates are significantly below 0.5 on a 5 percent level for revised data, none of the estimates differs significantly on any common level when using real-time data. As discussed in section 2.3, an estimate of $\alpha = 0.45$ already represents a considerable loss difference of about 18 percent and therefore the potential economic significance should not be overlooked. Apart from the differences between the α estimates in the two realization variants, the decrease in the number of rationality rejections is remarkable for the revised realizations. There is only one rejection in the sets that do not include the second factor (I11, $p = 2$). Among the rejections in the sets including $f2$, more than half only hold on a 10 percent level.

Table 3.29: GMM estimates of the loss function - CEE (I)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
I1	1	0.404 (0.080)	4.760 (0.029)	0.258 (0.083)	2.416 (0.120)
	1.5	0.422 (0.085)	3.386 (0.066)	0.173 (0.076)	2.493 (0.114)
	2	0.402 (0.092)	2.590 (0.108)	0.133 (0.082)	2.373 (0.123)
I2	1	0.544 (0.083)	8.098 (0.004)	0.198 (0.069)	4.386 (0.036)
	1.5	0.345 (0.091)	9.002 (0.003)	0.133 (0.067)	4.642 (0.031)
	2	0.302 (0.104)	6.787 (0.009)	0.149 (0.086)	4.569 (0.033)
I3	1	0.485 (0.083)	0.210 (0.647)	0.292 (0.083)	0.673 (0.412)
	1.5	0.501 (0.095)	0.099 (0.753)	0.229 (0.082)	0.573 (0.449)
	2	0.504 (0.108)	0.010 (0.921)	0.181 (0.085)	0.822 (0.365)
I4	1	0.380 (0.079)	4.995 (0.082)	0.233 (0.078)	2.271 (0.321)
	1.5	0.400 (0.086)	3.361 (0.186)	0.137 (0.059)	2.940 (0.230)
	2	0.385 (0.098)	2.745 (0.253)	0.089 (0.049)	2.799 (0.247)
I5	1	0.641 (0.086)	15.799 (0.000)	0.117 (0.060)	6.428 (0.040)
	1.5	0.768 (0.102)	22.167 (0.000)	0.105 (0.065)	5.331 (0.070)
	2	0.974 (0.096)	17.615 (0.000)	0.103 (0.080)	4.867 (0.088)
I6	1	0.467 (0.087)	1.041 (0.594)	0.265 (0.086)	0.854 (0.652)
	1.5	0.453 (0.098)	1.067 (0.586)	0.218 (0.084)	0.938 (0.626)
	2	0.438 (0.109)	0.974 (0.614)	0.186 (0.085)	1.079 (0.583)
I7	1	0.312 (0.067)	14.932 (0.001)	0.221 (0.076)	4.193 (0.123)
	1.5	0.335 (0.071)	15.943 (0.000)	0.133 (0.058)	3.378 (0.185)
	2	0.350 (0.083)	9.728 (0.008)	0.087 (0.049)	2.919 (0.232)
I8	1	0.379 (0.084)	9.625 (0.008)	0.234 (0.072)	4.567 (0.102)
	1.5	0.324 (0.090)	9.046 (0.011)	0.193 (0.069)	5.328 (0.070)
	2	0.281 (0.099)	6.860 (0.032)	0.216 (0.086)	4.679 (0.096)
I9	1	0.461 (0.083)	1.482 (0.477)	0.295 (0.083)	0.707 (0.702)
	1.5	0.459 (0.096)	1.538 (0.463)	0.220 (0.083)	1.246 (0.536)
	2	0.429 (0.096)	1.630 (0.443)	0.184 (0.085)	1.560 (0.458)
I10	1	0.207 (0.052)	16.435 (0.001)	0.102 (0.065)	6.784 (0.079)
	1.5	0.191 (0.057)	13.549 (0.004)	0.063 (0.046)	5.749 (0.124)
	2	0.164 (0.063)	10.679 (0.014)	0.035 (0.038)	5.489 (0.139)
I11	1	0.360 (0.068)	8.160 (0.043)	0.184 (0.075)	4.468 (0.215)
	1.5	0.356 (0.073)	5.016 (0.171)	0.124 (0.071)	5.016 (0.171)
	2	0.369 (0.083)	2.959 (0.398)	0.135 (0.066)	6.263 (0.099)
I12	1	0.719 (0.071)	27.473 (0.000)	0.184 (0.070)	4.589 (0.205)
	1.5	0.846 (0.077)	17.773 (0.000)	0.121 (0.060)	4.611 (0.203)
	2	0.186 (0.065)	9.357 (0.025)	0.081 (0.061)	4.682 (0.197)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

Table 3.30 presents analogous results for JF's forecast errors. In accordance with the CEE's results, the second column on the left side of the table shows that the number of rejections of rationality is a lot higher for real-time realizations than for the revised data, which are shown in the last column on the right side. The degree of asymmetry in the loss function indicated by the values of $\hat{\alpha}$ is more pronounced in the latter, with most of the estimates significantly below 0.5.

Table 3.30: GMM estimates of the loss function - JF (I)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
I1	1	0.323 (0.082)	3.204 (0.073)	0.312 (0.095)	1.923 (0.165)
	1.5	0.335 (0.100)	2.243 (0.134)	0.192 (0.080)	1.780 (0.182)
	2	0.284 (0.126)	1.714 (0.191)	0.118 (0.063)	1.564 (0.211)
I2	1	0.327 (0.080)	4.705 (0.030)	0.246 (0.068)	2.340 (0.126)
	1.5	0.264 (0.094)	4.913 (0.027)	0.173 (0.065)	1.933 (0.164)
	2	0.310 (0.136)	3.001 (0.083)	0.138 (0.069)	1.275 (0.259)
I3	1	0.369 (0.082)	0.603 (0.437)	0.334 (0.091)	0.737 (0.391)
	1.5	0.368 (0.099)	1.145 (0.285)	0.209 (0.075)	1.215 (0.270)
	2	0.299 (0.127)	1.335 (0.248)	0.125 (0.060)	1.358 (0.244)
I4	1	0.281 (0.072)	3.622 (0.163)	0.262 (0.080)	2.504 (0.286)
	1.5	0.300 (0.082)	2.137 (0.344)	0.156 (0.068)	2.441 (0.295)
	2	0.272 (0.110)	1.531 (0.465)	0.102 (0.056)	1.767 (0.413)
I5	1	0.522 (0.080)	24.114 (0.000)	0.163 (0.063)	4.950 (0.084)
	1.5	0.679 (0.104)	22.214 (0.000)	0.064 (0.033)	5.008 (0.082)
	2	0.721 (0.100)	11.812 (0.003)	0.339 (0.081)	15.414 (0.000)
I6	1	0.320 (0.082)	2.398 (0.302)	0.285 (0.089)	1.363 (0.506)
	1.5	0.307 (0.094)	1.779 (0.411)	0.164 (0.067)	1.707 (0.426)
	2	0.263 (0.120)	1.550 (0.461)	0.091 (0.048)	1.781 (0.410)
I7	1	0.508 (0.075)	16.631 (0.000)	0.306 (0.103)	2.750 (0.253)
	1.5	0.649 (0.098)	15.583 (0.000)	0.221 (0.083)	5.618 (0.060)
	2	0.985 (0.013)	6.452 (0.040)	0.227 (0.080)	7.008 (0.030)
I8	1	0.264 (0.079)	6.925 (0.031)	0.248 (0.068)	2.299 (0.317)
	1.5	0.204 (0.085)	6.826 (0.033)	0.138 (0.056)	2.964 (0.227)
	2	0.789 (0.047)	5.185 (0.075)	0.065 (0.035)	3.390 (0.184)
I9	1	0.348 (0.082)	1.316 (0.518)	0.328 (0.091)	1.154 (0.561)
	1.5	0.362 (0.099)	1.231 (0.540)	0.212 (0.076)	1.190 (0.552)
	2	0.304 (0.126)	1.257 (0.533)	0.133 (0.060)	1.177 (0.555)
I10	1	0.150 (0.053)	10.965 (0.012)	0.207 (0.064)	4.165 (0.244)
	1.5	0.734 (0.109)	18.614 (0.000)	0.084 (0.033)	4.665 (0.198)
	2	0.765 (0.113)	8.743 (0.033)	0.379 (0.097)	11.597 (0.009)
I11	1	0.250 (0.066)	7.156 (0.067)	0.259 (0.072)	2.733 (0.435)
	1.5	0.251 (0.089)	5.008 (0.171)	0.132 (0.048)	3.013 (0.390)
	2	0.486 (0.156)	5.088 (0.165)	0.062 (0.028)	3.154 (0.368)
I12	1	0.683 (0.073)	26.408 (0.000)	0.202 (0.063)	4.106 (0.250)
	1.5	0.738 (0.112)	15.780 (0.001)	0.099 (0.037)	4.280 (0.233)
	2	0.701 (0.109)	10.697 (0.013)	0.382 (0.093)	10.597 (0.014)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

For revised realizations, all rejections of rationality on a 5 percent level can be observed for a curvature parameter of $p = 2$, along with two rejections on a 10 percent level in set I5 and one in set I7. For real-time data, rejections are robust for the entire instrument sets, apart from sets I1 and I11, which show only one rejection on a 10 percent level for $p = 1$. Regardless of the type of realization, rejections strongly depend on the factor that is used in the instrument set under consideration. Once again, rationality cannot be rejected in any of the sets containing only the third factor (I3, I6 and I9), while many rejections can be found in sets built around the second factor.

Although a certain amount of the efficiency rejections under a symmetric loss function is removed by allowing for a more flexible loss function, there is still evidence for inefficient use of the information that is contained in the factors derived from the 18 variables described above. For the post-crisis period, the results analogous to tables 3.28 to 3.30 are presented in tables A28 to A30 in the appendix.

3.7 Variable Pre-Selection using the LASSO Estimator

With the intention of further improving the results presented in the preceding section, the LASSO (least absolute shrinkage selection operator) estimator is applied to pre-select the variables that best explain the growth rate of the labor force. The LASSO was initially proposed by Tibshirani (1996). In the context of economic forecasting, it has been used by Bai and Ng (2008) and more recently by Li and Chen (2014).[18]

In a setting with many predictors, it can be useful to use model selection or shrinkage methods to increase the predictive accuracy in terms of a reduced variance compared to an OLS estimation at the cost of a slightly biased estimator. In other words, it is often desirable to find a subset of the available predictors with the strongest effects on the target variable. The traditional methods for improving the OLS estimates are subset selection and ridge regression, which both have their drawbacks. Tibshirani (1996) argues that subset selection is not necessarily robust to changes in the data, although its results can easily be interpreted. Both characteristics originate in the discrete choice to either keep a variable in the subset or drop it. In contrast, ridge regression that shrinks the coefficients provides the advantage of a higher robustness, but has the drawback that none of the coefficients is set to zero, which makes its interpretation more challenging. The LASSO has been designed to overcome the disadvantages of these methods and combine their advantages by shrinking some coefficients while setting others to zero.

The general idea of the LASSO is to minimize the sum of squared residuals resulting from the regression of the target variable y_t on the entire set of standardized predictors $x_{j,t}$, $j \in 1, ..., k$ constrained by a penalty term:

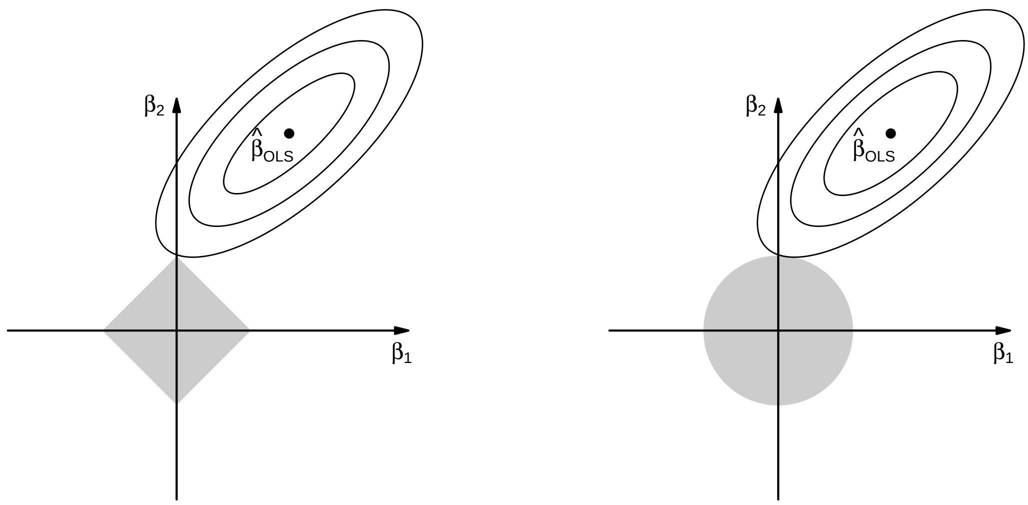
$$\hat{\boldsymbol{\beta}}^{\boldsymbol{LASSO}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \sum_{t=1}^{T} \left(y_t - \sum_{j=1}^{k} x_{j,t}\beta_j \right)^2 \text{ subject to} \sum_{j=1}^{k} |\beta_j| \leq \delta. \tag{14}$$

If the penalty parameter δ is small enough, all $\hat{\beta}_j$ will be set to zero, while a choice of a large δ where the OLS estimate is contained in the constraint region results in $\hat{\boldsymbol{\beta}}^{\boldsymbol{LASSO}} = \hat{\boldsymbol{\beta}}^{\boldsymbol{OLS}}$. For values of δ between these extremes, δ controls for the amount of shrinkage induced by the LASSO. The desirable property of setting certain coefficients to zero and thus, similar to subset selection, reduce the number of predictors, directly stems from the use of the L_1 norm in the

[18] Signal processing is another field where LASSO is used. Here it is referred to as basis pursuit (Chen et al., 1998).

LASSO penalty constraint. This can be seen clearly in comparison to the ridge regression, in which the sum of squared residuals is minimized subject to a L_2 penalty constraint, $\sum_{j=1}^{k} \beta_j^2 \leq \delta$. For $k = 2$, figure 3.11 shows that the constraint regions for the solution of $\hat{\boldsymbol{\beta}}$ is a circle in the case of the ridge regression and a diamond for the LASSO. Without a constraint, or with δ sufficiently large, the solution that minimizes the sum of squared residuals is $\hat{\boldsymbol{\beta}}^{\boldsymbol{OLS}}$. Once the OLS estimate lies out of the constraint region, the minimization problem is solved at the first point where the ellipses representing iso-sum-of-squared-residuals regions around $\hat{\boldsymbol{\beta}}^{\boldsymbol{OLS}}$ touch the circle or the diamond. In contrast to the ridge regression constraint, the LASSO constraint has corners at the axis. If the solution is taken in one of these corners, one of the coefficients is set to one, while the other shrinks to zero. In higher dimensions, the constraint of the ridge regression is a hypersphere that still does not have any corners, whereas the constraint region for the LASSO becomes a polytope that has many faces, flat edges and corners and thus, various possibilities for coefficients to be zero.

Figure 3.11: Constraint regions of LASSO and ridge regression



The minimization problem of the LASSO can also be expressed in Lagrangian form:

$$\hat{\boldsymbol{\beta}}^{\boldsymbol{LASSO}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \left\{ \sum_{t=1}^{T} \left(y_t - \sum_{j=1}^{k} x_{j,t}\beta_j \right)^2 + \lambda \sum_{j=1}^{k} |\beta_j| \right\}. \tag{15}$$

Although the solution of (14) has no closed form, efficient algorithms exist to compute the entire path of solutions while varying the penalty parameter λ in (15), as Hastie et al. (2009) point out. For the computation of the LASSO we rely on the R package "glmnet" described in Friedman et al. (2010) with λ chosen by fivefold cross-validation. In accordance with to Breiman and Spector (1992), who recommend fivefold and tenfold cross-validation in practice, tenfold cross-validation has also been applied, leading to almost identical results.

The subset of variables obtained when using the LASSO differs with respect to the type of realization. Not surprisingly, the growth rate of labor costs, along with the growth rate of real investments, are in the selected subset for both variants. For revised realizations, the growth rates of building permissions, CPI, DAX, the oil price, as well as the ifo business climate and sentiment indices have been further selected. The subset considering the real-time data only contains three more variables: the growth rates of the unemployment rate, government consumption and short-term interest.

Analogous to the proceeding in section 3.6, the variables selected by the LASSO as well as their squares are used to estimate the factors. Compared to the version without pre-selection, the share of variance explained by the first three factors increases considerably to 56 percent for the subset best suited for revised data and 69 percent for the real-time realization subset. After the extraction, the first three factors are employed to form the twelve instrument sets, now labeled J1 to J12, following the same structure as in section 3.6. The sets are then used to perform efficiency tests under symmetric and asymmetric loss.

Table 3.31: Regression tests of efficiency (J)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
J1	17.096 (0.000)	21.552 (0.000)	5.183 (0.029)	13.556 (0.001)
J2	0.381 (0.541)	0.348 (0.559)	0.951 (0.336)	0.216 (0.645)
J3	0.226 (0.637)	1.694 (0.202)	3.465 (0.071)	0.185 (0.670)
J4	8.275 (0.001)	18.794 (0.000)	15.066 (0.000)	12.838 (0.000)
J5	0.472 (0.628)	4.151 (0.025)	4.331 (0.022)	0.719 (0.495)
J6	0.298 (0.744)	8.821 (0.001)	1.716 (0.196)	0.119 (0.888)
J7	17.129 (0.000)	25.014 (0.000)	12.805 (0.000)	43.862 (0.000)
J8	1.711 (0.196)	1.771 (0.186)	29.639 (0.000)	34.416 (0.000)
J9	0.724 (0.492)	2.312 (0.115)	4.773 (0.015)	4.009 (0.028)
J10	9.839 (0.000)	13.248 (0.000)	16.594 (0.000)	45.656 (0.000)
J11	8.379 (0.000)	21.780 (0.000)	9.943 (0.000)	25.823 (0.000)
J12	2.046 (0.127)	3.479 (0.027)	4.123 (0.014)	5.370 (0.004)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Beginning with the results under the implied assumption of symmetric loss as depicted in table 3.31, it is interesting to note that an inclusion of the first factor in the instrument set, i.e. J1, J3, J7, J10 and J11, leads to a rejection of forecast efficiency regardless of the forecasting institution or the variant of realizations under consideration. For the other two factors, results are less robust concerning the inclusion of a specific factor. While rejections of efficiency only appear for CEE forecasts and within the scope of real-time realizations when the first factor is included, efficiency can be rejected for the same institutions' forecasts using revised realizations in all but two sets (J2 and J6), with one rejection merely on a 10 percent level. For JF's forecast errors, there are six instrumental sets in the results for both variants of realizations in which efficiency cannot be rejected: sets J2 and J3 for both types, along with J5 and J6 for revised and J8 and J9 for real-time realizations. Comparing the efficiency results to the results for the factors without variable pre-selection, the number of rejections drops slightly for all combinations of institutions and realizations, and even considerably in the case of real-time realizations and CEE forecast errors.

Table 3.32: GMM estimates of the loss function - CEE (J)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
J1	1	0.417 (0.074)	4.436 (0.035)	0.262 (0.080)	3.085 (0.079)
	1.5	0.426 (0.082)	3.008 (0.083)	0.123 (0.063)	3.760 (0.053)
	2	0.408 (0.094)	2.257 (0.133)	0.077 (0.060)	3.287 (0.070)
J2	1	0.545 (0.079)	3.222 (0.073)	0.231 (0.077)	4.681 (0.030)
	1.5	0.561 (0.075)	1.896 (0.169)	0.136 (0.063)	3.247 (0.072)
	2	0.542 (0.084)	0.606 (0.436)	0.091 (0.053)	2.523 (0.112)
J3	1	0.488 (0.083)	0.227 (0.634)	0.221 (0.073)	2.488 (0.115)
	1.5	0.516 (0.094)	0.275 (0.600)	0.194 (0.078)	1.782 (0.182)
	2	0.526 (0.104)	0.313 (0.576)	0.198 (0.094)	1.509 (0.219)
J4	1	0.394 (0.068)	4.182 (0.124)	0.214 (0.082)	2.848 (0.241)
	1.5	0.399 (0.080)	2.967 (0.227)	0.118 (0.057)	3.448 (0.178)
	2	0.385 (0.098)	2.342 (0.310)	0.077 (0.045)	3.164 (0.206)
J5	1	0.546 (0.082)	3.480 (0.176)	0.192 (0.075)	4.077 (0.130)
	1.5	0.597 (0.073)	3.182 (0.204)	0.133 (0.063)	3.068 (0.216)
	2	0.587 (0.077)	2.130 (0.345)	0.095 (0.052)	2.695 (0.260)
J6	1	0.468 (0.087)	1.617 (0.446)	0.196 (0.072)	2.900 (0.235)
	1.5	0.442 (0.092)	2.254 (0.324)	0.195 (0.079)	1.871 (0.392)
	2	0.442 (0.102)	2.112 (0.348)	0.188 (0.086)	1.518 (0.468)
J7	1	0.329 (0.065)	10.847 (0.004)	0.180 (0.064)	7.245 (0.027)
	1.5	0.325 (0.079)	11.512 (0.003)	0.079 (0.039)	4.712 (0.095)
	2	0.631 (0.086)	9.535 (0.009)	0.056 (0.030)	3.389 (0.184)
J8	1	0.547 (0.080)	3.448 (0.178)	0.232 (0.077)	4.545 (0.103)
	1.5	0.591 (0.073)	2.817 (0.245)	0.093 (0.046)	5.054 (0.080)
	2	0.584 (0.083)	3.236 (0.198)	0.082 (0.044)	2.571 (0.276)
J9	1	0.512 (0.082)	0.793 (0.673)	0.207 (0.068)	3.555 (0.169)
	1.5	0.535 (0.093)	0.713 (0.700)	0.190 (0.075)	2.624 (0.269)
	2	0.534 (0.105)	0.503 (0.778)	0.204 (0.092)	2.044 (0.360)
J10	1	0.524 (0.064)	18.791 (0.000)	0.184 (0.075)	5.416 (0.144)
	1.5	0.740 (0.060)	14.226 (0.003)	0.097 (0.055)	4.126 (0.248)
	2	0.741 (0.070)	11.242 (0.010)	0.040 (0.026)	4.683 (0.197)
J11	1	0.383 (0.065)	11.554 (0.009)	0.141 (0.050)	5.835 (0.120)
	1.5	0.415 (0.082)	11.621 (0.009)	0.075 (0.049)	5.044 (0.169)
	2	0.693 (0.085)	10.070 (0.018)	0.038 (0.032)	5.034 (0.169)
J12	1	0.502 (0.079)	6.636 (0.084)	0.114 (0.051)	7.388 (0.061)
	1.5	0.438 (0.082)	8.040 (0.045)	0.056 (0.032)	5.611 (0.132)
	2	0.366 (0.084)	4.928 (0.177)	0.637 (0.102)	31.768 (0.000)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

Table 3.32 shows the results for CEE's forecast errors, once asymmetry in the loss function is allowed for. Efficiency is rejected less often in comparison to the results under symmetry, although the rejections still agglomerate in instrument sets built around the first factor (i.e. in sets J1 and J7 for both types of realizations and only in sets J10 and J11 for real-time realizations). More efficiency rejections can be observed in sets J2 and J12 for both variant of realizations, and a single rejection can be observed in set J8 for revised data only. Considering the different curvature parameters in the instrument sets, rejections are rather stable in the results on real-time data and mostly hold on a 5 percent significance level. The opposite holds true for revised realizations. Only three of the rejections are significant on a 5 percent level, and J1 is the only instrument set in which efficiency can be rejected for all three curvature parameters.

Table 3.33: GMM estimates of the loss function - JF (J)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
J1	1	0.325 (0.081)	3.063 (0.080)	0.302 (0.092)	2.486 (0.115)
	1.5	0.327 (0.098)	2.278 (0.131)	0.185 (0.074)	2.173 (0.140)
	2	0.277 (0.121)	1.843 (0.175)	0.113 (0.057)	1.817 (0.178)
J2	1	0.375 (0.082)	1.834 (0.176)	0.351 (0.093)	0.001 (0.975)
	1.5	0.434 (0.097)	0.876 (0.349)	0.243 (0.083)	0.191 (0.662)
	2	0.430 (0.123)	0.407 (0.524)	0.154 (0.069)	0.524 (0.469)
J3	1	0.379 (0.081)	0.362 (0.547)	0.293 (0.082)	1.262 (0.261)
	1.5	0.462 (0.085)	0.665 (0.415)	0.196 (0.070)	1.227 (0.268)
	2	0.530 (0.084)	0.955 (0.329)	0.148 (0.070)	0.789 (0.374)
J4	1	0.279 (0.065)	3.803 (0.149)	0.276 (0.094)	2.226 (0.329)
	1.5	0.298 (0.077)	2.225 (0.329)	0.170 (0.076)	1.981 (0.371)
	2	0.262 (0.109)	1.708 (0.426)	0.104 (0.057)	1.625 (0.444)
J5	1	0.365 (0.086)	1.839 (0.399)	0.291 (0.090)	0.942 (0.624)
	1.5	0.409 (0.100)	0.862 (0.650)	0.189 (0.075)	0.983 (0.612)
	2	0.405 (0.115)	0.350 (0.839)	0.117 (0.058)	1.047 (0.592)
J6	1	0.339 (0.082)	3.235 (0.198)	0.246 (0.084)	2.994 (0.224)
	1.5	0.448 (0.094)	2.830 (0.243)	0.120 (0.057)	3.857 (0.145)
	2	0.520 (0.093)	2.383 (0.304)	0.077 (0.043)	3.260 (0.196)
J7	1	0.644 (0.074)	27.407 (0.000)	0.284 (0.090)	3.783 (0.151)
	1.5	0.752 (0.091)	17.195 (0.000)	0.122 (0.059)	5.242 (0.073)
	2	0.975 (0.018)	5.799 (0.055)	0.063 (0.038)	4.847 (0.089)
J8	1	0.317 (0.075)	2.824 (0.244)	0.293 (0.087)	3.168 (0.205)
	1.5	0.527 (0.093)	3.503 (0.173)	0.181 (0.066)	2.686 (0.261)
	2	0.613 (0.094)	2.993 (0.224)	0.113 (0.050)	2.077 (0.354)
J9	1	0.395 (0.082)	0.961 (0.619)	0.289 (0.077)	1.277 (0.528)
	1.5	0.473 (0.080)	0.809 (0.667)	0.192 (0.068)	1.319 (0.517)
	2	0.529 (0.075)	0.934 (0.627)	0.149 (0.070)	0.807 (0.668)
J10	1	0.074 (0.034)	16.376 (0.001)	0.279 (0.089)	4.516 (0.211)
	1.5	0.388 (0.096)	22.251 (0.000)	0.152 (0.066)	3.890 (0.274)
	2	0.355 (0.120)	16.780 (0.001)	0.092 (0.048)	2.922 (0.404)
J11	1	0.193 (0.078)	13.816 (0.003)	0.246 (0.082)	4.209 (0.240)
	1.5	0.120 (0.083)	12.531 (0.006)	0.110 (0.049)	4.855 (0.183)
	2	0.074 (0.067)	9.481 (0.024)	0.055 (0.030)	4.554 (0.208)
J12	1	0.199 (0.058)	7.813 (0.050)	0.238 (0.077)	3.563 (0.313)
	1.5	0.263 (0.065)	5.098 (0.165)	0.131 (0.050)	3.308 (0.347)
	2	0.339 (0.082)	4.830 (0.185)	0.073 (0.033)	2.920 (0.404)

Note: Standard errors for $\hat{\alpha}$ and p-values for the J-test statistics are shown in parentheses.

Considering the α estimates, the pronounced difference between real-time and revised results found in the preceding section persists. $\hat{\alpha}$ does not significantly differ from 0.5 in any of the real-time results when the accompanying J-statistic does not lead to a rejection of forecast rationality. In about half the cases the value of $\hat{\alpha}$ is even greater than 0.5. All estimates subject to revised data are significantly below 0.5.

The results of JF's forecast errors, presented in table 3.33, show even fewer rejections of forecast rationality than the CEE's. Only two rejections can be found in the results for revised realizations, both in set J7 and both only on a 10 percent significance level. There are some more rejections in the results for real-time realizations. All but one (in set J12) can be found in sets that are built around the first factor. Specifically, there is one rejection on a 10 percent level in set J1, along with three rejections each in sets J7, J10 and J11. Although there are some more estimates of α that are significantly below 0.5 for JF's real-time data results, the estimates mostly do not to differ significantly. The estimates for revised data are again all significantly below 0.5.

In comparison to the results without variable pre-selection, it is noteworthy that the number of rejections of forecast rationality drops in every combination of realization variant and forecasting institution. This tendency toward less rejections of forecast rationality is supported by Bai and Ng's (2008) finding that estimating factors after pre-selecting the variables leads to more precise forecasts, as this reduces the influence of uninformative predictors. However, the reduction in the number of rejections may also indicate that pre-selecting the variables could lead to omissions of potentially valuable information that otherwise might further improve the forecasts. As in the previous sections, the results including the post-crisis years can be found in tables A31 to A33 in the appendix.

3.8 Summary

In this chapter, the efficiency of the employment growth forecasts of the German Council of Economic Experts (CEE) and the Joint Forecast (JF) of the leading German economic research institutes have been analyzed in detail. The analysis has been conducted under symmetric and quadratic loss as well as under a potentially asymmetric loss function using the flexible EKT approach. In section 3.5, we analyzed the weak efficiency as well as the information efficiency with respect to 21 instrumental variables. Herein, the variables were looked at individually and we were able to find efficiency rejections for various variables. Subsection 3.5.9 at the end of section 3.5 provides an extended summary of these results.

Sections 3.6 and 3.7 used the information contained in the individual variables jointly, by extracting three factors that are then used as instrumental variables. While in section 3.6 the factors were extracted right away, the LASSO approach was used in section 3.7 to pre-select a subset of variables beforehand. We find strong evidence against efficiency under symmetric loss; when we allowed for asymmetry in the loss function, many of the inefficiencies dissolved.

Under asymmetric loss, efficiency can be rejected more often using the factors without pre-selection than for the factors extracted from the pre-selected subset of variables. Instrument sets that contain the first factor are most likely to show rejections of forecast efficiency in both approaches. The second factor only appears to contain unused information when variables are not pre-selected, while the information in the third factor seems to be used rather efficiently, regardless of whether or not variables are pre-selected.

In all three sections, we found indications for asymmetry in the preferences of both institutions. The forecasters seem to favor positive forecast errors (underpredicting employment growth) over negative forecast errors. However, the degree of asymmetry in the forecasters loss function also depends on the choice of the realization of the target variable, as the results indicate a higher degree of asymmetry in the loss function when revised realizations are considered.

4 A Monte-Carlo Analysis of the EKT Loss Function

4.1 Motivation

Forecast evaluation exercises that utilize sophisticated statistical methods increasingly are becoming standard. In particular, the EKT approach introduced in section 2.3, has been widely used in the literature (see section 2.4 for an overview). Because many macroeconomic time series only are observed quarterly of even just annually, the number of observation in a forecast evaluation exercise is often rather limited. Hence, knowledge of the finite sample properties of those forecast evaluation procedures is very important for the interpretation of their results.[19]

There exists quite a number of methodological and applied papers on forecast evaluation that also report the results of less extensive Monte-Carlo exercises for evaluating finite sample properties. EKT themselves, Capistrán (2005), Christodoulakis and Mamatzakis (2009) and Naghi (2015) are to be mentioned in this respect. Most of these papers only focus on test size and neglect test power, with Capistrán (2005) being the notable exception. Missing to date, however, is a more systematic comprehensive investigation of the finite sample properties of the EKT procedure, that likewise looks at the estimation of the asymmetry parameter as well as the size and power of the test for efficiency. The main aim of this study is to provide such an extensive Monte-Carlo investigation of the finite-sample properties of the EKT procedure.

Our analysis explores a wide range of different scenarios in order to shed light on the EKT procedure's ability to detect asymmetry in the loss function as well as the size and power properties of the associated test for forecast efficiency. The scenarios are specifically designed to investigate the effects of the forecast errors' statistical properties (i.e. variance, autocorrelation, fat tails) simultaneously with variations of the information set available to the forecaster (i.e. inducing omitted and irrelevant variables). In addition, we induce a single large outlier to mimic a major crisis analogous to the Great Recession. Moreover, while most Monte-Carlo exercises are based on parameter values which are quite arbitrarily chosen or inspired by previous studies in the literature, we also use the forecast errors from a predictive regression equation estimated with real data to obtain more realistic parameter values and thereby move closer to evaluating the properties of the EKT procedure in applied situations.

The following investigation begins in section 4.2 with a review of previous work upon which we build. Section 4.3 describes the design of the Monte-Carlo experiment. In the following sections the results are presented using a specific graphical device which we call the *fishbone plot*. The presentation starts with the discussion of the results from a baseline scenario in section 4.4, moves on to the implementation of various extensions in section 4.5 and concludes with the switch to the more realistic scenario based on real data in section 4.6. Section 4.7 summarizes the main results.

[19] This chapter is based on Hoss and Krüger (2015).

4.2 Previous Work

Apart from EKT themselves, Capistrán (2005), Christodoulakis and Mamatzakis (2009) and more recently Naghi (2015) perform Monte-Carlo simulations using the EKT approach. The setting of the short Monte-Carlo experiment by EKT will be described in the following section, as we build on their analysis. Their main finding is that under a piecewise linear loss function, the size of the t-test, testing the hypothesis that the α estimates equal their true values, is well controlled if only a constant is used as an instrument. There are some size distortions in the piecewise quadratic case as well as in cases when two additional instruments are included. Furthermore EKT find the J-test to be slightly undersized, especially for high degrees of asymmetry. In all cases, size distortions are reduced by increasing the sample size. Power is not scrutinized in their work.

Capistrán (2005) conducts a Monte-Carlo experiment which focuses on the power of the J-test. He simulates the target variable and one instrument from a bivariate normal distribution, drawing the actual data from a distribution conditional on the instrument. He generates the forecasts as the quantile of this conditional distribution, fixing the degree of asymmetry and directly controlling the bias. Then he obtains the J-statistic by estimating the asymmetry parameter using the EKT approach and finally calculates the power of the J-test at a test size of 5 percent. Overall, he finds good power for the J-test, although he remarks that for a sample size of $T = 100$ and a high degree of asymmetry, the test has no power for certain values of the induced bias. Moreover, he observes that the power in the highly asymmetric cases is mostly below the power reached in the symmetric case.

Christodoulakis and Mamatzakis (2009) conduct another Monte-Carlo experiment that uses the same design as EKT but focuses on how skew-normal error terms affect the size of the J-test. They find the J-test to be robust in this regard. The power of the J-test is not subject to their analysis.

Finally, Naghi (2015) finds that the EKT approach fails to provide correct estimates for the asymmetry parameter when the true underlying loss function is not contained in the family of loss functions proposed by EKT and the approach is still used to evaluate the forecasts. To demonstrate this, Naghi uses the Linex loss function (see equation (8) in section 2.2). In these cases, and especially for a strong departure from symmetry, the J-test shows size distortions. In a second Monte-Carlo experiment Naghi replaces the piecewise linear function with a data generating process that contains a linear and a nonlinear component, while maintaining a linear forecast equation.

Our experimental design builds on these previous investigations and is particularly oriented toward EKT's design, which is a flexible and yet straightforward approach to drawing inference on the asymmetry in the forecasters' preferences. The next section outlines the experimental design in greater detail.

4.3 Design of the Experiment

The experiment is designed with particular emphasis on the validity of the parameter estimates (in particular of α) as well as the size and power properties of the J-test, considering different specifications of the weighting matrix $\boldsymbol{S}$ used for the GMM estimation and variations in the error term.

We thus analyze five specification variants of $\boldsymbol{S}$ for our experiment, following EKT's data generating process that is described below. Building on this, we select two of the variants for further analysis, focusing on how changes in the forecast error, such as higher variance and autocorrelation, affect the properties of the GMM estimation.

For the five baseline specifications of $\boldsymbol{S}$, we use the identity matrix, a diagonal matrix with iid elements that allows for heteroscedasticity and finally three variants that use heteroscedasticity and autocorrelation consistent (HAC) covariance matrices, as proposed by Newey and West (1987). The three HAC variants differ in the selection of the bandwidth. Starting with a fixed bandwidth of 1, we continue computing the bandwidth using the approaches suggested by Andrews (1991) and Newey and West (1994).[20]

Furthermore, we use the continuously updated estimator proposed by Hansen et al. (1996) whenever the estimation of $\boldsymbol{S}$ is necessary. As shown by Newey and Smith (2004), this estimator appears to have better finite-sample properties compared to the more common two-stage GMM. For the numerical optimization, we use the quasi-Newton method simultaneously proposed by Broyden, Fletcher, Goldfarb and Shanno (BFGS, 1970), as this is a frequently used robust method that works well in one-dimensional optimization problems.

In the following, we construct a forecast error series for each of the variants described above. Herein, we limit our forecast horizon to $h = 1$. We mainly adopt the setup of the data generating process that EKT use in their Monte-Carlo study and define our data generating process with the following steps:

Step 1: We start by simulating a set of three instrumental variables of length T, which are normally distributed, $w_{1,t} \sim N(1,1)$, $w_{2,t} \sim N(-1,1)$ and $w_{3,t} \sim N(0,1)$. Then the target variable is calculated as

$$y_{t+1} = \boldsymbol{\theta}' \boldsymbol{w}_t + u_t, \tag{16}$$

where $\boldsymbol{w}_t$ contains a constant as well as $w_{1,t}$ and $w_{2,t}$. The parameters are adopted from EKT and are fixed at $\boldsymbol{\theta} = (1, 0.5, 0.5)'$. $u_t \sim N(0, 0.5)$ is a normal iid variate.

Step 2: The sample is split into an in-sample and an out-of sample set, where R represents the number of observations that are available to the forecaster prior to the first forecast, while P is the number of forecasts made, as described in Step 3. The latter are used to compute forecast errors in order to evaluate these forecasts.

Step 3: Given a set of variables from the information set available to the forecaster and the forecaster's true preferences towards asymmetry, α_0, the forecasts are made by estimating $\hat{\boldsymbol{\theta}}$

[20] We use the quadratic spectral kernel that is described in Andrews (1991) for all variants. Preliminary simulations with a Bartlett kernel showed generally inferior results compared to the quadratic spectral kernel. Size distortions where much larger with the Bartlett kernel.

using quantile regression (Koenker and Basset, 1978) for $p = 1$ and expectile regression (Newey and Powell, 1987) methods for $p = 2$.[21] Considering the growing information set, $\hat{\boldsymbol{\theta}}$ is estimated recursively, successively adding one observation to the information set for each new forecast,

$$f_{t+1} = \hat{\boldsymbol{\theta}}_{\boldsymbol{t}}{}' \boldsymbol{w}_{\boldsymbol{t}}. \tag{17}$$

The forecast errors are denoted by $e_{t+1} = y_{t+1} - f_{t+1}$, with $t + 1 = 1, ..., P$ and are computed as the difference between the target variable (16) and the forecasts (17).

Step 4: Here the GMM estimation of the EKT loss function is performed. The instrumental variables thus are held fixed for all replications. From the GMM estimation, we record $\hat{\alpha}$ as well as the J-test statistic and its p-value.

Steps 1 to 4 are executed for $M = 5000$ replications with the sample sizes $R \in \{50, 100, 100, 200\}$ and $P \in \{50, 50, 100, 100\}$. In order to cover a wide range from extremely asymmetric preferences, over values close to symmetry up to the symmetric case, the true value of the asymmetry parameter is set to $\alpha_0 \in \{0.2, 0.4, 0.5, 0.6, 0.8\}$. The information used to estimate $\hat{\boldsymbol{\theta}}$ in step 3 varies from (A) only a constant, to (B) $\boldsymbol{w}_{\boldsymbol{t}} = (1, w_{1,t})'$, (C) $\boldsymbol{w}_{\boldsymbol{t}} = (1, w_{1,t}, w_{2,t})'$, (D) $\boldsymbol{w}_t = (1, w_{1,t}, w_{2,t}, w_{3,t})'$ and (E) $\boldsymbol{w}_t = (1, w_{1,t}, w_{2,t}, w_{1,t} \cdot w_{2,t})'$. From the results, we analyze the distribution of $\hat{\alpha}$ as well as the size and power properties of the J-test. Note that these information sets comprise omitted information (A, B) as well as irrelevant information which are either uncorrelated (D) or correlated (E) to the other variables. Only set C specifies the correct information set. We expect mainly a large power of the J-test for the sets with the omitted information and may find size distortions for the sets with the irrelevant information.

The procedure described above is first carried out for each of the five baseline scenarios with respect to different specifications of the weighting matrix $\boldsymbol{S}$. Out of these, we can select two of the weighting matrix scenarios. The analysis is then expanded to check the EKT approach's robustness toward a higher degree of uncertainty in the target variable, the presence of autocorrelation in one or both instruments, and hence in the target variable, as well as its robustness in the presence of outliers.

4.4 Baseline Simulations

This section presents the results of the five baseline experiments. In section 4.5, we focus on two scenarios for a more detailed analysis of the estimates of α along with the size and power of the J-test in the presence of more noise and fat tails in the error term distribution as well as autocorrelation and outliers in the target variable and the instruments. Finally, we simulate a multivariate model that is parametrized to reflect the relationship between the unemployment rate, GDP growth and inflation in order to analyze the properties of the EKT test in a more realistic setting. These results are discussed in section 4.6.

To avoid the excessive use of tables that normally arises in Monte-Carlo studies, we have chosen a concise graphical representation of our results, summarizing the results of each scenario in one

[21] For the computation, we rely on the R packages "quantreg" described in Koenker (2005) and "expectreg" respectively.

figure in the following way: every figure is divided into five panels, A to E, which are structured identically and represent the different sets of information used in the forecast regression. Each panel is based on the same sequence of random draws, assured by fixing the seed of the random number generator. On the left side of each panel, we find the results for the Lin-Lin case $p = 1$, while on the right side the results for the Quad-Quad case $p = 2$ are shown. For both curvature parameters, five values of the true asymmetry parameter, denoted by horizontal lines, and four combinations of in-sample size and out-of-sample size are given, with sample sizes becoming larger from left to right in each block.

The mean values of the α estimates across all Monte-Carlo replications are plotted as bullets with an interval of two standard deviations extending vertically. The horizontal line represents α_0. The rejection frequencies of the J-test are represented by gray vertical bars. Since both the values of α and the J-test rejections take values between zero and unity, a common scale is used. The dotted horizontal line at 0.05 marks the size of the test. Because of the optical appearance of the α results, we have named this plot the *fishbone plot*.

For the first scenario, the identity matrix is used as weighting matrix in the GMM optimization and its results are depicted in figure 4.1. In panel A, which shows the results when only a constant is used in the forecasting regression, we find that the values of α are estimated quite well both with respect to bias and precision. However, when the sample size increases, the estimates do not systematically get closer to the true values. We also observe very high rejection frequencies for the J-test, which points to good power of the test.

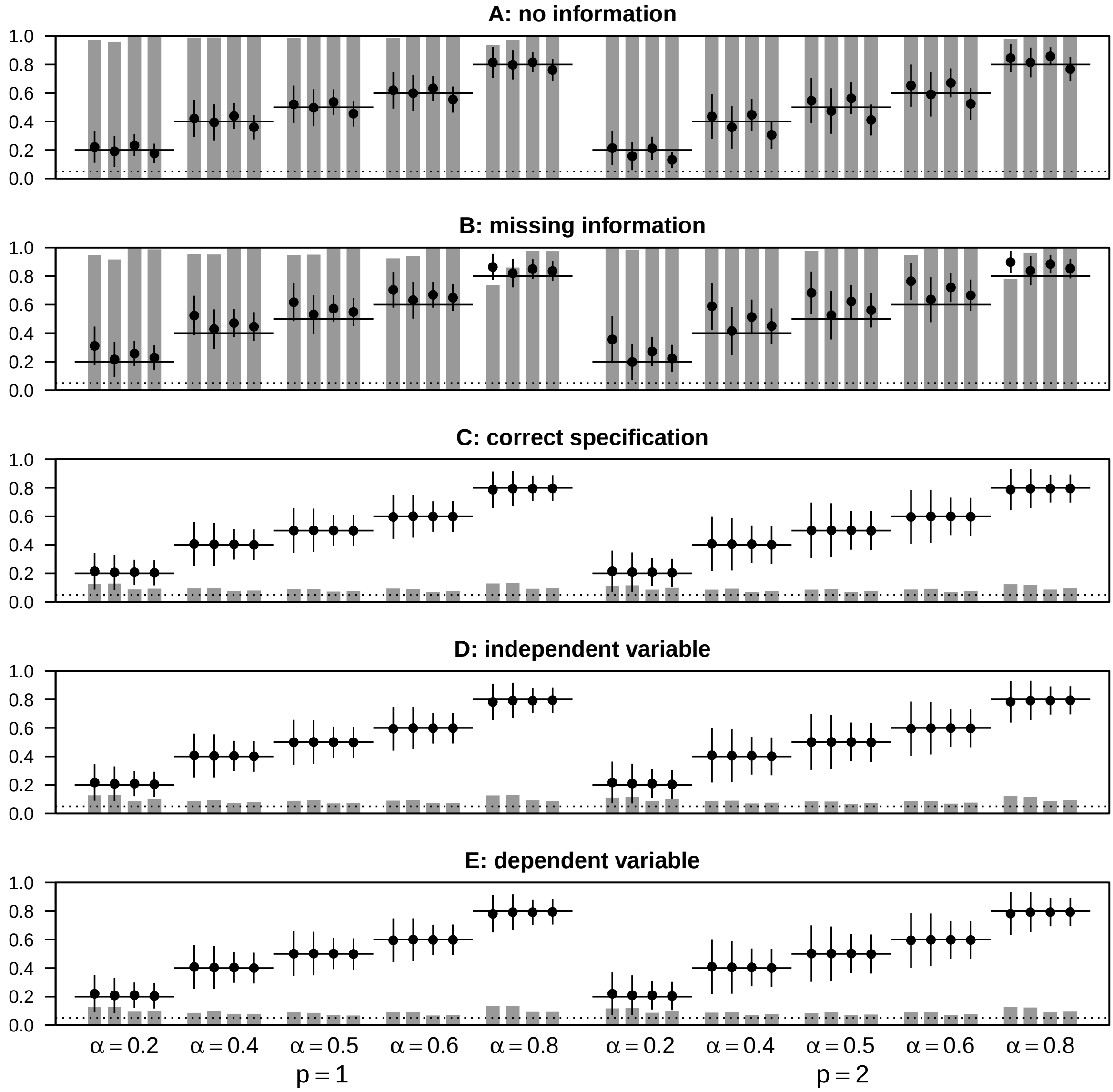
A similar pattern can be observed in panel B, for which the forecasts have been made using a constant and $w_{1,t}$. The relevant information in $w_{2,t}$ is thus contained in the information set used by the J-test for the forecast evaluation, but is neglected in the construction of the forecasts (equation (17)).[22] Here, a tendency to systematically overestimate α can be observed, in contrast to panel A. Moreover the J-test has slightly less power for smaller sample sizes. This outcome is quite reasonable regarding the enlarged information set compared to the previous case.

Panel C exhibits the results for the correctly specified forecasts, i.e. a constant together with $w_{1,t}$ and $w_{2,t}$ in the forecast equation. Here, we find the mean values of α to be very close to the true values, although we observe some variation nevertheless. The J-test shows a slight tendency to overreject the null of rationality of the forecasts. However, the size distortions tend to decrease with increasing sample sizes and for values of α closer to symmetry, i.e. $\alpha_0 \in \{0.4, 0.5, 0.6\}$.

This pattern of results remains robust when a further variable outside of the information set is included in the forecast regression. Herein we regard variables that are either independent of or dependent on $w_{1,t}$ and $w_{2,t}$, by adding $w_{3,t}$ to the forecast regression in panel D and including $w_{1,t} \cdot w_{2,t}$ in panel E.

[22] Keep in mind that the variables used in the data generating process (equation (16)) are fixed for all experiments.

Figure 4.1: Fishbone plot using the identity matrix for $\boldsymbol{S}$



The left side of figure 4.2 depicts the results of the second scenario, in which we allow for heteroscedasticity in $\boldsymbol{S}$. Looking at panels A and B, we find a tendency to underestimate small values of α and to overestimate the larger ones. Thus, we observe a systematic bias away from the symmetric case. For $p = 2$ the variation of the α's is larger than before for all values of α_0 we consider, whereas for $p = 1$, the increase in variation can be especially observed in the more symmetric cases, i.e. $\alpha_0 \in \{0.4, 0.5, 0.6\}$. The power of the J-test is somewhat reduced, although it remains high and is more pronounced for smaller sample sizes and a higher degree of asymmetry.

The overall pattern of panels C to E is very similar to the one in the first scenario, as far as the mean estimates and the variation of α are concerned. Regarding the rejections of the J-test, the size properties improve and the rejection frequencies are closer to the nominal value of 0.05.

The results for the HAC weighting matrix for a bandwidth fixed at 1 are presented on the right side of figure 4.2. Regarding the α estimates and the power of the J-test in panels A and B, we find the same pattern as in the previous figure. In panels C to E, the size properties of J-test appear to be further improved, albeit by a small margin.

Turning to the results using a HAC weighting matrix with a bandwidth choice according to Andrews, as shown on the left side of figure 4.3, the picture remains essentially unchanged, except for a slightly larger variation in the values of the α estimates.

In the fifth scenario, we use a HAC weighting matrix as well, but have chosen a bandwidth according to Newey and West. Using the Newey and West bandwidth selection, we observe occasional convergence problems in the GMM estimation of the loss function, especially when $p = 2$. We also observe α estimates outside the interval of zero and unity. Therefore, we have eliminated those cases on the right side of figure 4.3. We find the same pattern for the power of the J-test as in panels A and B, with a slightly larger power for small sample sizes. However, in panels C to E, size distortions appear to be larger now with a more pronounced tendency to overreject the null hypothesis of forecast rationality.

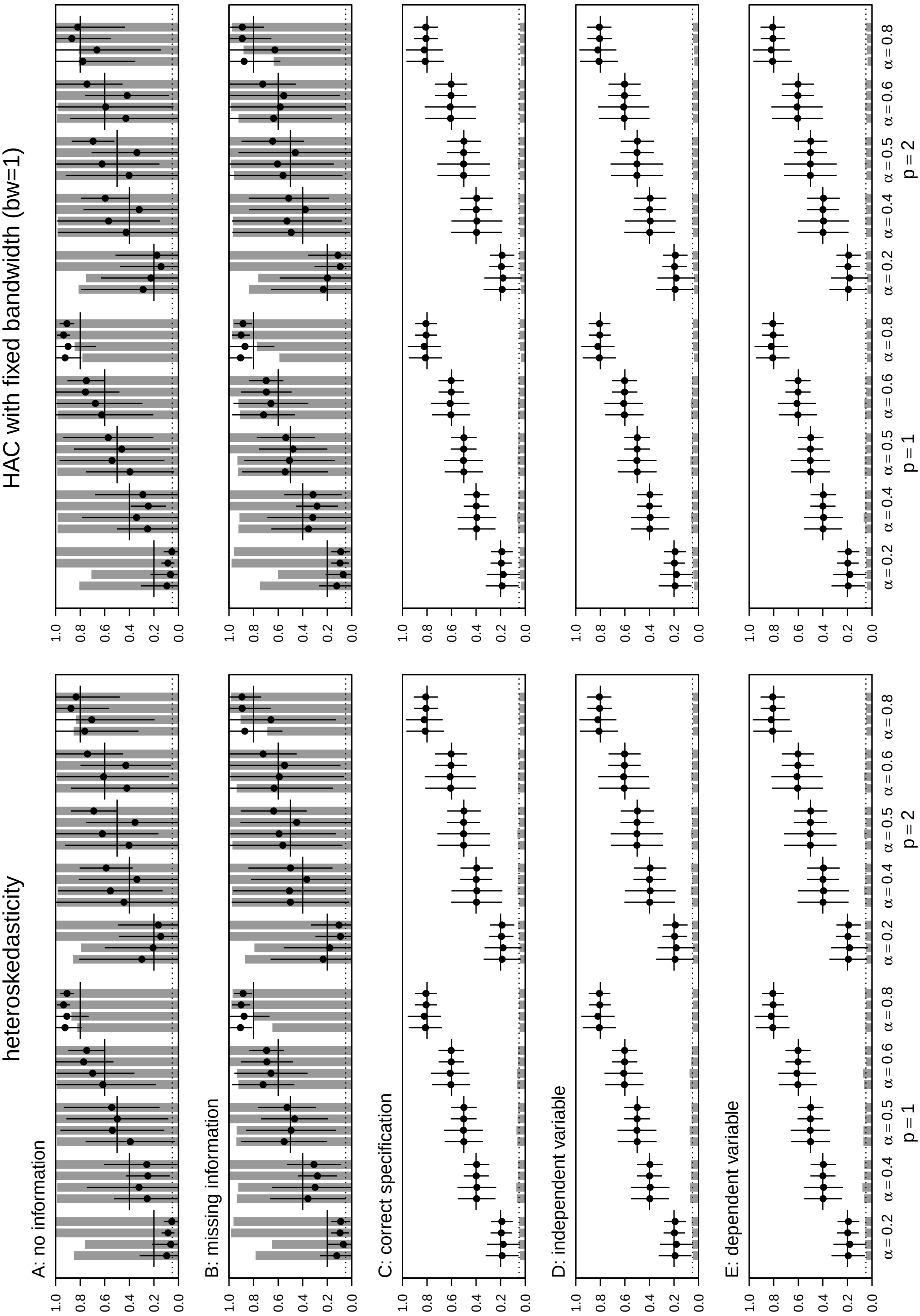
To summarize the findings of this section, we see that the GMM estimation of the EKT loss function leads to quite favorable results. The asymmetry parameter α is precisely estimated and precision increases with increasing sample size when the information set is correctly specified, or contains irrelevant information. The estimates are further away from the nominal values when information is omitted, although the general tendency is also met here. The J-test is quite capable of detecting inefficient use of information and reaches very high rejection frequencies in most of the cases considered. Exceptions can only be found for the smaller sample sizes. If the information set comprises irrelevant information, we observe a slight tendency to overreject. This tendency is more pronounced when a greater asymmetry is induced in the loss function (cases $\alpha = 0.2$ and $\alpha = 0.8$). The differences in the Lin-Lin ($p = 1$) and Quad-Quad ($p = 2$) cases appear to be minor. The choice of the weighting matrix is somewhat more crucial, but not really decisive in this setting.

In order to reach a more compact presentation of the results, we have selected two of the five weighting matrices for further investigation. This can be well founded in the results. First, we stick to the identity matrix because it is the simplest choice and does not require the estimation of any parameters. We neglect the heteroskedasticity variant because forecast evaluation exercises are usually performed in a time series context and thus HAC weighting matrices are the more natural choice. Second, we stick to the HAC variant with the bandwidth choice according to Andrews because the results are about the same as for the fixed bandwidth and are superior to the results obtained with the bandwidth choice according to Newey and West. Moreover, no convergence problems have been observed when using the Andrews bandwidth choice.

4.5 Variations of the Setting

So far, all results have been obtained under the clinical setup of a data generating process with observations that have been drawn independently and an error term that has induced a

Figure 4.2: Fishbone plots - baseline results I



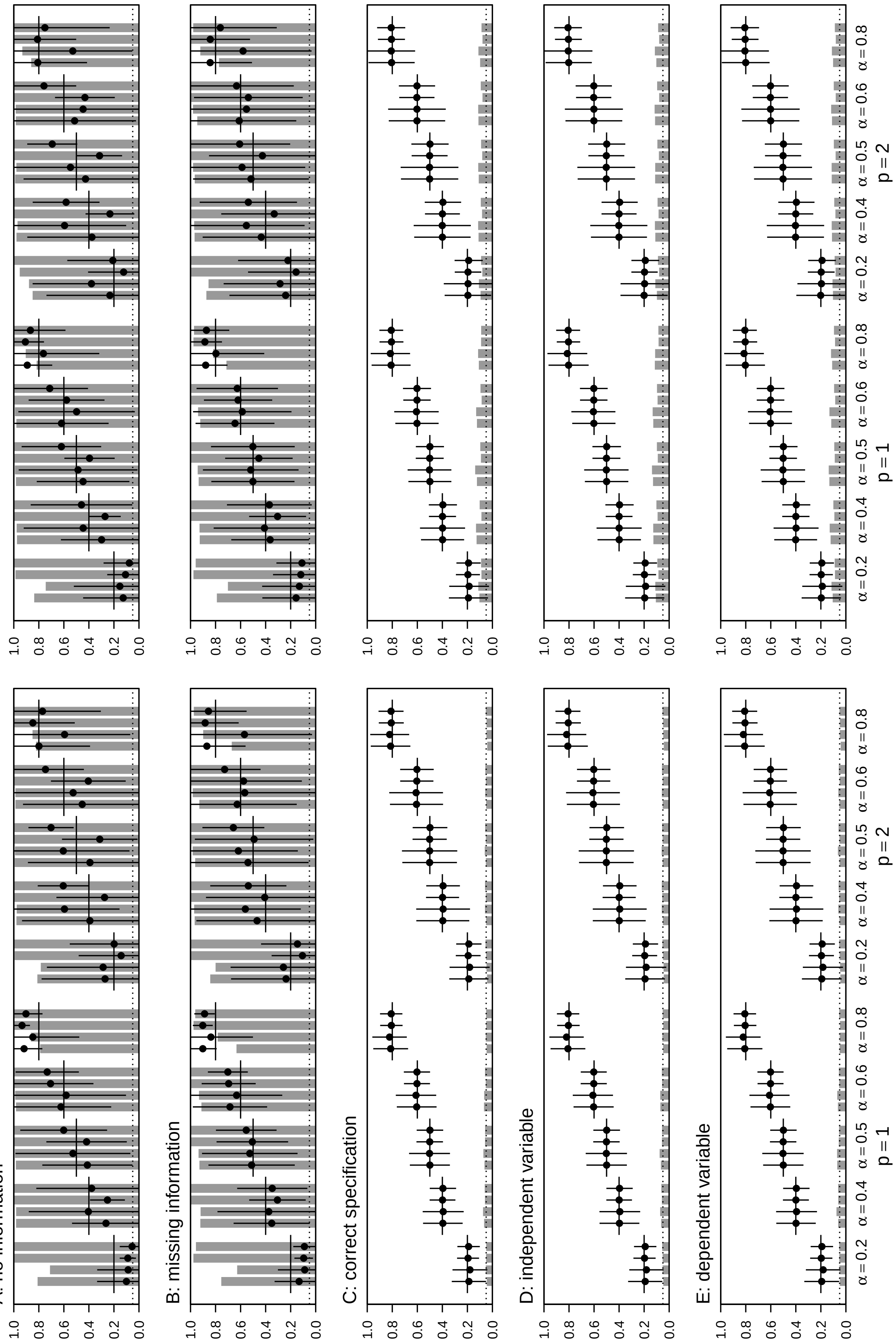


Figure 4.3: Fishbone plots - baseline results II

moderate amount of noise in the data. In order to learn more about the robustness of the EKT test's properties in a more realistic data environment, we now gradually change the DGP.

In subsections 4.5.1 and 4.5.2, we start by only varying the distribution of the error term used to compute the target variable, while still drawing the $w_{i,t}$ variables independently. In subsection 4.5.1, we increase the variance of the normally distributed error term and thus induce more noise into the process. In subsection 4.5.2, we draw the error terms from a t distribution with 4 degrees of freedom to investigate how the behavior of the J-test changes in the presence of a fat tailed error term distribution which assigns a higher probability to more extreme events.

Subsequently, the $w_{i,t}$ variables are generated as AR(1) processes to induce autocorrelation into the target variable, in subsections 4.5.3 and 4.5.4. While in subsection 4.5.3, the focus lies on how the autocorrelation structure affects the outcomes of the EKT test, in subsection 4.5.4, a shock in form of a single outlier is added to one of the $w_{i,t}$ variables in order to analyze the test's behavior when an unexpected crisis disturbs the regular path of the time series.

Later, in section 4.6, we simulate a multivariate model in order to mimic the relationship between the unemployment rate and lagged GDP growth and lagged inflation rate and conduct our experiment with these data both with and without an outlier.

4.5.1 Higher Error Term Variance

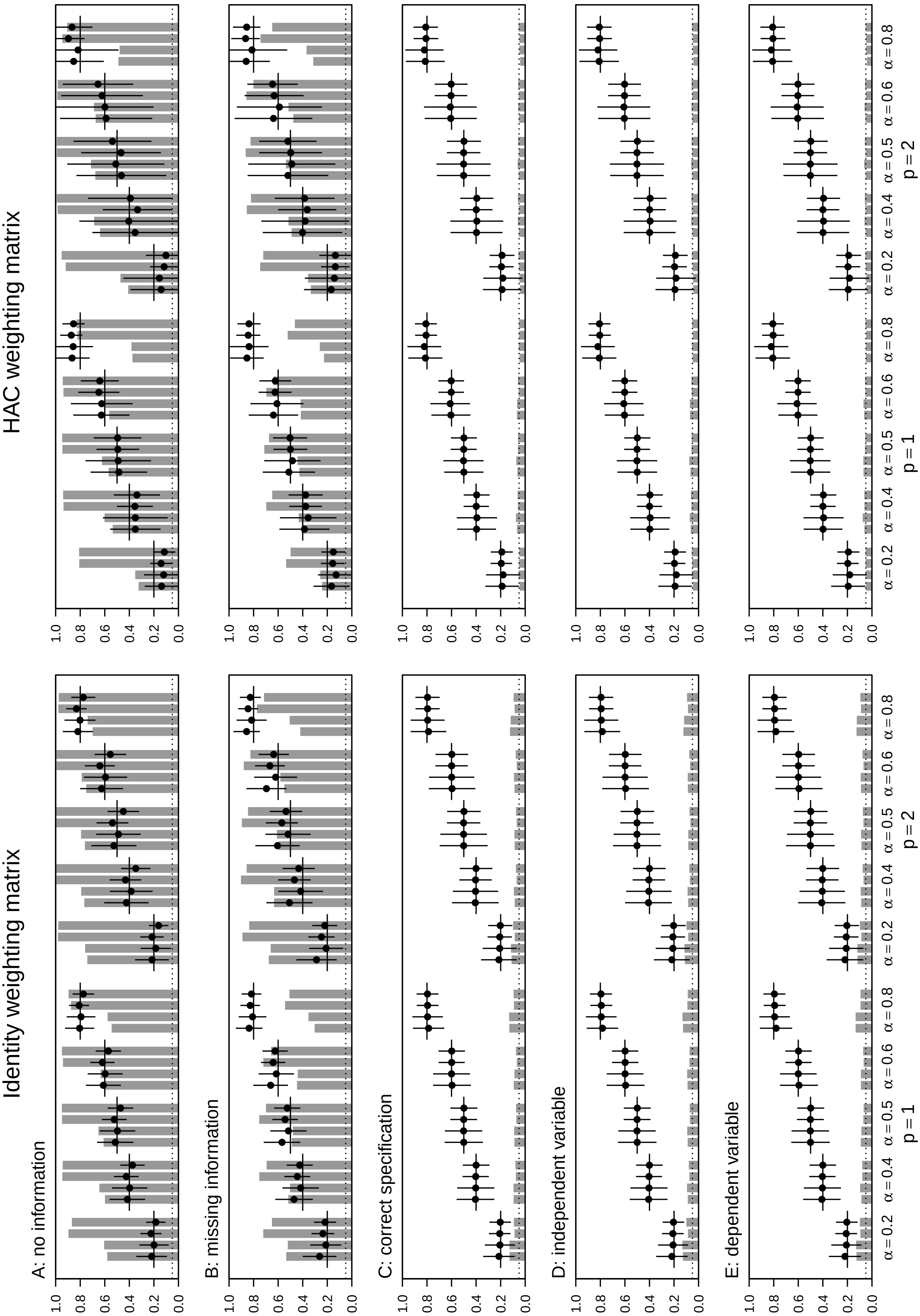
Figure 4.4 shows the results for a data generating process with a more noisy error term. The left side presents the results for the identity weighting matrix and the right side presents those for the HAC weighting matrix with a bandwidth choice according to Andrews. In comparison to the baseline experiment, the variance is increased and set to $\text{Var}(u_t) = 2$, which approximately doubles the standard deviation. Starting with the results under the identity weighting matrix and panels A and B, we find the asymmetry parameters again well estimated in mean, with slightly higher standard errors than in the baseline variant with $\text{Var}(u_t) = 0.5$. Furthermore, we observe a considerable loss of power, especially for $p = 1$ and smaller sample sizes. While in panel A the frequency of the J-test rejections almost returns to unity for larger sample sizes, this is not the case for panel B. In panels C to E, we find that the higher error term variance seems to have no observable influence on either the estimates of α or the size of the J-test.

The results under the HAC weighting matrix, presented on the right hand side of figure 4.4, point in the same direction, as differences in comparison to the baseline results can only be observed in panels A and B, while the results in panels C to E are essentially the same as in the baseline variant. Again, in panel A and B, the asymmetry parameters are slightly biased toward the direction of the asymmetry. The loss of power is more pronounced than in the setting with the identity matrix, especially in panel B. For the α's representing a higher degree of asymmetry, the test has even less power than for the more symmetric cases.

4.5.2 Fat Tails

In order to analyze how the properties of the J-test change when more extreme events occur more often, we draw the error term from a t distribution instead of a normal distribution. We chose

Figure 4.4: Fishbone plots - higher error term variance



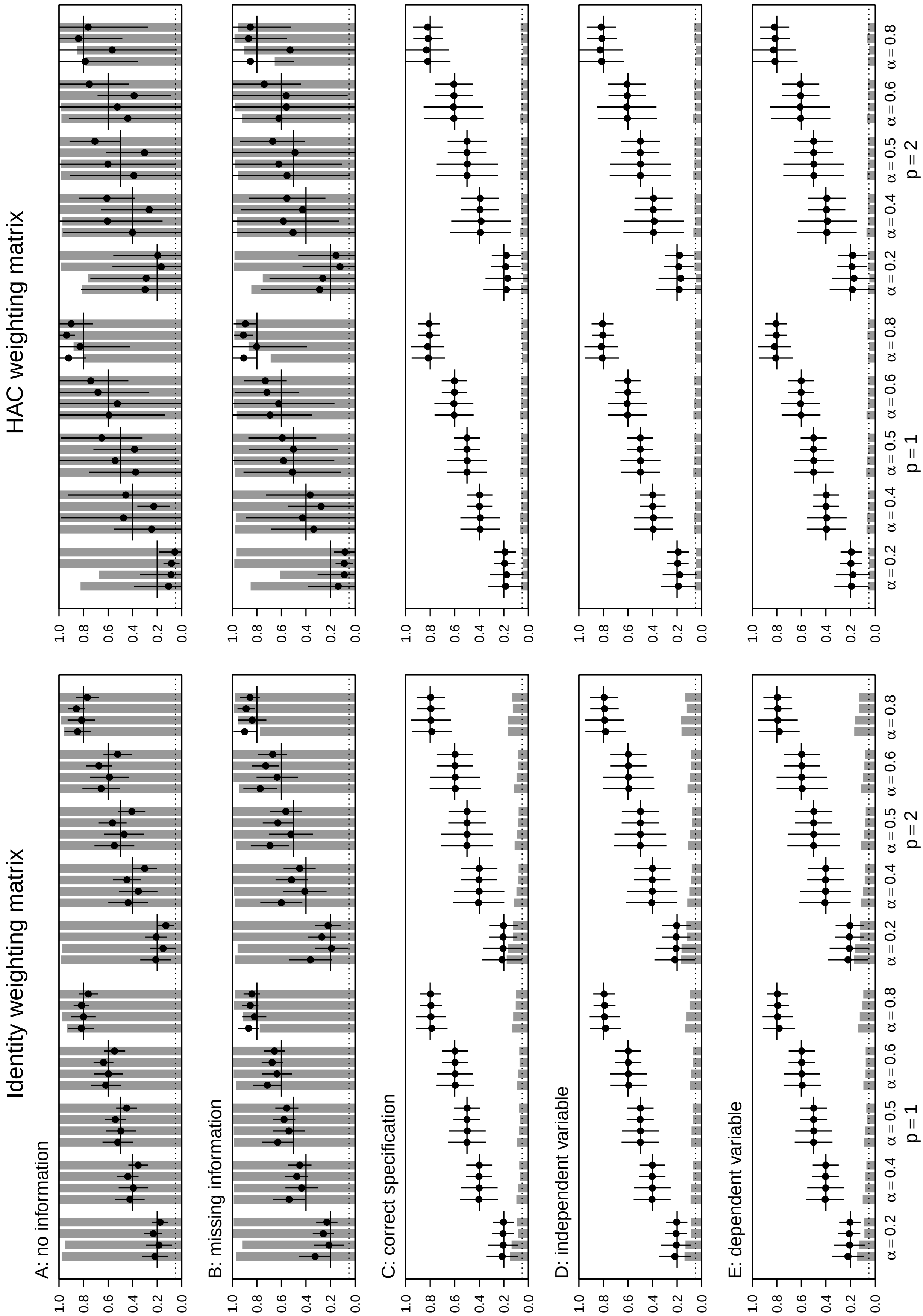


Figure 4.5: Fishbone plots - fat-tailed error term distribution

a $t(4)$ distribution and calibrate the obtained values $\tilde{u}_t$ in a way that the error term variance remains at $\operatorname{Var}(u_t) = 0.5$.[23]

The results are depicted in figure 4.5, in which the results for the identity weighting matrix again are shown on the left side and those obtained with the HAC weighting matrix on the right side. Starting with the left side and compared with the results under the normally distributed error terms in figure 4.1, we find very similar results in panels A and B. Thus, the α's are well estimated in A, while there is a tendency toward an overestimation in B. Again, the J-test has power close to unity in panel A and loses a bit of its power for smaller sample sizes. In panels C to E, the estimates of α again are very close the their true values. While for $p = 1$ the variation of the α's is almost identical to the scenario with normally distributed error terms, the variation is now slightly higher for $p = 2$. The size distortions in these three panels are slightly larger as well, especially for $p = 2$ and values of α_0 that represent a strong departure from symmetry, i.e. $\alpha_0 \in \{0.2, 0.8\}$.

Turning to the right side of figure 4.5 and the results obtained with the HAC weighing matrix, only minor differences can be found, in comparison with the corresponding baseline scenario on the left side of figure 4.3. While the mean estimates of the α's are essentially the same in all five panels, the variation is slightly higher in the fat-tailed variant. Moreover, in panels A and B, the power of the J-test is now a little higher in some cases than in the earlier scenario. For panels C to E, we find some small size distortions when error terms are t-distributed, which did not occur for the same experiment using normally distributed error terms.

These rather small size distortions disappear when we move from the $t(4)$ to the $t(5)$ distribution for the error terms (see table B1 in the appendix for these results). In sum, we find the EKT testing procedure to be almost unaffected by the fat-tailed error terms induced in this scenario.

4.5.3 Autocorrelation

In the next scenario, the variables $w_{i,t}$, and thus the target variable y_t, are simulated as AR(1)-processes

$$w_{i,t} = (1 - \rho_i) \cdot \mu_i + \rho_i w_{i,t-1} + \varepsilon_{i,t}, \tag{18}$$

where the autoregressive parameter is $\rho_i \in \{0.8, 0.5, 0.65\}$ and the constant of the processes is defined in order to leave the means of the variables unchanged compared to the previous scenarios, i.e. $\mu_i \in \{1, -1, 0\}$. In order to keep the variance of the three variables equal to unity, the error term variance of each process is set to $\operatorname{Var}(\varepsilon_{i,t}) = 1 - \rho_i^2$. For the simulation of the process, the first 500 observations have been discarded to remove initial value effects. The results are depicted in figure 4.6 in the same manner as before. The sequence of the panels on both sides of the figure has been changed slightly, as we are interested in the effects of leaving out either variable $w_{2,t}$ with the lower autocorrelation (panel B1), or $w_{1,t}$ with the higher autocorrelation (B2) on the construction of the forecasts. In order to keep the figure compact, we have omitted the additional variable $w_{1,t} \cdot w_{2,t}$, which has been shown in panel E in the other figures, as the differences between panels D and E have been negligible so far.

[23] This is seen from $u_t = 0.5 \cdot \tilde{u}_t$ with $\tilde{u}_t \sim t(4)$ and $\operatorname{Var}(\tilde{u}_t) = 4/(4-2) = 2$.

Beginning with the left side of the figure, we find the asymmetry to be slightly underestimated in panel A and B2 and slightly overestimated in panel B1. In all cases, the bias is more pronounced for $p = 2$ and becomes smaller with larger sample size. Similar to the baseline variant, the power of the test in panels A and B1 is close to unity or equals unity in most cases. The exception here is panel B1 and $\alpha_0 = 0.8$ for both curvature parameters and smaller sample sizes, for which the rejection frequency of the J-test is somewhat reduced. For panel B2, we observe a loss of power for smaller sample sizes, particularly for $p = 1$ and $\alpha_0 = 0.2$. Once more, panels C and D, which show the results for the correctly specified forecasts, appear to be robust to the changes we made to the DGP. The slight size distortions remain existent.

When the autocorrelation is taken into account in the weighting matrix, in the first three panels, we observe asymmetry parameters that are further away from the true values of α with even higher standard deviations than in the baseline scenario. Again, the reduction of the test power is more pronounced for smaller sample sizes and close to unity for larger sample sizes. Comparing the results in panels B1 and B2, we find that for both variants of the weighting matrix, the test loses more power if the more highly autocorrelated variable is neglected when constructing the forecasts. In panels C and D, results for the α values again are essentially the same as in the baseline variant. The size distortions are now entirely absent.

4.5.4 Outlier

In the context of the effects of a major economic crisis on the behavior of macroeconomic variables and the aggravated task of producing accurate forecasts associated with this, we are interested in the robustness of the EKT approach when there is a single shock in the form of an outlier in one of the time series $w_{i,t}$. To model the outlier, we follow Franses et al. (2014), who distinguish four types of outliers: additive outliers [AO], innovative outliers [IO], transient change [TC] and level shifts. For our analysis, we focus on the second type of outliers, as using the IO rather than the AO on the one hand allows the shock to affect subsequent observations, which appears to be more realistic for macroeconomic shocks, but on the other hand implies the assumption that the dynamics of the shock are identical to the dynamics of the AR(1)-process. Although modeling the outlier into the series by using TC would allow us to relax this assumption, we maintain it for simplicity. Moreover, we lack knowledge regarding the question if and how the AR(1) parameter of the process and the shock differs.

Two additional aspects are of interest when inducing an outlier into our simulated time series. These are the point in time at which the shock occurs, τ, and the magnitude of the shock ζ. Both aspects are especially relevant in the forecasting context, as the task of producing an accurate forecast becomes more challenging when the crisis is more severe. If the crisis arises at an early point of the time series, the forecaster has more time to adapt. To mimic the recent crisis, we set the magnitude of the crisis to minus three standard deviations of the time series, that is $\zeta = -3$ as $\mathrm{Var}(w_{i,t}) = 1\,\forall i$. This reflects the approximate slump of US GDP during the crisis. The shock is built into the model at time $\tau = P - 10$, where P is the out-of-sample sample size and thus

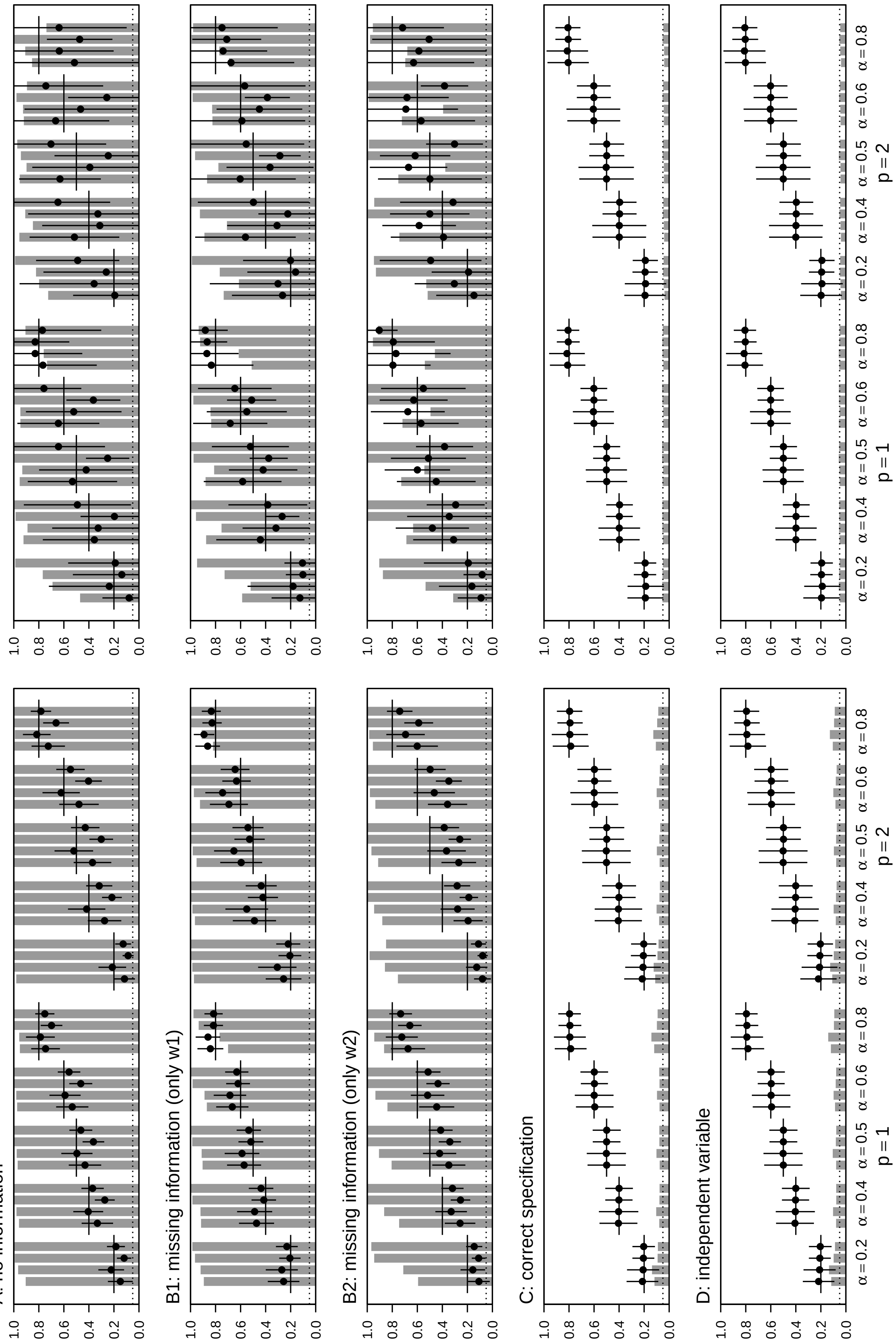
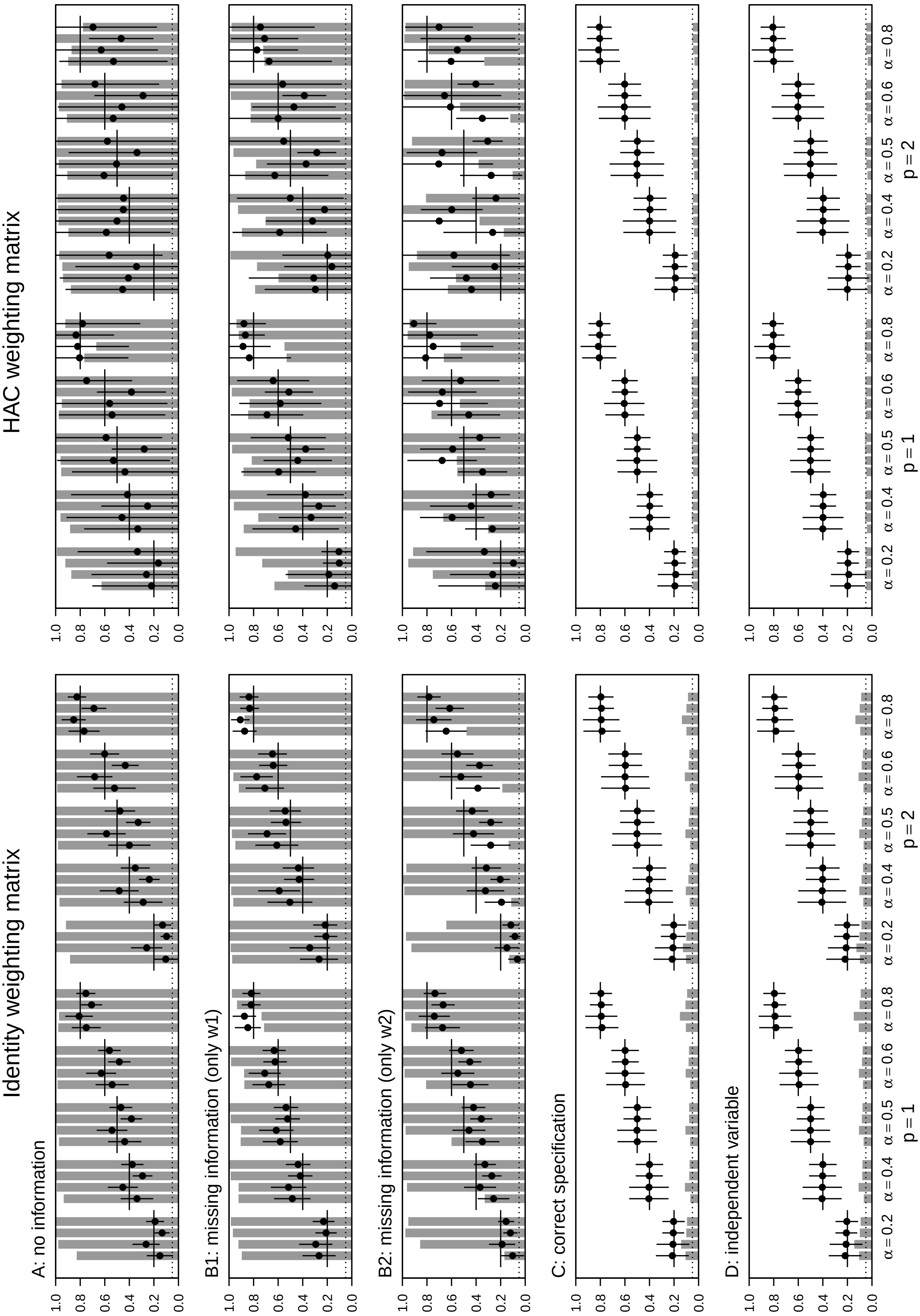


Figure 4.6: Fishbone plots - serial correlation

Figure 4.7: Fishbone plots - outlier



equals the number of forecasts in each simulation step. Including the IO in the AR(1)-process, we specify the following process,

$$w_{i,t} = (1 - \rho_i) \cdot \mu_i + \rho_i w_{i,t-1} + \varepsilon_{i,t} + \zeta I\,(t = \tau)\,. \tag{19}$$

For the scenario with a crisis, the shock is induced into the variable $w_{1,t}$ as described above, while all other variables and parameters are left unchanged with respect to the pure autocorrelation scenario of the previous section.

The results are shown in figure 4.7, using the familiar arrangement with respect to the identity and HAC weighting matrices. In panel A on the left side of the figure, we find the mean estimates of α to be slightly farther away from their true values than in the AR(1) scenario. For $\alpha_0 = 0.2$ and small samples, we also observe that rationality is rejected less frequently. No apparent difference to the AR(1) scenario can be detected for the results displayed in panel B1, which concern the misspecified forecasts and the variable $w_{2,t}$ without shock has been omitted from the forecast equation. Turning to panel B2, where the shocked variable $w_{1,t}$ has been omitted, the estimates of α remain essentially unchanged, while the test dramatically loses power in small sample sizes for all values of α. For $p = 2$, this loss of power is even more pronounced, while the biases of α are about the same. In all three panels, the power properties in larger samples remain very good. Similar to earlier results, no changes in the mean estimates, standard deviation, and power properties can be recognized when all relevant information has been included in the forecast equation, as in panels C and D.

In panel A on the right hand side of the figure, we notice a slight unsystematic change in the mean estimates of α along with a small increase in the standard deviations. A somewhat puzzling result concerns the power, as, especially for small sample sizes and $\alpha_0 = 0.2$, the test gains rather than loses power, in comparison with the pure AR(1) scenario. The mean estimates of α in panel B2 slightly differ from those in the AR(1) scenario and tend to be farther away from their true values. We also note that the standard deviations exhibit a tendency to increase which, however, is rather unsystematic across sample sizes and different degrees of asymmetry. Considering the power of the test, the results differ depending on the curvature parameter p. For $p = 1$ there is a slight increase of power, especially for small sample size, while we observe a loss of power for $p = 2$ both in the smaller and larger sample sizes. The results in panel B1, C and D again remain essentially unchanged. In our view, the most plausible explanation for this is that the outlier induced in $w_{1,\tau}$ is translated into an outlier of $y_{\tau+1}$ by the construction of this series. Thus, if the model is correctly specified, the outlier will be contained in the forecaster's information set.

4.6 A More Realistic Setting

The objective for the final two scenarios of the analysis is to estimate realistic parameters fitting quarterly US macroeconomic data. Herein, the year-to-year growth of the unemployment rate has been taken as the target variable (y_{t+1}), while forecasts have been produced using a linear forecasting model with a constant, lagged GDP growth ($w_{1,t}$) and lagged CPI inflation ($w_{2,t}$)

with all growth rates used in percentage points.[24] Stationarity has been tested using Elliott, Rothenberg and Stock's (1996) unit root test and we have found a clear rejection of the unit root null for all three series.

Two factors motivate the choice of these three variables. On the one hand, (un)employment, GDP and inflation are probably the most frequently used indicators for macroeconomic activity as reflected in section 2A of the Federal Reserve Act, for example, which determines the monetary objectives of the FED.[25] On the other hand, the certainty of the relationship between either GDP or inflation and the unemployment rate differs for both variables. The linear relationship between the unemployment rate and GDP and its lags has appeared to be a rather stable over the last four decades, as shown by Ball et al. (2013) also for quarterly data, for which deviations from long-run levels of log GDP and the unemployment rate have been used as variables. In contrast, the relationship between unemployment and inflation has been less clear in this period, as Putnam and Azzarello (2015) have stated recently. Using a forecast equation based on two variables with a potentially different relevance for the target variable gives us the opportunity to analyze the different results when either the stronger or the weaker variable is omitted from the equation.

To conduct our Monte-Carlo experiment, we first need to estimate the linear relationship between the variables to enable us to calculate the target variable y_{t+1}. We then fit a VAR(1) model for GDP growth and inflation to obtain realistic parameters underlying the simulated process. We estimate the following model by OLS

$$y_{t+1} = \underset{(1.961)}{15.497} - \underset{(0.345)}{5.970} \cdot w_{1,t} + \underset{(0.262)}{0.707} \cdot w_{2,t} + \hat{u}_{t+1},\ T = 182,\ R^2 = 0.72. \tag{20}$$

The standard errors are given in parentheses. Both explanatory variables are statistically significant and the overall fit is quite good. We observe a considerably larger influence of $w_{1,t}$ than of $w_{2,t}$ with respect to both coefficient magnitude and t-statistic. This leads to the fortunate situation that the omission of one of these variables should more seriously affect the results than the omission of the other. Therefore, we have chosen $\boldsymbol{\theta} = (15.5, -5.97, 0.71)'$ to specify the linear relationship. For the bivariate relationship of GDP growth and inflation we obtain the following VAR(1) model,

$$\begin{pmatrix} w_{1,t} \\ w_{2,t} \end{pmatrix} = \begin{pmatrix} \underset{(0.24)}{1.09} \\ \underset{(0.14)}{-0.45} \end{pmatrix} + \begin{pmatrix} \underset{(0.04)}{0.81} & \underset{(0.03)}{-0.11} \\ \underset{(0.03)}{0.13} & \underset{(0.02)}{1.01} \end{pmatrix} \begin{pmatrix} w_{1,t-1} \\ w_{2,t-1} \end{pmatrix} + \hat{\boldsymbol{\varepsilon}}_t, \tag{21}$$

again with standard deviations in parentheses and the following estimated covariance matrix of the innovations

$$\hat{\boldsymbol{V}}\left(\hat{\boldsymbol{\varepsilon}_t}\right) = \begin{pmatrix} 1.08 & -0.05 \\ -0.05 & 0.37 \end{pmatrix}.$$

[24] The data have been retrieved from the FRED database of the Federal Reserve Bank of St. Louis. We use the quarterly series of the real GDP (GDP96), the CPI for for all urban consumers and all items (CPIAUCSL) and the civilian unemployment rate (UNRATE) from 1970:Q1 to 2007:Q4. See: https://research.stlouisfed.org.

[25] "The Board of Governors of the Federal Reserve System and the Federal Open Market Committee shall maintain long run growth of the monetary and credit aggregates commensurate with the economy's long run potential to increase production, so as to promote effectively the goals of maximum employment, stable prices, and moderate long-term interest rates." See http://www.federalreserve.gov/aboutthefed/section2a.htm.

Using these parameters, we conduct the Monte-Carlo experiment as described in section 4.3. Herein, we draw the error terms, which are used to compute the target variable y_{t+1}, from a normal distribution with mean zero and variance equal to the residual variance obtained from equation (20), i.e. $u_t \sim N(0, 69.4)$. Figure 4.8 compares the kernel density estimate of the residuals (solid line) with the fitted normal distribution (dashed line). Both density estimates are quite close; this justifies the normality assumption.

Figure 4.8: Densities of the residuals and the normal distribution

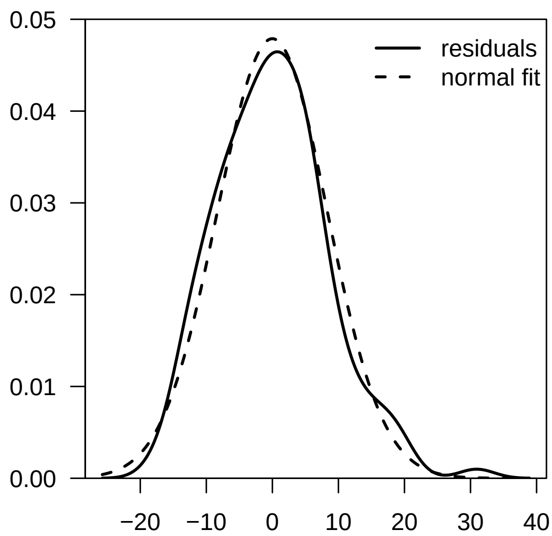


Figure 4.9 displays the results of the Monte-Carlo experiment with panels A to C remaining as described above. Here, Panel D shows the results for the forecast equation that employs the interaction term of $w_{1,t}$ and $w_{2,t}$ as an additional variable in the information set. As in the previous figures, we show the results with the identity weighting matrix on the left side. Starting with panel A, the estimates of α here are also close to their true values, although we observe a slight tendency toward overestimation for $p = 2$. Compared to our earlier results, the variation of the α estimates is rather small. The power of the J-test is close to unity.

Panels B1 and B2 show the results for the forecasts, which have been established with only a constant and one of the respective variables $w_{1,t}$ or $w_{2,t}$ that mimic US growth and inflation. Regarding equation (20) and the coefficients used to compute the target variable, it is not surprising that omitting $w_{2,t}$ (in B1) leads us to results that are closer to the correctly specified model than results neglecting $w_{1,t}$ (in B2). Accordingly, in panel B1 we find the α's to be very close to their true values in panel B1, while a slight tendency toward overestimation can be observed in panel B2. Considering the power of the J-test in panel B1 and especially for small sample sizes, the test has almost no power against the misspecification of the forecasts, as reflected in the rejection frequencies close to the nominal size. The power is higher for larger sample sizes, although still smaller than 0.5. In contrast, we observe J-test rejection frequencies equal or close to unity in panel B2. Thus, the omission of the more relevant variable $w_{1,t}$ (t-statistic 17.3) is clearly recognized by the J-test in panel B2 and leads to high rejection rates. By contrast, omitting the less relevant variable $w_{2,t}$ (t-statistic 2.7) is associated with a substantially smaller test power, as observed in panel B1.

Similar to the earlier scenarios, the α's are very precisely estimated and the J-test's size is close to the nominal size of 5 percent in panels C and D. Nevertheless, we find small size distortions that are more pronounced for values of α_0 and reflect stronger asymmetric preferences, i.e. $\alpha_0 \in \{0.2, 0.8\}$.

The results shown on the right side of figure 4.9 for which the HAC covariance matrix has been used are somewhat puzzling in the first and third panel. On the one hand, the α's are detached from their true values and show high variation. On the other hand, the power of the J-test varies strongly with no systematic pattern for different sample sizes. This holds especially for $p = 2$, while the power is rather high for $p = 1$. The drop in the power for larger sample sizes is particularly difficult to explain.

We checked the correct convergence of the numerical optimization and found a few irregular terminations. These cases have been discarded from the rest of the analysis. An explanation for these bewildering results may lie in the flexible bandwidth selection according to Andrews in the GMM optimization. Apparently, this leads to bandwidth choices greater than one in some cases and results are thus inferior compared to the optimal choice of one. Conducting the same experiment with a fixed bandwidth, the drops in power become more prevalent when the bandwidth increases (see tables B2 and B3 in the appendix for results with bandwidths fixed at values from 1 to 4).

The remaining panels B1, C and D show results quite similar to the corresponding panels on the left side. Again, the α estimates are close to their true values, although now there is slightly more variation in the estimates. Considering the power of the J-test in B1, the general pattern is similar, but the test now has even less power for all sample sizes. The small size distortions in panels C and D, observed in the result with the identity weighting matrix disappear here and, in some cases, the test is now slightly undersized.

As a final variant of our analysis, we induce a shock in the first variable of the VAR(1) model to mimic the recent crisis with the simulated data. In doing so, we follow the procedure explained in subsection 4.5.4 and induce a shock with a magnitude ζ of approximately three standard deviations of the variable in question (GDP growth). During the last recession, GDP growth was as low as -4.15 percent in 2009:Q2, which is about 6.86 percentage points below the average growth rate of the period from 1970:Q1 to 2015:Q2. As the standard deviation for this period has been 2.21, this equals a downturn of the economy of about three standard deviations. Hence, to model the crisis, we chose a magnitude of $\zeta = -6$ and, as above, induce the crisis at point $\tau = P - 10$.

The associated results are presented in figure 4.10. Starting with the left side of the figure, the α's are again very close to their true values in all panels even in the presence of the outlier. For $p = 2$ in panels A and B2, they are even closer than before. Furthermore, we again observe only a small variation in the α's in panels A and B2. Regarding the power of the J-test, however, we now find a loss of power for small sample sizes, especially for $p = 2$. In panel B1, the general pattern of the results is the same as in figure 4.9, with the exception that there is now a little

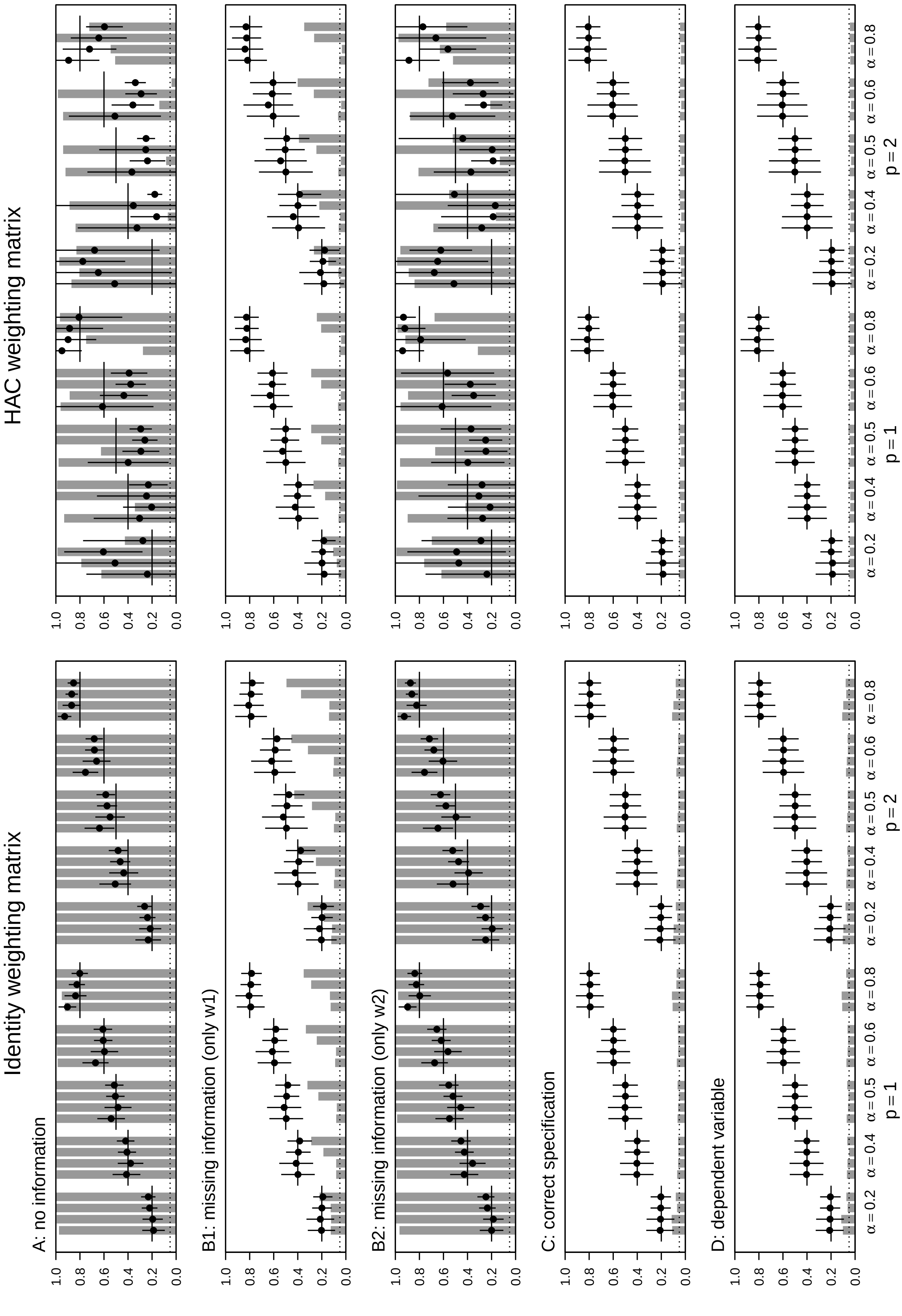
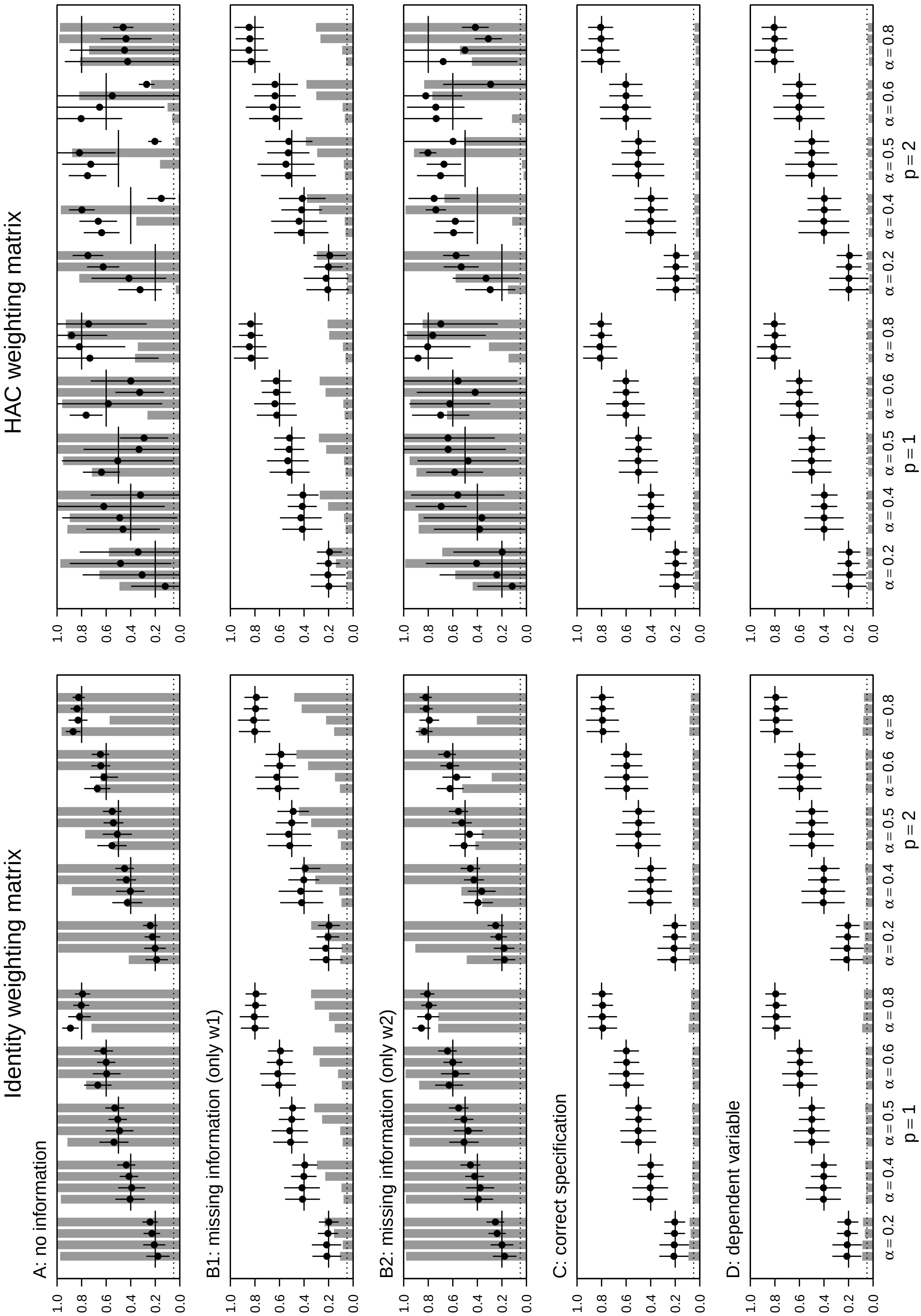


Figure 4.9: Fishbone plots - simulation of the US economy

Figure 4.10: Fishbone plots - simulation of the US economy with outlier



more power in small sample sizes. Apparently, the outlier hardly affects the results in panels C and D, in which the results are almost identical to the previous scenario without an outlier.

On the right side of figure 4.10 and in panels A and B2, we are confronted with the same puzzling pattern in the results as in figure 4.9. The α results are rather distant from their true values, the estimates exhibit high variation and the rejection frequencies of the J-test vary rather non-systematically across the different sample sizes. The pattern we observed for the scenario without an outlier in which the quality of the results decreases with an increasing fixed bandwidth also exists here (see tables B4 and B5 in the appendix).

For the remaining three panels, we observe results similar to the variant with the identity weighting matrix, with slightly more variation in the α estimates, slightly less power in the J-test in B1 and slightly lower than nominal rejection frequencies in C and D.

4.7 Summary

In this chapter the results of an extensive Monte-Carlo investigation of the EKT procedure for forecast evaluation have been reported. Summarizing the main findings for the baseline setting, we first find that the asymmetry induced in the loss function is well recognized by the GMM estimator. Overall, the mean of the asymmetry parameter α is very well estimated to be close to the true values and we observe decreasing variation in the α estimates with increasing sample size. This assertion generally holds even when misspecification is induced into the simulation setting. Second, regarding the ability of the J-test to test the rationality of the forecasts, we find satisfying size properties with rejection frequencies close to the nominal level of 5 percent for correctly specified forecasts and also for forecasts with additional, but irrelevant, variables in the information set. The J-test also shows very good power against misspecified forecasts. Power increases naturally with the sample size and becomes close to unity in most scenarios for the larger sample sizes. Third, the different variants of weighting matrices used in the GMM estimation are also problematic in the case of the EKT procedure, as they are in general (see e.g. Podivinsky, 1999). Using the identity matrix as a simple and numerically robust choice for the GMM optimization, we find that the J-test has a slight tendency to be oversized in small samples. These size distortions almost disappear when using the HAC weighting matrix with the bandwidth choice according to Andrews, or a fixed bandwidth.

Fourth, we modify the baseline scenario in different directions to assess the robustness of the findings. Increasing the error term variance reduces power somewhat when forecasts are misspecified, but hardly affects the α estimates and the size properties for correctly specified forecasts. Using a fat-tailed distribution for the error terms leads to a slightly higher variation in the α estimates and more size distortions, which are especially pronounced when the identity matrix is used as a weighting matrix. Autocorrelation in the DGP leads to slightly more biased estimates of α and reduced power of the J-test in smaller samples. The power reduction is stronger when the more highly autocorrelated variable is omitted from the forecast equation and when the HAC weighting matrix is used. Inducing an outlier into one of the variables used for forecasting only slightly affects the results when this variable is omitted from the forecast equation. Here, the J-test loses some power.

Fifth, when generating the data from a more realistic scenario with parameters estimated using real data, we find the general pattern of results rather robust. We find surprisingly small variation in the α estimates. Quite remarkably, the J-test loses most of its power when we omit the less relevant variable from the forecast equation. The estimates of α remain nevertheless quite precise. In the setting based on real data, the results using the HAC weighting matrix and the flexible bandwidth choice according to Andrews appear rather inconclusive. When the forecasts are misspecified, this may likely be caused by a suboptimal bandwidth choice by the algorithm. Fixing the bandwidth leads to better results for small bandwidths which deteriorate as the bandwidths increase.

5 Asymmetric Loss and Higher Moments in the ECB's SPF

5.1 Motivation

In the previous chapters, the asymmetry in the forecaster's preferences was determined by a given series of forecast errors. So far we have neglected to specifically control for the time varying variance and higher moments of these forecast errors. In financial decision making, for example, it has long been recognized that asset returns are not normally distributed and that higher moments, such as skewness and kurtosis, can play a decisive role when optimizing portfolio allocations (see e.g. Jondeau and Rockinger (2006)). In this chapter, we adapt this thought of the importance of higher moments and translate it into the forecast evaluation setup. Especially with respect to the recent recession, we are interested to see how accounting for the changes in the moments of the forecast errors affects the form of the forecaster's underlying loss function.[26]

While EKT's approach allows for asymmetry in the loss function and nests the symmetric cases of the mean absolute and the mean squared error, it concentrates exclusively on the first moment of the forecast errors. In the spirit of their approach, we expand the loss function to include higher moments and evaluate how this changes the form of the loss function. The prevalent procedures for doing so are the Taylor series expansion, which Jondeau and Rockinger (2006) apply to utility functions to maximize portfolio allocations' expected end-of-period wealth, and the exploitation of an assumed Gram-Charlier distribution of the target variable at hand, as suggested by Christodoulakis (2005), also in the context of forecast optimality of financial forecasts. These approaches require the loss function to be differentiable as many times as the number of moments to be included, which rules out the family of Lin-Lin and Quad-Quad loss used by EKT.

An obvious first choice for a loss function that does not vanish when differentiating is the Linex loss function, proposed by Varian (1975). However, the parametrization of the Linex loss function, differs from the one in the EKT approach and this function only nests the symmetric special case asymptotically. Thus, we propose the Linex-Linex loss function that overcomes both obstacles and apply both functions in an EKT-based setting to back out the asymmetry parameter of the loss function, while simultaneously testing forecast rationality.

We apply our approach to the quarterly business cycle forecasts obtained from the ECB's Survey of Professional Forecasters, a dataset that is particularly interesting because it is relatively new and therefore has not gathered a lot of attention in the forecast evaluation literature so far. Furthermore, the (unbalanced) panel structure of the survey is particularly appealing. On the one hand, we use the time series dimension to conduct our analysis for the individual forecasters and have the possibility of comparing the results across individuals. On the other hand, the cross sectional dimension provides us with an additional opportunity for computing the higher moments. Moreover, the forecasts in the survey concern the euro area, which is a monetary but not a political union, composed of countries that follow their own fiscal interest to a certain degree. Hence, producing accurate macroeconomic forecasts for this area, such as those provided by the professional forecasters in the survey, seems to be a rather difficult task. This is corroborated

[26] This chapter is based on Hoss (2015).

by Camacho et al. (2006, 2008) and Eickmeier (2009), who find the business cycles of the euro area's member states to be heterogeneous. More recently, Lee and Mercurelli (2014) have argued that the monetary union has accelerated the convergence process of Germany, France and Italy, while Barigozzi et al. (2014) observe that northern and southern European countries still respond differently to the ECB's monetary policy with respect to inflation and unemployment.

The remainder of this chapter is structured in the following way: in the next section we examine the Linex loss function more closely and introduce the Linex-Linex loss function. Section 5.3 provides a discussion of the Taylor expansion of a loss function and the introduction of higher moments to the loss function. Subsequently, the optimality conditions for the EKT-based GMM test to back out the asymmetry parameter are addressed in section 5.4. Section 5.5 introduces the ECB's Survey of Professional Forecasters and presents the datasets used. Sections 5.6 and 5.7 provide the different approaches used to compute the moments of the forecast errors as well as the GMM results for the respective approaches. Section 5.8 summarizes the main results.

5.2 A Closer Look at Linex Based Loss Functions

In section 2.2, the Linex loss function was introduced as an example of an asymmetric loss function (see equation (8) on page 17), that is approximately exponential for forecast errors of one sign and approximately linear for errors of the other sign. Later in section 3.2, this property was used to measure potential asymmetries in the CEE and JF employment forecasts. In this section, the Linex loss function and further functions that are based on the Linex will be discussed primarily because of their favorable property of being differentiable at any order without vanishing.

For the rest of this chapter we will employ the version of the Linex that is utilized by Batchelor and Peel (1998) and Clatworthy et al. (2012),

$$L(e_{t+h}; \alpha) = \frac{1}{\alpha^2} \left(\exp\left(\alpha e_{t+h}\right) - \alpha e_{t+h} - 1 \right). \tag{22}$$

While Batchelor and Peel slightly modify the Linex loss function we have looked at so far by dividing it by α^2, Clatworthy et al. also set the scale parameter $\beta = 1$. This version of the Linex is parsimonious because, given the forecast errors, the loss function only depends on α. Both studies argue that the Linex converges to the quadratic loss as $\alpha \to 0$. Clatworthy et al. point out that the convergence is non-intuitive and becomes clear when employing L'Hospital's rule.

In their 2005 paper, EKT argue that "linex loss only nests symmetric loss as a limiting case in the parameter space where loss is not defined. Obtaining symmetry only for a parameter on the boundary creates serious estimation problems and means that linex loss is not well-suited for our purpose" (p. 1110). Here, the Linex loss function is needed to introduce the higher moments to the loss function. When interpreting the asymmetry parameter α, it is important to keep in mind that the function is not defined for a perfectly symmetric case. Nevertheless, symmetry can be approximated for α close to zero. To overcome this difficulty, we need another function that actually nests a symmetric special case. Therefore, we will briefly discuss the Double Linex loss function first and then propose a modification to the Linex that we call the Linex-Linex.

One alternative loss function that not only is limiting in a symmetric special case but actually nests symmetric loss is the Double Linex loss function

$$\begin{aligned} L(e_{t+h};\, \alpha, \beta) &= (\exp(\alpha e_{t+h}) - \alpha e_{t+h} - 1) + (\exp(-\beta e_{t+h}) + \beta e_{t+h} - 1) \\ &= \exp(\alpha e_{t+h}) + \exp(-\beta e_{t+h}) + (\beta - \alpha) e_{t+h} - 2 \end{aligned} \tag{23}$$

proposed by Granger (1999), where $\alpha > 0$ and $\beta > 0$. The Double Linex loss function is symmetric for $\alpha = \beta$. Granger proposed the Double Linex loss function as an example that shows how loss functions, or cost functions in his terminology, can be combined in various ways and still remain a loss function (see section 2.2). The concept of the Double Linex loss is picked up by Christodoulakis (2005) and Demetrescu (2006). Although both mention the general asymmetric form with $\alpha \neq \beta$, their main findings are for the symmetric special case. Christodoulakis finds a closed form solution for the optimal forecast after a Gram-Charlier expansion but does not apply his analytical solution to data. Demetrescu tests an extended Gauss-Newton algorithm under asymmetric loss but skips symmetric Double Linex loss for the second non-linear part of his simulations, as convergence results were unsatisfactory.

These discouraging findings in literature as well as our own rather disappointing results concerning the Double Linex loss function require a further modification of the Linex loss function. Herein, we followed the idea of the Lin-Lin and the Quad-Quad loss function that both have a symmetric special case but are asymmetric in general. For lack of a better name, this modified Linex loss function is simply referred to as Linex-Linex loss

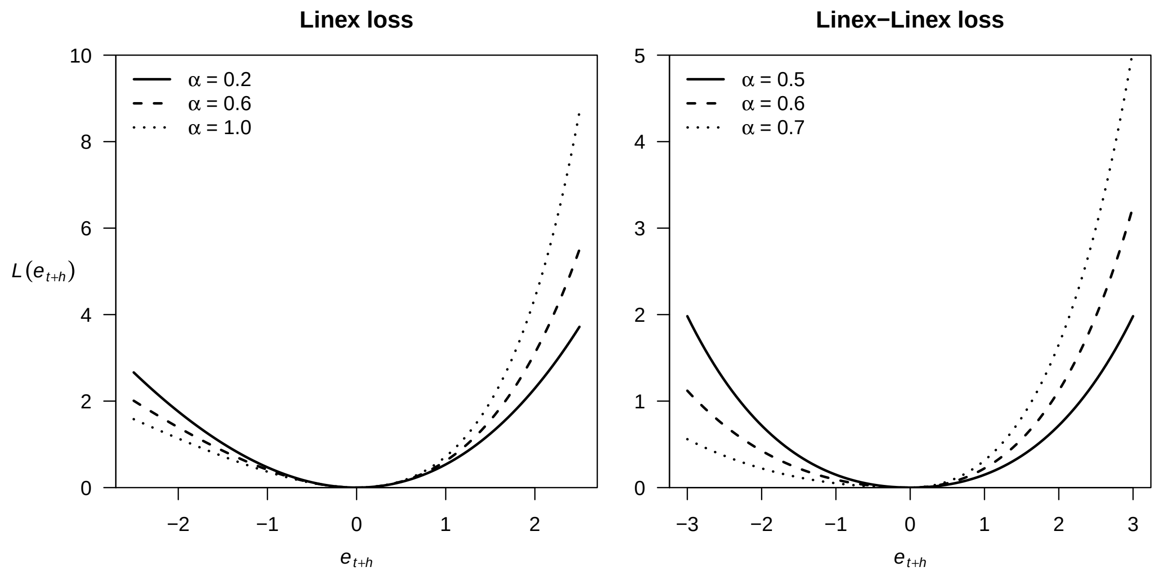
$$L(e_{t+h};\, \alpha) = \exp\left([\alpha - I(e_{t+h} < 0)] \cdot e_{t+h}\right) - [\alpha - I(e_{t+h} < 0)] \cdot e_{t+h} - 1, \tag{24}$$

where the shape parameter α is substituted by $\alpha - I(e_{t+h} < 0)$. Herein, $I(.)$ is the indicator function equal to unity when the condition $e_{t+h} < 0$ holds and equal to zero elsewhere. Unlike Linex loss, where $\alpha \in \mathbb{R} \setminus \{0\}$, the asymmetry parameter for Linex-Linex lies between zero and one, $\alpha \in (0, 1)$, and, in analogy to Lin-Lin and Quad-Quad loss, is symmetric for $\alpha = 0.5$. Values of α below 0.5 correspond to higher costs associated with an overestimation of the target variable (negative forecast errors), whereas for $\alpha > 0.5$, the opposite holds and an underestimation (positive errors) is weighted more heavily. As Linex-Linex loss nests a symmetric special case instead of only a limiting feature, dividing by α^2 to ensure convergence is not necessary.

To assure that Linex-Linex is indeed a loss function, we reconsider the properties that Granger (1999) stipulates for a loss function (see sections 2.1 and 2.2). While both Linex and Linex-Linex fulfill the essential requirements of a loss function and both are differentiable to any order, neither is homogeneous.

Before turning toward the discussion of expanding the loss function in order to include higher moments, figure 5.1 above shows the Linex and the Linex-Linex loss function for different asymmetry parameters α. For both Linex type loss functions and in contrast to the Lin-Lin and the Quad-Quad loss functions the relative costs of an underprediction versus an overprediction of the same magnitude do not only depend on the value of α but also depend on the absolute value of the forecast error. That is, the relative costs differ more strongly for higher absolute values of the forecast error. Thus, both functions reflect asymmetric preferences of the forecaster that are stronger for higher absolute forecast errors.

Figure 5.1: Linex and Linex-Linex loss functions



5.3 Higher Moments in the Loss Function

The idea to include higher moments in a loss function originates from financial decision making where two phenomena can be observed that are not consistent with the assumptions of normality (see e.g. Jondeau and Rockinger (2006)). First, the frequent occurrence of extreme outcomes indicates a fat-tailed rather than a normal distribution of asset returns. Second, financial crashes occur more often than financial booms and hence indicate that returns are distributed asymmetrically. Both arguments are well established in the literature (see e.g. Mandelbrot (1963), Kon (1984) and Longin (1996) for the former and Fama (1965), Singleton and Wingender (1986) and Peiró (1999) for the latter). Jondeau and Rockinger (2006) provide a literature overview of early studies, such as Levy (1969), Samuelson (1970), Levy and Markowitz (1979) and Kroll et al. (1984), that introduce higher moments to the decision-maker's expected utility function. However, these authors' results concerning the usefulness of the higher moment are predominantly inconclusive.

Jondeau and Rockinger (2006) use an approach that allows the role of higher moments in the utility function to be captured and which depends on a Taylor series expansion to approximate the expected utility function. When using Taylor approximations, one practical issue is the order or truncation k. Non-truncated Taylor series are power series with infinite summands if the function of concern does not vanish after a certain number of differentiations. Hence, truncation is necessary to make the approach feasible for applications. However, Brockett and Garven (1998) point out that including more moments does not have to result in a better approximation and that there is no general rule for determining the optimal truncation order. In keeping with literature, in which Taylor series are usually truncated at $k = 4$ at the most (see e.g. Markowitz

(1952) for $k = 2$, Levy (1969) for $k = 3$ and Benishay (1992) and Jondeau and Rockinger (2006) for $k = 4$), we use truncations up to order four in this study.

The second to fourth moment can be related to the forecaster's preferences toward risk aversion, prudence and temperance, respectively. Moments higher than fourth hardly have garnered attention in the literature. The reason for this certainly lies in the difficulty of clearly interpreting fifth moments and above in economic terms (see Lajeri-Chaherli (2004) for an exception that defines a measure of "edginess" based on the fifth moment). Since Pratt (1964) and Arrow (1971), it has become common to regard decision-makers as being risk-averse and thus to assume that an increasing variance reduces their expected utility. To describe preferences regarding a positive third moment, Kimball (1990) proposes the term "prudence" and defines it as the "propensity to prepare and forearm oneself in the face of uncertainty, in contrast to 'risk aversion,' which is how much one dislikes uncertainty" (p. 54). Temperance is referred to as an aversion to large fourth moments and thus an increased probability of extreme events (see e.g. Eeckhoudt et al. (1995) or Gollier and Pratt (1996)). Scott and Horvath (1980) generalize this preference structure to all higher moments and argue that a financial decision-maker's expected utility increases with large and positive odd moments, while it decreases with large even moments. In spite of this finding, these authors refrain from interpreting moments higher than the fourth and also concede that the fourth moment is not as important as the second and third because a smaller variance also implies a smaller kurtosis.

While the expected utility is maximized in the studies mentioned above, (e.g. over the expected end-of-period wealth of allocations of risky asset returns in Jondeau and Rockinger (2006)), in this study, the analogous concept is adopted to minimize the expected loss function over the forecaster's expectation of the forecast error. The central idea is to use a Taylor series expansion to incorporate higher moments of the forecast errors into the loss functions discussed in the previous section. Truncating the Taylor series at orders two, three and four respectively, provides us with the opportunity to analyze how capturing the forecasters preferences or aversions toward the degree of the variance, skewness and kurtosis in the forecast errors affects the parameters of the loss functions. In this and the next section, we focus on a truncation order of four, for the discussion of the methods, whereas in sections 5.6 and 5.7, results for truncation orders from two to four are presented.

Further studies that discuss Taylor series expansions in the context of loss functions are Elliott and Timmermann (2004), Patton and Timmermann (2007a) and Clatworthy et al. (2012). The first two base their argumentation on Rudin (1964), who states that a loss function $L(.)$ can be described by a power series if it depends exclusively on the forecast error and is analytic everywhere, except in a finite number of points. In contrast to the studies mentioned above, these authors argue that all moments of the forecast error distribution different from zero matter when expanding the loss function as a Taylor series. While Patton and Timmermann (2007a) mainly use that argument to point out that not only the loss function but also the density function of the forecast errors have to be symmetric in order to get unbiased optimal forecasts, Elliott and Timmermann's (2004) primary interest lies in optimal forecast combinations under asymmetric loss. Herein, Elliott and Timmermann rather focus on variance of the forecast error, than on its higher moments such as skewness and kurtosis and use loss functions that vanish after differentiating twice. Clatworthy et al. (2012) expand the Linex loss function and truncate the Taylor series at $k = 3$ in order to test whether earnings forecasts of financial analysts depend on the variance and skewness of past forecast errors.

Similar to these three studies, we conduct a Taylor expansion around $\mathrm{E}_t(e_{t+h}) = e^\star_{t+h}$. As the forecaster's preferences can be asymmetric and therefore producing biased forecasts can be rational, $e^\star_{t+h}$ does not necessarily have to equal zero. The fourth-order Taylor approximation of the loss function around the expected loss is

$$\begin{aligned}\mathrm{E}_t\left[L\left(e_{t+h}\right)\right] &= L(e^\star_{t+h}) + L^{(1)}(e^\star_{t+h}) \cdot \mathrm{E}_t\left[e_{t+h} - e^\star_{t+h}\right] \\ &+\frac{1}{2}L^{(2)}(e^\star_{t+h}) \cdot \mathrm{E}_t\left[\left(e_{t+h} - e^\star_{t+h}\right)^2\right] \\ &+\frac{1}{3!}L^{(3)}(e^\star_{t+h}) \cdot \mathrm{E}_t\left[\left(e_{t+h} - e^\star_{t+h}\right)^3\right] \\ &+\frac{1}{4!}L^{(4)}(e^\star_{t+h}) \cdot \mathrm{E}_t\left[\left(e_{t+h} - e^\star_{t+h}\right)^4\right] + O(e^4_{t+h}),\end{aligned} \tag{25}$$

with Taylor remainder $O(e^4_{t+h})$. As $\mathrm{E}_t\left[e_{t+h} - e^\star_{t+h}\right] = \mathrm{E}_t\left[e_{t+h}\right] - e^\star_{t+h} = 0$, the second summand vanishes. The expectations in the next three summands are the central but non-standardized moments of the forecast error and will be denoted σ^2_{t+h} (variance), s^3_{t+h} (skewness) and k^4_{t+h} (kurtosis) henceforth. Thus, equation (25) can be rewritten as

$$\mathrm{E}_t\left[L\left(e_{t+h}\right)\right] \approx L(e^\star_{t+h}) + \frac{1}{2}L^{(2)}(e^\star_{t+h}) \cdot \sigma^2_{t+h} + \frac{1}{3!}L^{(3)}(e^\star_{t+h}) \cdot s^3_{t+h} + \frac{1}{4!}L^{(4)}(e^\star_{t+h}) \cdot k^4_{t+h}. \tag{26}$$

In contrast to the maximization of the expected utility, the forecaster's objective here is to minimize the expected loss. For a given forecast error, a risk-averse and temperate forecaster thus expects loss to be larger when σ^2_{t+h} and k^4_{t+h} are large. This should be reflected by $L^{(2)}(e^\star_{t+h}) > 0$ and $L^{(4)}(e^\star_{t+h}) > 0$. The forecaster's preferences regarding the third moment depend on his or her asymmetry preferences. A forecaster averse to positive forecast errors (an underestimation of the target variable) is likely to move mass to the right in his or her underlying forecast error distribution in order to arm himself or herself against the uncertainty regarding that direction. Hence, such a forecaster will have a preference for negative skewness, i.e. $L^{(3)}(e^\star_{t+h}) > 0$, in oder to minimize his or her expected loss. For a negatively skewed unimodal distribution, the mean is left of the median. Thus, negative skewness might also partly account for a negative bias in the forecast errors. The opposite holds for a forecaster averse to negative forecast errors and with a preference for positive skewness.

In equation (26), we insert the Linex loss function (22) and its second to fourth derivative, $L^{(i)}\left(e_{t+h};\alpha\right) = \frac{\alpha^i}{\alpha^2}\exp\left(\alpha e_{t+h}\right)$ for $i \geq 2$ and the expected loss becomes

$$\mathrm{E}_t\left[L\left(e_{t+h};\alpha\right)\right] \approx \frac{1}{\alpha^2}\left[\exp(\alpha e^\star_{t+h}) \cdot \left(1 + \frac{\sigma^2_{t+h}}{2}\alpha^2 + \frac{s^3_{t+h}}{6}\alpha^3 + \frac{k^4_{t+h}}{24}\alpha^4\right) - \alpha e^\star_{t+h} - 1\right]. \tag{27}$$

Following the same procedure for Linex-Linex loss (24), the expected loss is approximated by

$$\begin{aligned} \mathrm{E}_t\left[L\left(e_{t+h};\alpha\right)\right] \approx\ & \exp([\alpha - I]\cdot e^{\star}_{t+h}) \times \left(1 + \frac{\sigma^2_{t+h}}{2}[\alpha - I]^2 + \frac{s^3_{t+h}}{6}[\alpha - I]^3 + \frac{k^4_{t+h}}{24}[\alpha - I]^4\right) \\ & - [\alpha - I]\cdot e^{\star}_{t+h} - 1, \end{aligned} \tag{28}$$

with I abbreviating the indicator function $I(e_{t+h} < 0)$. Given his asymmetry preferences expressed by α, the forecaster's optimization problem under these two loss functions consists of selecting the expected forecast error at each point in time t, in order to minimize the expected loss.

As the Taylor approximations of both loss functions contain the central moments of the forecast errors, the practical question of how to apply this approach without actually observing these moments arises. We overcome this obstacle by choosing several different approaches to reflect the moments. These approaches either are based on past values of the target variable, past forecast errors or the discrete probability forecasts of each individual forecaster. In sections 5.6 and 5.7, we present these approaches along with the corresponding results.

An alternative approach to the Taylor series expansion, that exclusively allows four moments to be included in the loss function, makes use of the assumption of a Gram-Charlier distribution of the target variable. While the Gram-Charlier distribution differs from the normal distribution in terms of its skewness and kurtosis, it nests the normal distribution as a special case. A profound discussion of Gram-Charlier distributions can be found in Jondeau and Rockinger (2001). Alongside Jondeau and Rockinger, the Gram-Charlier distribution has been discussed by Christodoulakis (2005) and Christodoulakis and Peel (2006). Christodoulakis (2005) offers a closed form solution for optimal forecasts under Linex loss and a first order condition for optimal forecast under Double Linex loss, with a closed form solution for the symmetric special case. Christodoulakis and Peel (2006) capture expected utility under an exponential utility function by Gram-Charlier.

Although, both the Taylor and the Gram-Charlier approach are feasible methods for including four moments, we focus on the Taylor approximation approach in the remainder of this study. This allows us to reduce the number of moments easily by choosing a different truncation order. All results presented in sections 5.6 and 5.7 relating to four moments also have been calculated using the Gram-Charlier approach. As these results do not differ considerably from the results using the Taylor series expansion, they are not shown here.

5.4 Optimality Condition

After having presented the Linex and Linex-Linex loss functions and having discussed the inclusion of higher moments of the forecast errors via a Taylor series approximation in both functions in the previous two sections, this section discusses how these loss functions can be used in an EKT-based setting to obtain information on the forecaster's asymmetry preferences. In accordance with the EKT framework, we assume that the forecaster establishes a forecast in order to minimize the expected loss, given a family of loss functions, the forecaster's preferences toward asymmetry measured by α and the information set Ω_t at time t. Similar to the EKT approach, we estimate the parameter of $\hat{\alpha}$ that minimizes the loss function for a given the series of forecast errors and the information set over time by using the following first order condition

$$\mathrm{E}\left[L^{(1)}(e_{t+h};\alpha)\,|\,\Omega_t\right] = 0, \tag{29}$$

where $L^{(1)}(.)$ is the first derivative of the loss function with respect to e_{t+h}. Forecasts are optimal if the forecast errors fulfill condition (29) above and there is no information left in Ω_t that can be exploited to further reduce the forecast errors. To test for optimality we use the orthogonality condition

$$\mathrm{E}\left[L^{(1)}(e_{t+h};\alpha)\cdot \boldsymbol{w}_t\right] = \mathbf{0}, \tag{30}$$

where $\boldsymbol{w}_t$ is any subset of instrumental variables from the information set Ω_t which are available when the forecast is established. The specific orthogonality condition for the derivative of the fourth-order Taylor approximation of the Linex loss function is

$$\mathrm{E}\left[\frac{1}{\alpha}\left(\exp(\alpha e_{t+h})\cdot\left(1+\frac{\sigma^2_{t+h}}{2}\alpha^2+\frac{s^3_{t+h}}{6}\alpha^3+\frac{k^4_{t+h}}{24}\alpha^4\right)-1\right)\cdot \boldsymbol{w}_t\right] = \mathbf{0}, \tag{31}$$

while the moment condition for the derivative of the fourth-order Taylor approximation of the Linex-Linex loss function is

$$\begin{aligned}\mathrm{E}\Bigg[[\alpha-I]\Bigg(\exp([\alpha-I]\,e_{t+h})\cdot\Bigg(1+\frac{\sigma^2_{t+h}}{2}\,[\alpha-I]^2\\ +\frac{s^3_{t+h}}{6}\,[\alpha-I]^3+\frac{k^4_{t+h}}{24}\,[\alpha-I]^4\Bigg)-1\Bigg)\quad\cdot \boldsymbol{w}_t\Bigg]\quad = \mathbf{0}.\end{aligned} \tag{32}$$

For the derivatives of the original loss functions (22) and (24) without higher moments, the orthogonality conditions reduce to equations (33) for Linex and (34) for Linex-Linex loss:

$$\mathrm{E}\left[\frac{1}{\alpha}\left(\exp(\alpha e_{t+h})-1\right)\cdot \boldsymbol{w}_t\right] = \mathbf{0} \tag{33}$$

$$\mathrm{E}\Big[[\alpha-I]\left(\exp([\alpha-I]\,e_{t+h})-1\right)\cdot \boldsymbol{w}_t\Big] = \mathbf{0}. \tag{34}$$

These moment conditions can be used to apply GMM to estimate the parameter α and check the validity of the orthogonality with $\boldsymbol{w}_t$ by the J-test. An objective value of the GMM estimation close to zero implies a small J-statistic and indicates that information is used efficiently. Hence, the procedure provides the estimate of an asymmetry parameter which is consistent with forecast rationality and reflects the forecaster's preferences concerning the asymmetry of the loss function.

After an introduction to the ECB's Survey of Professional Forecasters and a description of the dataset that provides the data for the realizations and the instrumental variables in the next section, the subsequent sections apply the methods discussed in the current and in the previous sections to these data.

5.5 The ECB's Survey of Professional Forecasters

In this section we present the two datasets to which the methods discussed above will be applied. We start by introducing the ECB's Survey of Professional Forecasters (SPF), the source of the forecasts we wish to analyze, and then present the Real Time Database (RTD), which is provided in the ECB's Statistical Data Warehouse.

5.5.1 History and Conceptional Design of the ECB's SPF

Before introducing the ECB's SPF, we briefly present its North American counterpart, or perhaps role model, the US SPF, the oldest quarterly survey of macroeconomic forecasts in the United States. It was started by the American Statistical Association and the National Bureau of Economic Research in 1968 and has been conducted by the Federal Reserve Bank of Philadelphia since 1990. In contrast to the ECB's SPF, which focuses on three main variables discussed below, the US SPF provides a long list of economic variables, including real and nominal output growth, (un)employment and different inflation measures, as well as industrial production, housing starts and several financial variables. Currently, about 170 professional forecasters from academic and commercial institutions participate in the survey. Apart from the quarterly projections for each of the next five quarters and the annual forecasts for up to three years, forecasters are also asked to provide a long-term forecast for the next 10 years for some variables, such as inflation and GDP growth. In addition, forecasters are invited to estimate the probability of a decline in GDP over the next five quarters. They also provide their probability forecasts for GDP growth, inflation and the unemployment rate for the next two years (or four in the case of unemployment). An early description of the survey is provided in Zarnowitz (1968). The data, a detailed documentation of the dataset, as well as a comprehensive academic bibliography, are provided on the Philadelphia Fed's website.[27] Some of the more recent studies that analyze the US SPF's forecasts are Elliott et al. (2008), Giannone et al. (2008), Rudebusch and Williams (2009), Capistrán and Timmermann (2009), Clements (2010, 2014) and Wang and Lee (2014). These studies focus predominantly on GDP growth and inflation forecasts, analyzing the forecasts' unbiasedness and rationality, their usefulness for nowcasting, potential inconsistencies between point and probability forecasts as well as the possibility of asymmetry in the forecasters' preferences.

The ECB's SPF was started in the first quarter of 1999 primarily to gain insight into the private sector's forecasts and expectations for macroeconomic developments in the euro area as a whole (see Garcia (2003) and Bowles et al. (2010)) and has been conducted quarterly ever since.[28] While only those features most relevant to the present analysis can be presented here, a detailed description of the survey can be found in Garcia (2003). In order to be considered for participation in the survey, professional forecasters must meet the following main criteria. On the one hand, forecasters are required to have a certain expertise considering euro area macroeconomics as well

[27] See https://philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters.

[28] The survey results are published on the ECB's website in the Monthly Bulletin (through 2014 Q4), and in the Economic Bulletin (from 2015:Q1). See http://www.ecb.europa.eu/stats/prices/indic/forecast/html/index.en.html.

as a track record of several years of publishing forecasts. On the other hand, the forecasting institutions are supposed to be independent from other participating institutions. Apart from these general requirements, the composition of the forecasters can be viewed under a regional and an occupational aspect. As some of the participants remain anonymous, the exact composition is unknown and can only be approximated by the list of known forecasters published on the ECB's website. Bowles et al. (2010) state that about 60 percent of the forecasters have a background in the financial sector and come predominantly from the original 12 euro area countries, as well as Denmark, Sweden and the United Kingdom, with a particularly strong representation of German (inside the euro area) and British (outside) forecasters. Moreover, Bowles et al. emphasize that neither economical nor sectoral weights are imposed, but that forecasters are selected by their capability and their willingness to contribute to the survey on a regular basis.[29]

Since the SPF has been conducted, the ECB has issued two special questionnaires. They are to be found in the appendices of the Monthly Bulletins for April 2009 and January 2014. In the latter questionnaire, more than 80 percent of the forecasters or forecasting institutions state that they do not prepare forecasts especially for the survey, but send their latest available forecast. The questionnaire also displays an interesting result regarding the forecasters' models. About one-fourth of the point forecasts with horizons one year ahead or less are said to be "essentially judgment-based," while 60 percent report their forecasts to be "model-based with judgmental adjustments." To the questions if and how the recent crisis has affected the forecasting model, the majority of the forecasters responded that the crisis did have an impact on their forecasts and they mainly adjusted with a "higher degree of judgment" in the forecast. Hence, the SPF's forecasts can be regarded essentially as judgmental forecasts. Another question addressed in the questionnaire is whether forecasters compute their forecasts directly for the entire euro area or do so by aggregating country-specific forecasts. While about one third of the respondents use both approaches, one of which is used to cross-check the results, more than 60 percent of the remaining respondents only compute one forecast for the entire euro area. This is remarkable considering that Marcellino et al. (2003) recommend using country-specific information for predicting inflation and growth in the euro area.

The main survey questions refer to the forecasters' expectations for the euro area regarding real GDP growth, the annual inflation of the Harmonized Index of Consumer Prices (HICP) and the seasonally adjusted unemployment rate for the current year, the following two calendar years and one and two years ahead. Forecasters are asked to give their point forecasts, along with their estimations of the probability that the outcome of each variable falls within certain intervals, i.e. discrete density forecasts. They are also asked for their expectations on the oil price (Brent crude oil), interest rate (main refinancing operations) and the USD/EUR exchange rate. Attached to each survey questionnaire are some basic reference data concerning the latest information on the three main variables.[30]

Our main focus in this study lies on the analysis of the point forecasts, although in section 5.7 the probability forecasts are used as well in order to approximate the moments of the forecast

[29] All authors currently work or used to work in the Euro Area Macroeconomic Developments Division of the ECB.

[30] For example, in the 2013:Q1 survey, the reference data are the annual HICP inflation for December 2012 (2.2%), the annual real GDP growth for 2012:Q3 (-0.6%) and the unemployment rate for November 2012 (11.8%).

errors. Concerning the forecast horizon, we focus on the one-year-ahead forecasts of these three variables, as a fixed horizon is better suited to control for the forecasters' information sets. If, for example, a forecaster is asked to forecast the real GDP growth for a certain year in each quarter of that year, this forecaster will posses more information each quarter and therefore find it easier to produce an accurate forecast. By contrast, the forecasting exercise does not get easier if the forecaster is asked for a one-year-ahead forecast for each quarter. Hence, forecasts made in different quarters can be compared with one another. The table below shows the survey rounds together with the horizons of the one-year-ahead or four-quarters-ahead forecasts for each variable, using 2013 as an example. The seemingly different forecast horizons across the three variables result from the discrepancy in the availability of the most current data for each variable. When the survey is conducted in the first quarter, for example, the most recent GDP realization stems from the third quarter of the previous year. Thus, the forecasters are asked to provide a one-year-ahead forecast for the third quarter of the current year. For HICP inflation and the unemployment rate, the latest realizations available in the first quarter date back respectively to December and November of the previous year.

Survey	GDP growth	HICP inflation	Unemployment rate
2013:Q1	2013:Q3	Dec 2013	Nov 2013
2013:Q2	2013:Q4	Mar 2014	Feb 2014
2013:Q3	2014:Q1	Jun 2014	May 2014
2013:Q4	2014:Q2	Oct 2014	Sep 2014

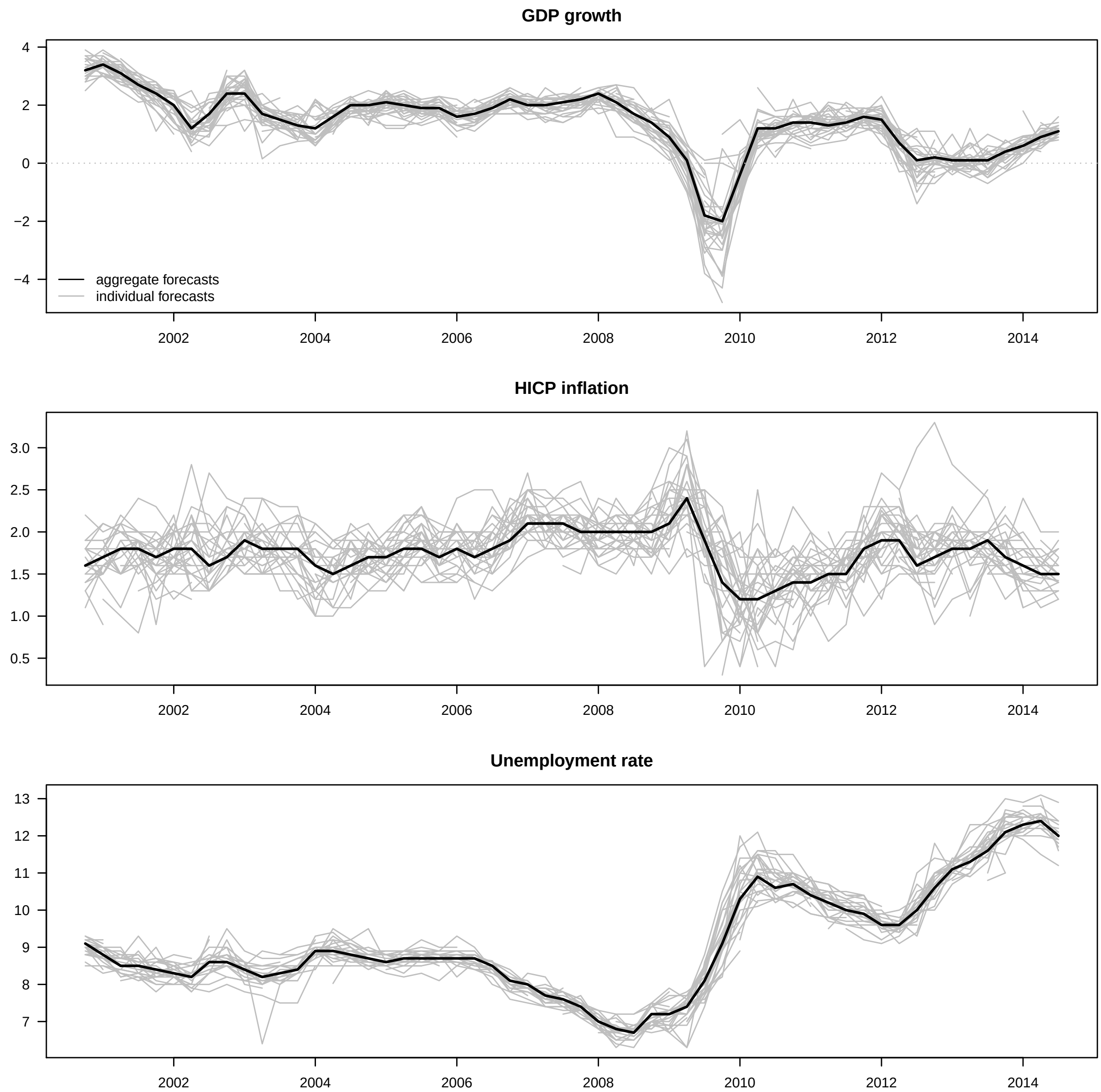
Since the SPF depends on the forecasting institutions' willingness to participate, the number of responses varies over time and for the different variables. For the period under consideration here, i.e. the forecasts made for inflation and unemployment in the survey rounds from 2000:Q1 to 2013:Q4 and for GDP growth from 2000:Q2 to 2014:Q1, the number of forecasters responding to the survey varies between 41 and 59 for GDP growth, between 41 and 64 for HICP inflation and between 36 and 59 for the unemployment rate. The mean and median number of responses are around 50 for GDP growth and inflation and around 46 for the unemployment rate. For the GMM tests conducted in section 5.6, only forecasters that participated at least in 40 of the 56 survey rounds in the observed period are considered. That leaves us with 31 (GDP), 33 (HICP) and 28 (unemployment) individual forecasters, plus the mean point forecast calculated from all forecasts provided in each particular survey round, henceforth referred to as the aggregate forecast. The table below briefly summarizes the number of forecasts these individual forecasters have provided for each variable:

	GDP growth	HICP inflation	Unemployment rate
Min	41	40	42
Mean	48.7	48.2	48.2
Median	49	49	48
Max	55	55	55

Figure 5.2 shows the forecasts of all individual forecasters with 40 or more survey responses (gray lines), along with the mean aggregate forecast of all forecasters (black line) for the three

variables at hand. Some gray lines show gaps that are due to survey rounds in which individual forecasters did not participate. The forecasts for all three variables reflect the recent crisis well. However, the forecasters' delay in actually detecting the crisis is notable. The gray lines also indicate an increased disagreement among the forecasters during the crisis.

Figure 5.2: Individual and aggregate forecasts



Forecasters predominantly predicted negative GDP growth rates for the last two quarters of 2009 and the first quarter of 2010 before returning to predictions of positive growth rates. Between the end of 2012 and the first half of 2013, forecasters returned to predicting growth rates close to zero. For HICP inflation, the forecasters predicted values close to the inflation target of 2 percent in most periods. Exceptions are the first quarter in 2009, for which the forecasters expected higher inflation, and the years 2010 and 2011, for which inflation expectations were rather low. The predicted value of the unemployment rate had been rather stable, around 9 percent, before it increased considerably between the second half of 2008 and 2010. After a short period of decreasing unemployment expectations, the predicted values of the unemployment rate increased until the end of the observed period.

5.5.2 Realization Data

In order to estimate the asymmetry parameter of the forecasters' loss functions and to test their forecast rationality, data for the realized values of the three variables are needed to calculate the forecast errors and to construct instrument sets that represent a subset of the forecasters' information. As argued by Croushore and Stark (2003) and Croushore (2011), among others, the information available to the forecaster at the time the forecast is produced can differ considerably from the information available after data have (potentially) been revised several times. Hence, only the information which was available at the time when each forecast was made is considered to be relevant to the instrument sets. Regarding the target variable, however, it could also be argued that the forecaster aims to forecast the revised instead of the real-time realization, as data revisions may be foreseeable. Thus, both realization variants are analyzed in the following sections and the differences between the real-time and the revised realization data are illustrated in figure 5.3 and discussed below.

All data used here, apart from the forecasts obtained from the SPF, can be downloaded from the Real Time Database, provided by the ECB's Statistical Data Warehouse and described in Giannone et al. (2012).[31] The real-time realizations are the first vintages of the data, with the exception of the HICP inflation, for which the first available realizations are flash estimates of the inflation.[32] For this variable the second vintage is used. For revised realizations, we use the most current vintage of each variable, i.e. the data published in the ECB's Monthly Bulletin in December 2014. For the instrument sets, the data available at the time of each survey round have been used. These data are published in the Monthly Bulletin in January, April, July and October of each year. While GDP growth is reported quarterly, HICP inflation and the unemployment rate each are reported in a monthly series. Quarterly means have been calculated to summarize the information available in each month. Considering the availability of the real-time data vintages since January 2001, as well as the publication lags of the data, which are about two quarters for GDP growth and two and three months for inflation and the unemployment rate respectively, we use the period of realization data from 2000:Q4 to 2014:Q3 and from 2001:Q1 to 2014:Q3 for the instruments.

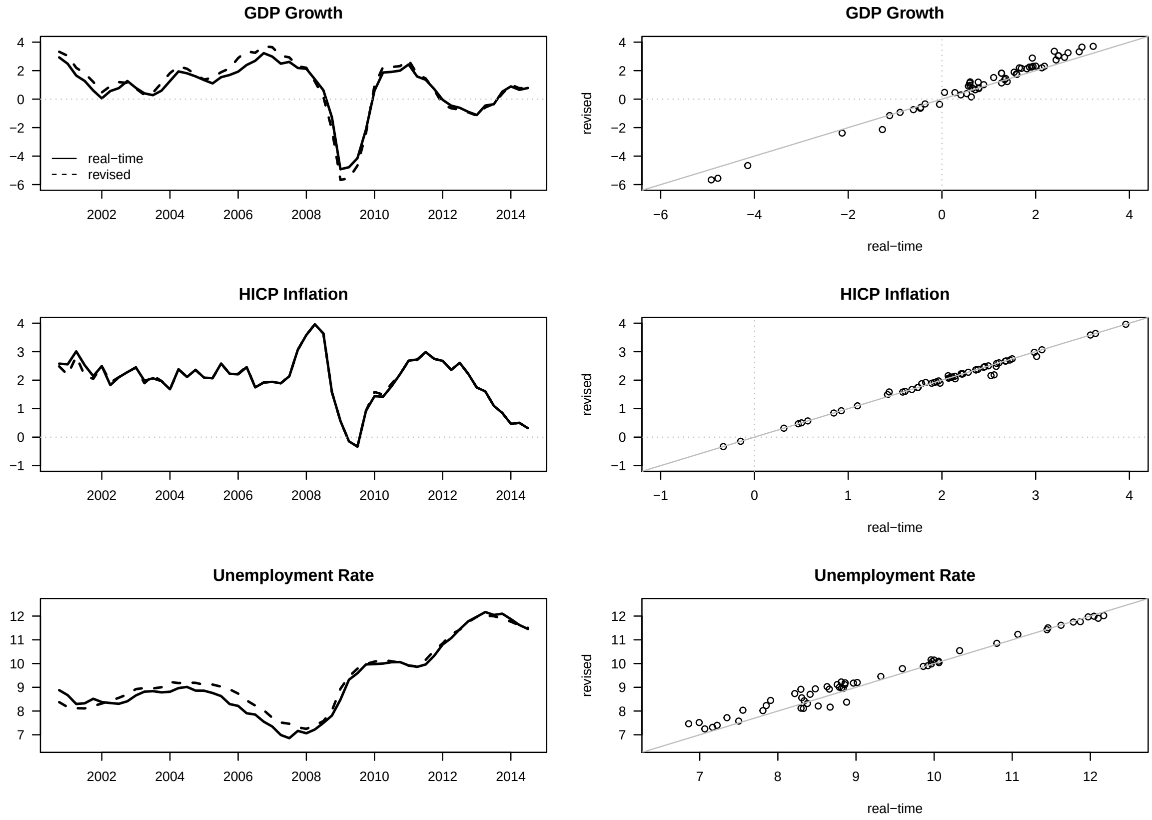
The first column of figure 5.3 depicts the time series of the real-time (solid line) and revised (dashed line) realizations from 2000:Q4 to 2014:Q3. The magnitude and direction of data revisions differ considerably between the three variables. The smallest revisions can be observed for HICP inflation, for which data hardly has been revised at all, except for the five downward revisions in the last quarter of 2000 and in 2001, and three upward revisions in the first quarters

[31] See https://www.ecb.europa.eu/stats/shared/download/stats/sdw/docu/databases/rtdb/rtdb_csv.html.

[32] Eurostat releases an early estimate of HICP inflation at the end of each reference month. The flash estimate is a combination of early HICP information from some euro area countries and one-step-ahead forecasts from countries that were not able to provide preliminary information. See http://ec.europa.eu/eurostat/statistics-explained/index.php/Inflation_%E2%80%93_methodology_of_the_euro_area_flash_estimate.

of 2010. Referring to Giannone et al. (2012), until 2009, the only source of revisions in the euro area had been a change in the weights for the aggregation of the national inflation indices for 2003.

Figure 5.3: Real-time and revised realizations



Compared to the inflation data, revisions of the real GDP growth and the unemployment rate have been more pronounced in the observed period. Until the last quarter of 2007, GDP was manly revised upward, followed by downward revisions during the crisis, especially in 2008:Q4 and the first two quarters of 2009, for which GDP was revised downward by approximately 0.8 percentage points. After five quarters of moderate upward revisions starting with 2010:Q1, GDP revisions recently have become less distinct, although the majority of the revisions are downward.

Turning toward the unemployment rate revisions, three phases of revision tendencies can be observed. While in the first six quarters all revisions are downward, from 2002:Q2 until 2010:Q3 all revisions are upward, with the highest values in 2005 and 2006 (around 0.6 percentage points). From 2010:Q4 on, revisions vary unsystematically around zero, with only small absolute values. For all three series, one has to bear in mind that the most recent quarters may not be fully revised yet.

In the second column of figure 5.3, the real-time realizations (x-axis) are plotted against the revised data (y-axis). This different visualization of the data revisions, on one hand, confirms the minor importance of data revisions for HICP inflation, as well as the overall tendency toward an upward revision for the unemployment rate, as the corresponding scatter plots lie on or above the 45° line, respectively. For GDP growth, on the other hand, it reveals a tendency for data revisions to be downward whenever GDP growth is negative and to be upward when it is positive.

5.5.3 SPF Forecast Errors

Table 5.1 below shows the mean error and the $RMSE$ for the aggregate forecasts and the individual forecasts produced by forecasters with at least 40 responses to the survey. Additionally, table C1 in the appendix provides analogous results for the median error and the MAE. Across the three variables, there are some individual forecasters who appear to slightly outperform the others in the observed period, while others perform better than average for one variable and on average or worse for the others. As an example, individual forecaster number twenty-two (id 22) performs above average in terms of the (symmetric) $RMSE$ measure for GDP and unemployment, while this forecaster's inflation forecasts are only marginally below average. Id 37 and id 39's forecasts are better than average for all three variables, while id 38 and id 48 both perform below average for growth, inflation and unemployment. According to Genre et al. (2013), the error statistics of different forecasters can only be compared with caution, as there are missing values in the individual series of forecasts, marking the survey rounds in which a forecaster did not respond. Hence, a forecaster that did not respond during the recent crisis, for which the forecasting exercise was more difficult than usual, is likely to perform better on average than a forecaster who participated during the crisis. This problem, although somewhat mitigated by only selecting for forecasters that participated at least 40 times, remains valid. Therefore, the $RMSE$ can only be regarded as indicating that the accuracy of the forecasts differs across forecasters.

Alongside the $RMSE$ statistic, it is of interest to see whether there is a systematic bias in the forecast errors. The negative mean error of GDP growth and the positive mean error of HICP support the results of Genre et al. (2013) and Andrade and Bihan (2013), who find that there seems to be a tendency to overpredict GDP growth along with a tendency to underpredict HICP inflation. For unemployment forecasts, the sign of the mean errors differs across forecasters and is closer to zero than for the other variables. This reflects the inconclusive results in Genre et al. (2013) concerning a bias in unemployment forecasts that is consistent over time. For the majority of the forecasters, mean errors are nevertheless rather positive and thus point toward a systematic underestimation, as detected by Andrade and Bihan (2013).

Considering the different error statistics of real-time and revised realizations, we find, on one hand, that the tendency for upward revision of GDP leads to a reduction of the bias, as the mean errors indicate. On the other hand, the pronounced downward revisions during the crisis cause an increase in the $RMSE$ statistics for revised data. At the same time, the small size of revisions in the HICP inflation data still leads to smaller revised error statistics in all but two cases, while the changing direction of data revisions for the unemployment rate is reflected in a non-systematic relation between real-time and revised error statistics. In spite of the unsystematic change in the $RMSE$ statistic for unemployment forecasts, the overall tendency toward upward revisions leads to an increased bias for most forecasters.

Table 5.1: Forecast error statistics

Forecaster	GDP growth				HICP inflation				Unemployment rate			
	Mean		*RMSE*		*Mean*		*RMSE*		*Mean*		*RMSE*	
	Real-time	Revised	Real-time	Revised	Real-time	Revised	Real-time	Revised	Real-time	Revised	Real-time	Revised
aggregate	-0.613	-0.453	1.264	1.544	0.274	0.262	0.881	0.871	0.119	0.258	0.673	0.677
id 1	–	–	–	–	0.215	0.200	0.918	0.907	–	–	–	–
id 4	-0.666	-0.517	1.026	1.273	0.369	0.352	0.758	0.746	0.130	0.279	0.692	0.689
id 5	-0.722	-0.540	1.456	1.749	0.302	0.282	0.914	0.902	0.192	0.338	0.798	0.775
id 7	-0.522	-0.392	1.297	1.531	0.527	0.515	0.930	0.930	-0.035	0.091	0.747	0.752
id 14	-0.503	-0.341	1.319	1.583	0.297	0.290	1.016	1.008	-0.101	0.043	0.677	0.721
id 15	-0.722	-0.583	1.466	1.737	0.241	0.228	0.925	0.919	0.204	0.334	0.736	0.746
id 16	-0.418	-0.279	1.361	1.665	0.230	0.220	0.927	0.913	0.083	0.231	0.742	0.744
id 20	-0.569	-0.398	1.275	1.553	0.319	0.306	0.881	0.875	0.085	0.237	0.728	0.688
id 22	-0.362	-0.198	0.995	1.247	0.395	0.383	0.854	0.830	-0.208	-0.071	0.610	0.626
id 23	-0.466	-0.328	1.329	1.573	0.266	0.255	1.053	1.045	0.150	0.285	0.773	0.780
id 24	-0.530	-0.374	1.393	1.670	0.348	0.336	0.912	0.907	0.078	0.220	0.698	0.700
id 26	-0.542	-0.391	1.272	1.496	0.248	0.232	0.936	0.924	0.285	0.402	0.756	0.776
id 29	-0.913	-0.803	1.581	1.863	0.195	0.189	1.094	1.086	–	–	–	–
id 31	-0.662	-0.477	1.368	1.639	0.319	0.307	0.873	0.863	0.124	0.284	0.799	0.783
id 33	-0.749	-0.556	1.405	1.692	0.010	-0.004	0.842	0.836	0.207	0.354	0.829	0.828
id 37	-0.529	-0.343	1.194	1.414	0.287	0.274	0.803	0.798	0.217	0.350	0.618	0.624
id 38	-0.296	-0.156	1.359	1.660	0.170	0.157	0.981	0.976	0.114	0.245	0.780	0.791
id 39	-0.445	-0.257	1.120	1.372	0.201	0.186	0.805	0.790	0.053	0.189	0.598	0.612
id 41	-0.676	-0.495	1.450	1.747	0.326	0.312	0.869	0.858	0.258	0.423	0.812	0.792
id 42	-0.666	-0.544	1.278	1.565	0.304	0.293	0.909	0.898	0.296	0.432	0.746	0.784
id 47	-0.289	-0.170	1.407	1.723	0.177	0.181	0.921	0.919	0.027	0.218	0.762	0.728
id 48	-0.805	-0.700	1.604	1.882	0.357	0.332	1.083	1.068	0.286	0.388	0.711	0.740
id 52	-0.465	-0.356	1.279	1.566	0.352	0.351	0.987	0.988	-0.073	0.091	0.781	0.803
id 54	-0.628	-0.446	1.348	1.608	0.252	0.246	0.941	0.930	0.013	0.176	0.886	0.884
id 56	-0.766	-0.585	1.464	1.744	0.372	0.354	0.855	0.841	0.284	0.464	0.750	0.690
id 82	-0.217	-0.108	1.428	1.754	0.317	0.314	1.031	1.024	–	–	–	–
id 85	-0.395	-0.163	0.885	1.057	0.403	0.387	0.706	0.696	–	–	–	–
id 89	-0.682	-0.514	1.406	1.678	0.133	0.118	0.928	0.917	0.390	0.528	0.712	0.760
id 90	-0.638	-0.450	1.481	1.755	0.425	0.410	0.864	0.849	0.126	0.281	0.719	0.736
id 93	–	–	–	–	0.039	0.035	0.924	0.916	–	–	–	–
id 94	-0.425	-0.271	1.237	1.472	0.135	0.122	1.013	1.001	0.044	0.173	0.712	0.747
id 95	-0.606	-0.454	1.263	1.552	0.221	0.208	0.899	0.885	0.135	0.278	0.688	0.670
id 96	-0.560	-0.395	1.153	1.439	0.115	0.104	0.897	0.886	0.059	0.212	0.665	0.670

5.5.4 Previous Studies that Consider the SPF

Along with the increasing number of survey rounds and thus a longer time series of observations that are available for conducting empirical tests, the number of empirical studies that evaluate the survey forecasts is growing. Bowles et al. (2010) is one of the first studies that evaluates the ECB's SPF. The authors analyze the GDP growth and unemployment forecasts between 1999:Q1 and 2008:Q4 and detect indications of bias in individual as well as aggregate forecasts. However, they argue that the small number of observations corrupts the general validity of the results. Being (former) ECB employees, the authors posses knowledge of the individual forecasters' backgrounds. Thus, they have the opportunity to look for significant differences between the forecast errors of financial and non-financial institutions in terms of their $RMSE$ statistics. However, they do not find any. Their study further emphasizes the different measures of uncertainty derived from the point forecasts and the density forecasts. They conclude that the aggregate uncertainty, measured by the standard deviation of the aggregate probability forecasts, should be preferred to the disagreement among forecasters measured by the standard deviation of the point forecasts. Moreover, they stress that due to possible limitations on individual forecasters for "fully internalis[ing] the overall level of macroeconomic uncertainty," (p. 25) other measures of uncertainty might be favorable. Again, the authors emphasize the limited general validity of their findings due to the short time period.

Another study discussing the matter of uncertainty in the ECB's SPF is Paloviita and Viren (2014). They use measures for uncertainty similar to Bowles et al. for the growth and inflation forecasts between 1999:Q1 and 2012:Q4 and find that the disagreement among forecasters is more sensitive to economic crisis compared to the individual forecasters' uncertainty measured by the variance of their probability forecasts. Most recently, Abel et al. (2015) discuss measures of uncertainty obtained form the survey respondents' probability forecasts. They find these measures to have been rising since 2007 and even observe a common behavior in the uncertainty measures and the consensus point forecasts for GDP growth and the unemployment rate.

Andrade and Bihan (2013) confirm the bias found by Bowles et al. in the one-year-ahead forecasts for an extended period from 1999:Q1 to 2012:Q4. While they find GDP growth to be systematically overestimated, inflation and unemployment are both systematically underestimated. In Genre et al. (2013), the primary focus is on various aggregation techniques in individual GDP, inflation and unemployment forecasts between 1999:Q3 and 2011:Q3 and whether they can beat the simple equal-weighted average. Although they find aggregation methods that are superior for single variables or horizons, none of the methods can outperform the simple average across variables and horizons. As a side result, they find negative bias in the GDP forecast and positive bias in the inflation forecasts, but only minimal evidence for a time-consistent bias in unemployment forecasts.

Pierdzioch et al. (2013) apply the EKT approach and find asymmetry in the survey's individual and pooled oil price forecasts between 2002:Q4 and 2010:Q4, with higher loss associated with an overprediction of the oil price. Even after controlling for asymmetry in the loss function, they find evidence against forecast rationality.

5.6 Estimating the Forecasters' Loss Functions

In this section, the methods discussed in sections 5.2 to 5.4 will be applied to the data presented in the previous section. Before turning to the results in subsections 5.6.2 to 5.6.4, we briefly will discuss the methods used to represent the higher moments of the forecast errors.

5.6.1 Computing Moments of the Forecast Errors

In order to be able to estimate the loss function's asymmetry parameters while allowing for higher moments, we address several variants for approximating the moments of the forecast errors. In this section, we use three rather traditional variants for representing the moments. We have to keep in mind that the purpose of including of higher moments is to account for the increased difficulty of forecasting a variable in economically turbulent times. Therefore, the most obvious way to calculate the higher moments if they are not observed directly is to simply use the lagged observations of the target variable over a fixed period, which ends at the point in time when the forecaster produces the forecast. Here, a period of five years has been used.[33] An evident disadvantage of this method is having to assign equal weights to each observation. Thus, we also calculate the moments using exponential weights as an alternative. Herein, the weights decrease exponentially for observations with higher lags. The exponentially weighted moments are calculated as

$$\begin{aligned} \mu_{t+h} &= \sum_{k=t-l-1}^{t} \omega_k y_k, \\ \sigma^2_{t+h} &= \sum_{k=t-l-1}^{t} \omega_k \left(y_k - \mu_{t+h}\right)^2, \\ s^3_{t+h} &= \sum_{k=t-l-1}^{t} \omega_k \left(y_k - \mu_{t+h}\right)^3, \\ k^4_{t+h} &= \sum_{k=t-l-1}^{t} \omega_k \left(y_k - \mu_{t+h}\right)^4, \end{aligned} \tag{35}$$

where l is the number of lags used and the weights ω_k are the standardized weights $\tilde{\omega}_k = \lambda^{t-k}$, with a choice of $\lambda = 0.9$. The standardization ensures that $\sum_{k=t-l-1}^{t} \omega_k = 1$. Fixing $\omega_k = \frac{1}{l}$ leads to the equally weighted version of the moments. Both approaches focus on the time dimension of the panel of forecasts.

As a third variant, the moments of the forecast errors have been calculated using the individual forecasts made in the survey round prior to each particular forecast. This exploits the cross-sectional dimension of the survey forecasts and has the advantage that only the most recent information is required to calculate the moments. Moreover, this variant allows the moments to

[33] Results were similar when using a shorter period of three years.

be interpreted as measures of disagreement between the forecasters. Thus, we have computed the moments as follows:

$$\begin{aligned}
\mu_{t+h} &= \frac{1}{n}\sum_{i=1}^{n} e_{i,t+h-1} = \bar{y}_{t+h-1} - \bar{f}_{t+h-1} = y_{t+h-1} - \bar{f}_{t+h-1}, \\
\sigma^2_{t+h} &= \frac{1}{n}\sum_{i=1}^{n} \left(e_{i,t+h-1} - \mu_{t+h}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} \left(f_{i,t+h-1} - \bar{f}_{t+h-1}\right)^2, \\
s^3_{t+h} &= \frac{1}{n}\sum_{i=1}^{n} \left(e_{i,t+h-1} - \mu_{t+h}\right)^3 = -\frac{1}{n}\sum_{i=1}^{n} \left(f_{i,t+h-1} - \bar{f}_{t+h-1}\right)^3, \\
k^4_{t+h} &= \frac{1}{n}\sum_{i=1}^{n} \left(e_{i,t+h-1} - \mu_{t+h}\right)^4 = \frac{1}{n}\sum_{i=1}^{n} \left(f_{i,t+h-1} - \bar{f}_{t+h-1}\right)^4, \qquad (36)
\end{aligned}$$

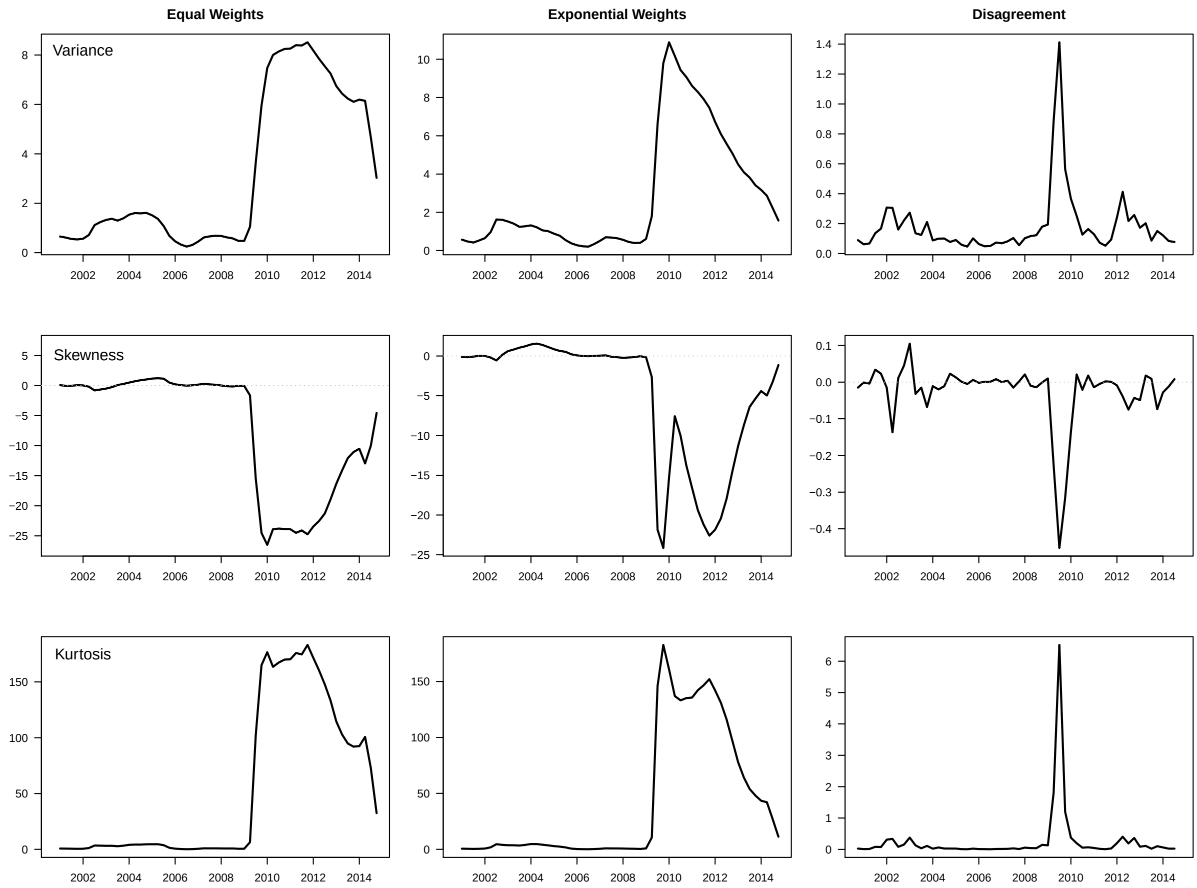
where n is the number of individual forecasters i included. The fact that the realizations of the target variable cancel out in these equations allows the moments to be calculated after the forecasts of the previous survey round are published, which always happens before the next survey round. Although there are mixed findings on whether a high discrepancy among forecasters is a good measure for macroeconomic uncertainty (see Bowles et al. (2010) and Paloviita and Viren (2014)), this method's potential benefit is in its use of only the most current information as opposed to several years of lagged information. In what follows, these three variants will be referred to as equally weighted moments, exponentially weighted moments and moments based on the disagreement among the individual forecasters. We will begin the following subsections with a short descriptive analysis of each variable's second to fourth moment computed with respect to the three methods described above. Then we will discuss the GMM results using these moments.

5.6.2 GDP Growth

Our discussion of the results in this and the following subsections is structured according to the variables we wish to analyze, starting with GDP growth. After shortly discussing the moments used, we turn to the GMM estimation results for each variable. The estimation is implemented using the iterative GMM discussed in Hansen et al. (1996) with a quadratic spectral kernel and bandwidth choice according to Andrews (1999). In our setup, we use the iterative GMM instead of the continuously updating estimator, as the former has proved to be more robust with respect to different starting values.

Figure 5.4 shows the development of the second to fourth moment over time. Herein, each column depicts another variant for computing the moments. These are the equally weighted past values of the target variable in the first column, the exponentially weighted past values in the second column and the disagreement between the forecasters in the third column. In the first row, we find the variance to be rather moderate until the second quarter of 2009, at which point it starts to increase considerably for all three variants. However, the peak of this increase differs for the variants. The variant using the equal weights continues to increase until 2011:Q4 before it starts to decline, while the exponentially weighted variant reaches its maximum in 2010:Q1. This nicely illustrates how the exponential weights discount past weights. Thus, this method is affected by the large negative growth rates during the crisis for a shorter period. For disagreement, the third proxy of the variance, the peak is already in 2009:Q3, suggesting that forecasters disagreed the most right before the growth rate started to rise again. The skewness is initially close to zero for all three variants, although the third variant shows some variation between 2002 and 2004. We observe a decrease in skewness that starts in 2009:Q2. Apart from decreasing instead of increasing, the patterns for the three variants are similar to their variance patterns, i.e. the disagreement proxy reaches its turning point first and then returns to its previous pattern, while the two other variants remain at their negative levels longer. The sharp rise followed by another drop in the exponentially weighted variant in the first half of 2010 is caused by a series of negative growth rates of similar magnitude, which lead to (exponentially weighted) mean values close to the actual values. The proxies for the variance are shown in the last row of figure 5.4. Here, we observe behavior similar to that of the variance, but with less volatility before the crisis, and, in the case of the disagreement proxy, after the crisis.

Figure 5.4: Proxies for the moments of the forecast error - GDP growth



The results concerning the form of the forecasters' loss functions are summarized in tables 5.2 to 5.7, and each variable of interest is addressed with two tables. The first table for each variable shows the results using the Linex loss function (see equation (22)) and the second shows the results that were obtained under the Linex-Linex function (equation (24)). The tables are all structured as follows. The upper half of every table exhibits the results using real-time realizations to calculate the forecast errors, while the lower part contains the revised realization results. The first column presents the base results with respect to the specific loss function addressed in the table, whereas the remaining columns depict the results when the second to fourth moment is included. First, the results for the equally weighted past observations are presented, followed by the exponentially weighted past observations and the disagreement of the forecasters. For each of the two realization variants there are three types of results in the rows of the table, starting with the GMM estimates of α averaged across the instrument sets that are described below. Second, the percentage of estimates that lead to a rejection of the hypothesis of symmetry is shown and finally the percentage of rejections of the rationality hypothesis, tested with a J-test. Both percentages of rejections refer to a significance level of 5 percent.

For each of the individual forecasters, as well as the mean aggregate of all individual forecasters, the following 16 sets of instruments are tested. The first three instrument sets (1 to 3) test for weak efficiency, while sets 4 to 15 represent further information that should be contained in the forecaster's information set. Set 0 merely contains a constant and sets 1 to 3 are formed with a constant and variants of the lagged forecast errors (that is (1) first lags, (2) first and second lags and (3) first lags and squared first lags), whereas the first lags refer to the most recent forecast errors known to the forecaster and second lags refer to the penultimate errors accordingly. Analogous to the construction of sets 1 to 3, a constant along with the first lag, the first and second lag and the first lag and its squares of the variable are tested for each variable. Sets 4 to 6 are formed around the lagged GDP growth, sets 7 to 9 around lagged HICP inflation and sets 10 to 12 around the lagged unemployment rate. The remaining three sets (13 to 15) each contain a constant, first lags of two of the three variables, and their interaction term. Again, the lags of the three variables are defined using the latest information available to the forecasters, i.e. the last two quarters. For the monthly data, quarterly averages are calculated.

In order to keep the tables' dimensions at a reasonable size, we report only the results of the aggregate or consensus forecast, the average across all individual results and the results of three exemplary individual forecasters, id 14, id 22 and id 48. These specific forecasters have been selected because they are well suited to demonstrate the different asymmetry preferences across forecasters. The results of the asymmetry parameters for all individual forecasters as well as the percentages of rationality rejections for the single instruments (averaged across forecasters) can be found in tables C2 to C10 in the appendix. We particularly emphasize the difference between the aggregate forecast and the average of the individual results. The results of the former imply an identical loss function for all individual forecasters, as the loss parameter for the time series of average forecast errors is calculated, while the latter focuses on the form of the loss function on average, as the asymmetry parameters $\hat{\alpha}$ are first estimated for each forecaster individually and then their mean is calculated. According to previous studies that argue in favor of testing forecast rationality separately for each individual forecaster instead of pooling or aggregating the data (see e.g. Hirsch and Lovell (1969), Figlewski and Wachtel (1983) and Bonham and Cohen (2001)), we expect rationality to be rejected more frequently for the aggregate forecasts, compared to the average share of rationality rejections.

Starting with the GDP growth forecasts, table 5.2 exhibits the results using the Linex loss function. Looking at the first column of the table, the loss function of all forecasters seems to be rather asymmetric, with a higher weight on positive forecast errors (underestimation), as the positive estimates of $\hat{\alpha}$ indicate. This complements the results of Andrade and Bihan (2013), who find GDP to be constantly overestimated, which could be explained by the forecasters' attempt to avoid positive errors. Krüger and Hoss (2012) give a possible economic explanation for this direction of the asymmetry in the loss function by arguing that underpredicting growth could unduly constrain decision-makers.

Table 5.2: Results under Linex loss - GDP growth

	Forecaster	**Base**	**Equal Weights**			**Expon. Weights**			**Disagreement**		
			σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4
					Real-time						
	aggregate	2.175	0.301	0.592	0.245	0.381	0.775	0.356	1.657	1.687	1.589
	average	1.454	0.218	0.314	0.220	0.281	0.377	0.268	1.132	1.152	1.112
$\hat{\alpha}$ estimates	id 14	1.083	0.185	0.312	0.196	0.224	0.305	0.224	0.897	0.905	0.874
	id 22	1.519	0.169	0.194	0.180	0.236	0.287	0.241	1.182	1.197	1.150
	id 48	1.039	0.173	0.230	0.180	0.187	0.275	0.191	0.871	0.880	0.871
	aggregate	100	81	63	88	81	69	81	100	100	100
	average	69	61	44	61	65	59	64	72	72	72
$H_0 : \alpha \leq 0$	id 14	44	25	13	25	25	31	25	44	44	44
	id 22	56	75	69	75	75	63	75	56	56	63
	id 48	50	56	38	56	50	50	56	50	50	50
	aggregate	60	73	40	67	60	53	60	67	73	67
	average	45	44	46	43	41	44	41	46	47	46
$H_0 : J = 0$	id 14	20	13	7	13	13	20	20	13	13	13
	id 22	47	33	40	33	27	27	27	53	60	53
	id 48	40	40	33	40	33	33	33	40	40	40
					Revised						
	aggregate	1.119	0.142	0.220	0.148	0.223	0.315	0.237	1.036	0.969	1.058
	average	0.654	0.120	0.157	0.124	0.152	0.184	0.152	0.568	0.569	0.576
$\hat{\alpha}$ estimates	id 14	0.630	0.120	0.137	0.128	0.152	0.185	0.153	0.483	0.483	0.470
	id 22	0.337	0.042	0.046	0.045	0.042	0.046	0.045	0.222	0.224	0.204
	id 48	0.623	0.174	0.363	0.185	0.255	0.335	0.251	0.565	0.568	0.566
	aggregate	63	81	81	81	81	81	81	69	63	63
	average	51	51	49	50	53	51	53	52	52	51
$H_0 : \alpha \leq 0$	id 14	38	31	25	25	44	44	44	38	38	38
	id 22	50	38	38	38	38	38	38	50	50	50
	id 48	69	56	63	50	63	63	63	63	63	63
	aggregate	53	67	53	60	53	53	60	60	60	60
	average	48	46	45	45	46	47	46	49	49	48
$H_0 : J = 0$	id 14	60	33	47	33	40	53	40	47	47	47
	id 22	27	33	33	33	40	40	40	27	27	20
	id 48	27	27	13	27	27	20	27	27	27	27

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets, while $H_0 : \alpha \leq 0$ and $H_0 : J = 0$ show the rejection frequencies across instruments on a 5 percent level.

Regarding the types of realizations, although the asymmetry is stronger for the real-time results, it persists for revised realizations. This reduction of the degree of asymmetry between the two types of realizations is plausible because of the tendency for upward revision of GDP growth, which shifts negative forecast errors toward zero. It is striking that the aggregate forecasts of the different forecasters are represented best by a more asymmetric loss function than the average of the individual forecasters. However, single large forecast errors of an individual forecaster can influence the aggregate forecast error considerably and thus increase the asymmetry in the respective loss function. In accordance with this higher degree of asymmetry, a higher share of parameters has values significantly above zero.

Regardless of the variant of moments that is used, including the variance appears to affect the asymmetry parameter and leads to less asymmetry in the loss function. We account for the forecasters' risk aversion by assigning more weight and thus larger loss to periods with an increased uncertainty. The effect is most pronounced for the equally weighted moments and weakest for the moments based on the disagreement among the forecasters. Because a skewed forecast error distribution can be a reason for bias in the forecast errors, including the third moment in the loss function could be expected to further reduce the degree of asymmetry. Accounting for the forecasters' temperance by including the fourth moment should influence the loss function in similar way as including the variance. However, the further inclusion of the third and fourth moment do not bear the expected effect. Although, including the fourth moment seems to reinforce a slightly more symmetric loss function in some cases.

Considering the percentage of rationality rejections, there is a difference between the individual forecasters, as some produce forecasts with less evidence against rationality than others (see e.g. id 14 for real-time and id 48 for revised realizations). For the aggregate forecasts, more than half of the estimation results suggest a rejection of forecast rationality. Across forecasters, there is no systematic pattern between the proposed extensions of the loss function and the percentage of rationality rejections. Regarding the sixteen instrument sets, we observe the number of rejections to be higher for sets that include lagged HICP inflation (i.e. sets 7 to 9, 13 and 15), while there are almost no rejections for sets 4 (lagged GDP), 10 and 12 (lagged unemployment rate).

Turning to table 5.3, which contains the analogous results under a Linex-Linex loss function, the general direction of asymmetry in the loss function persists, although results differ from the Linex results discussed above in a number of aspects. First, due to the different parametrization of the two loss functions, the asymmetry estimates now lie between zero and one, with 0.5 representing symmetric loss and values above (below) 0.5 representing asymmetry in terms of higher weights for positive (negative) forecast errors. Second, the tendency toward a less asymmetric loss function after the inclusion of a second moment persists, but is more moderate for all versions of the moments and weakest for the moments based on the disagreement between the forecasters. Moreover, for equally weighted moments after the inclusion of the third moment, the degree of asymmetry in the loss function is now as high as it is in the variant without moments.

Considering the H_0 of symmetry, the number of rejections is slightly higher here in comparison to the Linex results above. Given the smaller range of possible parameters and hence smaller absolute deviations from the symmetric case, this might indicate smaller standard errors of the α

estimates and thus a more efficient estimation. Compared to the Linex loss function, the number of rejections of the forecast rationality is considerably smaller under the Linex-Linex loss function for both types of realizations. The rationality of the aggregate forecasts can be rejected more frequently than the individual forecasters' rationality for real-time realizations. Regarding the instrument sets, the sets that only contain GDP growth (sets 4 to 6) or the unemployment rate (sets 10 and 12) now show the fewest rejections of rationality. Hence, the information in these variables seems to be adequately incorporated in the forecasts.

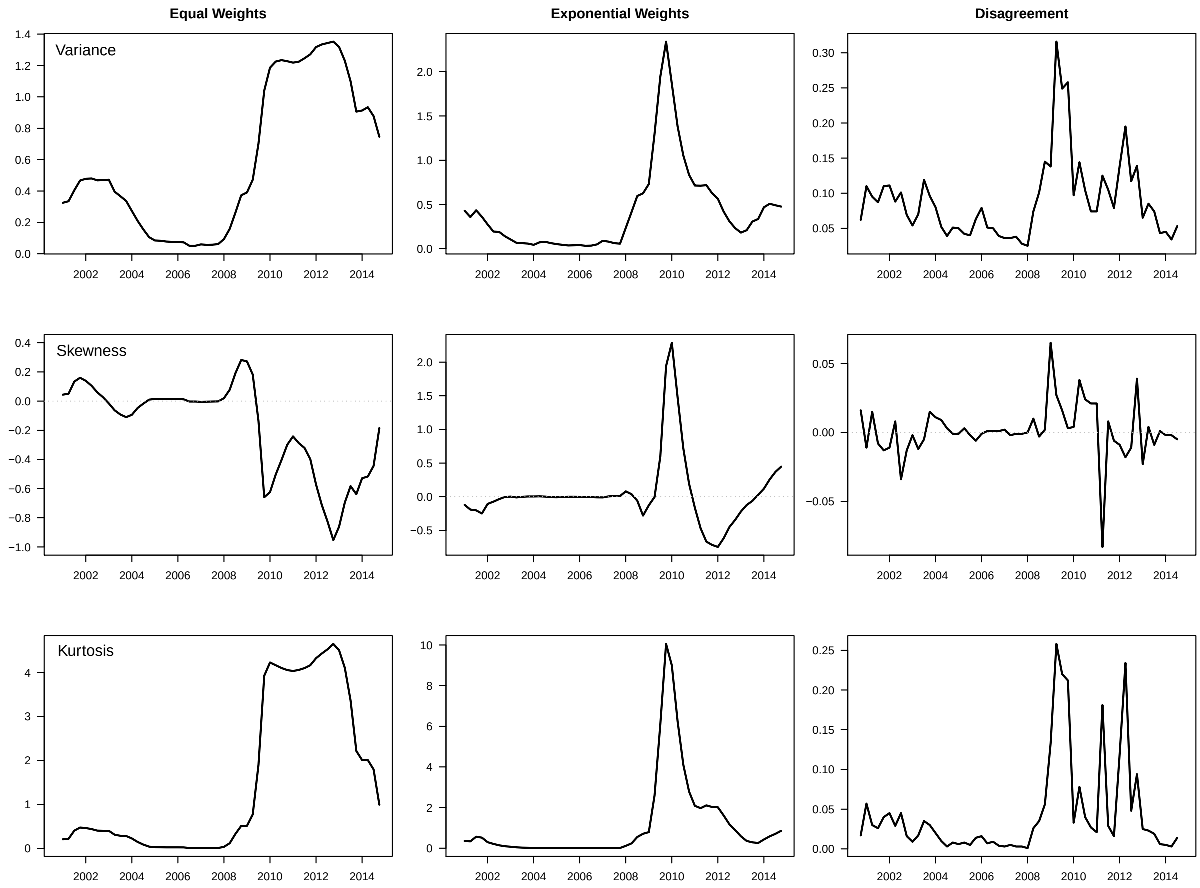
Table 5.3: Results under Linex-Linex loss - GDP growth

	Forecaster	**Base**	**Equal Weights**			**Expon. Weights**			**Disagreement**		
			σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4
					Real-time						
	aggregate	0.730	0.699	0.723	0.693	0.694	0.709	0.687	0.710	0.710	0.710
	average	0.705	0.644	0.716	0.649	0.642	0.669	0.644	0.694	0.696	0.695
$\hat{\alpha}$ estimates	id 14	0.694	0.644	0.708	0.644	0.640	0.664	0.629	0.663	0.663	0.662
	id 22	0.699	0.657	0.693	0.664	0.649	0.668	0.653	0.692	0.692	0.692
	id 48	0.756	0.617	0.766	0.627	0.594	0.618	0.602	0.750	0.750	0.750
	aggregate	100	100	100	100	94	100	100	100	100	100
	average	85	73	94	87	71	88	80	85	86	85
$H_0: \alpha \leq 0.5$	id 14	88	63	100	69	63	75	56	81	81	81
	id 22	81	69	100	100	56	100	100	88	88	88
	id 48	88	50	100	69	50	56	56	94	94	94
	aggregate	47	40	40	47	33	33	27	47	47	47
	average	25	23	30	21	23	23	23	24	24	24
$H_0: J = 0$	id 14	0	7	7	7	7	7	7	7	7	7
	id 22	13	7	27	7	0	0	0	13	13	13
	id 48	20	13	7	20	13	20	20	13	13	13
					Revised						
	aggregate	0.630	0.642	0.634	0.652	0.628	0.650	0.649	0.629	0.629	0.629
	average	0.626	0.594	0.636	0.609	0.595	0.615	0.606	0.622	0.623	0.624
$\hat{\alpha}$ estimates	id 14	0.617	0.573	0.607	0.580	0.549	0.610	0.559	0.615	0.616	0.616
	id 22	0.618	0.584	0.606	0.600	0.583	0.609	0.607	0.616	0.616	0.616
	id 48	0.697	0.652	0.835	0.634	0.624	0.642	0.614	0.692	0.693	0.692
	aggregate	75	69	100	100	63	88	88	75	75	75
	average	59	57	78	73	54	69	69	62	63	63
$H_0: \alpha \leq 0.5$	id 14	56	44	50	38	38	50	38	56	56	56
	id 22	50	50	81	69	44	63	63	50	50	50
	id 48	63	75	81	69	63	69	63	69	69	69
	aggregate	7	40	13	27	27	27	13	7	7	7
	average	16	23	27	27	22	23	23	15	16	15
$H_0: J = 0$	id 14	7	27	20	13	13	13	13	7	7	7
	id 22	0	13	13	20	13	13	13	0	0	0
	id 48	0	0	0	0	0	0	0	0	0	0

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets, while $H_0: \alpha \leq 0.5$ and $H_0: J = 0$ show the rejection frequencies across instruments on a 5 percent level.

5.6.3 HICP Inflation

Figure 5.5: Proxies for the moments of the forecast error - HICP inflation



Before turning to the GMM estimates for HICP inflation, figure 5.5 shows the proxies for the second to fourth moment for this variable. In the first row, the variance, again peaks for all three variants as a reaction to the crisis. This peak can be observed first for the variant based on the forecasters' disagreement. The other two variants nicely illustrate the effect of the weighting, as the exponentially weighted variant reaches its turning point in 2009:Q4, while the equally weighted variant increases until 2012:Q4. Before and after the crisis, we find the third variant to be most volatile. However, we observe a period of elevated variance for the first variant between 2000 and 2004. The proxies for the skewness reflect the period of high inflation between 2007:Q4 and 2008:Q3, followed by a sharp decrease until 2009:Q3 and a subsequent increase until 2011:Q3. Especially the proxy based on the forecasters' disagreement has turning points at the same time as inflation or even leads it. Nevertheless, we find the first two proxies to be lagging because of the use of past observations. This can be seen best for the second variant and the period 2009:Q3 to 2010:Q3, during which inflation is positive and the exponentially weighted mean value is negative. Thus this skewness variant has large positive values in this period. For the first variant, skewness is permanently negative after 2009:Q3. This can be explained with the equally weighted mean inflation, which persists on a higher level than most of the more recent inflation values. As for GDP growth, the kurtosis shows a pattern similar to the variance for all three variants. Again, the peaks are more pronounced than they are for the variance; otherwise there is less volatility.

Table 5.4: Results under Linex loss - HICP inflation

	Forecaster	**Base**	**Equal Weights**			**Expon. Weights**			**Disagreement**		
			σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4
					Real-time						
	aggregate	-1.079	-0.887	-0.838	-0.801	-0.777	-0.819	-0.763	-1.006	-1.005	-0.999
	average	-1.122	-0.784	-0.745	-0.710	-0.728	-0.772	-0.713	-0.987	-0.986	-0.972
$\hat{\alpha}$ estimates	id 14	0.082	-0.064	-0.091	-0.067	0.219	0.258	0.159	0.054	0.053	0.053
	id 22	-1.267	-0.950	-0.869	-0.792	-0.891	-0.939	-0.829	-1.155	-1.151	-1.138
	id 48	-0.976	-0.789	-0.758	-0.735	-0.747	-0.779	-0.738	-0.908	-0.908	-0.902
	aggregate	69	69	69	69	69	69	69	69	69	69
	average	62	64	64	65	62	64	64	64	64	64
$H_0:\ \alpha \geq 0$	id 14	6	13	13	13	6	6	6	6	6	6
	id 22	81	75	75	88	81	81	81	81	81	81
	id 48	81	81	81	81	81	81	81	81	81	81
	aggregate	87	80	80	80	87	87	87	87	87	87
	average	73	73	73	73	73	73	73	73	73	73
$H_0:\ J = 0$	id 14	93	93	87	93	93	87	80	87	87	87
	id 22	80	80	80	80	73	80	73	80	80	80
	id 48	67	67	67	67	67	67	67	67	67	67
					Revised						
	aggregate	-1.085	-0.908	-0.866	-0.832	-0.792	-0.835	-0.782	-1.017	-1.016	-1.010
	average	-1.095	-0.787	-0.756	-0.716	-0.722	-0.758	-0.702	-0.975	-0.971	-0.959
$\hat{\alpha}$ estimates	id 14	0.168	0.006	-0.028	-0.025	0.294	0.285	0.274	0.136	0.143	0.131
	id 22	-1.259	-0.970	-0.892	-0.825	-0.906	-0.963	-0.862	-1.154	-1.150	-1.138
	id 48	-0.967	-0.798	-0.769	-0.750	-0.750	-0.784	-0.743	-0.902	-0.902	-0.897
	aggregate	69	69	69	69	69	69	69	69	69	69
	average	62	65	64	65	62	63	63	64	63	64
$H_0:\ \alpha \geq 0$	id 14	6	13	13	13	6	6	6	6	6	6
	id 22	81	75	75	88	81	81	81	81	81	81
	id 48	81	81	81	81	81	81	81	81	81	81
	aggregate	87	80	80	80	87	87	80	87	87	87
	average	67	67	67	67	67	60	67	67	67	67
$H_0:\ J = 0$	id 14	93	87	80	80	87	93	87	93	93	93
	id 22	80	87	87	87	73	80	73	80	80	80
	id 48	67	67	60	60	67	67	67	67	67	67

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets, while $H_0:\ \alpha \geq 0$ and $H_0:\ J = 0$ show the rejection frequencies across instruments on a 5 percent level.

The GMM results for the HICP inflation forecasts are exhibited in tables 5.4 and 5.5. Once again, the former shows the results under Linex loss and the latter shows the results obtained under the Linex-Linex loss function. Data revisions for the HICP inflation data analyzed here are minor. The results therefore hardly differ with respect to the choice of the variant of realizations and can be discussed jointly.

Beginning with the Linex loss results and the general tendency toward asymmetry in the loss function, we observe that loss again predominantly seems to be asymmetric, but compared to GDP forecasts, shows a change in the direction of the asymmetry. Forecasters appear to be averse to negative forecast errors (overprediction) when forecasting inflation. Although this finding seems to contradict the assumption that forecasters tend to be averse to high inflation and therefore might rather aim to avoid underestimating it, this is supported by Andrade and Bihan (2013), who find that inflation forecast errors are systematically positive. One reason for the unexpected direction of the forecasters' preferences in the loss function could be that inflation rates are predominantly moderate or low for the observed period. As a consequence, aversion to very low inflation or even deflation might be higher, compared to the fear of high inflation rates. See also Capistrán (2008), who makes a similar case.

Turning to the specific asymmetry parameters of the individual forecasters, there is one forecaster (id 14) whose loss function appears to be rather symmetric under Linex loss, as $\hat{\alpha}$ is close to zero and rarely differs significantly from zero. This holds regardless of the inclusion of further moments. The other forecasters appear to be producing their forecasts following an asymmetric loss function. The same holds for the aggregate forecasts. Similar to GDP forecasts, including the second moments reduces the degree of asymmetry in the loss function, as the asymmetry parameters closer to zero show. Although the absolute reduction of asymmetry, when including the second moment, is smaller than for GDP growth, it now differs for the three variants that represent the moments. The strongest reduction can be found for the equally and the exponentially weighted moments, while the reduction using the moments based on disagreement induce less reduction. However, the degree of asymmetry in id 14's loss function increases for exponentially weighted moments. The inclusion of further moments only leads to marginal changes in the loss parameter. The symmetry hypothesis, $H_0 : \alpha \geq 0$, is rejected on average in approximately two-thirds of the cases, with a broad variation across forecasters.

Considering the percentage of rationality rejections, none of the three individual forecasters whose results are shown in the table seems to perform better than the aggregate forecasts, while on average, rationality is rejected less often than for the aggregate forecast across all individual forecasters who participated in the survey at least 40 times. As for the GDP growth forecasts, rationality is rejected with different frequencies for different instrument sets. For example, sets 1 to 3, which test for weak efficiency, lead to a rejection for less than 15 percent (sets 1 and 2) and 30 percent (set 3) of the forecasters. Another interesting result concerns the percentage of rejections in the instrument sets that contain the lagged unemployment rate. While there are only a few rejections for sets 10 (first lags) and 12 (first lags and squared first lags), the inclusion of the second lags in set 11 leads to a share of rationality rejections of about 90 percent.

In table 5.5, which shows the analogue results given a Linex-Linex loss function, the overall tendency toward asymmetry is the same as under Linex loss. That is to say, forecasters are averse to falsely overpredicting inflation. For Linex-Linex loss, this is indicated by the estimates of $\hat{\alpha}$ below 0.5. Interestingly, under this loss function, the asymmetry parameters show less variation between the different forecasters, although for id 14, whose loss preferences appear to be symmetric under Linex loss, there are less rejections of the symmetry hypothesis and more rejections of rationality than on average. Including a second moment once again reduces the degree of asymmetry in the loss function, while the inclusion of further moments, as we have seen before, yields no additional gain. Between the three variants for calculating moments, the exponentially weighted moments now lead to the highest reduction of asymmetry.

Table 5.5: Results under Linex-Linex loss - HICP inflation

	Forecaster	**Base**	**Equal Weights** σ^2	s^3	k^4	**Expon. Weights** σ^2	s^3	k^4	**Disagreement** σ^2	s^3	k^4
					Real-time						
	aggregate	0.283	0.312	0.313	0.316	0.317	0.316	0.319	0.290	0.290	0.290
	average	0.303	0.326	0.327	0.330	0.329	0.329	0.334	0.308	0.309	0.309
$\hat{\alpha}$ estimates	id 14	0.297	0.364	0.367	0.371	0.404	0.403	0.455	0.306	0.306	0.307
	id 22	0.236	0.258	0.259	0.262	0.270	0.268	0.271	0.243	0.243	0.244
	id 48	0.275	0.296	0.298	0.300	0.303	0.302	0.304	0.282	0.282	0.282
	aggregate	75	69	69	69	63	63	63	75	75	75
	average	71	70	69	69	68	68	67	71	71	71
$H_0:\ \alpha \geq 0.5$	id 14	69	63	63	63	50	50	38	69	69	69
	id 22	88	81	81	81	81	81	81	81	81	81
	id 48	81	75	75	75	81	81	81	81	81	81
	aggregate	0	7	7	7	0	0	0	0	0	0
	average	10	12	12	12	11	11	11	10	10	9
$H_0:\ J = 0$	id 14	13	27	27	33	33	20	33	7	7	7
	id 22	0	7	13	13	0	0	0	0	0	0
	id 48	0	0	0	0	0	0	0	0	0	0
					Revised						
	aggregate	0.292	0.314	0.316	0.318	0.320	0.319	0.322	0.299	0.299	0.299
	average	0.306	0.330	0.331	0.334	0.334	0.335	0.338	0.312	0.312	0.312
$\hat{\alpha}$ estimates	id 14	0.327	0.396	0.399	0.401	0.455	0.477	0.472	0.336	0.336	0.336
	id 22	0.240	0.262	0.263	0.266	0.271	0.269	0.273	0.248	0.248	0.248
	id 48	0.279	0.302	0.303	0.306	0.306	0.305	0.308	0.286	0.286	0.286
	aggregate	69	69	69	69	63	63	63	69	69	69
	average	71	69	68	68	67	67	66	70	70	70
$H_0:\ \alpha \geq 0.5$	id 14	63	50	50	50	38	31	31	63	63	63
	id 22	81	81	81	81	81	81	81	81	81	81
	id 48	81	75	75	75	81	81	81	81	81	81
	aggregate	7	7	7	7	0	0	0	7	7	7
	average	10	12	12	12	11	12	12	10	10	10
$H_0:\ J = 0$	id 14	27	33	33	33	40	40	33	20	20	20
	id 22	0	7	13	13	0	0	0	7	7	7
	id 48	0	0	0	0	0	0	0	0	0	0

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets, while $H_0:\ \alpha \geq 0.5$ and $H_0:\ J = 0$ show the rejection frequencies across instruments on a 5 percent level.

A crucial difference between the results under Linex loss and the Linex-Linex loss function is how often the rationality hypothesis is rejected. While rationality on average has to be rejected for more than sixty percent of the estimates for the former, the share of rationality rejections is around ten percent for the latter. For some of the individual forecasters rationality is not rejected at all (see e.g. id 48). Focusing on the rationality rejection in certain instrument sets, there is no particular information set that indicates that forecasters used the information available inefficiently.

5.6.4 Unemployment Rate

Figure 5.6: Proxies for the moments of the forecast error - Unemployment rate

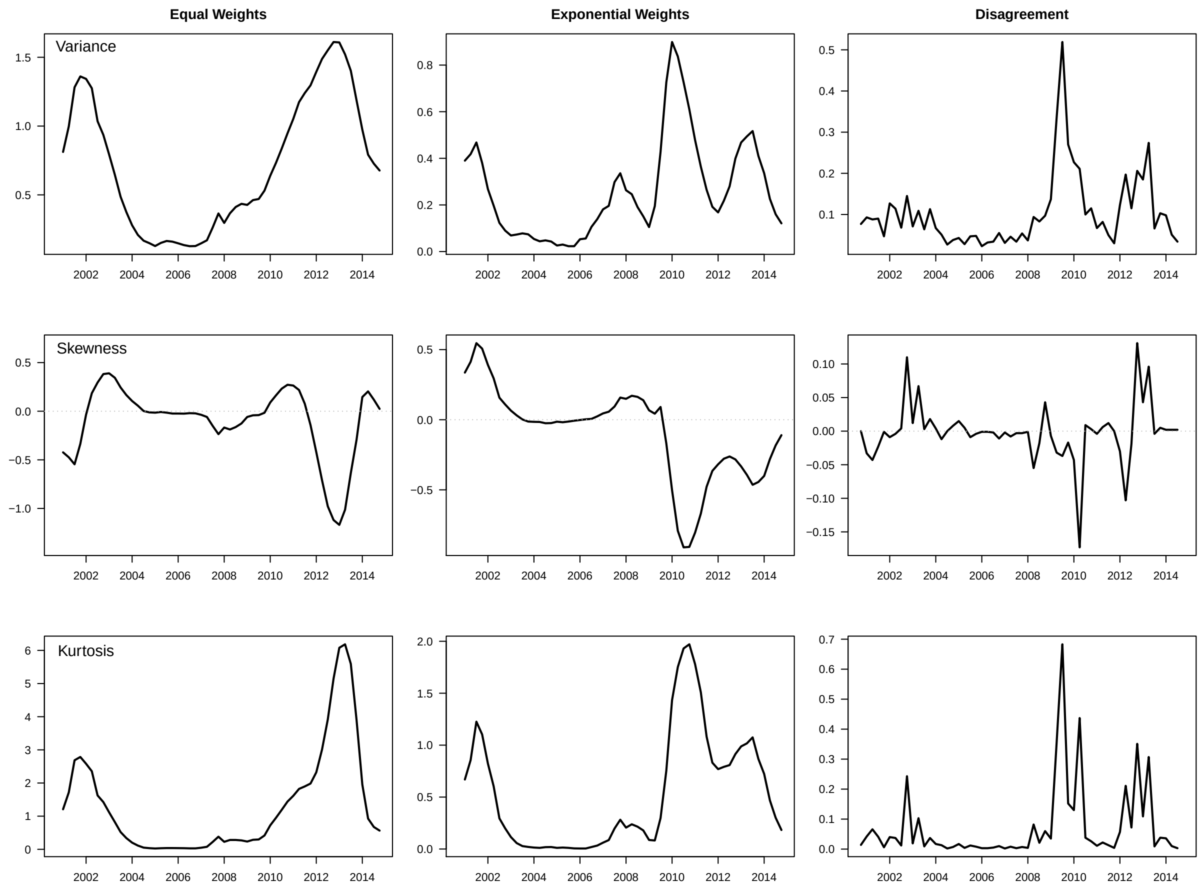


Figure 5.6 shows the proxies for the moments of the unemployment rate, which are organized in the same way as for the other variables. During the observed period, unemployment was rather constant until 2005, at which point it declined until 2007:Q3, before increasing during and after the crisis. However, the unemployment rate was rather constant in 2011 and stopped increasing in 2014:Q1 (see figure 5.3). The moments especially reflect the sharpest increase in unemployment in 2009. Once more, the turning points in the moments based on disagreement best match the periods with the most movement in the unemployment rate.

The GMM results for the unemployment rate, the third main variable that survey participants are asked to forecast, are presented in tables 5.6 (Linex loss) and 5.7 (Linex-Linex loss). Starting with table 5.6 and the results concerning the presence of asymmetry in the loss function, the selected individual forecasters nicely illustrate the different preferences in forecasters' loss functions.

Table 5.6: Results under Linex loss - Unemployment rate

	Forecaster	**Base**	**Equal Weights** σ^2	s^3	k^4	**Expon. Weights** σ^2	s^3	k^4	**Disagreement** σ^2	s^3	k^4
					Real-time						
	aggregate	-0.865	-0.210	-0.210	-0.209	-0.387	-0.314	-0.319	-0.627	-0.633	-0.618
	average	-0.620	-0.131	-0.130	-0.125	-0.257	-0.246	-0.222	-0.439	-0.441	-0.427
$\hat{\alpha}$ estimates	id 14	0.533	0.473	0.483	0.475	0.551	0.609	0.664	0.377	0.383	0.389
	id 22	1.542	0.555	0.536	0.522	1.093	1.008	0.949	1.087	1.073	1.047
	id 48	-1.349	-0.238	-0.236	-0.233	-0.533	-0.498	-0.470	-0.955	-0.949	-0.916
	aggregate	75	63	63	63	75	75	75	75	75	75
	average	39	38	38	38	42	43	44	39	39	40
$H_0: \alpha \geq 0.5$	id 14	0	0	0	0	0	0	0	0	0	0
	id 22	0	0	0	0	0	0	0	0	0	0
	id 48	50	44	44	44	50	50	56	50	50	50
	aggregate	27	40	40	40	40	40	40	27	27	27
	average	35	39	39	39	36	38	38	35	36	35
$H_0: J = 0$	id 14	67	40	27	27	53	60	53	60	60	53
	id 22	40	27	27	27	27	33	40	40	40	40
	id 48	33	27	27	27	27	27	27	33	33	33
					Revised						
	aggregate	-2.367	-0.739	-0.722	-0.683	-1.226	-1.193	-1.107	-1.767	-1.734	-1.676
	average	-2.086	-0.703	-0.688	-0.633	-1.072	-1.075	-0.951	-1.534	-1.518	-1.452
$\hat{\alpha}$ estimates	id 14	-0.352	-0.068	-0.067	-0.067	-0.083	-0.101	-0.068	-0.274	-0.273	-0.272
	id 22	-0.183	-0.074	-0.074	-0.074	-0.120	-0.129	-0.126	-0.144	-0.148	-0.147
	id 48	-2.906	-0.627	-0.634	-0.600	-1.447	-1.184	-1.086	-2.013	-1.892	-1.754
	aggregate	100	100	100	100	100	100	100	100	100	100
	average	79	82	82	82	83	84	85	81	81	81
$H_0: \alpha \geq 0.5$	id 14	6	6	6	6	6	6	6	6	6	6
	id 22	25	38	38	38	31	38	38	31	31	31
	id 48	94	100	100	100	100	100	100	88	88	88
	aggregate	13	27	27	27	20	33	27	13	13	13
	average	29	23	23	22	27	28	28	27	27	25
$H_0: J = 0$	id 14	7	0	0	0	13	13	13	13	13	13
	id 22	7	13	13	13	13	13	13	7	7	7
	id 48	20	7	7	13	20	13	20	13	13	13

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets, while $H_0: \alpha \geq 0$ and $H_0: J = 0$ show the rejection frequencies across instruments on a 5 percent level.

For real-time data, on the one hand, the asymmetry on average across forecasters and that of the aggregate forecasts, points in the same direction as that of inflation forecasts (i.e. forecasters putting higher weights on negative forecast errors and therefore showing a stronger aversion to overpredicting unemployment). Some of the individual forecasters on the other hand, apparently have a rather symmetric loss function (id 14), or even asymmetric preferences pointing in the opposite direction (id 22). The rejections of $H_0 : \alpha \geq 0$ underline the uncertainty concerning the asymmetry, as less than 40 percent of the hypotheses are rejected on average, across the individual forecasters. Even for the exemplary forecaster (id 48) whose forecast appears to have a higher degree of asymmetry at first glance, only 50 percent of the estimates lead to a rejection of the hypothesis.

Turning to the inclusion of further moments, the second moment once again moves the asymmetry parameter toward zero, but does so without significantly changing the number of times the symmetry hypothesis is rejected. Adding a third and fourth moment does not further change the loss parameter. The effect is strongest for equally weighted moments and less pronounced for the variant based on the forecasters' disagreement.

Regarding the rationality rejections across instruments, about one third of the J-statistics lead to a rejection of the hypothesis of forecast rationality. However, rationality has to be rejected for almost two-thirds of the estimates in the forecasts produced by id 14. Focusing on the rejections for single instrument sets, we find higher percentages of rationality rejections across forecasters for sets that contain GDP growth (sets 4 to 6), or second lags of the other variables (sets 2, 8, 11), and the sets that hold the interaction terms (sets 13 to 15). We find the smallest number of rejections in the sets that test for weak efficiency (sets 1 and 3) and the sets that contain the first lag of HICP inflation (set 7), in the first lag of the unemployment rate (set 10) and in the first lag of the unemployment rate and its square (set 12). These findings indicate that the forecasters use this information efficiently.

For revised realizations, the results presented above change slightly. In this case, all estimates are negative and the percentage of symmetry rejections toward this direction of asymmetry increases. Nevertheless, for some of the individual forecasters (id 14 and id 22), only a few of the asymmetry estimates are significantly below zero. Including the moments has the same effect as for real-time data.

The rationality of the forecasts is rejected less often for revised than for real-time data. This applies in particular to id 14 and id 22 whose loss functions appear to be symmetric. In their cases, rationality is merely rejected. This is remarkable, as id 14 has the highest share of rationality rejections for real-time realizations. With respect to the instrument sets, the rationality rejections are still most pronounced for the set that contains the first and second lag of the unemployment rate (set 11) and for the sets that hold an interaction term with lagged GDP growth (sets 13 and 14).

Finally, in table 5.7, the results for the unemployment rate under the Linex-Linex loss function are presented. The overall results are similar to those obtained under Linex loss. For real-time realizations, the asymmetry parameter for the aggregate forecasts and the average of the individual forecasters' parameters take values below 0.5, indicating asymmetric loss in term of a higher weight for negative forecast errors. Some of the individual forecasters produce forecasts suggesting that they have a symmetric loss function or even asymmetric preferences in the opposite direction.

As under Linex loss, less than half of the asymmetry parameters across forecasters and different instrument sets lead to a rejection of the hypothesis that $\alpha \geq 0.5$, although results vary considerably between individual forecasters. On average, the number of times the rationality hypothesis is rejected is approximately the same under Linex-Linex loss. However, the forecast rationality of individual forecaster id 14, who has the highest share of rationality rejections under Linex loss, now is rejected considerably less often.

Table 5.7: Results under Linex-Linex loss - Unemployment rate

	Forecaster	**Base**	**Equal Weights** σ^2	s^3	k^4	**Expon. Weights** σ^2	s^3	k^4	**Disagreement** σ^2	s^3	k^4
						Real-time					
	aggregate	0.425	0.444	0.444	0.444	0.442	0.445	0.446	0.427	0.427	0.427
	average	0.444	0.462	0.461	0.462	0.458	0.461	0.462	0.443	0.443	0.443
$\hat{\alpha}$ estimates	id 14	0.645	0.623	0.625	0.623	0.634	0.635	0.635	0.636	0.636	0.635
	id 22	0.652	0.629	0.621	0.625	0.649	0.650	0.650	0.636	0.636	0.636
	id 48	0.355	0.428	0.429	0.431	0.381	0.385	0.386	0.361	0.361	0.362
	aggregate	69	63	63	63	63	69	69	69	69	69
	average	45	42	42	42	40	40	39	46	46	46
$H_0: \alpha \geq 0.5$	id 14	0	0	0	0	0	0	0	0	0	0
	id 22	0	0	0	0	0	0	0	0	0	0
	id 48	81	69	63	56	63	63	56	81	81	75
	aggregate	27	40	40	40	27	27	27	27	27	27
	average	29	35	36	37	32	32	32	29	29	29
$H_0: J = 0$	id 14	20	13	13	13	13	13	13	20	20	20
	id 22	27	40	40	40	40	40	40	27	27	27
	id 48	33	27	20	27	47	47	47	33	33	33
						Revised					
	aggregate	0.323	0.360	0.360	0.358	0.352	0.353	0.355	0.329	0.330	0.330
	average	0.320	0.357	0.357	0.360	0.350	0.354	0.356	0.328	0.326	0.328
$\hat{\alpha}$ estimates	id 14	0.484	0.485	0.485	0.485	0.489	0.492	0.493	0.481	0.481	0.481
	id 22	0.488	0.502	0.503	0.503	0.503	0.503	0.503	0.488	0.488	0.488
	id 48	0.262	0.346	0.346	0.349	0.315	0.327	0.330	0.277	0.277	0.277
	aggregate	100	100	100	100	100	100	100	100	100	100
	average	87	84	84	84	84	83	83	87	87	87
$H_0: \alpha \geq 0.5$	id 14	6	13	13	13	13	13	13	6	6	6
	id 22	25	6	6	6	25	25	25	25	25	25
	id 48	100	100	100	100	88	88	88	100	100	100
	aggregate	27	33	33	27	33	33	33	33	33	33
	average	21	23	22	22	22	21	22	21	21	21
$H_0: J = 0$	id 14	0	7	7	7	0	0	0	0	0	0
	id 22	7	20	20	20	13	7	7	7	7	7
	id 48	7	7	7	7	7	7	7	7	7	7

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets, while $H_0: \alpha \geq 0.5$ and $H_0: J = 0$ show the rejection frequencies across instruments on a 5 percent level.

When we turn to the revised realization results, the asymmetry for the aggregate forecasts and the individual forecasters' average becomes more pronounced as we have seen under Linex loss. For some individual forecasters, we observe loss parameters that are even closer to symmetric loss (id 14 and id 22). For both real-time and revised realizations, the effect of including second, or higher, moments is still present for all variants of the moments, although it is weaker than under Linex loss.

5.7 Exploiting the Probability Forecasts

In this section we estimate the asymmetry parameters in the forecasters' loss functions again. The main difference from the previous section lies in the way the moments are calculated. Instead of approximating the forecast errors' moments using either equally or exponentially weighted past realizations or the cross-sectional forecasts of the individual forecasters, we now use the probability forecasts provided by each individual forecaster. A major benefit of this approach is the fact that the moments now can differ between the individual forecasters. In what follows, we use the term probability forecast. The forecasters, however, actually provide their estimate of the probability that the target variable will lie in certain intervals instead of providing a true density forecast with an underlying probability distribution.

5.7.1 Computing the Moments Using the Probability Forecasts

For each of the main variables in the survey, forecasters are asked to assess the probabilities of the variable in certain intervals. Admittedly, these probability forecasts are somewhat guided by the ECB, as the range and the width of intervals is stipulated in the questionnaire. Although this can be seen as restrictive, and one could argue that extreme forecasts are excluded in this way, the given intervals cover a wide range of possible outcomes and are adapted to the current economic situation if necessary.[34] Because fewer forecasters respond to the request to provide their probability forecasts in comparison to the single point forecasts, the number of individual forecasters with at least 40 responses to the questionnaire has been further reduced in the approach presented here. This leaves us with 23 forecasters for GDP growth and the unemployment rate and 24 forecasters for HICP inflation.

In order to calculate the moments, we face the questions of how to weight the probability assigned to each interval and how to treat the open intervals at both ends of the possible spectrum of intervals. Here, pragmatic approaches have been chosen; the mean of each interval has been used as weight and the most extreme intervals simply have been treated as closed intervals of the same width as the other intervals. Treating the open intervals as closed is not overly restrictive because most forecasters assign a probability equal or close to zero to these intervals. Furthermore, there are two possibilities for the mean values used to center the moments. These either could be centered around the mean calculated from the probability forecasts, or centered around the point forecasts. Both versions have been applied and have lead to similar results. Therefore, only the moments centered around the point forecasts are presented in subsections 5.7.2 to 5.7.4. The similarity in the results seems plausible, considering the fact that half of the respondents to the ECB's January 2014 special questionnaire state that their reported point forecasts refer to the mean of their probability distribution. Another questionnaire result that is

[34] Annex 2 of the *Description of the ECB Survey of Professional Forecasters Dataset*, which can be downloaded from the ECB's website, provides an overview of the changes in the intervals for the three variables over time.

interesting in this context is that about 80 percent of the forecasters compute their probability distribution for the sole purpose of reporting it in the SPF and do not use it otherwise.

We denote the probability a forecaster i assigns to interval κ at time $t+h$ by $p_{i,\kappa,t+h}$, and the mean of interval κ is m_κ. The number of intervals is K. Thus, the two possibilities for representing the first moment of the probability forecast are $\mu_{i,t+h} = \sum_{\kappa=1}^{K} m_\kappa \cdot p_{i,\kappa,t+h}$ and $\mu_{i,t+h} = f_{i,t+h}$. As argued above, we use the latter to calculate the rest of the moments. In equation (36) in the previous section, we showed the second and fourth moment to be identical with respect to forecast errors and forecasts when the realization of the target variable is the same for each forecaster (i.e. $y_{i,t+h} = y_{t+h}$), which is obviously true. The third moment only differs in sign. Hence, we compute the moments directly as moments of the forecasts:

$$\begin{aligned}
\sigma^2_{i,t+h} &= \sum_{\kappa=1}^{K} p_{i,\kappa,t+h} \left(m_\kappa - f_{i,t+h}\right)^2, \\
s^3_{i,t+h} &= -\sum_{\kappa=1}^{K} p_{i,\kappa,t+h} \left(m_\kappa - f_{i,t+h}\right)^3, \\
k^4_{i,t+h} &= \sum_{\kappa=1}^{K} p_{i,\kappa,t+h} \left(m_\kappa - f_{i,t+h}\right)^4.
\end{aligned} \tag{37}$$

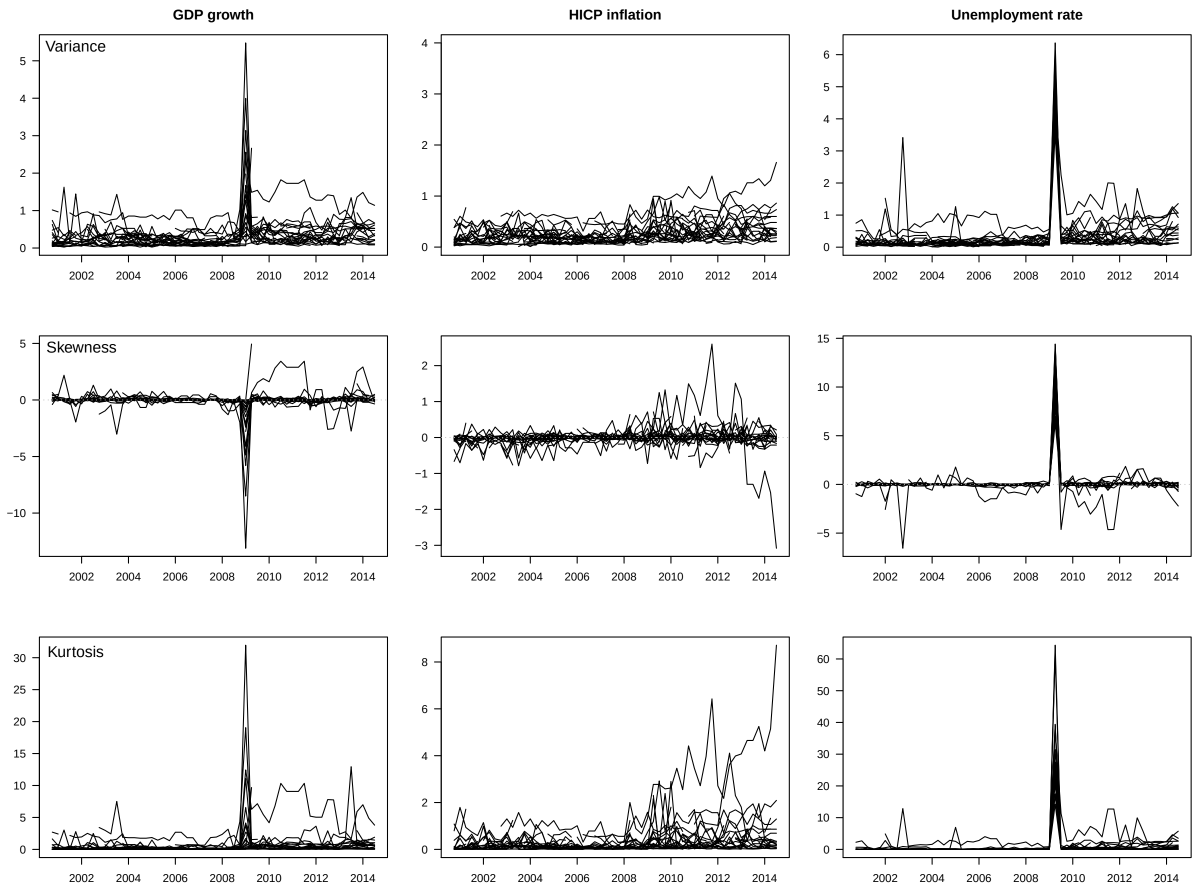
These moments will be used in the next three subsections to conduct an analysis of the forecasters' loss functions analogous to the one in section 5.6. Before discussing the results, figure 5.7 depicts the individual forecasters' second to fourth moments. The figure dedicates one column to each variable and each row to another moment. Gaps or breaks in the plotted lines are due to missing values, i.e. survey rounds, in which a particular forecaster did not respond.

Starting with GDP growth in the first column, we observe a turning point for all three moments in 2009:Q1. After this peak, the levels of variance and kurtosis remain higher than before the crisis. While skewness was close to zero for most forecasters before 2009:Q1, we find it to be more volatile afterward. For HICP inflation, the moments are less pronounced in terms of a considerable peak at some point. However, we observe an increase in the second and fourth moments and a higher volatility in the third moment after 2009:Q2. We also find some forecasters' moments to vary notably in this period. The last column of figure 5.7 contains the moments regarding the unemployment rate forecasts. Here, all three moments peak in 2009:Q2. Variance and kurtosis show a pattern similar to the other variables, with elevated levels after the peak. The skewness, however, is close to zero for most forecasters and in most periods apart from the peak. In addition, some forecasters have a rather negatively skewed forecast error distribution.

5.7.2 GDP Growth

The results for the individual forecasters are presented in tables 5.8 to 5.13. Again, the results for each variable under Linex loss are discussed first, followed by the results under the Linex-Linex loss function. Tables C11 to C16 in the appendix contain the results for each instrument set. The structure of the tables differs from the one in section 5.6, as the tables now present the results for all individual forecasters with sufficient observations. Each table shows the results with respect to real-time realizations in the upper and revised realizations in the lower panel. Both panels contain three blocks each, which show the mean estimates of α across instruments, the percentage of rejections of symmetry toward the direction that is most common for each variable and the percentage of rejections of the J-test for the rationality of the forecasts. Every block consists of four columns that contain the results with respect to the number of additional moments in the loss function, starting with no additional moments and ending with four.

Figure 5.7: Moments based on the probability forecasts



In the rows of each table, the results of the individual forecasters are presented, along with the average results across all individual forecasters. In contrast to the previous section, we abstain from including the aggregate forecast in the analysis here. As argued by Elliott and Timmermann (2004), once higher moments are included in the loss function, the difference in the moments between forecasters has to be taken into consideration for the aggregated loss function when aggregating the forecasts. This problem could be neglected as long as all moments were identical by construction. As the main focus of this study lies on the inclusion of additional moments in the loss function rather than on the aggregation problem, the aggregate forecast is left out in what follows. The focus on individual forecasters is also supported by Clements (2014), who finds evidence that the forecasters participating in the US SPF have heterogeneous asymmetric preferences when forecasting inflation. While in the previous section, the aggregate forecast as well as the average across forecasters were provided along with three selected forecasters in order to keep the tables compact, the reduced number of forecasters allows us to present all individual results here.

Starting with table 5.2 and the real-time results for the GDP forecasts under Linex loss, the positive asymmetry parameters in the first column indicate that all survey participants in this sample have a preference for negative forecast errors. However, the extent of the asymmetry preference varies considerably across forecasters. The preference persists when further moments are included, although the absolute value of the α estimates again decreases for mosts forecasters. Herein, the largest decrease can be observed for the inclusion of the second moment. The rejection frequencies of the null that $\hat{\alpha}$ is below zero are rather high for most forecasters and do not change systematically after including the moments. A similar observation can be made for the rejection frequencies of the rationality hypothesis, which varies between 13 and 67 percent for the individual forecasters, but does not change systematically when the moments are added to the loss function.

Comparing the real-time to the revised result, we need to keep in mind that GDP has been systematically revised upward for the observed time period. This tendency in the data revision leads to a reduction of the observed bias in the forecast and hence smaller estimates of α. Although the general patterns described for the real-time results remain valid with respect to the revised data, for most forecasters, the hypothesis $H_0 : \alpha \leq 0$ is rejected less often now and some of the forecasters even appear to have symmetric preferences.

Table 5.9 shows the analogue results for GDP growth under the Linex-Linex loss function. Unsurprisingly, the general findings remain valid. Overall, the forecasters apparently try to avoid positive forecast errors (i.e. an underestimation of GDP growth) and there is a tendency toward reducing the asymmetry in the loss function when the second moments are included. The effect of data revision causing a reduction in the degree of asymmetry also persists. Nevertheless, some differences can be observed in comparison to the results under the Linex loss function. First, the loss functions appear to be affected less by the inclusion of the moments, as far as the absolute change in the estimates is concerned. However, because the parametrization under both loss functions is different, this comparison has to be made with caution. Second, the rejection frequencies of the hypothesis of symmetry rise under Linex-Linex loss and even reach 100 percent for about half of the forecasters and for real-time data. Third and probably most interesting, the number of J-test rejections drops considerably, indicating that the Linex-Linex loss function may be better suited to reflect the forecasters' true loss function than the Linex loss function.

Our final focus, before turning to results on HICP inflation, is on single instruments that potentially contain information possibly neglected by the forecasters when making their forecasts. While the majority of the forecasters provides forecasts that use the lags and squared lags of GDP efficiently, there seems to be unused information left in lagged inflation. The corresponding results are provided in tables C11 and C12 in the appendix.

Table 5.8: Results under Linex loss - GDP growth

Forecaster	$\hat{\alpha}$ **estimates**				$H_0: \alpha \leq 0$				$H_0: J = 0$			
	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4
Real-time												
average	1.516	0.987	0.971	0.909	73	71	72	72	43	43	42	42
id 4	3.587	1.470	1.450	1.344	100	100	100	100	20	53	60	60
id 5	1.983	0.992	0.859	0.782	81	81	75	75	27	53	47	40
id 15	1.425	0.936	0.881	0.852	88	88	88	88	53	60	60	60
id 16	0.942	0.257	0.240	0.231	63	56	50	50	67	60	60	60
id 20	1.274	0.910	0.895	0.882	50	50	56	56	53	27	27	27
id 22	1.519	0.927	0.874	0.836	56	69	75	75	47	20	20	20
id 23	0.815	0.410	0.430	0.379	31	50	56	50	60	53	53	53
id 24	1.493	0.845	0.829	0.777	81	75	75	75	47	27	27	27
id 26	1.137	1.292	1.208	1.133	88	81	81	88	13	40	27	20
id 31	1.663	1.017	1.010	0.998	75	44	44	44	13	13	13	13
id 33	1.653	1.361	1.331	1.311	81	100	100	100	47	33	33	33
id 37	2.178	1.314	1.271	1.238	100	75	75	81	60	67	60	60
id 38	0.617	0.424	0.427	0.422	31	31	31	31	40	33	33	40
id 39	1.445	0.714	0.698	0.677	69	69	69	69	33	33	33	33
id 42	1.967	1.935	1.797	1.722	88	100	100	100	53	47	47	53
id 52	1.285	1.023	1.024	0.949	63	75	75	81	60	47	47	47
id 54	1.891	1.262	1.287	1.218	94	75	75	81	73	73	73	73
id 56	2.117	0.892	0.889	0.862	100	56	56	56	27	13	13	13
id 89	1.421	1.158	1.362	1.044	94	100	100	100	33	40	27	47
id 90	1.281	1.466	1.451	1.409	50	75	75	75	27	27	27	27
id 94	1.258	0.425	0.411	0.385	50	31	31	31	60	67	67	67
id 95	2.167	0.955	1.121	0.863	81	88	88	88	33	53	67	53
id 96	1.827	1.195	1.072	1.029	100	100	100	100	47	47	47	40
Revised												
average	0.687	0.453	0.437	0.424	53	49	48	48	47	43	41	43
id 4	1.590	0.716	0.710	0.696	81	75	75	75	33	47	47	47
id 5	1.054	0.369	0.310	0.295	69	38	38	38	47	40	40	40
id 15	0.862	0.624	0.591	0.56	63	81	75	75	67	73	73	73
id 16	0.452	0.008	0.011	0.008	50	31	31	31	53	80	80	80
id 20	0.688	0.527	0.533	0.530	38	44	44	44	27	27	27	27
id 22	0.337	0.252	0.188	0.187	50	50	44	44	27	20	13	13
id 23	0.559	0.061	0.064	0.062	50	38	38	38	67	53	53	53
id 24	0.821	0.398	0.394	0.384	63	44	44	44	27	47	47	47
id 26	0.399	0.476	0.418	0.404	31	19	13	13	13	40	33	33
id 31	0.846	0.408	0.412	0.412	56	44	44	44	13	20	20	20
id 33	0.886	0.832	0.827	0.820	56	69	69	69	47	40	40	40
id 37	1.152	0.915	0.894	0.903	69	63	63	63	53	60	60	60
id 38	-0.013	-0.156	-0.158	-0.158	25	13	13	13	40	27	27	33
id 39	0.202	0.144	0.146	0.149	44	38	38	38	33	27	27	27
id 42	0.841	0.791	0.772	0.763	56	44	50	50	67	40	40	40
id 52	0.450	0.733	0.712	0.703	56	94	94	94	67	67	67	67
id 54	0.758	0.482	0.389	0.376	56	63	63	63	67	60	47	47
id 56	1.045	0.390	0.389	0.386	75	19	19	19	33	7	7	7
id 89	0.786	0.698	0.768	0.675	56	69	69	69	67	40	33	47
id 90	0.454	0.569	0.564	0.559	31	38	38	38	40	13	13	20
id 94	0.586	0.384	0.393	0.376	38	38	38	38	73	73	73	73
id 95	1.019	0.519	0.540	0.491	63	75	75	75	60	47	47	47
id 96	0.911	0.502	0.432	0.424	69	75	69	69	53	47	40	40

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets, while $H_0: \alpha \leq 0$ and $H_0: J = 0$ show the rejection frequencies across instruments on a 5 percent level.

Table 5.9: Results under Linex-Linex loss - GDP growth

Forecaster	$\hat{\alpha}$ **estimates**				$H_0: \alpha \leq 0.5$				$H_0: J = 0$			
	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4
Real-time												
average	0.713	0.700	0.698	0.697	89	86	86	86	21	20	20	19
id 4	0.784	0.741	0.740	0.739	100	100	100	100	33	27	27	27
id 5	0.749	0.676	0.664	0.662	100	88	88	88	20	20	7	7
id 15	0.716	0.676	0.674	0.673	88	75	75	75	40	40	40	40
id 16	0.614	0.608	0.603	0.602	56	56	56	56	27	40	27	27
id 20	0.743	0.711	0.710	0.710	100	100	100	100	7	7	7	7
id 22	0.699	0.690	0.681	0.680	81	94	94	94	13	7	7	7
id 23	0.639	0.611	0.611	0.610	69	69	69	69	27	27	27	27
id 24	0.681	0.626	0.625	0.624	100	81	81	81	7	7	7	7
id 26	0.715	0.801	0.799	0.797	100	100	100	100	0	13	13	13
id 31	0.777	0.722	0.722	0.721	100	100	100	100	0	0	0	0
id 33	0.761	0.751	0.751	0.751	100	100	100	100	7	13	13	13
id 37	0.767	0.732	0.732	0.730	100	100	100	100	27	20	20	20
id 38	0.607	0.602	0.603	0.604	25	25	25	25	13	7	7	7
id 39	0.698	0.659	0.659	0.658	94	88	88	88	13	7	7	7
id 42	0.748	0.834	0.833	0.833	100	100	100	100	40	27	27	27
id 52	0.654	0.670	0.671	0.670	69	75	75	75	33	40	40	40
id 54	0.725	0.725	0.726	0.725	100	100	100	100	40	20	20	20
id 56	0.745	0.717	0.717	0.716	100	100	100	100	27	7	7	7
id 89	0.707	0.710	0.710	0.708	100	88	88	88	13	13	13	13
id 90	0.750	0.789	0.790	0.789	100	100	100	100	0	0	0	0
id 94	0.705	0.675	0.685	0.684	69	56	69	69	47	47	47	47
id 95	0.739	0.704	0.711	0.704	100	100	100	100	20	53	53	47
id 96	0.738	0.705	0.686	0.684	100	100	88	88	33	27	33	33
Revised												
average	0.635	0.625	0.626	0.625	62	58	59	59	11	11	12	13
id 4	0.694	0.660	0.660	0.659	100	88	88	88	13	0	0	0
id 5	0.716	0.612	0.61 0	0.610	100	69	69	69	0	13	20	20
id 15	0.661	0.629	0.628	0.628	75	69	69	69	27	13	13	13
id 16	0.582	0.569	0.574	0.574	44	31	38	44	7	27	40	40
id 20	0.657	0.645	0.645	0.644	69	75	75	75	0	7	7	7
id 22	0.618	0.579	0.577	0.577	50	31	31	31	0	0	0	0
id 23	0.561	0.559	0.576	0.576	38	38	44	44	7	0	7	7
id 24	0.598	0.591	0.591	0.591	63	56	56	56	0	0	0	0
id 26	0.615	0.762	0.761	0.758	75	100	100	100	0	7	7	7
id 31	0.682	0.571	0.571	0.571	81	13	13	13	0	7	7	7
id 33	0.69	0.672	0.672	0.671	81	81	81	81	0	7	7	7
id 37	0.672	0.656	0.657	0.656	63	75	88	81	20	20	20	20
id 38	0.602	0.606	0.608	0.608	44	38	38	38	27	40	40	40
id 39	0.602	0.555	0.555	0.555	38	13	13	13	0	0	0	0
id 42	0.641	0.709	0.683	0.682	63	88	81	81	7	7	13	13
id 52	0.608	0.635	0.635	0.634	44	63	63	63	13	33	27	27
id 54	0.592	0.62	0.62	0.619	63	56	56	56	33	13	13	13
id 56	0.665	0.626	0.626	0.626	88	69	69	69	27	0	0	7
id 89	0.658	0.639	0.639	0.639	75	69	69	69	7	0	0	0
id 90	0.644	0.669	0.668	0.668	56	69	69	69	7	7	7	7
id 94	0.606	0.606	0.612	0.612	38	38	38	38	40	40	40	40
id 95	0.663	0.628	0.642	0.637	75	75	75	75	7	0	7	7
id 96	0.626	0.616	0.613	0.612	50	56	56	56	7	7	7	7

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets, while $H_0: \alpha \leq 0.5$ and $H_0: J = 0$ show the rejection frequencies across instruments on a 5 percent level.

5.7.3 HICP Inflation

The results for HICP inflation under Linex loss are shown in table 5.10. As in section 5.6, the asymmetry in the forecasters' loss functions points to the opposite direction compared to the one in the GDP forecasts. This indicates that, as far as HICP inflation is concerned, forecasters are adverse to overestimating inflation (negative forecast errors). Again, the asymmetry is significant at a higher rate for the majority of the forecasters. Including additional higher moments in the loss function reduces the degree of asymmetry, with the strongest reduction induced by the second moments. However, the reduction is less pronounced than for GDP growth. The forecast rationality can be rejected frequently for the majority of the forecasters. As HICP inflation is a variable that is hardly revised at all, the differences between real-time and revised results are negligible.

Apart from the reparametrization of the α estimates, the results concerning the direction of asymmetry and the rejection frequencies of the symmetry hypothesis are similar under Linex-Linex loss, as presented in table 5.11. As for the GDP forecasts, the main difference in the results consists in the rejection frequencies of the J-test. While under Linex loss rationality is rejected on average in about forty-five percent of the cases across forecasters and instruments, we hardly observe any rationality rejections at all for most of the forecasters' predictions under Linex-Linex. Tables C13 and C14 in the appendix show that there are only a few instruments that contain information which is not used by all forecasters. In conclusion, it appears that either using Linex-Linex loss function the J-test for forecast rationality has no power against misspecified or irrational forecasts, or simply that the Linex-Linex loss function reflects the true asymmetry preferences better than the Linex.

5.7.4 Unemployment Rate

Finally, the results for the unemployment rate forecasts are presented in tables 5.12 and 5.13, showing the results under Linex and Linex-Linex loss respectively. Starting with table 5.12, we find the same direction of asymmetric preferences as for HICP inflation forecasts, that is, a preference for underestimating the unemployment rate and producing positive forecast errors. Nevertheless, in contrast to the HICP forecasts, the asymmetry is less pronounced for the unemployment rate, as indicated by the small absolute values of $\hat{\alpha}$ and the low rejection frequencies of the symmetry hypothesis. Some of the forecasters appear to have rather symmetric preferences. As for the other variables, the degree of asymmetry in the loss functions becomes further reduced for most forecasters when the additional moments of the forecast errors are taken into consideration.

As it did for the GDP forecasts, data revision changes the results for the unemployment rate forecasts. However, the direction of the revision for unemployment is not as systematic as it is for GDP. On average, the unemployment rate is revised upward, which leads to higher positive mean forecast errors for the revised forecast and thus increases the asymmetry in the forecasters' loss functions.

Table 5.10: Results under Linex loss - HICP inflation

Forecaster	$\hat{\alpha}$ **estimates**				$H_0: \alpha \geq 0$				$H_0: J = 0$			
	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4
					Real-time							
average	-1.161	-1.040	-1.060	-1.029	62	66	67	67	56	45	44	41
id 1	-1.095	-0.674	-0.718	-0.702	81	56	63	69	80	80	67	67
id 4	-1.338	-1.702	-1.823	-1.647	69	94	94	94	67	27	7	20
id 5	-1.294	-1.690	-1.791	-1.651	75	100	100	100	47	20	13	27
id 15	-1.181	-1.044	-1.077	-1.061	69	75	75	75	53	60	60	60
id 16	-0.857	-0.670	-0.653	-0.645	56	56	56	56	67	47	53	40
id 20	-1.615	-1.589	-1.834	-1.520	63	63	63	63	53	27	27	20
id 22	-1.398	-1.035	-1.233	-1.114	81	75	75	75	80	67	60	60
id 23	-0.910	-0.760	-0.709	-0.746	63	69	69	69	67	60	60	60
id 24	-1.021	-0.879	-0.887	-0.876	56	56	56	56	60	67	67	67
id 26	-1.104	-0.956	-0.955	-0.954	75	69	69	69	73	47	47	47
id 31	-1.746	-1.736	-1.652	-1.597	63	75	75	75	33	33	40	40
id 33	-0.134	-0.497	-0.497	-0.498	0	31	31	31	27	0	0	0
id 37	-1.052	-1.088	-1.039	-1.140	56	63	63	63	67	53	67	40
id 38	-0.809	-0.750	-0.747	-0.746	56	56	56	56	53	53	53	53
id 39	-0.669	-0.367	-0.365	-0.365	25	0	0	0	53	47	47	47
id 42	-1.122	-1.208	-1.188	-1.222	81	88	88	88	67	87	87	87
id 52	-1.365	-1.186	-1.179	-1.196	63	56	56	56	20	33	33	20
id 54	-1.714	-1.317	-1.294	-1.285	88	88	88	88	27	40	40	40
id 56	-2.147	-1.465	-1.453	-1.437	88	88	100	100	47	27	27	13
id 89	-0.599	-0.474	-0.482	-0.478	31	38	38	44	60	67	67	67
id 90	-1.762	-1.338	-1.346	-1.341	88	94	94	94	87	40	40	27
id 94	-0.850	-0.572	-0.573	-0.558	56	56	56	56	47	20	27	27
id 95	-1.057	-0.894	-0.895	-0.883	63	69	69	69	53	53	53	47
id 96	-0.952	-0.703	-0.708	-0.697	63	63	63	63	47	20	20	13
					Revised							
average	-1.151	-1.045	-1.069	-1.038	61	66	66	67	56	44	43	40
id 1	-1.087	-0.676	-0.721	-0.707	81	56	63	63	80	73	67	67
id 4	-1.313	-1.709	-1.825	-1.655	69	94	94	94	67	20	7	20
id 5	-1.254	-1.677	-1.777	-1.641	75	100	100	100	47	20	20	27
id 15	-1.166	-1.039	-1.063	-1.046	69	75	75	75	53	60	60	60
id 16	-0.848	-0.675	-0.656	-0.648	63	56	56	56	67	47	53	40
id 20	-1.594	-1.574	-1.820	-1.506	56	63	63	63	53	27	27	20
id 22	-1.384	-1.069	-1.267	-1.180	81	75	75	75	80	67	53	60
id 23	-0.894	-0.776	-0.720	-0.769	56	56	56	69	67	60	60	60
id 24	-1.008	-0.880	-0.889	-0.88	50	56	56	56	60	67	67	67
id 26	-1.145	-1.016	-1.033	-1.032	75	69	69	69	60	53	47	40
id 31	-1.721	-1.669	-1.700	-1.594	63	75	75	75	33	33	33	33
id 33	-0.084	-0.469	-0.468	-0.469	0	38	38	38	20	0	0	0
id 37	-1.03	-1.077	-1.046	-1.140	56	63	63	63	67	47	67	40
id 38	-0.832	-0.76	-0.756	-0.755	56	56	56	56	53	53	47	47
id 39	-0.662	-0.373	-0.370	-0.370	19	0	0	0	60	53	53	53
id 42	-1.119	-1.241	-1.218	-1.255	81	88	88	88	67	87	87	87
id 52	-1.359	-1.171	-1.177	-1.196	63	56	56	56	20	33	33	27
id 54	-1.679	-1.387	-1.320	-1.306	88	88	88	88	27	40	40	40
id 56	-2.155	-1.472	-1.461	-1.444	88	88	88	94	53	27	20	20
id 89	-0.587	-0.484	-0.492	-0.489	31	38	38	38	60	67	67	67
id 90	-1.781	-1.356	-1.370	-1.362	88	94	94	94	87	40	40	27
id 94	-0.856	-0.567	-0.568	-0.554	56	56	56	56	47	20	27	20
id 95	-1.048	-0.902	-0.898	-0.89	56	69	69	69	60	53	53	40
id 96	-0.946	-0.696	-0.699	-0.689	63	69	69	69	53	0	0	0

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets, while $H_0: \alpha \geq 0$ and $H_0: J = 0$ show the rejection frequencies across instruments on a 5 percent level.

Table 5.11: Results under Linex-Linex loss - HICP inflation

Forecaster	$\hat{\alpha}$ estimates				$H_0: \alpha \geq 0.5$				$H_0: J = 0$			
	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4
Real-time												
average	0.300	0.310	0.310	0.311	71	73	72	72	1	2	2	2
id 1	0.257	0.338	0.336	0.338	88	75	75	75	0	7	7	7
id 4	0.261	0.271	0.271	0.272	81	100	100	100	0	0	0	0
id 5	0.274	0.255	0.253	0.254	81	100	100	100	0	0	0	0
id 15	0.275	0.296	0.297	0.298	81	69	69	69	0	7	7	7
id 16	0.299	0.339	0.341	0.344	63	63	56	56	0	0	0	0
id 20	0.263	0.262	0.262	0.263	69	69	69	69	0	0	0	0
id 22	0.227	0.260	0.257	0.258	81	88	88	88	0	0	0	0
id 23	0.281	0.313	0.312	0.313	75	63	63	63	0	0	7	7
id 24	0.289	0.296	0.295	0.296	69	69	69	69	0	0	0	0
id 26	0.312	0.312	0.312	0.312	75	75	75	75	0	7	7	7
id 31	0.305	0.254	0.254	0.255	69	75	75	75	0	0	0	0
id 33	0.426	0.421	0.421	0.421	19	50	50	50	0	0	0	0
id 37	0.300	0.301	0.301	0.302	63	63	63	63	0	0	0	0
id 38	0.347	0.361	0.361	0.362	56	56	56	56	0	0	0	0
id 39	0.360	0.397	0.397	0.397	63	50	50	50	13	7	7	7
id 42	0.286	0.256	0.254	0.254	81	88	88	88	0	0	0	0
id 52	0.278	0.286	0.286	0.287	63	63	63	63	0	0	0	0
id 54	0.264	0.261	0.260	0.261	88	88	88	88	0	0	0	0
id 56	0.236	0.278	0.278	0.278	100	100	100	100	0	0	0	0
id 89	0.358	0.377	0.378	0.379	56	56	56	56	0	0	0	0
id 90	0.236	0.243	0.244	0.244	88	94	94	94	0	0	0	0
id 94	0.378	0.394	0.394	0.395	69	69	69	69	7	7	7	7
id 95	0.319	0.343	0.343	0.344	63	63	63	63	7	7	7	7
id 96	0.337	0.361	0.362	0.362	81	63	63	63	0	0	0	0
Revised												
average	0.304	0.312	0.312	0.313	70	72	72	72	1	1	1	1
id 1	0.209	0.343	0.342	0.343	88	75	75	75	0	7	7	7
id 4	0.264	0.271	0.270	0.271	81	100	100	100	0	0	0	0
id 5	0.277	0.254	0.252	0.253	81	100	100	100	0	0	0	0
id 15	0.282	0.299	0.300	0.300	75	69	69	69	0	0	0	0
id 16	0.301	0.340	0.342	0.345	63	56	56	56	0	0	0	0
id 20	0.265	0.264	0.263	0.264	69	69	69	69	0	0	0	0
id 22	0.228	0.261	0.258	0.259	81	94	94	94	0	0	0	0
id 23	0.288	0.316	0.315	0.316	63	63	63	63	0	0	7	7
id 24	0.291	0.296	0.296	0.297	69	69	69	69	0	0	0	0
id 26	0.314	0.313	0.313	0.314	75	81	81	81	0	0	0	0
id 31	0.307	0.253	0.254	0.254	63	69	69	69	0	0	0	0
id 33	0.435	0.424	0.424	0.424	19	44	44	44	0	0	0	0
id 37	0.304	0.304	0.304	0.304	63	56	56	56	0	0	0	0
id 38	0.349	0.362	0.362	0.362	56	56	56	56	0	0	0	0
id 39	0.375	0.401	0.401	0.401	56	50	50	50	7	7	7	7
id 42	0.282	0.255	0.254	0.254	81	88	88	88	0	0	0	0
id 52	0.278	0.286	0.286	0.287	63	63	63	63	0	0	0	0
id 54	0.265	0.269	0.269	0.269	88	81	81	81	0	0	0	0
id 56	0.238	0.284	0.285	0.285	100	94	94	94	0	0	0	0
id 89	0.362	0.380	0.382	0.382	56	56	56	56	0	0	0	0
id 90	0.237	0.244	0.244	0.245	88	94	94	94	0	0	0	0
id 94	0.381	0.398	0.398	0.398	69	69	69	69	7	7	7	7
id 95	0.323	0.344	0.343	0.344	63	63	63	63	7	7	7	7
id 96	0.339	0.364	0.364	0.365	81	69	69	69	0	0	0	0

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets, while $H_0: \alpha \geq 0.5$ and $H_0: J = 0$ show the rejection frequencies across instruments on a 5 percent level.

Table 5.12: Results under Linex loss - Unemployment rate

Forecaster	$\hat{\alpha}$ **estimates**				$H_0: \alpha \geq 0$				$H_0: J = 0$			
	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4
Real-time												
average	-0.756	-0.494	-0.500	-0.488	30	26	25	26	30	26	25	26
id 4	-0.478	0.045	0.062	0.058	19	0	0	0	27	13	13	13
id 5	-1.539	-0.551	-0.573	-0.585	31	0	0	0	53	53	53	53
id 15	-1.309	-0.414	-0.395	-0.399	38	6	6	6	47	33	33	33
id 16	-0.382	-0.098	-0.094	-0.081	19	13	13	13	40	40	40	40
id 20	-0.568	-0.568	-0.561	-0.562	13	25	25	25	40	33	40	33
id 22	0.883	0.555	0.459	0.437	0	0	0	0	40	27	33	33
id 23	-0.586	-0.322	-0.314	-0.314	25	13	13	13	20	27	20	20
id 24	-0.549	-0.243	-0.248	-0.240	19	19	19	19	40	20	20	20
id 26	-1.360	-1.699	-1.641	-1.590	63	69	69	69	40	20	20	20
id 31	-0.523	-0.437	-0.432	-0.430	38	38	38	38	27	40	40	40
id 33	-0.837	-0.577	-0.580	-0.581	19	6	6	6	7	7	7	7
id 37	-2.324	-1.500	-1.276	-1.435	81	63	63	69	53	40	40	40
id 38	-0.405	-0.311	-0.440	-0.435	25	25	31	31	33	33	33	33
id 39	-0.440	0.044	0.054	0.054	6	0	0	0	47	47	47	47
id 42	-1.301	-0.884	-0.877	-0.882	19	6	0	6	27	20	13	20
id 52	-0.009	-0.191	-0.198	-0.192	13	31	25	31	27	40	40	40
id 54	-0.178	-0.270	-0.270	-0.268	13	31	31	31	13	7	7	7
id 56	-1.295	-1.328	-1.580	-1.265	44	38	44	38	13	7	7	7
id 89	-1.994	-0.735	-0.735	-0.696	100	94	94	94	20	27	27	27
id 90	-0.736	-0.863	-0.858	-0.840	31	50	50	50	20	13	13	20
id 94	-0.403	-0.162	-0.147	-0.147	6	6	0	6	33	33	47	47
id 95	-0.753	-0.372	-0.382	-0.373	44	31	31	31	20	33	27	27
id 96	-0.031	0.060	0.099	0.094	13	6	0	6	13	33	27	27
Revised												
average	-1.767	-1.159	-1.184	-1.072	65	63	61	65	30	29	30	29
id 4	-1.960	-0.703	-0.681	-0.661	100	44	31	38	40	33	40	40
id 5	-2.915	-1.726	-1.773	-1.621	69	44	50	50	40	40	40	40
id 15	-2.567	-0.982	-0.878	-0.851	75	69	81	88	27	13	13	13
id 16	-1.109	-0.424	-0.379	-0.369	56	56	63	63	20	40	40	47
id 20	-1.252	-1.099	-1.049	-1.052	31	44	38	50	33	27	33	27
id 22	0.176	0.090	0.075	0.074	6	6	6	6	20	20	20	20
id 23	-1.095	-0.679	-0.665	-0.652	38	38	31	38	27	13	13	13
id 24	-1.719	-1.004	-1.017	-0.955	81	69	69	75	27	20	20	20
id 26	-2.244	-2.397	-2.202	-2.099	100	100	100	100	53	33	33	33
id 31	-1.288	-0.847	-0.856	-0.821	63	69	69	69	33	33	33	27
id 33	-1.438	-1.530	-1.185	-1.492	50	50	50	56	7	13	27	20
id 37	-3.982	-2.754	-3.023	-2.599	100	100	100	100	80	60	67	67
id 38	-1.242	-0.997	-1.012	-0.989	63	75	75	75	27	47	40	47
id 39	-1.755	-0.905	-0.907	-0.895	75	50	38	50	33	33	33	33
id 42	-1.965	-1.444	-1.455	-1.443	94	75	56	81	40	40	40	40
id 52	-0.542	-0.392	-0.377	-0.373	31	38	25	31	33	47	47	47
id 54	-0.622	-0.640	-0.637	-0.624	31	50	50	50	7	0	0	0
id 56	-3.589	-2.985	-3.869	-2.312	100	100	100	100	33	33	40	27
id 89	-3.356	-1.129	-1.152	-0.984	100	94	94	94	33	27	20	27
id 90	-1.962	-1.464	-1.423	-1.381	63	69	69	69	20	20	20	20
id 94	-0.963	-0.479	-0.465	-0.458	19	25	25	25	27	27	27	27
id 95	-2.187	-1.001	-1.064	-0.988	100	100	100	100	20	27	27	27
id 96	-1.259	-0.703	-0.730	-0.706	81	63	56	63	20	13	13	13

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets, while $H_0: \alpha \geq 0$ and $H_0: J = 0$ show the rejection frequencies across instruments on a 5 percent level.

Table 5.13: Results under Linex-Linex loss - Unemployment rate

Forecaster	$\hat{\alpha}$ **estimates**				$H_0: \alpha \geq 0.5$				$H_0: J = 0$			
	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4
Real-time												
average	0.410	0.410	0.408	0.408	42	40	41	40	21	17	17	16
id 4	0.434	0.482	0.480	0.479	31	6	6	6	27	13	13	13
id 5	0.287	0.379	0.379	0.380	75	31	31	31	20	20	20	20
id 15	0.343	0.408	0.415	0.417	56	25	25	19	40	33	40	40
id 16	0.446	0.441	0.449	0.448	31	31	31	31	33	40	40	40
id 20	0.418	0.405	0.405	0.405	38	44	44	44	13	27	27	27
id 22	0.656	0.639	0.635	0.631	0	0	0	0	0	0	0	0
id 23	0.394	0.432	0.429	0.429	38	6	6	6	27	7	0	0
id 24	0.454	0.455	0.454	0.454	13	25	25	25	33	7	7	7
id 26	0.316	0.271	0.272	0.272	94	88	88	88	27	0	0	0
id 31	0.403	0.424	0.424	0.424	44	50	50	50	33	33	33	33
id 33	0.380	0.405	0.403	0.403	19	25	25	25	0	0	0	0
id 37	0.369	0.332	0.331	0.331	75	88	88	88	33	20	13	13
id 38	0.402	0.391	0.388	0.388	44	44	44	44	40	27	27	20
id 39	0.514	0.532	0.529	0.528	0	0	0	0	20	7	7	7
id 42	0.284	0.314	0.311	0.311	88	50	56	56	13	0	0	0
id 52	0.523	0.464	0.453	0.451	13	19	25	19	20	27	27	27
id 54	0.478	0.407	0.402	0.402	13	44	44	44	20	13	13	13
id 56	0.343	0.284	0.284	0.284	56	81	81	81	20	0	0	0
id 89	0.278	0.324	0.323	0.326	100	100	100	100	7	7	7	7
id 90	0.404	0.371	0.370	0.370	50	50	50	50	7	7	7	7
id 94	0.448	0.462	0.457	0.457	19	19	19	19	40	47	47	47
id 95	0.397	0.396	0.390	0.390	50	56	56	56	13	40	33	33
id 96	0.494	0.482	0.478	0.477	6	6	6	6	7	27	27	20
Revised												
average	0.313	0.322	0.320	0.321	79	81	81	81	20	18	18	18
id 4	0.314	0.374	0.376	0.376	94	81	88	88	27	20	20	20
id 5	0.232	0.242	0.242	0.242	88	81	81	81	13	33	33	33
id 15	0.260	0.330	0.331	0.331	100	100	94	94	27	20	20	20
id 16	0.360	0.371	0.381	0.383	75	63	63	63	33	27	27	27
id 20	0.350	0.334	0.334	0.334	44	69	69	69	33	27	27	27
id 22	0.527	0.550	0.547	0.546	0	0	0	0	20	7	7	7
id 23	0.331	0.362	0.36	0.361	69	75	81	81	7	7	7	7
id 24	0.360	0.335	0.334	0.336	81	100	100	100	20	20	20	20
id 26	0.231	0.235	0.237	0.237	100	100	100	100	20	7	7	7
id 31	0.332	0.330	0.329	0.330	94	94	94	94	33	27	27	27
id 33	0.316	0.302	0.301	0.301	94	88	88	88	0	13	13	13
id 37	0.220	0.255	0.242	0.245	100	94	100	94	60	53	53	53
id 38	0.348	0.340	0.337	0.337	69	88	88	88	33	47	47	47
id 39	0.330	0.355	0.352	0.352	75	69	75	75	20	13	13	13
id 42	0.223	0.258	0.256	0.256	100	100	100	100	0	0	0	0
id 52	0.398	0.395	0.390	0.395	56	56	63	63	20	7	7	7
id 54	0.357	0.332	0.328	0.328	63	69	75	75	13	0	0	0
id 56	0.178	0.114	0.113	0.113	100	100	100	100	20	0	0	0
id 89	0.205	0.264	0.262	0.266	100	100	100	100	7	7	7	7
id 90	0.293	0.296	0.296	0.296	100	94	94	94	7	13	13	13
id 94	0.365	0.376	0.374	0.375	44	56	56	56	7	13	13	13
id 95	0.318	0.333	0.330	0.331	100	100	100	100	27	40	40	40
id 96	0.351	0.370	0.367	0.366	88	81	63	63	13	7	7	7

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets, while $H_0: \alpha \geq 0.5$ and $H_0: J = 0$ show the rejection frequencies across instruments on a 5 percent level.

These are again considerably smaller for Linex-Linex loss. In terms of unused information in the instrumental variables, it appears that there is information left in the lags of the GDP growth data, as all instrument sets that contain GDP growth show considerable rejection rates of forecast rationality. Apart from GDP, the squared lag of the unemployment rate appears to contain information neglected across forecasters as well (see tables C15 and C16 in the appendix).

5.8 Summary

In this chapter we have analyzed the GDP growth, HICP inflation and unemployment rate forecasts of the ECB's Survey of Professional Forecasters. We have focused primarily on how the loss function is affected by the inclusion of higher moments. First, the moments are calculated using past observations of the target variable or are based on the dispersion of the individual forecasters' past forecasts. Then, we compute the moments individually for each forecaster, using the discrete probability forecasts provided by the forecasters. In order to enable us to include these moments in the loss function, we propose a Taylor series expansion of the Linex loss function as well as the new Linex-Linex loss function presented here. The latter has been developed primarily to overcome two drawbacks of the Linex loss function. First, the parametrization of the Linex-Linex allows the forecasters' preferences toward asymmetry to be interpreted analogous to the EKT approach, i.e. $\alpha \in (0, 1)$. Second, the Linex-Linex nests a symmetric special case for $\alpha = 0.5$, whereas the Linex only approaches symmetry asymptotically.

Summarizing the main results and starting with the GDP growth forecasts, we find evidence for asymmetric preferences for most of the forecasters and under both loss functions. The direction of the asymmetry indicates that forecasters are averse to positive forecast errors and hence averse to an underestimation the GDP growth. The degree of asymmetry is reduced when the forecast errors are calculated with respect to the revised realizations. Including the moments also leads to less asymmetry in the loss function. Here, the most pronounced effect is induced by the second moment, indicating the risk aversion of the forecasters. The third and fourth moments hardly change the degree of asymmetry any further. This holds with respect to each method for computing the moments. A main difference between both loss functions is the considerably lower number of rationality rejections under the Linex-Linex loss function.

For the other two variables, the asymmetry points in the opposite direction, indicating that forecasters prefer to underestimate inflation and the unemployment rate. However, the indication for asymmetry is weaker for the unemployment rate than for the other two variables. Data revisions on the one hand, have a negligible effect on HICP inflation forecasts, as this variable is hardly revised, and, on the other hand, rather strengthen the evidence for asymmetry in the unemployment forecasts. For inflation and unemployment, the effect observed when including the moments is similar to the one observed for GDP, although the effect is slightly weaker for inflation than for the other two variables. Furthermore, the Linex-Linex loss function again appears to be better suited to reflect the forecasters' preferences adequately, as considerably fewer rejections of forecast rationality are detected under this loss function.

6 Conclusion

This dissertation consists of three studies concerned with evaluating macroeconomic forecasts when allowing for potential asymmetries in the forecasters' loss functions. Herein, the approach of Elliott, Komunjer and Timmermann (EKT) is central to the analysis of all three studies. In the third chapter this approach has been applied to German employment forecasts, focusing on the forecast efficiency for a broad spectrum of variables. The fourth chapter has provided an extensive Monte-Carlo study to gain insight into the size and power properties of the EKT test. An extension of the approach has been considered in the fifth chapter, where we analyze private sector forecasts for three macroeconomic variables with a specific interest in including higher moments of the forecast error into the loss function. In this concluding chapter, we highlight the main findings of these studies, illustrate connections between them and identify prospects for further research on the topic of forecast evaluation under asymmetric loss.

In chapter 3, the German Council of Economic Experts and the Joint Forecasts of the leading economic research institutes' annual employment growth forecasts have been analyzed under symmetric as well as asymmetric loss. For both institutions, we observed a preference for underestimating employment growth rather than overestimating it. Overestimation, or a negative forecast error, means that the labor force grows less than expected and policymakers may have reacted to the forecast by prematurely reducing the budget for unemployment care or for job creation programs. It seems reasonable that this forecast error is associated with higher costs than an error resulting from an overly pessimistic forecast that underpredicts the actual rise in employment.

Alongside the direction of asymmetry reflecting the institutions' loss preferences, we are particularly interested in the information efficiency. Confirming previous findings, we find less evidence against forecast efficiency when allowing for asymmetry in the loss function (see section 2.4). However, efficiency could be rejected even under a flexible loss function for some variables, indicating that these variables contain information that could be used by the forecasters to improve their forecasts in the future. These variables are labor force growth, GDP growth, industrial orders growth, the long-term and the short-term interest rate, as well as the growth rates of the ifo business climate index and its components. When including the post-crisis years in the analysis, we observe less efficiency rejections for the interest rates, suggesting that these variables have been scrutinized more closely after the financial crisis (see e.g. Drechsel and Scheufele (2012)). For non-financial variables, we find even more rejections of forecast efficiency, which can be explained by the increased forecast uncertainty during the crisis.

In order to combine the information contained in the individual instrumental variables, we extract three factors form the dataset. First, we conduct this extraction with the entire dataset. Then we use the LASSO approach to pre-select a subset of the variables before extracting the factors. In both variants, the results suggest that the first factor contains information suitable for improving future forecasts, as forecast efficiency with respect to this factor is mainly rejected. For the second factor this only holds without pre-selection. The results for the third factor suggest either that information is used efficiently, or there is no relevant information left in the third factor, as efficiency is hardly ever rejected here. Again, some of the inefficiencies dissolve after allowing for asymmetry in the forecasters' loss functions, whereas the direction of the asymmetry remains the same. A potential benefit of this study to forecast applications is the use of information that is bundled in factors with or without variable pre-selection. Pre-selecting variables also implies omitting the information in the variables that are not selected. Hence, the trade-off between a more extensive information set with possible redundancies, on the one hand, and a compact, pre-selected information set that potentially neglects useful information could be an interesting area for future research (see e.g. Bai and Ng (2008)).

Another important area that deserves more attention is the lack of a sufficient real-time database in Germany. While we used real-time as well as revised realizations of employment growth for the analysis, some of the instrumental variables are not available in real time and thus, the information set of the forecasters at the time they produce their forecast may not be adequately represented. Although many of the time series used as instruments here are rarely subject to revisions or are not revised at all (e.g. financial variables, price indices and business climate), revisions may affect variables such as the GDP or the unemployment rate. Hence, extending the real-time databases available in Germany is an important task to be undertaken in the future. Given the scope of the work needed to be done, a research institute could dedicate an entire department to this task and make real-time data its trademark.

The fourth chapter helps to fill a gap in the literature, providing an extensive Monte-Carlo study focusing on the size and power properties of the EKT test in finite samples. This is relevant for macroeconomic forecast evaluation exercises in particular, as the number of forecasts and forecast errors is usually rather limited due to the low frequencies compared to financial time series, for example. From our results, we can extract at least three main lessons for the application of the EKT forecast evaluation procedure. First, for a correctly specified model, the EKT test is, of course, very reliable across all scenarios and the correct degree of asymmetry is well detected. In this case, using the HAC weighting matrix with the Andrews bandwidth choice is preferable, as it shows almost no size distortions, whereas the identity weighting matrix is associated with a more pronounced tendency to overreject. Second, forecasts based on misspecified models are generally detected by the J-test, while the estimates of α remain quite precise. Here, using the identity weighting matrix nevertheless leads to higher power against the misspecification and leads to less variation in the α estimates. Third, using the HAC weighting matrix with the bandwidth choice according to Newey and West (1994) does not yield satisfactory results in our simulation setting. The scenario using real data to estimate realistic parameters for the simulations indicates that it may be favorable to use a fixed instead of the flexible bandwidth choice according to Andrews.

For future research on the EKT loss function, we can identify several interesting opportunities for gaining even more experience with the procedure. First, it would be beneficial to consider other relevant data situations for the simulation setting, e.g. the well-documented specifics of financial data. Evaluating volatility forecasts generated from GARCH models or stochastic volatility models would be another promising line of research. Here, the simulation designs of Patton and Timmermann (2007a, 2007b) could be a valuable point of departure. Second, the consequences of the recent financial crisis leading to the Great Recession and their effects on time series forecasts deserve much more attention, as pointed out by Ng and Wright (2013). In this respect, inducing outliers in a way that excludes them from the forecaster's information set when the forecast produced would be an opportunity for seeing how these outliers gradually diffuse into the information set. Finally, conducting a similar Monte-Carlo experiment to evaluate the multivariate extension of the EKT approach proposed by Komunjer and Owyang (2012) would be an opportunity for further research.

In chapter 5, we extend the EKT procedure to include higher moments of the forecast errors into the loss function. Therefore, we need loss functions that are different from the Lin-Lin and Quad-Quad function used by EKT. We employ the Linex and the Linex-Linex loss function, whereas the latter is proposed in analogy to the Lin-Lin and Quad-Quad loss function. The dataset used to apply our approach is the ECB's Survey of Professional Forecasts (SPF) and we focus on the survey participants' quarterly year-to-year point forecasts of GDP growth, inflation and the unemployment rate of the euro area. Regardless of the loss function, we find evidence for asymmetry in the forecasters' loss functions that is strongest for GDP growth and inflation forecasts and less evident in the forecasts for the unemployment rate. The direction of asymmetry suggests that the survey respondents tend to be optimistic about GDP growth and averse to deflation, as their forecast errors are most consistent with loss functions that put higher weights on an underestimation of GDP and an overestimation of inflation. Although the results for the unemployment rate forecasts are less conclusive, as the direction of the asymmetry parameter varies for different forecasters, we find a tendency toward optimism, i.e. underestimating the unemployment rate.

In order to include higher moments of the forecast errors we need to calculate these using four different methods, as the moments are not reported directly by the survey participants. Regardless of the method, we observe a reduction in the degree of asymmetry across forecasters and for both loss functions after including the moments. However, we find the strongest decrease to be caused by the second moment, whereas the third and fourth moment only induce minor changes to the form of the loss function. This holds for all three variables and thus may suggest that it is sufficient to account for the risk aversion of the forecasters when evaluating forecasts for macroeconomic target variables. Nevertheless, the forecast errors' skewness and kurtosis should not be neglected in research on other forecasts as the results may be specific to the euro area economy or the time period under investigation.

Considering the results with respect to real-time and revised realizations of the target variable, we find that data revisions affect the results differently for the three variables. This finding is in line with de Castro et al. (2013), who argue that fiscal data revisions are of special concern in Europe, as data are aggregates of the individual countries' data. However, inflation is hardly revised at all and thus the results are almost identical for both types of revision. As GDP growth tends to be revised upward and forecasts are generally optimistic, the degree of asymmetry in the loss function is lower for revised data. Data revisions have the opposite effect for the unemployment rate. Here, we observe a tendency for upward revision as well, but forecasts tend to be biased downward and hence, the asymmetry rather increases. We recommend the analysis of real-time and revised realization data in every forecast evaluation exercise and when real-time data is available, as it may not be obvious which of the two is the appropriate benchmark for evaluating forecast quality.

In terms of the loss functions we analyze, we advocate further exploration of the Linex-Linex loss function, as this reparametrization of the Linex provides several characteristics that are useful for forecast evaluation under asymmetric loss. It nests a symmetric special case, is interpretable in a similar way as the wide-spread family of Lin-Lin and Quad-Quad loss functions used by EKT and can be differentiated infinitely often without vanishing. In our application it appeared to be better suited for reflecting the forecasters' asymmetry preferences than the standard Linex loss function, as efficiency has been rejected less frequently when using the Linex-Linex. Hence, a Monte-Carlo analysis similar to the one in chapter 4 also seems to be promising in order to gain further insight on the size and power properties of the EKT test when using the Linex-Linex loss function.

The forecasts of the ECB's SPF together with the realization data and potential further instrumental variables in the Real Time Database that is provided by the ECB's Statistical Data Warehouse are datasets that appear well suited for further forecast evaluation studies. This study in particular could be extended in several ways: the study could be conducted using other forecast horizons in order to analyze whether asymmetry and forecast efficiency vary across horizons. Given more observations, it would also be interesting to test for time-varying preferences in a couple of years (see e.g. Wang and Lee (2014) for an application on US SPF and Greenbook forecasts and Giacomini and Rossi (2013) for an adequate choice of the rolling window size). Relaxing the parameter restrictions and estimating the effect of each moment individually could also provide further insight, but would also require more observations. Taking into account that a majority of the respondents to the special questionnaire states that their forecasts are not independent of one another, a multivariate procedure such as the one proposed by Komunjer and Owyang (2012) might lead to different results. Finally, the approach proposed by Bernhardt et al. (2006) could be used to test how herding and anti-herding behavior among forecasters (see Fritsche et al. (2015)) potentially explains bias in their forecasts.

References

Abel, J., R. Rich, J. Song, and J. Tracy (2015). The measurement and behavior of uncertainty: Evidence from the ECB Survey of Professional Forecasters. *Journal of Applied Econometrics*.

Adrian, T. and A. Estrella (2008). Monetary tightening cycles and the predictability of economic activity. *Economics Letters 99*(2), 260–264.

Ager, P., M. Kappler, and S. Osterloh (2009). The accuracy and efficiency of the consensus forecasts: A further application and extension of the pooled approach. *International Journal of Forecasting 25*(1), 167–181.

Álvarez, L. and A. Cabrero (2010). Does housing really lead the business cycle in Spain? In O. de Bandt, T. Knetsch, J. Peñalosa, and F. Zollino (Eds.), *Housing Markets in Europe*, pp. 61–84. Springer Berlin Heidelberg.

Andrade, P. and H. L. Bihan (2013). Inattentive professional forecasters. *Journal of Monetary Economics 60*(8), 967–982.

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica 59*(3), 817–858.

Andrews, D. W. K. (1999). Consistent moment selection procedures for generalized method of moments estimation. *Econometrica 67*(3), 543–563.

Aretz, K., S. M. Bartram, and P. F. Pope (2011). Asymmetric loss functions and the rationality of expected stock returns. *International Journal of Forecasting 27*(2), 413–437.

Armstrong, J. and F. Collopy (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting 8*(1), 69–80.

Armstrong, J. S. (2001). Evaluating forecasting methods. In *Principles of forecasting*, pp. 443–472. Springer.

Arrow, K. (1971). *Essays in the Theory of Risk-bearing.* Markham economics series. Markham Publishing.

Artis, M. and M. Marcellino (2001). Fiscal forecasting: The track record of the IMF, OECD and EC. *Econometrics Journal 4*(1), 20–36.

Ashley, R. (1988). On the relative worth of recent macroeconomic forecasts. *International Journal of Forecasting 4*(3), 363–376.

Auffhammer, M. (2007). The rationality of EIA forecasts under symmetric and asymmetric loss. *Resource and Energy Economics 29*(2), 102–121.

Bai, J. and S. Ng (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics 146* (2), 304–317. Honoring the research contributions of Charles R. Nelson.

Bai, J. and S. Ng (2010, 12). Instrumental variable estimation in a data rich environvent. *Econometric Theory 26*, 1577–1606.

Ball, L., J. T. Jalles, and P. Loungani (2015). Do forecasters believe in Okun's Law? An assessment of unemployment and output forecasts. *International Journal of Forecasting 31* (1), 176–184.

Ball, L. M., D. Leigh, and P. Loungani (2013). Okun's Law: Fit at fifty? Working Paper 18668, National Bureau of Economic Research.

Barigozzi, M., A. M. Conti, and M. Luciani (2014). Do euro area countries respond asymmetrically to the common monetary policy? *Oxford Bulletin of Economics and Statistics 76* (5), 693–714.

Batchelor, R. (2007). Bias in macroeconomic forecasts. *International Journal of Forecasting 23* (2), 189–203.

Batchelor, R. and P. Dua (1991). Blue chip rationality tests. *Journal of Money, Credit and Banking 23* (4), 692–705.

Batchelor, R. and D. A. Peel (1998). Rationality testing under asymmetric loss. *Economics Letters 61* (1), 49–54.

Benishay, H. (1992). The Pratt-Arrow requirement in a fourth degree polynomial utility function. *Journal of Accounting, Auditing & Finance 7* (1), 97–112.

Bernanke, B. S. and J. Boivin (2003). Monetary policy in a data-rich environment. *Journal of Monetary Economics 50* (3), 525–546.

Bernanke, B. S., T. Laubach, and F. S. Mishkin (2001). *Inflation targeting: lessons from the international experience.* Princeton University Press.

Bernhardt, D., M. Campello, and E. Kutsoati (2006). Who herds? *Journal of Financial Economics 80* (3), 657–675.

Blanchard, O. J. and M. W. Watson (1986). Are business cycles all alike? In *The American business cycle: Continuity and change*, pp. 123–180. University of Chicago Press.

Boivin, J. and S. Ng (2005). Understanding and comparing factor-based forecasts. *International Journal of Central Banking 1* (3), 117–151.

Boivin, J. and S. Ng (2006). Are more data always better for factor analysis? *Journal of Econometrics 132* (1), 169–194. Common Features.

Bonham, C. and R. Cohen (1995). Testing the rationality of price forecasts: Comment. *The American Economic Review 85* (1), 284–289.

Bonham, C. S. and R. H. Cohen (2001). To aggregate, pool, or neither: Testing the rational-expectations hypothesis using survey data. *Journal of Business & Economic Statistics 19*(3), 278–291.

Bowles, C., R. Friz, V. Genre, G. Kenny, A. Meyler, and T. Rautanen (2010). An evaluation of the growth and unemployment forecasts in the ECB Survey of Professional Forecasters. *OECD Journal: Journal of Business Cycle Measurement and Analysis 2010*(2), 1–28.

Breiman, L. and P. Spector (1992). Submodel selection and evaluation in regression. The X-random case. *International Statistical Review / Revue Internationale de Statistique 60*(3), 291–319.

Breitung, J. and S. Eickmeier (2006). Dynamic factor models. *AStA Advances in Statistical Analysis 90*(1), 27–42.

Brockett, P. and J. Garven (1998). A reexamination of the relationship between preferences and moment orderings by rational risk-averse investors. *The GENEVA Papers on Risk and Insurance Theory 23*(2), 127–137.

Broyden, C. G. (1970). The convergence of a class of double-rank minimization algorithms 1. general considerations. *IMA Journal of Applied Mathematics 6*(1), 76–90.

Camacho, M., G. Perez-Quiros, and L. Saiz (2006). Are European business cycles close enough to be just one? *Journal of Economic Dynamics and Control 30*(9–10), 1687–1706.

Camacho, M., G. Perez-Quiros, and L. Saiz (2008). Do European business cycles look like one? *Journal of Economic Dynamics and Control 32*(7), 2165–2190.

Campbell, B. and E. Ghysels (1995). Federal budget projections: A nonparametric assessment of bias and efficiency. *The Review of Economics and Statistics 77*(1), 17–31.

Capistrán, C. (2005). Bias in Federal Reserve inflation forecasts: Is the Federal Reserve irrational or just cautious? UC San Diego, Economics Working Paper Series.

Capistrán, C. (2008). Bias in Federal Reserve inflation forecasts: Is the Federal Reserve irrational or just cautious? *Journal of Monetary Economics 55*(8), 1415–1427.

Capistrán, C. and A. Timmermann (2009). Disagreement and biases in inflation expectations. *Journal of Money, Credit and Banking 41*(2-3), 365–396.

Caunedo, J., R. DiCecio, I. Komunjer, and M. Owyang (2013). Federal reserve forecasts: asymmetry and state-dependence. Technical report.

Chaussé, P. (2010). Computing generalized method of moments and generalized empirical likelihood with R. *Journal of Statistical Software 34*(11), 1–35.

Chen, S., D. Donoho, and M. Saunders (1998). Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing 20*(1), 33–61.

Christodoulakis, G. and D. Peel (2006). The relationship between expected utility and higher moments for distributions captured by the Gram-Charlier class. *Finance Research Letters 3*(4), 273–276.

Christodoulakis, G. A. (2005). Financial forecasts in the presence of asymmetric loss aversion, skewness and excess kurtosis. *Finance Research Letters 2*(4), 227–233.

Christodoulakis, G. A. and E. C. Mamatzakis (2008). An assessment of the EU growth forecasts under asymmetric preferences. *Journal of Forecasting 27*(6), 483–492.

Christodoulakis, G. A. and E. C. Mamatzakis (2009). Assessing the prudence of economic forecasts in the EU. *Journal of Applied Econometrics 24*(4), 583–606.

Christoffersen, P. F. and F. X. Diebold (1996). Further results on forecasting and model selection under asymmetric loss. *Journal of Applied Econometrics 11*(5), 561–571.

Christoffersen, P. F. and F. X. Diebold (1997). Optimal prediction under asymmetric loss. *Econometric Theory 13*, 808–817.

Clatworthy, M. A., D. A. Peel, and P. F. Pope (2012). Are analysts' loss functions asymmetric? *Journal of Forecasting 31*(8), 736–756.

Clements, M. and D. Hendry (1998). *Forecasting economic time series.* Cambridge University Press.

Clements, M. P. (2010). Explanations of the inconsistencies in survey respondents' forecasts. *European Economic Review 54*(4), 536–549.

Clements, M. P. (2014). US inflation expectations and heterogeneous loss functions, 1968-2010. *Journal of Forecasting 33*(1), 1–14.

Clements, M. P. (2015). Do US macroeconomic forecasters exaggerate their differences? *Journal of Forecasting 34*(8), 649–660.

Clements, M. P. and D. F. Hendry (2002). *A companion to economic forecasting.* Blackwells: Oxford.

Clements, M. P. and D. F. Hendry (2008). Economic forecasting in a changing world. *Capitalism and Society 3*(2).

Clements, M. P. and D. F. Hendry (2011). *The Oxford handbook of economic forecasting.* Oxford University Press.

Croushore, D. (2006). Forecasting with real-time macroeconomic data. In C. G. G. Elliott and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 1 of *Handbook of Economic Forecasting*, Chapter 17, pp. 961–982. Elsevier.

Croushore, D. (2011). Frontiers of real-time data analysis. *Journal of Economic Literature 49*(1), 72–100.

Croushore, D. and T. Stark (2003). A real-time data set for macroeconomists: Does the data vintage matter? *The Review of Economics and Statistics 85*(3), 605–617.

de Castro, F., J. J. Pérez, and M. Rodríguez-Vives (2013). Fiscal data revisions in europe. *Journal of Money, Credit and Banking 45*(6), 1187–1209.

Demetrescu, M. (2006). An extension of the Gauss-Newton algorithm for estimation under asymmetric loss. *Computational Statistics & Data Analysis 50*(2), 379–401.

Diebold, F. X. (1998). The past, present, and future of macroeconomic forecasting. *The Journal of Economic Perspectives 12*(2), 175–192.

Diebold, F. X. (2015). Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of diebold–mariano tests. *Journal of Business & Economic Statistics 33*(1), 1–9.

Diebold, F. X. and J. A. Lopez (1996). Forecast evaluation and combination. In G. Maddala and C. Rao (Eds.), *Statistical Methods in Finance*, Volume 14 of *Handbook of Statistics*, pp. 241–268. Elsevier.

Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics 13*(3), 253–263.

Döpke, J., U. Fritsche, and B. Siliverstovs (2010). Evaluating German business cycle forecasts under an asymmetric loss function. *OECD Journal: Journal of Business Cycle Measurement and Analysis 1*(6), 1–18.

Dovern, J. and J. Weisser (2011). Accuracy, unbiasedness and efficiency of professional macroeconomic forecasts: An empirical comparison for the G7. *International Journal of Forecasting 27*(2), 452–465.

Drechsel, K. and R. Scheufele (2012). The performance of short-term forecasts of the german economy before and during the 2008/2009 recession. *International Journal of Forecasting 28*(2), 428–445.

Eeckhoudt, L., C. Gollier, and T. Schneider (1995). Risk-aversion, prudence and temperance: A unified approach. *Economics Letters 48*(3-4), 331–336.

Ehrbeck, T. and R. Waldmann (1996). Why are professional forecasters biased? Agency versus behavioral explanations. *The Quarterly Journal of Economics 111*(1), pp. 21–40.

Eickmeier, S. (2009). Comovements and heterogeneity in the euro area analyzed in a non-stationary dynamic factor model. *Journal of Applied Econometrics 24*(6), 933–959.

Elliott, G. and C. W. Granger (2004). Evaluating significance: comments on "size matters". *The Journal of Socio-Economics 33*(5), 547–550.

Elliott, G., C. W. J. Granger, and A. Timmermann (2006). *Handbook of Economic Forecasting*, Volume 1. Elsevier.

Elliott, G., I. Komunjer, and A. Timmermann (2005). Estimation and testing of forecast rationality under flexible loss. *Review of Economic Studies 72*(4), 1107–1125.

Elliott, G., I. Komunjer, and A. Timmermann (2008). Biases in macroeconomic forecasts: Irrationality or asymmetric loss? *Journal of the European Economic Association 6*(1), 122–157.

Elliott, G., T. J. Rothenberg, and J. H. Stock (1996). Efficient tests for an autoregressive unit root. *Econometrica 64*(4), pp. 813–836.

Elliott, G. and A. Timmermann (2004). Optimal forecast combinations under general loss functions and forecast error distributions. *Journal of Econometrics 122*(1), 47–79.

Elliott, G. and A. Timmermann (2008). Economic forecasting. *Journal of Economic Literature 46*(1), 3–56.

Elliott, G. and A. Timmermann (2013). *Handbook of Economic Forecasting*, Volume 2A-B. Elsevier.

Estrella, A. and F. S. Mishkin (1998). Predicting U.S. recessions: Financial variables as leading indicators. *Review of Economics and Statistics 80*(1), 45–61.

Estrella, A. and M. R. Trubin (2006). The yield curve as a leading indicator: some practical issues. *Current Issues in Economics and Finance 12*(Jul), 1–7.

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business 38*(1), 34–105.

Figlewski, S. and P. Wachtel (1981). The formation of inflationary expectations. *The Review of Economics and Statistics 63*(1), 1–10.

Figlewski, S. and P. Wachtel (1983). Rational expectations, informational efficiency, and tests using survey data: A reply. *The Review of Economics and Statistics 65*(3), 529–531.

Fildes, R. and H. Stekler (2002). The state of macroeconomic forecasting. *Journal of Macroeconomics 24*(4), 435–468.

Fletcher, R. (1970). A new approach to variable metric algorithms. *The Computer Journal 13*(3), 317–322.

Franses, P. H., D. v. Dijk, and A. Opschoor (2014). *Time Series Models for Business and Economic Forecasting*. Cambridge University Press.

Friedman, J. H., T. Hastie, and R. Tibshirani (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software 33*(1), 1–22.

Fritsche, U., C. Pierdzioch, J.-C. Rülke, and G. Stadtmann (2014). A note on forecasting the euro: Do forecasters have an asymmetric loss function? *International Economic Journal 28*(2), 333–343.

Fritsche, U., C. Pierdzioch, J.-C. Rülke, and G. Stadtmann (2015). Forecasting the Brazilian real and the Mexican peso: Asymmetric loss, forecast rationality, and forecaster herding. *International Journal of Forecasting 31*(1), 130–139.

Garcia, J. A. (2003). An introduction to the ECB's Survey of Professional Forecasters. Technical Report 8, ECB Occasional Paper.

Genre, V., G. Kenny, A. Meyler, and A. Timmermann (2013). Combining expert forecasts: Can anything beat the simple average? *International Journal of Forecasting 29*(1), 108–121.

Giacomini, R. and B. Rossi (2013). Forecasting in macroeconomics. In N. Hashimzade and M. A. Thornton (Eds.), *Handbook of Research Methods and Applications in Empirical Macroeconomics*, pp. 381–410. Edward Elgar Publishing.

Giacomini, R. and H. White (2006). Tests of conditional predictive ability. *Econometrica 74*(6), 1545–1578.

Giannone, D., J. Henry, M. Lalik, and M. Modugno (2012). An area-wide real-time database for the euro area. *Review of Economics and Statistics 94*(4), 1000–1013.

Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics 55*(4), 665–676.

Gneiting, T. (2011). Quantiles as optimal point forecasts. *International Journal of Forecasting 27*(2), 197–207.

Goldfarb, D. (1970). A family of variable-metric methods derived by variational means. *Mathematics of computation 24*(109), 23–26.

Gollier, C. and J. W. Pratt (1996). Risk vulnerability and the tempering effect of background risk. *Econometrica 64*(5), 1109–1123.

Granger, C. W. (1999). Outline of forecast theory using generalized cost functions. *Spanish Economic Review 1*(2), 161–173.

Granger, C. W. and M. J. Machina (2006). Forecasting and decision theory. *Handbook of economic forecasting 1*, 81–98.

Granger, C. W. and P. Newbold (1973). Some comments on the evaluation of economic forecasts. *Applied Economics 5*(1), 35–47.

Granger, C. W. J. (1969). Prediction with a generalized cost of error function. *OR 20*(2), 199–207.

Granger, C. W. J. and P. Newbold (1986). *Forecasting Economic Time Series* (Second edition ed.). Orlando, Fla.; London; Sydney and Toronto: Harcourt, Brace, Jovanovich; Academic Press.

Granger, C. W. J. and M. H. Pesaran (2000). Economic and statistical measures of forecast accuracy. *Journal of Forecasting 19*(7), 537–560.

Guisinger, A. Y. and T. M. Sinclair (2015). Okun's Law in real time. *International Journal of Forecasting 31*(1), 185–187.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50*(4), pp. 1029–1054.

Hansen, L. P., J. Heaton, and A. Yaron (1996). Finite-sample properties of some alternative GMM estimators. *Journal of Business & Economic Statistics 14*(3), 262–280.

Hastie, T., R. Tibshirani, and J. Friedman (2009). *The elements of statistical learning: data mining, inference and prediction* (2 ed.). Springer.

Heilemann, U. (2002). Increasing the transparency of macroeconometric forecasts: a report from the trenches. *International Journal of Forecasting 18*(1), 85–105.

Henzel, S. and S. Rast (2013). Prognoseeigenschaften von Indikatoren zur Vorhersage des Bruttoinlandsprodukts in Deutschland. *Ifo Schnelldienst 66*(17), 39–46.

Hirsch, A. A. and M. C. Lovell (1969). *Sales Anticipations and Inventory Behavior*. Wiley.

Holden, K. and D. A. Peel (1990). On testing for unbiasedness and efficiency of forecasts. *The Manchester School 58*(2), 120–127.

Hoque, A., J. R. Magnus, and B. Pesaran (1988). The exact multi-period mean-square forecast error for the first-order autoregressive model. *Journal of Econometrics 39*(3), 327–346.

Hoss, J. (2014). German employment forecasts, asymmetric loss and information efficiency. Unpublished Manuscript, Darmstadt University of Technology.

Hoss, J. (2015). Evaluating the ECB's Survey of Professional Forecasters under asymmetric loss and higher moments. Unpublished Manuscript, Darmstadt University of Technology.

Hoss, J. and J. J. Krüger (2015). A Monte-Carlo analysis of the EKT loss function. Unpublished Manuscript, Darmstadt University of Technology.

Hyndman, R. J. and A. B. Koehler (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting 22*(4), 679–688.

Jiménez-Rodríguez, R. and M. Sánchez (2005). Oil price shocks and real GDP growth: empirical evidence for some OECD countries. *Applied Economics 37*(2), 201–228.

Jondeau, E. and M. Rockinger (2001). Gram-Charlier densities. *Journal of Economic Dynamics and Control 25*(10), 1457–1483.

Jondeau, E. and M. Rockinger (2006). Optimal portfolio allocation under higher moments. *European Financial Management 12*(1), 29–55.

Jones, C. S. and S. Tuzel (2013). New orders and asset prices. *Review of Financial Studies 26*(1), 115–157.

Kaldor, N. (1966). *Causes of the slow rate of economic growth of the United Kingdom: an inaugural lecture*. Cambridge University Press.

Kapetanios, G. and M. Marcellino (2010). Factor-GMM estimation with large sets of possibly weak instruments. *Computational Statistics & Data Analysis 54*(11), 2655–2675.

Keane, M. P. and D. E. Runkle (1990). Testing the rationality of price forecasts: New evidence from panel data. *The American Economic Review 80*(4), 714–735.

Kimball, M. S. (1990). Precautionary saving in the small and in the large. *Econometrica 58*(1), 53–73.

Klein, M. W., S. Schuh, and R. K. Triest (2003). Job creation, job destruction, and the real exchange rate. *Journal of International Economics 59*(2), 239–265.

Klinger, S. and E. Weber (2014). On GDP-employment decoupling in Germany. IAB Discussion Paper 201421, Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg [Institute for Employment Research, Nuremberg, Germany].

Koenker, R. (2005). *Quantile regression*. Number 38. Cambridge university press.

Koenker, R. and G. Bassett Jr (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, 33–50.

Komunjer, I. and M. T. Owyang (2012). Multivariate forecast evaluation and rationality testing. *Review of Economics and Statistics 94*(4), 1066–1080.

Kon, S. J. (1984). Models of stock returns–a comparison. *The Journal of Finance 39*(1), 147–165.

Krol, R. (2013). Evaluating state revenue forecasting under a flexible loss function. *International Journal of Forecasting 29*(2), 282–289.

Kroll, Y., H. Levy, and H. M. Markowitz (1984). Mean-variance versus direct utility maximization. *The Journal of Finance 39*(1), 47–61.

Krüger, J. J. (2014). A multivariate evaluation of German output growth and inflation forecasts. *Economics Bulletin 34*(3), 1410–1418.

Krüger, J. J. and J. Hoss (2012). German business cycle forecasts, asymmetric loss and financial variables. *Economics Letters 114*(3), 284–287.

Lajeri-Chaherli, F. (2004). Proper prudence, standard prudence and precautionary vulnerability. *Economics Letters 82*(1), 29–34.

Lee, K.-S. and F. Mercurelli (2014). Convergence in the core euro zone under the global financial crisis. *Journal of Economic Integration 29*, 20–63.

Levy, H. (1969). A utility function depending on the first three moments. *The Journal of Finance 24*(4), 715–719.

Levy, H. and H. M. Markowitz (1979). Approximating expected utility by a function of mean and variance. *The American Economic Review 69*(3), 308–317.

Li, J. and W. Chen (2014). Forecasting macroeconomic time series: Lasso-based approaches and their forecast combinations with dynamic factor models. *International Journal of Forecasting 30*(4), 996–1015.

Lim, T. (2001). Rationality and analysts' forecast bias. *The Journal of Finance 56*(1), 369–385.

Longin, F. M. (1996). The asymptotic distribution of extreme stock market returns. *The Journal of Business 69*(3), 383–408.

Ludvigson, S. C. and S. Ng (2009). Macro factors in bond risk premia. *Review of Financial Studies 22*(12), 5027–5067.

Lutz, C. and B. Meyer (2009). Economic impacts of higher oil and gas prices: The role of international trade for germany. *Energy Economics 31*(6), 882–887.

Magnus, J. R. and B. Pesaran (1989). The exact multi-period mean-square forecast error for the first-order autoregressive model with an intercept. *Journal of Econometrics 42*(2), 157–179.

Magnus, J. R. and B. Pesaran (1991, 6). The bias of forecasts from a first-order autoregression. *Econometric Theory 7*, 222–235.

Makridakis, S., A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen, and R. Winkler (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of Forecasting 1*(2), 111–153.

Makridakis, S., C. Chatfield, M. Hibon, M. Lawrence, T. Mills, K. Ord, and L. F. Simmons (1993). The M2-competition: A real-time judgmentally based forecasting study. *International Journal of Forecasting 9*(1), 5–22.

Makridakis, S. and M. Hibon (2000). The M3-Competition: results, conclusions and implications. *International Journal of Forecasting 16*(4), 451–476.

Mamatzakis, E. (2014). Revealing asymmetries in the loss function of WTI oil futures market. *Empirical Economics 47*(2), 411–426.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business 36*(4), 394–419.

Marcellino, M., J. H. Stock, and M. W. Watson (2003). Macroeconomic forecasting in the euro area: Country specific versus area-wide information. *European Economic Review 47*(1), 1–18.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance 7*(1), 77–91.

McKenzie, J. (2011). Mean absolute percentage error and bias in economic forecasting. *Economics Letters 113*(3), 259–262.

Mencher, A. G. (1971). On the social deployment of science. *Bulletin of the Atomic Scientists 27*(10), 34–38.

Mincer, J. A. (1969). *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance.* NBER Books. National Bureau of Economic Research, Inc.

Mincer, J. A. and V. Zarnowitz (1969). The evaluation of economic forecasts. In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, NBER Chapters, pp. 1–46. National Bureau of Economic Research, Inc.

Morgenstern, O. (1928). *Wirtschaftsprognose: Eine Untersuchung ihrer Voraussetzungen und Möglichkeiten*. Springer Vienna.

Moser, C., D. Urban, and B. W. di Mauro (2010). International competitiveness, job creation and job destruction - An establishment-level study of German job flows. *Journal of International Economics 80*(2), 302 – 317.

Naghi, A. A. (2015). A forecast rationality test that allows for loss function asymmetries. Working paper, University of Warwick.

Nardo, M. (2003). The quantification of qualitative survey data: A critical assessment. *Journal of Economic Surveys 17*(5), 645–668.

Nelder, J. A. and R. Mead (1965). A simplex method for function minimization. *The Computer Journal 7*(4), 308–313.

Newey, W. K. and J. L. Powell (1987). Asymmetric least squares estimation and testing. *Econometrica 55*(4), pp. 819–847.

Newey, W. K. and R. J. Smith (2004). Higher order properties of gmm and generalized empirical likelihood estimators. *Econometrica 72*(1), 219–255.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*(3), 703–708.

Newey, W. K. and K. D. West (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies 61*(4), 631–653.

Ng, S. and J. H. Wright (2013). Facts and challenges from the great recession for forecasting and macroeconomic modeling. *Journal of Economic Literature 51*(4), 1120–54.

Nobay, R. A. and D. A. Peel (2003). Optimal discretionary monetary policy in a model of asymmetric central bank preferences. *The Economic Journal 113*(489), 657–665.

Nordhaus, W. D. (1987). Forecasting efficiency: Concepts and applications. *The Review of Economics and Statistics 69*(4), 667–674.

Okun, A. M. (1962). Potential gnp: its measurement and significance. In *Proceedings of the business and economic statistics section of the American Statistical Association*, Volume 7, pp. 89–104. Washington, DC:.

Öller, L.-E. and B. Barot (2000). The accuracy of european growth and inflation forecasts. *International Journal of Forecasting 16*(3), 293–315.

Öller, L.-E. and A. Teterukovsky (2007). Quantifying the quality of macroeconomic variables. *International Journal of Forecasting 23*(2), 205–217.

Paloviita, M. and M. Viren (2014). Inflation and output growth uncertainty in individual survey expectations. *Empirica 41*(1), 69–81.

Patton, A. J. and A. Timmermann (2007a). Properties of optimal forecasts under asymmetric loss and nonlinearity. *Journal of Econometrics 140*(2), 884–918.

Patton, A. J. and A. Timmermann (2007b). Testing forecast optimality under unknown loss. *Journal of the American Statistical Association 102*(480), 1172–1184.

Peiró, A. (1999). Skewness in financial returns. *Journal of Banking & Finance 23*(6), 847–862.

Pesaran, M. H. and S. Skouras (2002). Decision-based methods for forecast evaluation. In M. P. Clements and D. F. Hendry (Eds.), *A Companion to Economic Forecasting*, pp. 241–267. Blackwell Publishing Ltd.

Pierdzioch, C., J.-C. Rülke, and G. Stadtmann (2012). On the loss function of the Bank of Canada: A note. *Economics Letters 115*(2), 155–159.

Pierdzioch, C., J.-C. Rülke, and G. Stadtmann (2013). Oil price forecasting under asymmetric loss. *Applied Economics 45*(17), 2371–2379.

Pierdzioch, C., J.-C. Rülke, and G. Stadtmann (2015). Central banks' inflation forecasts under asymmetric loss: Evidence from four Latin-American countries. *Economics Letters 129*, 66–70.

Pierdzioch, C. and G. Stadtmann (2011). Do exchange-rate forecasters herd? *Applied Economics Letters 18*(8), 739–741.

Podivinsky, J. (1999). Finite-sample properties of gmm estimators and tests. In L. Mátyás (Ed.), *Generalized Method of Moments Estimation*, pp. 128–148. Cambridge University Press, Cambridge (Mass).

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica 32*(1/2), 122–136.

Putnam, B. H. and S. Azzarello (2015). Evolving dynamics of the relationship between US core inflation and unemployment. *Review of Financial Economics 25*, 27–34. Special Issue on Changing Dynamics in Financial Economics.

Ramey, V. A. (2012). Government spending and private activity. In *Fiscal Policy after the Financial Crisis*, NBER Chapters, pp. 19–55. National Bureau of Economic Research, Inc.

Romer, C. D. and D. H. Romer (2000). Federal reserve information and the behavior of interest rates. *The American Economic Review 90*(3), 429–457.

Rudebusch, G. D. and J. C. Williams (2009). Forecasting recessions: The puzzle of the enduring power of the yield curve. *Journal of Business & Economic Statistics 27*(4), 492–503.

Rudin, W. (1964). *Principles of mathematical analysis* (2nd ed.), Volume 3. McGraw-Hill, New York.

Ruge-Murcia, F. J. (2003). Inflation targeting under asymmetric preferences. *Journal of Money, Credit and Banking 35*(5), 763–785.

Samuelson, P. A. (1970). The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. *The Review of Economic Studies 37*(4), 537–542.

Scharfstein, D. S. and J. C. Stein (1990). Herd behavior and investment. *The American Economic Review 80*(3), 465–479.

Scott, R. C. and P. A. Horvath (1980). On the direction of preference for moments of higher order than the variance. *The Journal of Finance 35*(4), 915–919.

Shanno, D. F. (1970). Conditioning of quasi-Newton methods for function minimization. *Mathematics of computation 24*(111), 647–656.

Sinai, A. (1992). Financial and real business cycles. *Eastern Economic Journal 18*(1), pp. 1–54.

Singleton, J. C. and J. Wingender (1986, 9). Skewness persistence in common stock returns. *Journal of Financial and Quantitative Analysis 21*, 335–341.

Stekler, H. (1991). Macroeconomic forecast evaluation techniques. *International Journal of Forecasting 7*(3), 375–384.

Stekler, H. (2007). The future of macroeconomic forecasting: Understanding the forecasting process. *International Journal of Forecasting 23*(2), 237–248.

Stock, J. H. and M. W. Watson (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association 97*(460), 1167–1179.

Stock, J. H. and M. W. Watson (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics 20*(2), 147–162.

Stock, J. H. and M. W. Watson (2006). Forecasting with many predictors. Volume 1 of *Handbook of Economic Forecasting*, Chapter 10, pp. 515–554. Elsevier.

Stock, J. H. and M. W. Watson (2011). Dynamic factor models. In M. P. Clements and D. F. Hendry (Eds.), *The Oxford Handbook of Economic Forecasting*. Oxford University Press.

Strauss, J. (2013). Does housing drive state-level job growth? Building permits and consumer expectations forecast a state's economic activity. *Journal of Urban Economics 73*(1), 77–93.

Swidler, S. and D. Ketcher (1990). Economic forecasts, rationality, and the processing of new information over time. *Journal of Money, Credit and Banking 22*(1), 65–76.

Theil, H. (1966). *Applied economic forecasting*. Rand McNally, Chicago.

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological) 58*(1), pp. 267–288.

Trueman, B. (1994). Analyst forecasts and herding behavior. *Review of Financial Studies 7*(1), 97–124.

Ulu, Y. (2007). Optimal prediction under LINLIN loss: Empirical evidence. *International Journal of Forecasting 23*(4), 707–715.

Ulu, Y. (2013). Multivariate test for forecast rationality under asymmetric loss functions: Recent evidence from MMS survey of inflation-output forecasts. *Economics Letters 119*(2), 168–171.

Varian, H. R. (1975). A bayesian approach to real estate assessment. *Studies in Bayesian econometrics and statistics in honor of Leonard J. Savage*, 195–208.

Verdoorn, J. (1949). On the factors determining the growth of labor productivity. *Italian economic papers 2*, 59–68.

Wang, Y. and T.-H. Lee (2014). Asymmetric loss in the Greenbook and the Survey of Professional Forecasters. *International Journal of Forecasting 30*(2), 235–245.

Weber, E. U. (1994). From subjective probabilities to decision weights: The effect of asymmetric loss functions on the evaluation of uncertain outcomes and events. *Psychological Bulletin 115*(2), 228–242.

Weiss, A. A. (1996). Estimating time series models using the relevant cost function. *Journal of Applied Econometrics 11*(5), 539–560.

West, K. D. (1996). Asymptotic inference about predictive ability. *Econometrica 64*(5), 1067–1084.

Zarnowitz, V. (1968). The New ASA-NBER Survey of Forecasts by Economic Statisticians. In *Supplement to NBER Report Four*, NBER Chapters, pp. 1–8. National Bureau of Economic Research, Inc.

Zarnowitz, V. (1985). Rational expectations and macroeconomic forecasts. *Journal of Business & Economic Statistics 3*(4), pp. 293–311.

Zeileis, A. (2004, 11). Econometric computing with HC and HAC covariance matrix estimators. *Journal of Statistical Software 11*(10), 1–17.

Zellner, A. (1986a). Bayesian estimation and prediction using asymmetric loss functions. *Journal of the American Statistical Association 81*(394), pp. 446–451.

Zellner, A. (1986b). Biased predictors, rationality and the evaluation of forecasts. *Economics Letters 21*(1), 45–48.

Ziliak, S. T. and D. N. McCloskey (2004). Size matters: the standard error of regressions in the American Economic Review. *The Journal of Socio-Economics 33*(5), 527–546.

Appendix A

Table A1: Forecast error measures for GDP growth

		Real-time Forecast Errors					Revised Forecast Errors				
		ME	*MAE*	*RMSE*	*Linex 1*	*Linex 2*	*ME*	*MAE*	*RMSE*	*Linex 1*	*Linex 2*
CEE	1971-1980	-0.710	1.550	2.078	20.426	1.322	-0.439	1.319	1.622	2.148	2.742
	1981-1990	0.220	0.740	1.049	0.464	0.992	0.390	0.966	1.202	0.489	1.490
	1991-2000	-0.190	0.870	1.149	0.533	1.609	-0.172	1.058	1.234	0.729	1.308
	2001-2010	-0.499	1.449	1.948	12.072	1.478	-0.371	1.719	2.194	17.800	2.843
	1971-2007	-0.229	1.041	1.450	6.012	1.212	-0.067	1.135	1.375	1.074	1.882
	1971-2012	-0.266	1.146	1.609	8.177	1.300	-0.122	1.257	1.601	5.170	2.011
JF	1971-1980	-0.810	1.750	2.362	34.604	1.582	-0.539	1.421	1.849	3.917	2.909
	1981-1990	0.370	1.070	1.219	0.849	0.993	0.540	1.282	1.420	0.885	1.742
	1991-2000	-0.400	0.880	1.028	0.849	0.460	-0.382	1.068	1.257	1.137	0.906
	2001-2010	-0.602	1.644	2.083	14.271	1.878	-0.474	1.940	2.353	21.189	3.481
	1971-2007	-0.299	1.219	1.588	10.066	0.981	-0.137	1.289	1.534	1.843	1.899
	1971-2012	-0.327	1.328	1.750	12.344	1.187	-0.183	1.418	1.757	6.626	2.171

Note: *Linex 1* refers to exponential weights for negative forecast errors and *Linex 2* to positive forecast errors.

Figure A1: Forecasts and forecast errors

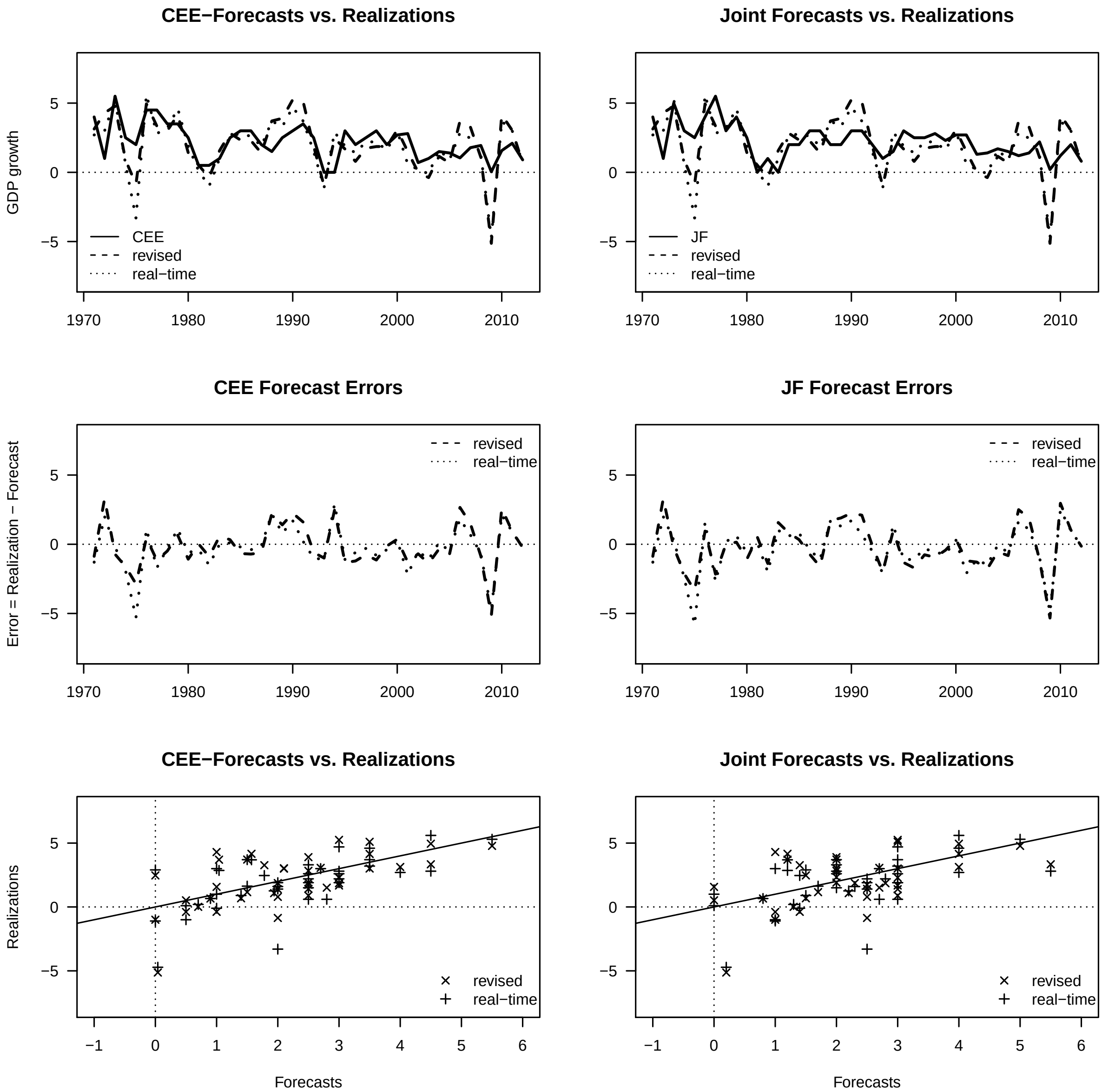


Table A2: Correlation table of forecasts, realizations and instruments

		CEE	JF	real-time	revised	lf	ur	rlc	gdp	inv	gov	ord	bp	rs	cpi	ppi	nlc	dax	ltr	str	ft	xrt	oil	bc	bs	be
1	CEE	1	0.46	0.77	0.66	0.69	-0.68	0.01	0.37	0.50	-0.32	0.11	-0.06	-0.24	-0.34	-0.26	0.42	-0.09	0.20	0.55	0.03	0.19	0.00	-0.12	0.11	-0.39
2	JF		1	0.30	0.36	0.41	-0.31	-0.16	0.04	0.19	-0.24	0.05	-0.24	0.09	-0.33	-0.16	0.12	-0.04	0.15	0.39	0.48	0.24	0.17	-0.11	0.09	-0.36
3	real-time			1	0.89	0.86	-0.82	0.00	0.55	0.67	-0.42	0.36	0.17	-0.22	-0.40	-0.41	0.53	0.13	0.29	0.49	0.04	0.13	0.02	0.24	0.40	-0.02
4	revised				1	0.98	-0.70	-0.01	0.61	0.64	-0.36	0.40	0.11	-0.10	-0.34	-0.39	0.29	0.18	0.34	0.49	0.03	0.15	0.01	0.30	0.46	0.01
5	lf					1	-0.65	0.01	0.56	0.56	-0.31	0.32	0.00	-0.18	-0.29	-0.32	0.25	0.08	0.36	0.51	0.02	0.13	0.08	0.16	0.34	-0.13
6	ur						1	0.13	-0.40	-0.58	0.44	-0.33	-0.22	0.02	0.55	0.57	-0.43	-0.07	-0.11	-0.41	0.04	-0.06	0.15	-0.32	-0.47	-0.04
7	rlc							1	0.51	0.30	0.53	0.17	0.12	-0.06	0.55	0.05	0.25	-0.16	0.11	0.17	-0.28	0.17	0.33	0.01	0.05	-0.04
8	gdp								1	0.82	-0.08	0.81	0.17	-0.06	0.16	-0.27	0.43	0.32	0.44	0.60	-0.11	0.16	0.33	0.59	0.69	0.32
9	inv									1	-0.28	0.67	0.37	-0.04	-0.10	-0.35	0.53	0.41	0.33	0.54	-0.13	0.19	0.13	0.53	0.62	0.28
10	gov										1	-0.25	-0.16	-0.17	0.65	0.34	-0.19	-0.32	-0.01	-0.04	-0.29	0.05	0.23	-0.36	-0.39	-0.23
11	ord											1	0.08	0.09	-0.04	-0.44	0.42	0.49	0.39	0.47	0.05	0.09	0.37	0.80	0.83	0.57
12	bp												1	-0.02	0.01	-0.17	0.11	0.10	-0.02	-0.02	-0.19	-0.01	-0.20	0.30	0.20	0.37
13	rs													1	-0.40	-0.18	-0.19	0.19	-0.19	-0.46	0.30	0.11	-0.11	0.41	0.33	0.42
14	cpi														1	0.64	-0.07	-0.12	0.21	0.21	-0.35	-0.08	0.30	-0.20	-0.22	-0.12
15	ppi															1	-0.28	-0.07	-0.09	-0.07	-0.11	-0.14	-0.12	-0.39	-0.44	-0.24
16	nlc																1	0.11	0.25	0.36	-0.02	0.16	0.15	0.32	0.41	0.12
17	dax																	1	0.21	0.10	0.01	-0.14	-0.11	0.60	0.48	0.61
18	ltr																		1	0.59	-0.08	0.04	0.41	0.25	0.26	0.17
19	str																			1	-0.09	0.05	0.32	0.13	0.35	-0.19
20	ft																				1	0.39	0.10	0.10	0.10	0.08
21	xrt																					1	0.13	0.04	0.12	-0.08
22	oil																						1	0.06	0.17	-0.09
23	bc																							1	0.93	0.87
24	bs																								1	0.62
25	be																									1

Table A3: Regression tests of efficiency (A)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
A0	0.732 (0.487)	2.406 (0.103)	6.919 (0.003)	7.011 (0.002)
A1	0.870 (0.357)	6.007 (0.019)	5.116 (0.029)	1.071 (0.307)
A2	0.515 (0.602)	6.604 (0.004)	4.084 (0.025)	3.133 (0.055)
A3	2.384 (0.106)	5.393 (0.009)	3.866 (0.030)	4.675 (0.015)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Table A4: GMM estimates of the loss function - CEE (A)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
A0	1	0.419 (0.077)	0	0.256 (0.074)	0
	1.5	0.425 (0.090)	0	0.214 (0.075)	0
	2	0.421 (0.103)	0	0.197 (0.083)	0
A1	1	0.431 (0.078)	0.030 (0.863)	0.161 (0.069)	4.227 (0.040)
	1.5	0.460 (0.092)	0.818 (0.366)	0.116 (0.060)	3.120 (0.077)
	2	0.450 (0.107)	1.571 (0.210)	0.095 (0.070)	2.458 (0.117)
A2	1	0.410 (0.077)	0.168 (0.919)	0.130 (0.068)	4.156 (0.125)
	1.5	0.443 (0.090)	1.120 (0.571)	0.114 (0.060)	3.072 (0.215)
	2	0.417 (0.103)	1.964 (0.374)	0.102 (0.052)	2.268 (0.322)
A3	1	0.437 (0.079)	3.113 (0.211)	0.137 (0.064)	5.768 (0.056)
	1.5	0.483 (0.092)	4.320 (0.115)	0.111 (0.063)	4.387 (0.112)
	2	0.538 (0.107)	4.869 (0.088)	0.089 (0.067)	3.175 (0.204)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A5: GMM estimates of the loss function - JF (A)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
A0	1	0.326 (0.073)	0	0.302 (0.084)	0
	1.5	0.349 (0.092)	0	0.219 (0.075)	0
	2	0.333 (0.125)	0	0.157 (0.069)	0
A1	1	0.327 (0.076)	0.533 (0.465)	0.402 (0.087)	5.737 (0.017)
	1.5	0.380 (0.094)	1.687 (0.194)	0.321 (0.080)	4.791 (0.029)
	2	0.496 (0.120)	1.788 (0.181)	0.255 (0.076)	3.141 (0.076)
A2	1	0.301 (0.078)	1.193 (0.551)	0.146 (0.071)	5.640 (0.060)
	1.5	0.360 (0.099)	2.276 (0.320)	0.073 (0.046)	4.633 (0.099)
	2	0.433 (0.131)	2.489 (0.288)	0.036 (0.031)	3.859 (0.145)
A3	1	0.328 (0.077)	0.604 (0.739)	0.251 (0.075)	4.613 (0.100)
	1.5	0.360 (0.089)	1.652 (0.438)	0.164 (0.065)	5.568 (0.062)
	2	0.416 (0.120)	2.362 (0.307)	0.326 (0.077)	4.743 (0.093)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A6: Regression tests of efficiency (B)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
B1	12.556 (0.001)	84.074 (0.000)	676.59 (0.000)	690.08 (0.000)
B2	31.142 (0.000)	34.900 (0.000)	22.909 (0.000)	30.119 (0.000)
B3	0.002 (0.962)	0.124 (0.727)	0.013 (0.908)	0.261 (0.613)
B4	12.829 (0.000)	88.250 (0.000)	195.19 (0.000)	628.50 (0.000)
B5	17.823 (0.000)	20.862 (0.000)	10.190 (0.000)	16.301 (0.000)
B6	10.167 (0.000)	4.050 (0.026)	9.450 (0.000)	4.956 (0.012)
B7	6.628 (0.003)	54.049 (0.000)	473.52 (0.000)	382.04 (0.000)
B8	21.397 (0.000)	47.751 (0.000)	11.090 (0.000)	19.764 (0.000)
B9	0.024 (0.977)	0.627 (0.539)	0.238 (0.790)	0.824 (0.446)
B10	36.137 (0.000)	315.85 (0.000)	485.87 (0.000)	966.61 (0.000)
B11	17.828 (0.000)	58.512 (0.000)	279.41 (0.000)	396.75 (0.000)
B12	15.637 (0.000)	23.279 (0.000)	10.951 (0.000)	13.746 (0.000)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Table A7: GMM estimates of the loss function - CEE (B)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
B1	1	0.232 (0.071)	13.349 (0.000)	0.091 (0.058)	7.649 (0.006)
	1.5	0.702 (0.089)	12.633 (0.000)	0.065 (0.041)	5.133 (0.023)
	2	0.199 (0.081)	6.498 (0.011)	0.044 (0.033)	3.551 (0.059)
B2	1	0.308 (0.078)	7.298 (0.007)	0.225 (0.074)	2.706 (0.100)
	1.5	0.310 (0.085)	4.900 (0.027)	0.180 (0.073)	1.867 (0.172)
	2	0.295 (0.090)	3.156 (0.076)	0.169 (0.083)	1.363 (0.243)
B3	1	0.426 (0.079)	0.059 (0.809)	0.266 (0.076)	0.059 (0.808)
	1.5	0.426 (0.086)	0.052 (0.820)	0.216 (0.077)	0.071 (0.790)
	2	0.428 (0.104)	0.000 (0.991)	0.176 (0.072)	0.245 (0.620)
B4	1	0.193 (0.071)	12.440 (0.002)	0.058 (0.054)	7.417 (0.025)
	1.5	0.165 (0.074)	11.509 (0.003)	0.058 (0.041)	5.516 (0.063)
	2	0.156 (0.077)	7.501 (0.024)	0.048 (0.032)	4.326 (0.115)
B5	1	0.330 (0.077)	8.942 (0.011)	0.087 (0.053)	9.614 (0.008)
	1.5	0.289 (0.083)	8.085 (0.018)	0.011 (0.027)	8.179 (0.017)
	2	0.240 (0.090)	8.198 (0.017)	0.487 (0.080)	32.611 (0.000)
B6	1	0.338 (0.077)	5.425 (0.066)	0.119 (0.059)	5.007 (0.082)
	1.5	0.273 (0.085)	6.979 (0.031)	0.085 (0.049)	4.736 (0.094)
	2	0.236 (0.097)	5.670 (0.059)	0.061 (0.044)	4.290 (0.117)
B7	1	0.177 (0.064)	17.457 (0.000)	0.036 (0.032)	10.757 (0.005)
	1.5	0.453 (0.092)	14.228 (0.001)	0.030 (0.023)	7.561 (0.023)
	2	0.462 (0.099)	11.609 (0.003)	0.028 (0.019)	4.389 (0.111)
B8	1	0.197 (0.072)	14.885 (0.001)	0.247 (0.068)	6.317 (0.042)
	1.5	0.751 (0.060)	17.944 (0.000)	0.188 (0.062)	5.494 (0.064)
	2	0.714 (0.073)	14.307 (0.001)	0.187 (0.071)	3.488 (0.175)
B9	1	0.424 (0.080)	0.104 (0.950)	0.264 (0.076)	0.413 (0.813)
	1.5	0.434 (0.083)	0.103 (0.950)	0.213 (0.073)	0.266 (0.875)
	2	0.441 (0.094)	0.058 (0.971)	0.177 (0.067)	0.317 (0.853)
B10	1	0.184 (0.070)	15.389 (0.002)	0.068 (0.047)	8.914 (0.030)
	1.5	0.581 (0.087)	13.727 (0.003)	0.069 (0.036)	6.434 (0.092)
	2	0.582 (0.103)	11.812 (0.008)	0.053 (0.028)	4.244 (0.236)
B11	1	0.210 (0.071)	12.094 (0.007)	0.071 (0.057)	8.963 (0.030)
	1.5	0.729 (0.087)	14.280 (0.003)	0.067 (0.039)	5.987 (0.112)
	2	0.197 (0.082)	6.287 (0.098)	0.042 (0.027)	5.049 (0.168)
B12	1	0.364 (0.061)	10.367 (0.016)	0.224 (0.075)	3.758 (0.289)
	1.5	0.276 (0.076)	8.323 (0.040)	0.158 (0.065)	4.483 (0.214)
	2	0.852 (0.095)	15.671 (0.001)	0.118 (0.054)	4.616 (0.202)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A8: GMM estimates of the loss function - JF (B)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
B1	1	0.504 (0.076)	16.445 (0.000)	0.530 (0.088)	14.520 (0.000)
	1.5	0.584 (0.094)	8.060 (0.005)	0.401 (0.083)	11.390 (0.001)
	2	0.641 (0.121)	3.823 (0.051)	0.304 (0.081)	6.015 (0.014)
B2	1	0.261 (0.076)	4.096 (0.043)	0.237 (0.089)	2.239 (0.135)
	1.5	0.246 (0.087)	3.147 (0.076)	0.154 (0.070)	1.951 (0.163)
	2	0.204 (0.097)	2.470 (0.116)	0.096 (0.053)	1.762 (0.184)
B3	1	0.331 (0.075)	0.057 (0.812)	0.305 (0.071)	0.007 (0.932)
	1.5	0.352 (0.096)	0.181 (0.670)	0.219 (0.068)	0.016 (0.899)
	2	0.358 (0.127)	0.201 (0.654)	0.157 (0.069)	0.031 (0.859)
B4	1	0.542 (0.077)	19.183 (0.000)	0.525 (0.085)	19.314 (0.000)
	1.5	0.550 (0.089)	11.275 (0.004)	0.039 (0.037)	5.918 (0.052)
	2	0.533 (0.108)	7.308 (0.026)	0.028 (0.024)	4.474 (0.107)
B5	1	0.255 (0.070)	5.014 (0.082)	0.230 (0.081)	2.890 (0.236)
	1.5	0.243 (0.075)	3.241 (0.198)	0.116 (0.053)	3.489 (0.175)
	2	0.203 (0.091)	2.457 (0.293)	0.071 (0.040)	2.877 (0.237)
B6	1	0.240 (0.071)	4.333 (0.115)	0.210 (0.072)	3.931 (0.140)
	1.5	0.304 (0.096)	5.833 (0.054)	0.059 (0.042)	5.904 (0.052)
	2	0.386 (0.131)	5.329 (0.070)	0.011 (0.022)	5.953 (0.051)
B7	1	0.123 (0.057)	13.132 (0.001)	0.483 (0.086)	14.567 (0.001)
	1.5	0.396 (0.092)	9.788 (0.007)	0.012 (0.020)	10.626 (0.005)
	2	0.379 (0.111)	5.743 (0.057)	0.142 (0.049)	6.613 (0.037)
B8	1	0.491 (0.074)	16.828 (0.000)	0.104 (0.061)	6.798 (0.033)
	1.5	0.582 (0.082)	13.690 (0.001)	0.037 (0.040)	7.086 (0.029)
	2	0.634 (0.093)	7.282 (0.026)	0.206 (0.069)	6.345 (0.042)
B9	1	0.328 (0.075)	0.070 (0.966)	0.302 (0.064)	0.015 (0.992)
	1.5	0.365 (0.079)	0.218 (0.897)	0.236 (0.052)	0.158 (0.924)
	2	0.430 (0.095)	0.590 (0.744)	0.200 (0.052)	0.557 (0.757)
B10	1	0.515 (0.075)	18.106 (0.000)	0.552 (0.092)	16.107 (0.001)
	1.5	0.462 (0.086)	9.920 (0.019)	0.020 (0.023)	7.818 (0.050)
	2	0.458 (0.112)	5.109 (0.164)	0.011 (0.013)	6.371 (0.095)
B11	1	0.152 (0.068)	10.776 (0.013)	0.475 (0.070)	15.015 (0.002)
	1.5	0.628 (0.093)	11.655 (0.009)	0.054 (0.039)	6.934 (0.074)
	2	0.768 (0.111)	11.139 (0.011)	0.037 (0.028)	4.765 (0.190)
B12	1	0.222 (0.072)	6.586 (0.086)	0.238 (0.066)	2.148 (0.542)
	1.5	0.646 (0.100)	12.438 (0.006)	0.133 (0.048)	2.397 (0.494)
	2	0.144 (0.087)	4.819 (0.186)	0.069 (0.035)	2.647 (0.449)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A9: Regression tests of efficiency (C)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
C1	5.685 (0.022)	5.735 (0.022)	8.233 (0.007)	7.864 (0.008)
C2	9.574 (0.004)	11.778 (0.001)	12.917 (0.001)	14.706 (0.000)
C3	3.016 (0.090)	3.725 (0.061)	0.842 (0.364)	1.900 (0.176)
C4	6.708 (0.003)	4.145 (0.024)	4.336 (0.020)	4.418 (0.019)
C5	7.535 (0.002)	12.015 (0.000)	9.658 (0.000)	14.168 (0.000)
C6	1.550 (0.225)	5.113 (0.011)	0.584 (0.563)	2.868 (0.069)
C7	5.584 (0.007)	7.723 (0.002)	9.664 (0.000)	14.430 (0.000)
C8	5.613 (0.007)	7.964 (0.001)	6.387 (0.004)	7.901 (0.001)
C9	1.562 (0.222)	2.978 (0.063)	0.709 (0.499)	1.117 (0.337)
C10	6.116 (0.002)	7.968 (0.000)	9.027 (0.000)	9.785 (0.000)
C11	4.368 (0.010)	10.846 (0.000)	2.663 (0.062)	4.481 (0.009)
C12	4.640 (0.007)	5.765 (0.002)	6.594 (0.001)	5.779 (0.002)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Table A10: GMM estimates of the loss function - CEE (C)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
C1	1	0.453 (0.078)	8.958 (0.003)	0.153 (0.062)	4.527 (0.033)
	1.5	0.686 (0.098)	15.832 (0.000)	0.063 (0.043)	5.287 (0.021)
	2	0.181 (0.080)	7.100 (0.008)	0.027 (0.051)	4.926 (0.026)
C2	1	0.318 (0.075)	6.621 (0.010)	0.210 (0.070)	1.930 (0.165)
	1.5	0.258 (0.081)	7.062 (0.008)	0.128 (0.065)	3.358 (0.067)
	2	0.222 (0.086)	5.774 (0.016)	0.152 (0.084)	2.762 (0.097)
C3	1	0.403 (0.075)	1.180 (0.277)	0.257 (0.074)	0.436 (0.509)
	1.5	0.399 (0.090)	1.933 (0.164)	0.199 (0.072)	0.555 (0.456)
	2	0.345 (0.098)	2.514 (0.113)	0.157 (0.069)	0.742 (0.389)
C4	1	0.486 (0.079)	12.517 (0.002)	0.082 (0.051)	6.051 (0.049)
	1.5	0.770 (0.100)	18.536 (0.000)	0.047 (0.033)	5.452 (0.065)
	2	0.156 (0.080)	7.820 (0.020)	0.457 (0.086)	23.379 (0.000)
C5	1	0.244 (0.073)	8.767 (0.012)	0.132 (0.063)	3.899 (0.142)
	1.5	0.191 (0.077)	8.895 (0.012)	0.104 (0.060)	4.016 (0.134)
	2	0.173 (0.083)	7.250 (0.027)	0.115 (0.071)	3.811 (0.149)
C6	1	0.403 (0.076)	1.484 (0.476)	0.260 (0.075)	0.818 (0.664)
	1.5	0.393 (0.090)	1.972 (0.373)	0.194 (0.068)	0.582 (0.748)
	2	0.349 (0.100)	2.448 (0.294)	0.136 (0.058)	0.979 (0.613)
C7	1	0.592 (0.079)	16.957 (0.000)	0.122 (0.054)	6.522 (0.038)
	1.5	0.675 (0.096)	14.091 (0.001)	0.037 (0.028)	6.343 (0.042)
	2	0.445 (0.102)	14.754 (0.001)	0.007 (0.018)	5.151 (0.076)
C8	1	0.303 (0.072)	7.176 (0.028)	0.219 (0.071)	1.905 (0.386)
	1.5	0.234 (0.077)	8.346 (0.015)	0.132 (0.065)	3.237 (0.198)
	2	0.186 (0.081)	7.167 (0.028)	0.128 (0.071)	2.940 (0.230)
C9	1	0.414 (0.078)	1.102 (0.576)	0.247 (0.073)	0.844 (0.656)
	1.5	0.411 (0.091)	2.023 (0.364)	0.195 (0.071)	0.789 (0.674)
	2	0.388 (0.101)	2.845 (0.241)	0.141 (0.064)	0.988 (0.610)
C10	1	0.602 (0.080)	16.724 (0.001)	0.138 (0.061)	5.877 (0.118)
	1.5	0.676 (0.092)	14.711 (0.002)	0.047 (0.033)	5.745 (0.125)
	2	0.129 (0.074)	9.819 (0.020)	0.016 (0.021)	5.665 (0.129)
C11	1	0.445 (0.078)	10.570 (0.014)	0.154 (0.060)	5.885 (0.117)
	1.5	0.526 (0.093)	14.657 (0.002)	0.053 (0.037)	5.585 (0.134)
	2	0.122 (0.070)	11.573 (0.009)	0.026 (0.026)	5.370 (0.147)
C12	1	0.343 (0.073)	10.749 (0.013)	0.141 (0.057)	5.793 (0.122)
	1.5	0.272 (0.078)	11.846 (0.008)	0.075 (0.043)	4.934 (0.177)
	2	0.207 (0.078)	11.231 (0.011)	0.039 (0.030)	4.773 (0.189)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A11: GMM estimates of the loss function - JF (C)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
C1	1	0.247 (0.069)	7.075 (0.008)	0.202 (0.070)	3.686 (0.055)
	1.5	0.544 (0.099)	7.823 (0.005)	0.079 (0.040)	5.076 (0.024)
	2	0.622 (0.113)	3.430 (0.064)	0.283 (0.079)	5.911 (0.015)
C2	1	0.225 (0.069)	5.876 (0.015)	0.243 (0.079)	2.148 (0.143)
	1.5	0.206 (0.084)	5.263 (0.022)	0.111 (0.052)	3.383 (0.066)
	2	0.699 (0.100)	4.543 (0.033)	0.055 (0.033)	3.265 (0.071)
C3	1	0.308 (0.073)	1.910 (0.167)	0.307 (0.084)	0.651 (0.420)
	1.5	0.311 (0.091)	2.153 (0.142)	0.209 (0.074)	0.864 (0.353)
	2	0.239 (0.110)	1.889 (0.169)	0.134 (0.064)	0.880 (0.348)
C4	1	0.121 (0.055)	10.496 (0.005)	0.144 (0.066)	4.775 (0.092)
	1.5	0.620 (0.099)	15.111 (0.001)	0.036 (0.026)	5.613 (0.060)
	2	0.697 (0.111)	7.024 (0.030)	0.012 (0.012)	4.966 (0.083)
C5	1	0.151 (0.063)	8.069 (0.018)	0.176 (0.077)	3.674 (0.159)
	1.5	0.659 (0.097)	12.615 (0.002)	0.070 (0.042)	4.365 (0.113)
	2	0.755 (0.080)	6.100 (0.047)	0.037 (0.025)	3.656 (0.161)
C6	1	0.308 (0.074)	1.836 (0.399)	0.303 (0.081)	0.675 (0.714)
	1.5	0.319 (0.090)	2.289 (0.318)	0.211 (0.073)	0.862 (0.650)
	2	0.313 (0.108)	2.527 (0.283)	0.153 (0.064)	1.164 (0.559)
C7	1	0.148 (0.058)	13.077 (0.001)	0.153 (0.060)	6.513 (0.039)
	1.5	0.463 (0.093)	8.950 (0.011)	0.070 (0.035)	6.412 (0.041)
	2	0.285 (0.108)	6.882 (0.032)	0.025 (0.019)	6.200 (0.045)
C8	1	0.217 (0.070)	5.631 (0.060)	0.245 (0.080)	2.191 (0.334)
	1.5	0.187 (0.084)	5.784 (0.055)	0.102 (0.051)	3.467 (0.177)
	2	0.112 (0.083)	5.759 (0.056)	0.047 (0.032)	3.673 (0.159)
C9	1	0.308 (0.073)	2.008 (0.366)	0.306 (0.083)	0.619 (0.734)
	1.5	0.333 (0.093)	2.329 (0.312)	0.211 (0.074)	0.840 (0.657)
	2	0.258 (0.113)	2.429 (0.297)	0.135 (0.062)	0.854 (0.652)
C10	1	0.182 (0.067)	9.742 (0.021)	0.153 (0.063)	5.394 (0.145)
	1.5	0.386 (0.093)	8.678 (0.034)	0.075 (0.040)	5.588 (0.133)
	2	0.378 (0.111)	6.808 (0.078)	0.032 (0.024)	5.646 (0.130)
C11	1	0.297 (0.072)	10.175 (0.017)	0.220 (0.068)	4.037 (0.257)
	1.5	0.417 (0.092)	8.644 (0.034)	0.060 (0.035)	6.152 (0.104)
	2	0.481 (0.114)	4.915 (0.178)	0.019 (0.015)	5.364 (0.147)
C12	1	0.289 (0.070)	12.881 (0.005)	0.243 (0.079)	2.580 (0.461)
	1.5	0.449 (0.098)	8.770 (0.033)	0.068 (0.043)	5.590 (0.133)
	2	0.538 (0.114)	5.475 (0.140)	0.018 (0.018)	5.716 (0.126)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A12: Regression tests of efficiency (D)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
D1	4.787 (0.035)	3.800 (0.058)	5.380 (0.026)	5.218 (0.028)
D2	4.779 (0.035)	1.857 (0.181)	1.557 (0.219)	1.304 (0.260)
D3	0.155 (0.696)	2.020 (0.164)	0.165 (0.687)	0.779 (0.383)
D4	3.182 (0.052)	2.540 (0.092)	3.275 (0.048)	2.992 (0.062)
D5	6.204 (0.005)	2.162 (0.129)	3.768 (0.032)	2.073 (0.139)
D6	1.044 (0.363)	13.891 (0.000)	1.011 (0.374)	16.328 (0.000)
D7	11.316 (0.000)	5.145 (0.010)	7.905 (0.001)	5.283 (0.009)
D8	4.809 (0.014)	1.246 (0.299)	1.613 (0.212)	0.776 (0.467)
D9	1.338 (0.275)	1.494 (0.238)	1.187 (0.317)	1.115 (0.339)
D10	3.477 (0.025)	2.013 (0.128)	2.811 (0.052)	2.058 (0.122)
D11	1.345 (0.276)	1.639 (0.198)	2.632 (0.065)	2.357 (0.088)
D12	5.463 (0.003)	1.203 (0.323)	2.531 (0.073)	0.581 (0.631)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Table A13: GMM estimates of the loss function - CEE (D)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
D1	1	0.432 (0.075)	5.962 (0.015)	0.167 (0.064)	4.897 (0.027)
	1.5	0.245 (0.080)	9.513 (0.002)	0.082 (0.046)	4.321 (0.038)
	2	0.601 (0.104)	8.690 (0.003)	0.044 (0.036)	3.560 (0.059)
D2	1	0.388 (0.077)	4.844 (0.028)	0.250 (0.074)	0.693 (0.405)
	1.5	0.372 (0.088)	4.318 (0.038)	0.215 (0.076)	0.994 (0.319)
	2	0.345 (0.098)	3.498 (0.061)	0.209 (0.083)	0.569 (0.450)
D3	1	0.420 (0.081)	0.130 (0.719)	0.252 (0.080)	0.075 (0.784)
	1.5	0.425 (0.098)	0.146 (0.703)	0.209 (0.073)	0.145 (0.703)
	2	0.401 (0.105)	0.363 (0.547)	0.168 (0.069)	0.474 (0.491)
D4	1	0.423 (0.076)	5.985 (0.050)	0.178 (0.065)	4.956 (0.084)
	1.5	0.262 (0.082)	9.353 (0.009)	0.065 (0.042)	5.917 (0.052)
	2	0.263 (0.092)	7.481 (0.024)	0.007 (0.027)	6.362 (0.042)
D5	1	0.337 (0.077)	7.381 (0.025)	0.176 (0.065)	3.853 (0.146)
	1.5	0.282 (0.083)	6.506 (0.039)	0.123 (0.060)	3.837 (0.147)
	2	0.263 (0.089)	4.559 (0.102)	0.086 (0.055)	3.975 (0.137)
D6	1	0.392 (0.081)	1.475 (0.478)	0.255 (0.082)	0.465 (0.793)
	1.5	0.388 (0.098)	1.177 (0.555)	0.199 (0.071)	0.422 (0.810)
	2	0.362 (0.103)	1.195 (0.550)	0.133 (0.059)	1.328 (0.515)
D7	1	0.328 (0.073)	20.942 (0.000)	0.013 (0.027)	10.706 (0.005)
	1.5	0.357 (0.078)	17.860 (0.000)	0.013 (0.015)	7.294 (0.026)
	2	0.377 (0.097)	9.531 (0.009)	0.011 (0.010)	4.531 (0.104)
D8	1	0.400 (0.077)	4.830 (0.089)	0.247 (0.069)	3.722 (0.156)
	1.5	0.430 (0.090)	4.937 (0.085)	0.227 (0.070)	2.650 (0.266)
	2	0.445 (0.102)	4.166 (0.125)	0.211 (0.074)	1.832 (0.400)
D9	1	0.421 (0.081)	1.048 (0.592)	0.256 (0.080)	0.846 (0.655)
	1.5	0.354 (0.095)	3.634 (0.163)	0.203 (0.074)	0.926 (0.630)
	2	0.279 (0.092)	3.633 (0.163)	0.142 (0.064)	1.245 (0.537)
D10	1	0.492 (0.079)	10.819 (0.013)	0.160 (0.060)	7.364 (0.061)
	1.5	0.323 (0.085)	12.500 (0.006)	0.063 (0.042)	7.090 (0.069)
	2	0.297 (0.091)	8.950 (0.030)	0.030 (0.032)	6.539 (0.088)
D11	1	0.435 (0.082)	6.336 (0.096)	0.100 (0.061)	6.278 (0.099)
	1.5	0.720 (0.104)	15.507 (0.001)	0.085 (0.048)	5.054 (0.168)
	2	0.156 (0.069)	7.322 (0.062)	0.054 (0.036)	4.273 (0.233)
D12	1	0.420 (0.084)	5.753 (0.124)	0.213 (0.077)	4.448 (0.217)
	1.5	0.386 (0.097)	5.938 (0.115)	0.166 (0.066)	4.496 (0.213)
	2	0.298 (0.099)	5.139 (0.162)	0.089 (0.051)	4.133 (0.247)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A14: GMM estimates of the loss function - JF (D)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
D1	1	0.309 (0.070)	6.503 (0.011)	0.251 (0.080)	2.622 (0.105)
	1.5	0.147 (0.078)	9.079 (0.003)	0.092 (0.050)	5.274 (0.022)
	2	0.747 (0.179)	9.769 (0.002)	0.042 (0.031)	4.387 (0.036)
D2	1	0.244 (0.067)	4.867 (0.027)	0.263 (0.079)	1.884 (0.170)
	1.5	0.253 (0.082)	3.628 (0.057)	0.156 (0.061)	2.358 (0.125)
	2	0.203 (0.094)	2.611 (0.106)	0.093 (0.046)	2.103 (0.147)
D3	1	0.314 (0.077)	0.843 (0.359)	0.267 (0.089)	1.268 (0.260)
	1.5	0.358 (0.103)	1.749 (0.186)	0.243 (0.086)	1.582 (0.208)
	2	0.470 (0.132)	1.445 (0.229)	0.208 (0.081)	1.201 (0.273)
D4	1	0.327 (0.072)	7.113 (0.029)	0.251 (0.080)	2.939 (0.230)
	1.5	0.150 (0.078)	8.891 (0.012)	0.092 (0.050)	5.188 (0.075)
	2	0.906 (0.161)	10.268 (0.006)	0.043 (0.031)	4.349 (0.114)
D5	1	0.240 (0.067)	5.041 (0.080)	0.252 (0.078)	2.029 (0.363)
	1.5	0.238 (0.082)	4.191 (0.123)	0.122 (0.053)	3.168 (0.205)
	2	0.419 (0.110)	4.065 (0.131)	0.045 (0.028)	3.952 (0.139)
D6	1	0.294 (0.079)	2.014 (0.365)	0.252 (0.094)	1.650 (0.438)
	1.5	0.223 (0.087)	2.945 (0.229)	0.139 (0.076)	2.510 (0.285)
	2	0.406 (0.134)	3.216 (0.200)	0.145 (0.067)	2.743 (0.254)
D7	1	0.018 (0.032)	17.392 (0.000)	0.251 (0.077)	5.688 (0.058)
	1.5	0.018 (0.038)	12.341 (0.002)	0.041 (0.027)	6.816 (0.033)
	2	0.158 (0.080)	8.278 (0.016)	0.015 (0.011)	5.658 (0.059)
D8	1	0.240 (0.066)	6.400 (0.041)	0.261 (0.074)	2.156 (0.340)
	1.5	0.284 (0.079)	6.251 (0.044)	0.157 (0.057)	3.395 (0.183)
	2	0.529 (0.142)	4.824 (0.090)	0.089 (0.042)	3.576 (0.167)
D9	1	0.310 (0.076)	0.912 (0.634)	0.269 (0.086)	1.467 (0.480)
	1.5	0.363 (0.101)	1.658 (0.436)	0.225 (0.078)	1.644 (0.439)
	2	0.444 (0.134)	1.619 (0.445)	0.190 (0.074)	1.378 (0.502)
D10	1	0.281 (0.069)	10.865 (0.012)	0.250 (0.080)	2.712 (0.438)
	1.5	0.214 (0.083)	11.915 (0.008)	0.120 (0.053)	6.862 (0.076)
	2	0.358 (0.113)	10.613 (0.014)	-0.006 (0.017)	9.808 (0.020)
D11	1	0.394 (0.077)	9.898 (0.019)	0.180 (0.076)	3.699 (0.296)
	1.5	0.721 (0.093)	19.694 (0.000)	0.075 (0.049)	5.259 (0.154)
	2	0.738 (0.089)	11.968 (0.007)	0.045 (0.033)	5.386 (0.146)
D12	1	0.210 (0.069)	6.777 (0.079)	0.221 (0.084)	3.119 (0.374)
	1.5	0.188 (0.079)	5.100 (0.165)	0.123 (0.065)	3.466 (0.325)
	2	0.112 (0.070)	3.870 (0.276)	0.068 (0.046)	3.144 (0.370)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A15: Regression tests of efficiency (E)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
E1	1.875 (0.179)	4.414 (0.042)	0.810 (0.374)	2.046 (0.160)
E2	3.206 (0.081)	4.506 (0.040)	4.770 (0.035)	5.024 (0.031)
E3	3.496 (0.069)	16.876 (0.000)	0.005 (0.943)	1.815 (0.186)
E4	1.177 (0.319)	6.732 (0.003)	0.952 (0.395)	5.472 (0.008)
E5	1.661 (0.203)	2.932 (0.065)	2.423 (0.102)	2.883 (0.068)
E6	7.344 (0.002)	7.284 (0.002)	7.226 (0.002)	0.943 (0.399)
E7	2.020 (0.146)	3.081 (0.057)	0.428 (0.655)	1.391 (0.261)
E8	4.877 (0.013)	9.902 (0.000)	3.772 (0.032)	8.027 (0.001)
E9	7.667 (0.002)	9.913 (0.000)	17.713 (0.000)	3.749 (0.033)
E10	7.023 (0.001)	7.743 (0.000)	2.088 (0.118)	4.427 (0.009)
E11	2.670 (0.062)	8.093 (0.000)	0.577 (0.634)	2.798 (0.054)
E12	2.112 (0.116)	12.404 (0.000)	1.493 (0.233)	3.317 (0.031)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Table A16: GMM estimates of the loss function - CEE (E)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
E1	1	0.409 (0.077)	1.415 (0.234)	0.257 (0.073)	0.018 (0.892)
	1.5	0.415 (0.091)	1.429 (0.232)	0.208 (0.074)	0.274 (0.601)
	2	0.375 (0.101)	1.507 (0.220)	0.159 (0.072)	0.707 (0.400)
E2	1	0.397 (0.076)	1.803 (0.179)	0.258 (0.074)	0.224 (0.636)
	1.5	0.394 (0.090)	1.627 (0.202)	0.205 (0.075)	0.639 (0.424)
	2	0.377 (0.100)	1.511 (0.219)	0.177 (0.083)	0.980 (0.322)
E3	1	0.383 (0.082)	3.178 (0.075)	0.243 (0.079)	0.537 (0.464)
	1.5	0.640 (0.087)	8.572 (0.003)	0.208 (0.073)	0.035 (0.851)
	2	0.648 (0.099)	7.113 (0.008)	0.166 (0.067)	0.337 (0.562)
E4	1	0.402 (0.077)	1.544 (0.462)	0.257 (0.074)	0.026 (0.987)
	1.5	0.407 (0.092)	2.004 (0.367)	0.209 (0.075)	0.331 (0.848)
	2	0.392 (0.103)	1.876 (0.391)	0.163 (0.071)	0.754 (0.686)
E5	1	0.397 (0.077)	1.859 (0.395)	0.254 (0.074)	0.398 (0.820)
	1.5	0.391 (0.091)	2.021 (0.364)	0.200 (0.074)	0.828 (0.661)
	2	0.366 (0.101)	1.644 (0.440)	0.170 (0.079)	1.366 (0.505)
E6	1	0.353 (0.082)	5.749 (0.056)	0.211 (0.075)	2.258 (0.323)
	1.5	0.216 (0.090)	8.734 (0.013)	0.155 (0.060)	2.469 (0.291)
	2	0.209 (0.099)	6.112 (0.047)	0.107 (0.050)	2.831 (0.243)
E7	1	0.399 (0.078)	1.539 (0.463)	0.257 (0.073)	0.984 (0.611)
	1.5	0.385 (0.089)	1.777 (0.411)	0.200 (0.073)	0.578 (0.749)
	2	0.353 (0.093)	1.589 (0.452)	0.148 (0.067)	0.795 (0.672)
E8	1	0.383 (0.078)	1.886 (0.389)	0.253 (0.071)	0.275 (0.872)
	1.5	0.394 (0.086)	1.643 (0.440)	0.204 (0.067)	0.610 (0.737)
	2	0.393 (0.093)	1.621 (0.445)	0.181 (0.073)	1.139 (0.566)
E9	1	0.224 (0.076)	9.425 (0.009)	0.175 (0.079)	2.746 (0.253)
	1.5	0.624 (0.090)	9.437 (0.009)	0.151 (0.068)	2.255 (0.324)
	2	0.251 (0.090)	5.707 (0.058)	0.112 (0.055)	1.917 (0.383)
E10	1	0.366 (0.079)	3.246 (0.355)	0.255 (0.075)	0.273 (0.965)
	1.5	0.368 (0.079)	1.942 (0.584)	0.207 (0.073)	0.668 (0.881)
	2	0.366 (0.086)	1.489 (0.685)	0.170 (0.068)	1.299 (0.729)
E11	1	0.393 (0.082)	3.914 (0.271)	0.219 (0.079)	1.590 (0.662)
	1.5	0.627 (0.085)	10.185 (0.017)	0.212 (0.073)	0.559 (0.906)
	2	0.244 (0.096)	7.381 (0.061)	0.153 (0.063)	0.779 (0.854)
E12	1	0.375 (0.083)	3.181 (0.365)	0.242 (0.080)	0.454 (0.929)
	1.5	0.296 (0.093)	6.731 (0.081)	0.187 (0.070)	0.748 (0.862)
	2	0.299 (0.097)	4.727 (0.193)	0.127 (0.058)	1.590 (0.662)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A17: GMM estimates of the loss function - JF (E)

IV set	p	Real-Time $\hat{\alpha}$	Real-Time J	Revised $\hat{\alpha}$	Revised J
E1	1	0.325 (0.074)	0.144 (0.704)	0.307 (0.083)	0.065 (0.799)
	1.5	0.345 (0.093)	0.339 (0.560)	0.216 (0.075)	0.237 (0.626)
	2	0.297 (0.118)	0.512 (0.474)	0.142 (0.065)	0.421 (0.516)
E2	1	0.321 (0.073)	0.339 (0.560)	0.298 (0.084)	0.513 (0.474)
	1.5	0.325 (0.091)	0.910 (0.340)	0.192 (0.071)	0.950 (0.330)
	2	0.270 (0.113)	1.113 (0.291)	0.115 (0.057)	1.169 (0.280)
E3	1	0.289 (0.076)	3.791 (0.052)	0.262 (0.090)	3.146 (0.076)
	1.5	0.419 (0.101)	3.653 (0.056)	0.152 (0.075)	3.880 (0.049)
	2	0.554 (0.122)	2.925 (0.087)	0.078 (0.058)	3.295 (0.069)
E4	1	0.324 (0.074)	0.490 (0.783)	0.307 (0.081)	0.063 (0.969)
	1.5	0.304 (0.091)	2.382 (0.304)	0.208 (0.073)	0.376 (0.828)
	2	0.201 (0.102)	2.863 (0.239)	0.129 (0.061)	0.761 (0.683)
E5	1	0.320 (0.074)	0.344 (0.842)	0.283 (0.081)	1.425 (0.490)
	1.5	0.323 (0.088)	0.906 (0.636)	0.190 (0.068)	2.285 (0.319)
	2	0.266 (0.105)	1.755 (0.416)	0.131 (0.053)	3.725 (0.155)
E6	1	0.279 (0.075)	4.405 (0.111)	0.247 (0.088)	3.579 (0.167)
	1.5	0.315 (0.093)	4.596 (0.100)	0.142 (0.067)	4.165 (0.125)
	2	0.309 (0.106)	4.525 (0.104)	0.085 (0.050)	3.527 (0.171)
E7	1	0.301 (0.074)	1.422 (0.491)	0.305 (0.080)	0.071 (0.965)
	1.5	0.279 (0.087)	2.470 (0.291)	0.188 (0.065)	0.811 (0.667)
	2	0.307 (0.117)	3.124 (0.210)	0.105 (0.047)	1.712 (0.425)
E8	1	0.306 (0.074)	1.274 (0.529)	0.296 (0.081)	0.511 (0.775)
	1.5	0.314 (0.086)	1.062 (0.588)	0.198 (0.066)	1.014 (0.602)
	2	0.274 (0.105)	1.146 (0.564)	0.123 (0.053)	1.398 (0.497)
E9	1	0.343 (0.077)	4.480 (0.106)	0.318 (0.083)	3.813 (0.149)
	1.5	0.355 (0.099)	3.736 (0.154)	0.272 (0.078)	5.153 (0.076)
	2	0.643 (0.127)	5.162 (0.076)	0.047 (0.033)	3.677 (0.159)
E10	1	0.278 (0.072)	3.111 (0.375)	0.251 (0.070)	1.204 (0.752)
	1.5	0.226 (0.076)	3.256 (0.354)	0.126 (0.045)	2.496 (0.476)
	2	0.441 (0.138)	4.348 (0.226)	0.064 (0.034)	3.321 (0.345)
E11	1	0.309 (0.078)	4.246 (0.236)	0.259 (0.084)	3.715 (0.294)
	1.5	0.463 (0.098)	3.752 (0.290)	0.148 (0.064)	5.450 (0.142)
	2	0.545 (0.122)	4.068 (0.254)	0.291 (0.080)	6.177 (0.103)
E12	1	0.298 (0.078)	3.792 (0.285)	0.269 (0.087)	3.105 (0.376)
	1.5	0.401 (0.098)	3.778 (0.286)	0.187 (0.072)	4.244 (0.236)
	2	0.490 (0.125)	3.787 (0.285)	0.191 (0.065)	4.686 (0.196)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A18: Regression tests of efficiency (F)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
F1	5.781 (0.021)	1.100 (0.301)	8.961 (0.005)	2.432 (0.127)
F2	1.447 (0.237)	2.028 (0.163)	2.905 (0.097)	3.029 (0.090)
F3	0.951 (0.335)	5.596 (0.023)	3.420 (0.072)	6.715 (0.013)
F4	4.907 (0.013)	3.809 (0.031)	9.469 (0.000)	5.971 (0.006)
F5	1.459 (0.247)	1.362 (0.270)	1.455 (0.248)	2.025 (0.148)
F6	2.365 (0.107)	4.307 (0.020)	4.765 (0.014)	6.413 (0.004)
F7	2.954 (0.064)	1.127 (0.334)	5.598 (0.007)	2.028 (0.145)
F8	17.397 (0.000)	5.678 (0.007)	6.905 (0.003)	3.625 (0.037)
F9	4.672 (0.015)	6.828 (0.003)	3.516 (0.039)	6.003 (0.005)
F10	12.002 (0.000)	3.780 (0.019)	5.521 (0.003)	2.993 (0.044)
F11	3.339 (0.029)	3.862 (0.017)	4.041 (0.014)	4.385 (0.010)
F12	4.784 (0.007)	3.010 (0.044)	4.746 (0.007)	3.417 (0.028)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Table A19: GMM estimates of the loss function - CEE (F)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
F1	1	0.392 (0.077)	2.688 (0.101)	0.218 (0.067)	0.792 (0.374)
	1.5	0.400 (0.091)	3.152 (0.076)	0.154 (0.058)	1.486 (0.223)
	2	0.402 (0.103)	3.839 (0.050)	0.137 (0.076)	2.116 (0.146)
F2	1	0.457 (0.085)	1.599 (0.206)	0.313 (0.074)	9.577 (0.002)
	1.5	0.490 (0.103)	2.901 (0.089)	0.297 (0.076)	7.720 (0.005)
	2	0.513 (0.115)	3.376 (0.066)	0.296 (0.085)	6.002 (0.014)
F3	1	0.431 (0.077)	2.526 (0.112)	0.247 (0.072)	2.179 (0.140)
	1.5	0.427 (0.089)	3.047 (0.081)	0.210 (0.074)	1.411 (0.235)
	2	0.396 (0.097)	1.718 (0.190)	0.201 (0.084)	0.551 (0.458)
F4	1	0.338 (0.076)	6.022 (0.049)	0.127 (0.058)	5.383 (0.068)
	1.5	0.324 (0.088)	5.519 (0.063)	0.093 (0.045)	3.827 (0.148)
	2	0.297 (0.094)	5.582 (0.061)	0.083 (0.066)	3.442 (0.179)
F5	1	0.438 (0.082)	4.604 (0.100)	0.265 (0.074)	9.234 (0.010)
	1.5	0.510 (0.102)	6.367 (0.041)	0.028 (0.050)	7.477 (0.024)
	2	0.554 (0.113)	5.989 (0.050)	-0.002 (0.058)	6.532 (0.038)
F6	1	0.420 (0.076)	5.170 (0.075)	0.209 (0.069)	3.724 (0.155)
	1.5	0.303 (0.083)	7.140 (0.028)	0.083 (0.051)	5.379 (0.068)
	2	0.269 (0.089)	5.957 (0.051)	0.024 (0.040)	4.981 (0.083)
F7	1	0.404 (0.076)	4.230 (0.121)	0.224 (0.065)	1.186 (0.553)
	1.5	0.440 (0.088)	3.749 (0.153)	0.138 (0.058)	2.822 (0.244)
	2	0.469 (0.101)	4.064 (0.131)	0.076 (0.054)	3.867 (0.145)
F8	1	0.549 (0.087)	6.091 (0.048)	0.409 (0.076)	15.768 (0.000)
	1.5	0.613 (0.100)	6.929 (0.031)	0.370 (0.081)	11.403 (0.003)
	2	0.601 (0.109)	5.853 (0.054)	0.343 (0.088)	8.462 (0.015)
F9	1	0.576 (0.074)	10.301 (0.006)	0.331 (0.073)	8.837 (0.012)
	1.5	0.658 (0.077)	11.853 (0.003)	0.320 (0.074)	8.367 (0.015)
	2	0.676 (0.097)	11.346 (0.003)	0.306 (0.081)	7.375 (0.025)
F10	1	0.512 (0.085)	7.834 (0.050)	0.009 (0.024)	9.897 (0.019)
	1.5	0.558 (0.106)	7.780 (0.051)	0.000 (0.030)	8.999 (0.029)
	2	0.585 (0.116)	7.190 (0.066)	-0.023 (0.038)	6.655 (0.084)
F11	1	0.489 (0.078)	8.755 (0.033)	0.092 (0.049)	7.206 (0.066)
	1.5	0.465 (0.091)	7.107 (0.069)	0.123 (0.055)	3.565 (0.312)
	2	0.338 (0.096)	5.805 (0.121)	0.137 (0.078)	2.819 (0.420)
F12	1	0.635 (0.073)	8.377 (0.039)	0.415 (0.076)	15.831 (0.001)
	1.5	0.689 (0.073)	7.955 (0.047)	0.346 (0.073)	11.115 (0.011)
	2	0.681 (0.092)	6.847 (0.077)	0.314 (0.083)	8.540 (0.036)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A20: GMM estimates of the loss function - JF (F)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
F1	1	0.260 (0.070)	3.434 (0.064)	0.255 (0.079)	1.602 (0.206)
	1.5	0.298 (0.090)	2.719 (0.099)	0.142 (0.057)	2.476 (0.116)
	2	0.276 (0.117)	1.562 (0.211)	0.081 (0.042)	2.468 (0.116)
F2	1	0.352 (0.081)	2.058 (0.151)	0.310 (0.098)	1.445 (0.229)
	1.5	0.416 (0.109)	2.776 (0.096)	0.228 (0.089)	2.392 (0.122)
	2	0.460 (0.145)	2.349 (0.125)	0.170 (0.082)	2.208 (0.137)
F3	1	0.332 (0.072)	3.207 (0.073)	0.279 (0.079)	1.687 (0.194)
	1.5	0.385 (0.092)	4.079 (0.043)	0.206 (0.065)	3.691 (0.055)
	2	0.480 (0.140)	2.638 (0.104)	0.206 (0.070)	3.286 (0.070)
F4	1	0.238 (0.068)	4.321 (0.115)	0.210 (0.076)	3.559 (0.169)
	1.5	0.246 (0.083)	3.823 (0.148)	0.107 (0.047)	3.691 (0.158)
	2	0.207 (0.097)	2.607 (0.272)	0.978 (0.150)	42.194 (0.000)
F5	1	0.283 (0.074)	7.288 (0.026)	0.256 (0.088)	3.207 (0.201)
	1.5	0.112 (0.079)	9.809 (0.007)	0.046 (0.042)	5.754 (0.056)
	2	0.043 (0.067)	7.597 (0.022)	0.017 (0.024)	5.682 (0.058)
F6	1	0.286 (0.069)	6.528 (0.038)	0.257 (0.079)	2.468 (0.291)
	1.5	0.175 (0.071)	6.876 (0.032)	0.106 (0.049)	4.302 (0.116)
	2	0.075 (0.051)	5.843 (0.054)	0.048 (0.031)	3.993 (0.136)
F7	1	0.275 (0.071)	5.414 (0.067)	0.270 (0.079)	1.993 (0.369)
	1.5	0.322 (0.088)	3.170 (0.205)	0.145 (0.058)	3.837 (0.147)
	2	0.278 (0.114)	1.542 (0.462)	0.078 (0.042)	3.433 (0.180)
F8	1	0.437 (0.087)	6.597 (0.037)	0.370 (0.095)	4.958 (0.084)
	1.5	0.505 (0.107)	6.103 (0.047)	0.292 (0.088)	5.097 (0.078)
	2	0.482 (0.143)	4.719 (0.094)	0.210 (0.082)	3.843 (0.146)
F9	1	0.450 (0.077)	11.524 (0.003)	0.367 (0.082)	7.696 (0.021)
	1.5	0.524 (0.101)	11.206 (0.004)	0.299 (0.072)	9.299 (0.010)
	2	0.730 (0.166)	9.004 (0.011)	0.293 (0.073)	9.852 (0.007)
F10	1	0.319 (0.077)	10.438 (0.015)	0.244 (0.081)	4.701 (0.195)
	1.5	0.417 (0.102)	9.276 (0.026)	0.009 (0.020)	7.765 (0.051)
	2	0.419 (0.125)	9.474 (0.024)	0.002 (0.008)	6.736 (0.081)
F11	1	0.113 (0.049)	11.187 (0.011)	0.222 (0.076)	3.014 (0.389)
	1.5	0.171 (0.071)	7.144 (0.067)	0.094 (0.043)	4.266 (0.234)
	2	0.276 (0.110)	7.163 (0.067)	0.044 (0.024)	3.994 (0.262)
F12	1	0.512 (0.085)	9.895 (0.019)	0.464 (0.101)	7.362 (0.061)
	1.5	0.583 (0.114)	8.336 (0.040)	0.381 (0.100)	6.859 (0.077)
	2	0.682 (0.166)	7.829 (0.050)	0.303 (0.095)	5.892 (0.117)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A21: Regression tests of efficiency (G)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
G1	0.016 (0.901)	0.475 (0.495)	0.001 (0.972)	1.064 (0.309)
G2	0.015 (0.904)	0.220 (0.642)	0.073 (0.788)	0.198 (0.659)
G3	0.042 (0.840)	0.036 (0.850)	0.018 (0.894)	0.104 (0.749)
G4	0.460 (0.635)	2.032 (0.145)	1.274 (0.292)	4.416 (0.019)
G5	0.268 (0.767)	0.105 (0.901)	0.050 (0.951)	0.210 (0.811)
G6	0.428 (0.655)	0.106 (0.899)	0.061 (0.941)	0.050 (0.952)
G7	0.045 (0.956)	1.096 (0.345)	0.319 (0.729)	1.299 (0.285)
G8	0.971 (0.388)	0.115 (0.891)	0.271 (0.764)	0.645 (0.530)
G9	0.423 (0.658)	1.704 (0.195)	0.040 (0.961)	1.340 (0.274)
G10	1.589 (0.208)	0.818 (0.492)	0.235 (0.872)	0.848 (0.476)
G11	2.088 (0.118)	5.497 (0.003)	3.467 (0.026)	16.039 (0.000)
G12	0.204 (0.893)	1.053 (0.380)	1.213 (0.318)	2.811 (0.052)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Table A22: GMM estimates of the loss function - CEE (G)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
G1	1	0.427 (0.079)	0.400 (0.527)	0.261 (0.077)	0.062 (0.803)
	1.5	0.437 (0.093)	0.038 (0.845)	0.215 (0.075)	0.035 (0.852)
	2	0.430 (0.104)	0.007 (0.932)	0.177 (0.071)	0.203 (0.652)
G2	1	0.416 (0.077)	0.208 (0.648)	0.255 (0.074)	0.010 (0.921)
	1.5	0.425 (0.091)	0.015 (0.901)	0.213 (0.075)	0.040 (0.842)
	2	0.421 (0.104)	0.000 (0.994)	0.198 (0.084)	0.053 (0.819)
G3	1	0.416 (0.077)	0.230 (0.632)	0.249 (0.072)	0.261 (0.610)
	1.5	0.424 (0.091)	0.008 (0.930)	0.206 (0.073)	0.279 (0.597)
	2	0.422 (0.103)	0.006 (0.937)	0.186 (0.079)	0.553 (0.457)
G4	1	0.386 (0.081)	2.248 (0.325)	0.198 (0.073)	1.485 (0.476)
	1.5	0.439 (0.096)	1.790 (0.409)	0.169 (0.070)	1.415 (0.493)
	2	0.450 (0.102)	0.889 (0.641)	0.157 (0.071)	1.163 (0.559)
G5	1	0.414 (0.077)	0.203 (0.904)	0.250 (0.074)	0.225 (0.894)
	1.5	0.426 (0.092)	0.122 (0.941)	0.186 (0.069)	0.804 (0.669)
	2	0.420 (0.105)	0.422 (0.810)	0.133 (0.068)	1.615 (0.446)
G6	1	0.411 (0.077)	0.502 (0.778)	0.251 (0.072)	0.698 (0.705)
	1.5	0.419 (0.090)	0.126 (0.939)	0.208 (0.072)	0.513 (0.774)
	2	0.387 (0.098)	0.719 (0.698)	0.190 (0.075)	0.580 (0.748)
G7	1	0.428 (0.082)	0.444 (0.801)	0.256 (0.076)	0.618 (0.734)
	1.5	0.433 (0.092)	0.059 (0.971)	0.209 (0.072)	0.144 (0.931)
	2	0.424 (0.101)	0.056 (0.973)	0.157 (0.062)	0.936 (0.626)
G8	1	0.418 (0.077)	0.426 (0.808)	0.286 (0.075)	2.414 (0.299)
	1.5	0.414 (0.092)	1.559 (0.459)	0.200 (0.071)	2.591 (0.274)
	2	0.418 (0.105)	1.113 (0.573)	0.100 (0.057)	2.502 (0.286)
G9	1	0.418 (0.077)	0.292 (0.864)	0.308 (0.069)	5.005 (0.082)
	1.5	0.419 (0.086)	0.332 (0.847)	0.284 (0.070)	7.897 (0.019)
	2	0.413 (0.098)	0.703 (0.704)	0.027 (0.024)	4.215 (0.122)
G10	1	0.431 (0.081)	2.698 (0.441)	0.267 (0.079)	0.496 (0.920)
	1.5	0.452 (0.095)	4.646 (0.200)	0.210 (0.075)	0.228 (0.973)
	2	0.503 (0.106)	4.377 (0.224)	0.168 (0.069)	0.505 (0.918)
G11	1	0.372 (0.081)	5.599 (0.133)	0.254 (0.076)	0.421 (0.936)
	1.5	0.471 (0.095)	5.437 (0.142)	0.219 (0.073)	1.492 (0.684)
	2	0.514 (0.103)	5.090 (0.165)	0.180 (0.067)	2.462 (0.482)
G12	1	0.405 (0.078)	1.427 (0.699)	0.249 (0.073)	0.260 (0.967)
	1.5	0.430 (0.093)	1.076 (0.783)	0.196 (0.070)	0.716 (0.869)
	2	0.451 (0.105)	1.305 (0.728)	0.194 (0.078)	1.375 (0.711)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A23: GMM estimates of the loss function - JF (G)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
G1	1	0.327 (0.075)	0.977 (0.323)	0.302 (0.087)	0.660 (0.417)
	1.5	0.365 (0.096)	1.542 (0.214)	0.167 (0.065)	1.992 (0.158)
	2	0.455 (0.132)	1.526 (0.217)	0.075 (0.040)	2.711 (0.100)
G2	1	0.321 (0.074)	1.069 (0.301)	0.300 (0.084)	0.224 (0.636)
	1.5	0.321 (0.090)	1.156 (0.282)	0.190 (0.067)	0.890 (0.346)
	2	0.294 (0.117)	0.657 (0.418)	0.105 (0.047)	1.276 (0.259)
G3	1	0.309 (0.073)	1.006 (0.316)	0.275 (0.084)	2.000 (0.157)
	1.5	0.324 (0.089)	0.696 (0.404)	0.186 (0.069)	1.310 (0.252)
	2	0.283 (0.113)	0.796 (0.372)	0.121 (0.055)	1.108 (0.292)
G4	1	0.293 (0.076)	3.362 (0.186)	0.240 (0.089)	3.444 (0.179)
	1.5	0.439 (0.092)	2.833 (0.243)	0.092 (0.055)	4.654 (0.098)
	2	0.471 (0.097)	1.706 (0.426)	0.212 (0.067)	4.616 (0.099)
G5	1	0.317 (0.075)	1.206 (0.547)	0.302 (0.085)	0.511 (0.774)
	1.5	0.314 (0.090)	1.274 (0.529)	0.192 (0.068)	1.019 (0.601)
	2	0.278 (0.102)	0.718 (0.699)	0.109 (0.049)	1.185 (0.553)
G6	1	0.310 (0.073)	0.982 (0.612)	0.255 (0.081)	2.619 (0.270)
	1.5	0.317 (0.085)	0.740 (0.691)	0.175 (0.066)	1.577 (0.455)
	2	0.273 (0.102)	0.810 (0.667)	0.114 (0.051)	1.189 (0.552)
G7	1	0.345 (0.077)	1.767 (0.413)	0.343 (0.089)	2.243 (0.326)
	1.5	0.409 (0.096)	1.776 (0.411)	0.232 (0.071)	2.809 (0.246)
	2	0.433 (0.114)	1.481 (0.477)	0.143 (0.052)	3.344 (0.188)
G8	1	0.322 (0.074)	1.563 (0.458)	0.352 (0.086)	2.373 (0.305)
	1.5	0.319 (0.091)	1.356 (0.508)	0.223 (0.071)	2.447 (0.294)
	2	0.391 (0.115)	2.350 (0.309)	0.110 (0.048)	1.801 (0.406)
G9	1	0.321 (0.074)	1.210 (0.546)	0.292 (0.074)	2.365 (0.306)
	1.5	0.326 (0.080)	0.592 (0.744)	0.200 (0.060)	1.349 (0.509)
	2	0.285 (0.093)	0.807 (0.668)	0.132 (0.048)	0.900 (0.638)
G10	1	0.318 (0.077)	1.670 (0.644)	0.326 (0.088)	2.068 (0.558)
	1.5	0.308 (0.094)	2.246 (0.523)	0.194 (0.068)	2.531 (0.470)
	2	0.339 (0.120)	2.446 (0.485)	0.105 (0.046)	3.432 (0.330)
G11	1	0.307 (0.077)	1.953 (0.582)	0.271 (0.086)	2.026 (0.567)
	1.5	0.342 (0.093)	1.953 (0.582)	0.168 (0.064)	2.040 (0.564)
	2	0.354 (0.114)	2.776 (0.428)	0.098 (0.045)	2.915 (0.405)
G12	1	0.299 (0.073)	2.036 (0.565)	0.272 (0.084)	1.950 (0.583)
	1.5	0.314 (0.090)	1.819 (0.611)	0.163 (0.062)	1.790 (0.617)
	2	0.285 (0.113)	2.958 (0.398)	0.087 (0.041)	2.064 (0.559)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A24: Regression tests of efficiency (H)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
H1	25.041 (0.000)	5.093 (0.030)	17.805 (0.000)	9.993 (0.003)
H2	27.812 (0.000)	11.725 (0.001)	14.107 (0.001)	13.715 (0.001)
H3	14.204 (0.001)	0.201 (0.656)	14.525 (0.000)	0.848 (0.363)
H4	12.321 (0.000)	5.750 (0.007)	8.558 (0.001)	7.034 (0.003)
H5	11.791 (0.000)	5.958 (0.006)	8.650 (0.001)	6.343 (0.004)
H6	7.598 (0.002)	3.031 (0.060)	7.568 (0.002)	4.821 (0.014)
H7	18.799 (0.000)	3.608 (0.037)	8.520 (0.001)	7.072 (0.002)
H8	29.306 (0.000)	6.831 (0.003)	8.176 (0.001)	9.460 (0.000)
H9	15.025 (0.000)	0.888 (0.420)	9.154 (0.001)	2.118 (0.134)
H10	16.219 (0.000)	16.228 (0.000)	6.250 (0.001)	10.201 (0.000)
H11	10.575 (0.000)	10.206 (0.000)	5.872 (0.002)	7.054 (0.001)
H12	8.365 (0.000)	5.758 (0.002)	5.541 (0.003)	5.126 (0.004)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Table A25: GMM estimates of the loss function - CEE (H)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
H1	1	0.286 (0.074)	11.406 (0.001)	0.092 (0.052)	8.132 (0.004)
	1.5	0.226 (0.077)	11.168 (0.001)	0.061 (0.038)	5.167 (0.023)
	2	0.634 (0.104)	9.839 (0.002)	0.042 (0.028)	3.513 (0.061)
H2	1	0.267 (0.073)	11.221 (0.001)	0.112 (0.056)	6.700 (0.010)
	1.5	0.231 (0.076)	9.365 (0.002)	0.080 (0.045)	4.319 (0.038)
	2	0.211 (0.084)	6.728 (0.009)	0.055 (0.036)	3.101 (0.078)
H3	1	0.401 (0.076)	4.811 (0.028)	0.210 (0.067)	4.212 (0.040)
	1.5	0.459 (0.092)	6.363 (0.012)	0.085 (0.049)	5.062 (0.024)
	2	0.530 (0.104)	5.183 (0.023)	0.044 (0.041)	3.794 (0.051)
H4	1	0.249 (0.073)	12.051 (0.002)	0.095 (0.054)	7.766 (0.021)
	1.5	0.204 (0.077)	11.254 (0.004)	0.058 (0.037)	5.914 (0.052)
	2	0.187 (0.085)	8.752 (0.013)	0.021 (0.021)	5.965 (0.051)
H5	1	0.260 (0.071)	10.969 (0.004)	0.113 (0.058)	6.893 (0.032)
	1.5	0.233 (0.079)	9.222 (0.010)	0.081 (0.045)	4.984 (0.083)
	2	0.223 (0.088)	6.838 (0.033)	0.053 (0.035)	4.626 (0.099)
H6	1	0.308 (0.074)	10.312 (0.006)	0.126 (0.058)	6.458 (0.040)
	1.5	0.261 (0.085)	11.630 (0.003)	0.072 (0.043)	5.167 (0.076)
	2	0.578 (0.108)	8.276 (0.016)	0.056 (0.036)	3.964 (0.138)
H7	1	0.366 (0.075)	17.440 (0.000)	0.029 (0.033)	10.549 (0.005)
	1.5	0.377 (0.089)	14.095 (0.001)	0.016 (0.015)	7.142 (0.028)
	2	0.343 (0.098)	9.987 (0.007)	0.007 (0.008)	5.544 (0.063)
H8	1	0.116 (0.055)	20.468 (0.000)	0.031 (0.030)	10.519 (0.005)
	1.5	0.430 (0.090)	19.006 (0.000)	0.021 (0.017)	6.796 (0.033)
	2	0.416 (0.100)	12.133 (0.002)	0.022 (0.016)	4.271 (0.118)
H9	1	0.430 (0.076)	5.272 (0.072)	0.234 (0.059)	8.911 (0.012)
	1.5	0.436 (0.088)	5.351 (0.069)	0.046 (0.028)	5.771 (0.056)
	2	0.434 (0.097)	5.006 (0.082)	0.031 (0.022)	3.846 (0.146)
H10	1	0.159 (0.058)	21.116 (0.000)	0.030 (0.032)	10.295 (0.016)
	1.5	0.350 (0.087)	17.511 (0.001)	0.016 (0.016)	7.029 (0.071)
	2	0.360 (0.098)	10.663 (0.014)	0.009 (0.009)	5.320 (0.150)
H11	1	0.248 (0.070)	14.640 (0.002)	0.043 (0.037)	10.157 (0.017)
	1.5	0.307 (0.081)	12.063 (0.007)	0.018 (0.015)	6.826 (0.078)
	2	0.307 (0.091)	9.070 (0.028)	0.006 (0.005)	5.471 (0.140)
H12	1	0.221 (0.067)	15.749 (0.001)	0.065 (0.044)	8.911 (0.031)
	1.5	0.290 (0.081)	12.291 (0.006)	0.020 (0.021)	7.039 (0.071)
	2	0.309 (0.093)	9.314 (0.025)	0.004 (0.009)	5.965 (0.113)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A26: GMM estimates of the loss function - JF (H)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
H1	1	0.203 (0.065)	9.154 (0.002)	0.219 (0.077)	4.613 (0.032)
	1.5	0.184 (0.076)	7.244 (0.007)	0.087 (0.048)	5.641 (0.018)
	2	0.155 (0.086)	4.087 (0.043)	0.049 (0.033)	4.130 (0.042)
H2	1	0.213 (0.066)	7.276 (0.007)	0.230 (0.079)	3.807 (0.051)
	1.5	0.192 (0.078)	6.562 (0.010)	0.100 (0.052)	4.884 (0.027)
	2	0.157 (0.088)	4.097 (0.043)	0.055 (0.038)	3.694 (0.055)
H3	1	0.248 (0.066)	6.071 (0.014)	0.232 (0.077)	4.616 (0.032)
	1.5	0.236 (0.080)	6.335 (0.012)	0.080 (0.046)	6.378 (0.012)
	2	0.163 (0.085)	3.867 (0.049)	0.048 (0.033)	4.323 (0.038)
H4	1	0.186 (0.063)	9.168 (0.010)	0.214 (0.080)	4.632 (0.099)
	1.5	0.182 (0.079)	7.009 (0.030)	0.088 (0.049)	5.587 (0.061)
	2	0.155 (0.087)	4.060 (0.131)	0.049 (0.033)	4.232 (0.120)
H5	1	0.203 (0.065)	7.730 (0.021)	0.231 (0.081)	3.724 (0.155)
	1.5	0.189 (0.080)	6.441 (0.040)	0.099 (0.054)	4.752 (0.093)
	2	0.165 (0.090)	4.932 (0.085)	0.048 (0.038)	3.881 (0.144)
H6	1	0.231 (0.067)	6.896 (0.032)	0.211 (0.078)	5.348 (0.069)
	1.5	0.223 (0.082)	6.612 (0.037)	0.070 (0.044)	7.109 (0.029)
	2	0.168 (0.087)	4.039 (0.133)	0.043 (0.032)	5.422 (0.066)
H7	1	0.348 (0.066)	15.617 (0.000)	0.214 (0.071)	5.272 (0.072)
	1.5	0.293 (0.078)	10.318 (0.006)	0.040 (0.026)	8.064 (0.018)
	2	0.235 (0.089)	6.390 (0.041)	0.008 (0.009)	7.325 (0.026)
H8	1	0.147 (0.057)	11.987 (0.002)	0.227 (0.074)	4.408 (0.110)
	1.5	0.271 (0.077)	10.573 (0.005)	0.061 (0.034)	6.593 (0.037)
	2	0.188 (0.085)	6.313 (0.043)	0.016 (0.014)	5.780 (0.056)
H9	1	0.331 (0.071)	7.393 (0.025)	0.250 (0.074)	4.887 (0.087)
	1.5	0.314 (0.080)	6.010 (0.050)	0.172 (0.053)	9.487 (0.009)
	2	0.275 (0.097)	4.508 (0.105)	0.123 (0.042)	7.851 (0.020)
H10	1	0.083 (0.045)	13.999 (0.003)	0.220 (0.073)	5.239 (0.155)
	1.5	0.262 (0.077)	10.094 (0.018)	0.050 (0.029)	7.564 (0.056)
	2	0.191 (0.085)	6.160 (0.104)	0.012 (0.011)	6.045 (0.109)
H11	1	0.346 (0.069)	13.137 (0.004)	0.220 (0.072)	5.272 (0.153)
	1.5	0.273 (0.077)	9.754 (0.021)	0.026 (0.024)	9.807 (0.020)
	2	0.222 (0.089)	6.862 (0.076)	-0.001 (0.007)	8.947 (0.030)
H12	1	0.105 (0.049)	13.789 (0.003)	0.218 (0.074)	4.953 (0.175)
	1.5	0.223 (0.074)	8.939 (0.030)	0.052 (0.030)	7.517 (0.057)
	2	0.170 (0.083)	5.433 (0.143)	0.010 (0.009)	6.013 (0.111)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A27: Absolute number of rejections of the efficiency hypothesis

set	variable	CEE				JF			
		symmetric		asymmetric		symmetric		asymmetric	
		real-time	revised	real-time	revised	real-time	revised	real-time	revised
A	unbiasedness	0 (0)	1 (1)	-	-	0 (0)	1 (1)	-	-
	weak efficiency	0 (0)	3 (3)	1 (3)	0 (1)	3 (3)	1 (2)	2 (8)	0 (0)
B	*lf*	3 (3)	3 (3)	5 (7)	9 (9)	3 (3)	3 (3)	7 (8)	7 (9)
	ur	3 (3)	3 (3)	4 (6)	8 (9)	3 (3)	3 (3)	3 (3)	4 (6)
	rlc	1 (1)	1 (1)	0 (2)	1 (3)	1 (1)	1 (1)	0 (2)	0 (2)
C	*gdp*	3 (3)	3 (3)	7 (9)	9 (9)	3 (3)	3 (3)	5 (8)	8 (9)
	inv	3 (3)	3 (3)	0 (2)	9 (9)	3 (3)	3 (3)	0 (2)	6 (9)
	gov	0 (1)	0 (0)	3 (3)	0 (0)	1 (3)	0 (1)	0 (1)	0 (0)
D	*ord*	2 (3)	3 (3)	5 (8)	8 (9)	1 (3)	2 (3)	3 (6)	9 (9)
	bp	3 (3)	1 (1)	0 (0)	4 (7)	0 (0)	0 (0)	0 (0)	3 (6)
	rs	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)	1 (1)	0 (0)	0 (0)
E	*cpi*	0 (0)	0 (0)	0 (0)	0 (0)	2 (3)	1 (1)	0 (0)	0 (0)
	ppi	1 (2)	2 (2)	0 (0)	0 (0)	2 (3)	2 (3)	0 (0)	0 (0)
	nlc	2 (3)	2 (2)	0 (0)	6 (9)	3 (3)	1 (1)	1 (4)	0 (4)
F	*dax*	2 (3)	3 (3)	0 (1)	1 (5)	1 (1)	2 (2)	1 (1)	0 (3)
	ltr	1 (1)	1 (2)	9 (9)	3 (7)	1 (1)	0 (1)	0 (4)	5 (7)
	str	1 (1)	2 (3)	3 (5)	4 (7)	3 (3)	3 (3)	3 (5)	6 (8)
G	*ft*	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)	0 (3)	0 (0)
	xrt	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
	oil	0 (0)	0 (0)	1 (2)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
H	*bc*	3 (3)	3 (3)	5 (9)	9 (9)	3 (3)	3 (3)	5 (8)	8 (8)
	bs	3 (3)	3 (3)	5 (8)	9 (9)	3 (3)	3 (3)	2 (6)	8 (9)
	be	3 (3)	3 (3)	4 (7)	6 (9)	0 (1)	1 (1)	5 (8)	7 (7)

Note: Number of rejections of the efficiency hypothesis on a 5 (10) percent level, due to a high F-statistic (J-statistic) under symmetric (asymmetric) loss with 3 (9) tests for each variable. Unbiasedness was only tested once and only under symmetric loss.

Table A28: Regression tests of efficiency (I)

IV set	Real-Time		Revised	
	CEE	JF	CEE	JF
I1	2.708 (0.108)	3.409 (0.072)	4.281 (0.045)	4.721 (0.036)
I2	1.198 (0.280)	1.156 (0.289)	0.138 (0.712)	0.414 (0.524)
I3	7.728 (0.008)	0.523 (0.474)	3.642 (0.064)	0.475 (0.495)
I4	2.224 (0.122)	2.586 (0.089)	2.514 (0.095)	3.562 (0.038)
I5	19.917 (0.000)	3.497 (0.041)	8.849 (0.001)	3.094 (0.057)
I6	5.453 (0.008)	0.327 (0.723)	3.474 (0.041)	0.361 (0.699)
I7	14.943 (0.000)	14.239 (0.000)	15.075 (0.000)	19.861 (0.000)
I8	2.064 (0.141)	8.900 (0.001)	6.771 (0.003)	8.331 (0.001)
I9	7.003 (0.003)	0.433 (0.652)	3.169 (0.053)	0.610 (0.548)
I10	6.297 (0.001)	7.434 (0.001)	6.732 (0.001)	9.526 (0.000)
I11	11.329 (0.000)	12.808 (0.000)	11.821 (0.000)	17.164 (0.000)
I12	4.282 (0.011)	1.919 (0.143)	4.240 (0.011)	4.515 (0.009)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Table A29: GMM estimates of the loss function - CEE (I)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
I1	1	0.392 (0.079)	2.055 (0.152)	0.245 (0.076)	1.295 (0.255)
	1.5	0.353 (0.084)	2.609 (0.106)	0.151 (0.067)	2.321 (0.128)
	2	0.323 (0.091)	2.386 (0.122)	0.116 (0.074)	2.402 (0.121)
I2	1	0.424 (0.079)	1.290 (0.256)	0.264 (0.074)	0.005 (0.946)
	1.5	0.424 (0.094)	0.859 (0.354)	0.219 (0.077)	0.167 (0.683)
	2	0.396 (0.101)	0.716 (0.397)	0.183 (0.082)	0.418 (0.518)
I3	1	0.519 (0.080)	7.621 (0.006)	0.213 (0.066)	4.409 (0.036)
	1.5	0.534 (0.095)	5.372 (0.020)	0.137 (0.059)	5.191 (0.023)
	2	0.526 (0.107)	3.418 (0.064)	0.244 (0.078)	4.986 (0.026)
I4	1	0.366 (0.080)	3.272 (0.195)	0.206 (0.070)	1.787 (0.409)
	1.5	0.353 (0.089)	4.936 (0.085)	0.091 (0.046)	3.670 (0.160)
	2	0.348 (0.097)	4.628 (0.099)	0.046 (0.032)	3.545 (0.170)
I5	1	0.414 (0.083)	1.720 (0.423)	0.204 (0.069)	2.515 (0.284)
	1.5	0.486 (0.099)	3.255 (0.196)	0.136 (0.069)	3.389 (0.184)
	2	0.520 (0.106)	3.349 (0.187)	0.183 (0.082)	2.900 (0.235)
I6	1	0.512 (0.086)	8.467 (0.014)	0.124 (0.057)	5.473 (0.065)
	1.5	0.495 (0.094)	8.144 (0.017)	0.112 (0.058)	5.295 (0.071)
	2	0.470 (0.106)	5.938 (0.051)	0.227 (0.079)	5.091 (0.078)
I7	1	0.236 (0.076)	11.135 (0.004)	0.122 (0.058)	6.213 (0.045)
	1.5	0.217 (0.082)	12.631 (0.002)	0.081 (0.043)	4.280 (0.118)
	2	0.239 (0.089)	8.547 (0.014)	0.053 (0.038)	3.392 (0.183)
I8	1	0.430 (0.081)	1.241 (0.538)	0.264 (0.072)	1.035 (0.596)
	1.5	0.428 (0.087)	0.845 (0.655)	0.220 (0.078)	0.172 (0.917)
	2	0.372 (0.081)	0.854 (0.653)	0.184 (0.083)	0.459 (0.795)
I9	1	0.385 (0.079)	13.149 (0.001)	0.250 (0.068)	4.407 (0.110)
	1.5	0.341 (0.085)	10.522 (0.005)	0.200 (0.065)	5.128 (0.077)
	2	0.310 (0.097)	8.021 (0.018)	0.205 (0.076)	5.211 (0.074)
I10	1	0.354 (0.073)	4.776 (0.189)	0.168 (0.065)	3.772 (0.287)
	1.5	0.296 (0.067)	4.532 (0.209)	0.130 (0.062)	3.018 (0.389)
	2	0.247 (0.067)	4.440 (0.218)	0.101 (0.059)	2.997 (0.392)
I11	1	0.459 (0.080)	15.266 (0.002)	0.143 (0.058)	5.500 (0.139)
	1.5	0.149 (0.068)	14.730 (0.002)	0.068 (0.042)	5.592 (0.133)
	2	0.348 (0.084)	16.379 (0.001)	0.018 (0.045)	5.917 (0.116)
I12	1	0.504 (0.082)	8.777 (0.032)	0.214 (0.064)	4.240 (0.237)
	1.5	0.527 (0.089)	8.216 (0.042)	0.157 (0.062)	5.165 (0.160)
	2	0.537 (0.094)	9.632 (0.022)	0.207 (0.075)	5.547 (0.136)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A30: GMM estimates of the loss function - JF (I)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
I1	1	0.304 (0.078)	1.857 (0.173)	0.283 (0.088)	1.113 (0.291)
	1.5	0.287 (0.093)	2.031 (0.154)	0.166 (0.072)	1.689 (0.194)
	2	0.243 (0.110)	1.754 (0.185)	0.101 (0.054)	1.624 (0.203)
I2	1	0.330 (0.077)	0.741 (0.389)	0.312 (0.088)	0.511 (0.475)
	1.5	0.338 (0.097)	0.616 (0.433)	0.208 (0.077)	0.556 (0.456)
	2	0.291 (0.118)	0.639 (0.424)	0.133 (0.063)	0.660 (0.417)
I3	1	0.382 (0.073)	8.604 (0.003)	0.258 (0.074)	3.098 (0.078)
	1.5	0.413 (0.094)	6.084 (0.014)	0.098 (0.040)	4.048 (0.044)
	2	0.415 (0.130)	3.851 (0.050)	0.048 (0.028)	4.004 (0.045)
I4	1	0.260 (0.072)	2.871 (0.238)	0.227 (0.078)	3.043 (0.218)
	1.5	0.258 (0.088)	3.700 (0.157)	0.129 (0.059)	4.506 (0.105)
	2	0.224 (0.102)	2.845 (0.241)	0.085 (0.045)	3.477 (0.176)
I5	1	0.312 (0.081)	0.952 (0.621)	0.258 (0.093)	2.277 (0.320)
	1.5	0.353 (0.101)	2.076 (0.354)	0.185 (0.078)	2.145 (0.342)
	2	0.333 (0.126)	2.180 (0.336)	0.118 (0.061)	1.654 (0.437)
I6	1	0.369 (0.072)	11.542 (0.003)	0.192 (0.070)	4.094 (0.129)
	1.5	0.311 (0.082)	10.148 (0.006)	0.053 (0.027)	4.688 (0.096)
	2	0.278 (0.098)	7.720 (0.021)	0.021 (0.013)	4.352 (0.113)
I7	1	0.122 (0.074)	12.798 (0.002)	0.248 (0.087)	4.646 (0.098)
	1.5	0.780 (0.099)	19.633 (0.000)	0.054 (0.046)	6.057 (0.048)
	2	0.916 (0.109)	7.781 (0.020)	0.267 (0.077)	10.477 (0.005)
I8	1	0.334 (0.078)	0.727 (0.695)	0.310 (0.087)	0.499 (0.779)
	1.5	0.326 (0.092)	0.756 (0.685)	0.190 (0.067)	0.933 (0.627)
	2	0.284 (0.115)	1.723 (0.423)	0.103 (0.048)	1.686 (0.431)
I9	1	0.224 (0.064)	10.024 (0.007)	0.293 (0.072)	3.155 (0.206)
	1.5	0.234 (0.077)	7.689 (0.021)	0.145 (0.045)	4.650 (0.098)
	2	0.217 (0.097)	6.045 (0.049)	0.646 (0.116)	24.281 (0.000)
I10	1	0.225 (0.058)	4.611 (0.203)	0.184 (0.065)	4.685 (0.196)
	1.5	0.199 (0.067)	4.588 (0.205)	0.091 (0.038)	3.741 (0.291)
	2	0.393 (0.128)	4.651 (0.199)	0.050 (0.025)	3.998 (0.262)
I11	1	0.122 (0.057)	10.376 (0.016)	0.225 (0.068)	3.442 (0.328)
	1.5	0.116 (0.068)	9.321 (0.025)	0.080 (0.034)	4.210 (0.240)
	2	0.082 (0.069)	8.806 (0.032)	0.037 (0.020)	4.471 (0.215)
I12	1	0.330 (0.070)	9.586 (0.022)	0.253 (0.060)	2.962 (0.398)
	1.5	0.217 (0.065)	7.731 (0.052)	0.117 (0.034)	4.143 (0.246)
	2	0.240 (0.084)	5.393 (0.145)	0.047 (0.022)	4.127 (0.248)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A31: Regression tests of efficiency (J)

	Real-Time		Revised	
IV set	CEE	JF	CEE	JF
J1	26.278 (0.000)	28.669 (0.000)	8.131 (0.007)	10.795 (0.002)
J2	0.772 (0.385)	0.071 (0.791)	0.146 (0.704)	0.009 (0.927)
J3	0.530 (0.471)	0.460 (0.502)	0.814 (0.372)	0.308 (0.582)
J4	12.388 (0.000)	23.226 (0.000)	5.242 (0.010)	11.007 (0.000)
J5	5.079 (0.011)	0.204 (0.817)	8.071 (0.001)	2.783 (0.075)
J6	0.426 (0.656)	0.259 (0.773)	1.018 (0.371)	0.705 (0.500)
J7	19.530 (0.000)	28.702 (0.000)	11.045 (0.000)	8.063 (0.001)
J8	0.536 (0.590)	0.678 (0.514)	2.393 (0.105)	4.757 (0.014)
J9	1.860 (0.170)	3.066 (0.058)	12.069 (0.000)	28.912 (0.000)
J10	14.030 (0.000)	9.836 (0.000)	6.592 (0.001)	9.616 (0.000)
J11	13.742 (0.000)	28.604 (0.000)	12.505 (0.000)	17.890 (0.000)
J12	0.766 (0.520)	0.976 (0.415)	1.313 (0.285)	0.838 (0.482)

Note: F-statistics with HAC covariance matrix with p-values in parentheses

Table A32: GMM estimates of the loss function - CEE (J)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
J1	1	0.357 (0.078)	3.588 (0.058)	0.220 (0.074)	2.208 (0.137)
	1.5	0.354 (0.086)	2.685 (0.101)	0.126 (0.062)	3.031 (0.082)
	2	0.334 (0.094)	2.095 (0.148)	0.088 (0.066)	3.003 (0.083)
J2	1	0.445 (0.081)	0.651 (0.420)	0.215 (0.067)	1.157 (0.282)
	1.5	0.438 (0.092)	0.530 (0.467)	0.197 (0.072)	0.433 (0.511)
	2	0.416 (0.105)	0.583 (0.445)	0.190 (0.083)	0.185 (0.667)
J3	1	0.499 (0.079)	4.412 (0.036)	0.249 (0.077)	1.750 (0.186)
	1.5	0.493 (0.079)	2.288 (0.130)	0.173 (0.071)	1.688 (0.194)
	2	0.465 (0.089)	0.663 (0.415)	0.121 (0.070)	1.834 (0.176)
J4	1	0.340 (0.071)	3.554 (0.169)	0.190 (0.070)	2.156 (0.340)
	1.5	0.330 (0.086)	3.343 (0.188)	0.087 (0.043)	3.651 (0.161)
	2	0.319 (0.097)	3.049 (0.218)	0.049 (0.030)	3.416 (0.181)
J5	1	0.427 (0.082)	0.743 (0.690)	0.155 (0.060)	3.024 (0.220)
	1.5	0.457 (0.087)	0.855 (0.652)	0.096 (0.055)	3.770 (0.152)
	2	0.470 (0.100)	1.275 (0.529)	0.056 (0.063)	4.296 (0.117)
J6	1	0.501 (0.082)	4.575 (0.101)	0.230 (0.079)	1.610 (0.447)
	1.5	0.473 (0.083)	2.364 (0.307)	0.166 (0.072)	1.628 (0.443)
	2	0.460 (0.091)	0.712 (0.701)	0.119 (0.062)	1.784 (0.410)
J7	1	0.273 (0.071)	7.414 (0.025)	0.123 (0.057)	5.522 (0.063)
	1.5	0.254 (0.083)	8.602 (0.014)	0.078 (0.039)	4.246 (0.120)
	2	0.234 (0.092)	7.589 (0.022)	0.046 (0.030)	3.538 (0.170)
J8	1	0.454 (0.081)	0.723 (0.697)	0.210 (0.068)	1.350 (0.509)
	1.5	0.460 (0.076)	0.593 (0.743)	0.161 (0.066)	1.632 (0.442)
	2	0.433 (0.083)	0.607 (0.738)	0.143 (0.082)	2.171 (0.338)
J9	1	0.502 (0.080)	4.484 (0.106)	0.263 (0.075)	1.876 (0.391)
	1.5	0.564 (0.069)	3.721 (0.156)	0.169 (0.069)	1.702 (0.427)
	2	0.554 (0.078)	3.619 (0.164)	0.120 (0.059)	1.782 (0.410)
J10	1	0.345 (0.064)	3.979 (0.264)	0.134 (0.061)	4.872 (0.181)
	1.5	0.306 (0.075)	5.481 (0.140)	0.084 (0.046)	4.326 (0.228)
	2	0.261 (0.082)	6.049 (0.109)	0.050 (0.042)	4.216 (0.239)
J11	1	0.593 (0.073)	20.929 (0.000)	0.204 (0.070)	2.600 (0.457)
	1.5	0.677 (0.067)	19.269 (0.000)	0.105 (0.049)	3.364 (0.339)
	2	0.690 (0.084)	16.327 (0.001)	0.060 (0.040)	3.829 (0.281)
J12	1	0.530 (0.072)	5.720 (0.126)	0.116 (0.054)	5.666 (0.129)
	1.5	0.612 (0.062)	9.062 (0.028)	0.079 (0.042)	4.111 (0.250)
	2	0.587 (0.072)	11.793 (0.008)	0.048 (0.034)	3.976 (0.264)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Table A33: GMM estimates of the loss function - JF (J)

		Real-Time		Revised	
IV set	p	$\hat{\alpha}$	J	$\hat{\alpha}$	J
J1	1	0.286 (0.078)	2.603 (0.107)	0.270 (0.089)	1.491 (0.222)
	1.5	0.282 (0.092)	2.075 (0.150)	0.160 (0.070)	1.858 (0.173)
	2	0.241 (0.107)	1.720 (0.190)	0.097 (0.052)	1.740 (0.187)
J2	1	0.340 (0.079)	0.325 (0.568)	0.309 (0.088)	0.000 (0.989)
	1.5	0.350 (0.098)	0.209 (0.648)	0.221 (0.079)	0.063 (0.801)
	2	0.320 (0.117)	0.158 (0.691)	0.148 (0.067)	0.241 (0.623)
J3	1	0.327 (0.079)	1.788 (0.181)	0.308 (0.088)	0.074 (0.786)
	1.5	0.377 (0.095)	0.738 (0.390)	0.206 (0.076)	0.565 (0.452)
	2	0.369 (0.117)	0.295 (0.587)	0.129 (0.061)	0.911 (0.340)
J4	1	0.243 (0.061)	3.312 (0.191)	0.211 (0.078)	3.020 (0.221)
	1.5	0.246 (0.079)	2.624 (0.269)	0.124 (0.062)	3.483 (0.175)
	2	0.214 (0.098)	1.985 (0.371)	0.082 (0.048)	2.541 (0.281)
J5	1	0.310 (0.076)	0.515 (0.773)	0.204 (0.084)	2.757 (0.252)
	1.5	0.325 (0.093)	0.273 (0.872)	0.140 (0.069)	2.293 (0.318)
	2	0.292 (0.116)	0.258 (0.879)	0.109 (0.058)	1.467 (0.480)
J6	1	0.302 (0.082)	1.819 (0.403)	0.262 (0.087)	0.838 (0.658)
	1.5	0.370 (0.097)	1.138 (0.566)	0.158 (0.066)	1.288 (0.525)
	2	0.428 (0.112)	1.184 (0.553)	0.091 (0.048)	1.453 (0.484)
J7	1	0.555 (0.079)	30.159 (0.000)	0.230 (0.088)	3.135 (0.209)
	1.5	0.699 (0.095)	21.345 (0.000)	0.083 (0.050)	4.928 (0.085)
	2	0.970 (0.025)	8.084 (0.018)	0.029 (0.025)	5.345 (0.069)
J8	1	0.350 (0.077)	0.396 (0.820)	0.261 (0.081)	1.670 (0.434)
	1.5	0.352 (0.086)	0.201 (0.904)	0.120 (0.050)	2.938 (0.230)
	2	0.304 (0.110)	0.285 (0.867)	0.066 (0.034)	2.968 (0.227)
J9	1	0.305 (0.075)	1.957 (0.376)	0.282 (0.095)	0.958 (0.619)
	1.5	0.263 (0.080)	2.798 (0.247)	0.184 (0.074)	1.067 (0.587)
	2	0.566 (0.103)	5.594 (0.061)	0.115 (0.057)	1.136 (0.567)
J10	1	0.267 (0.063)	3.744 (0.290)	0.133 (0.062)	5.733 (0.125)
	1.5	0.212 (0.080)	6.614 (0.085)	0.078 (0.039)	4.383 (0.223)
	2	0.935 (0.045)	12.685 (0.005)	0.588 (0.114)	20.777 (0.000)
J11	1	0.526 (0.078)	25.447 (0.000)	0.268 (0.089)	1.517 (0.678)
	1.5	0.572 (0.076)	15.116 (0.002)	0.121 (0.057)	3.437 (0.329)
	2	0.597 (0.085)	8.329 (0.040)	0.040 (0.028)	4.449 (0.217)
J12	1	0.301 (0.064)	8.484 (0.037)	0.275 (0.089)	3.873 (0.276)
	1.5	0.609 (0.070)	19.325 (0.000)	0.190 (0.075)	2.788 (0.425)
	2	0.648 (0.087)	12.399 (0.006)	0.129 (0.061)	1.353 (0.717)

Note: Shown in parentheses are standard errors for $\hat{\alpha}$ and p-values for the J-test statistics.

Appendix B

Figure B1: Fishbone plots - fat-tailed error term distribution - $t(5)$

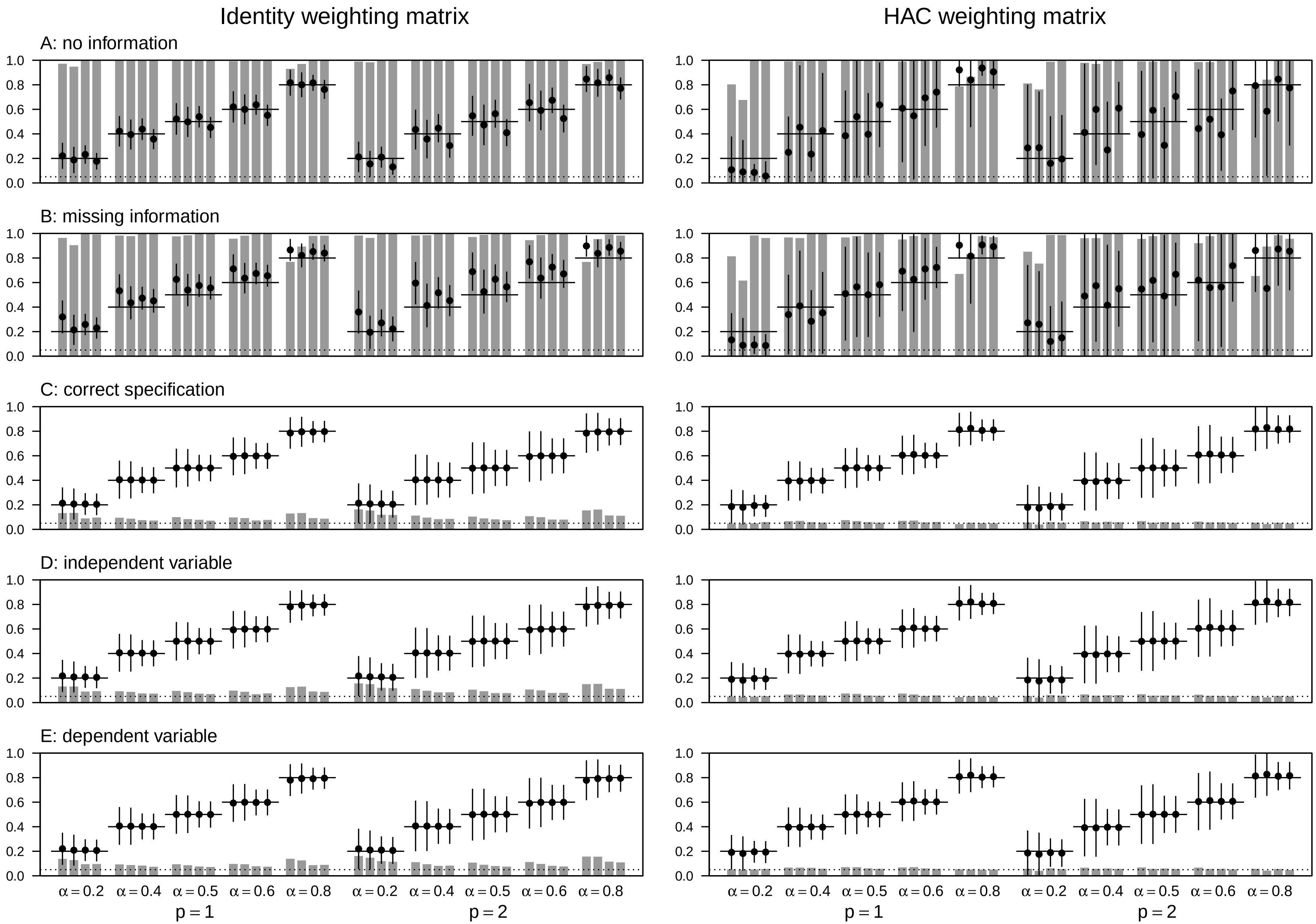


Figure B2: Fishbone plots - realistic scenario with fixed bandwidth I

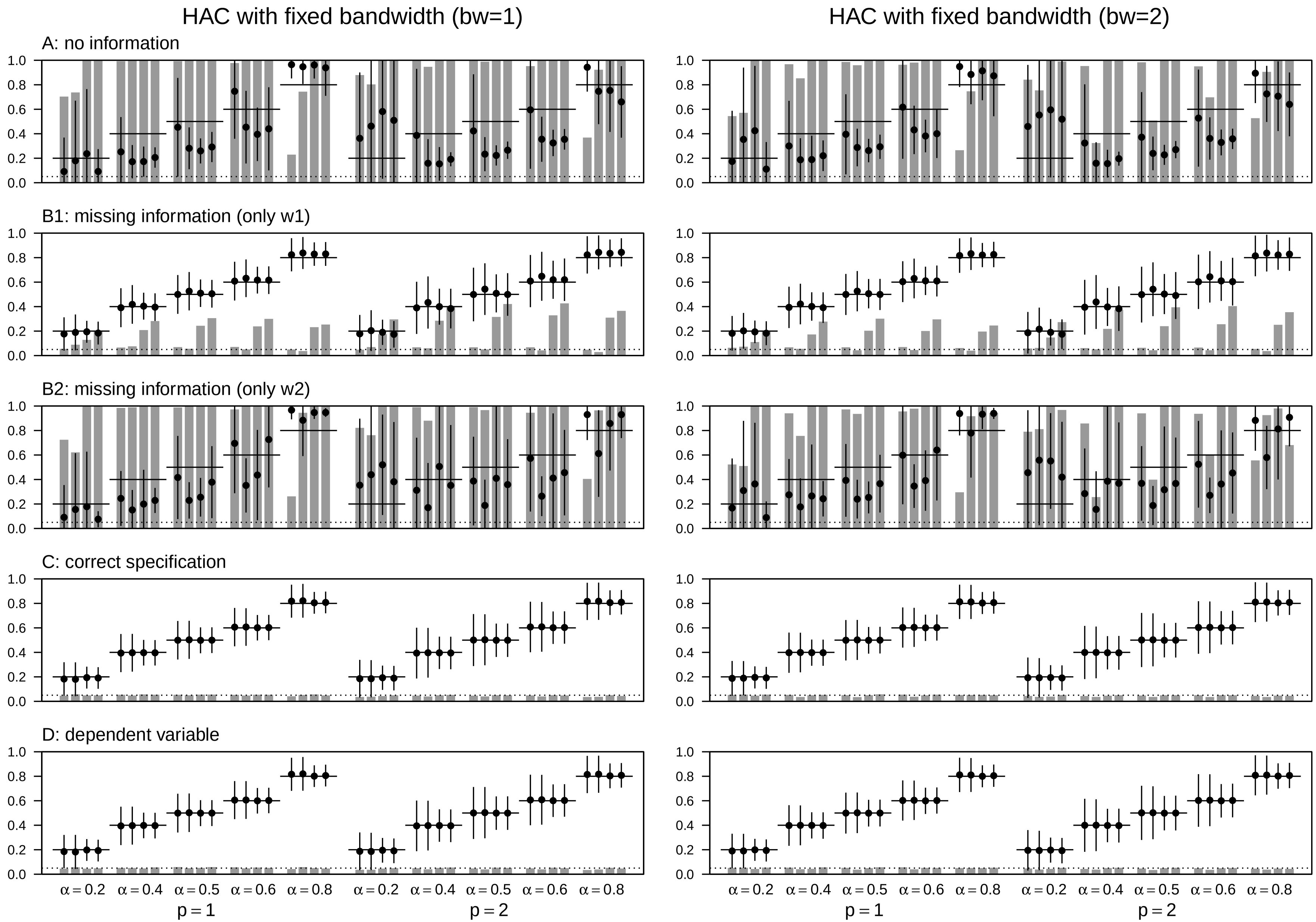


Figure B3: Fishbone plots - realistic scenario with fixed bandwidth II

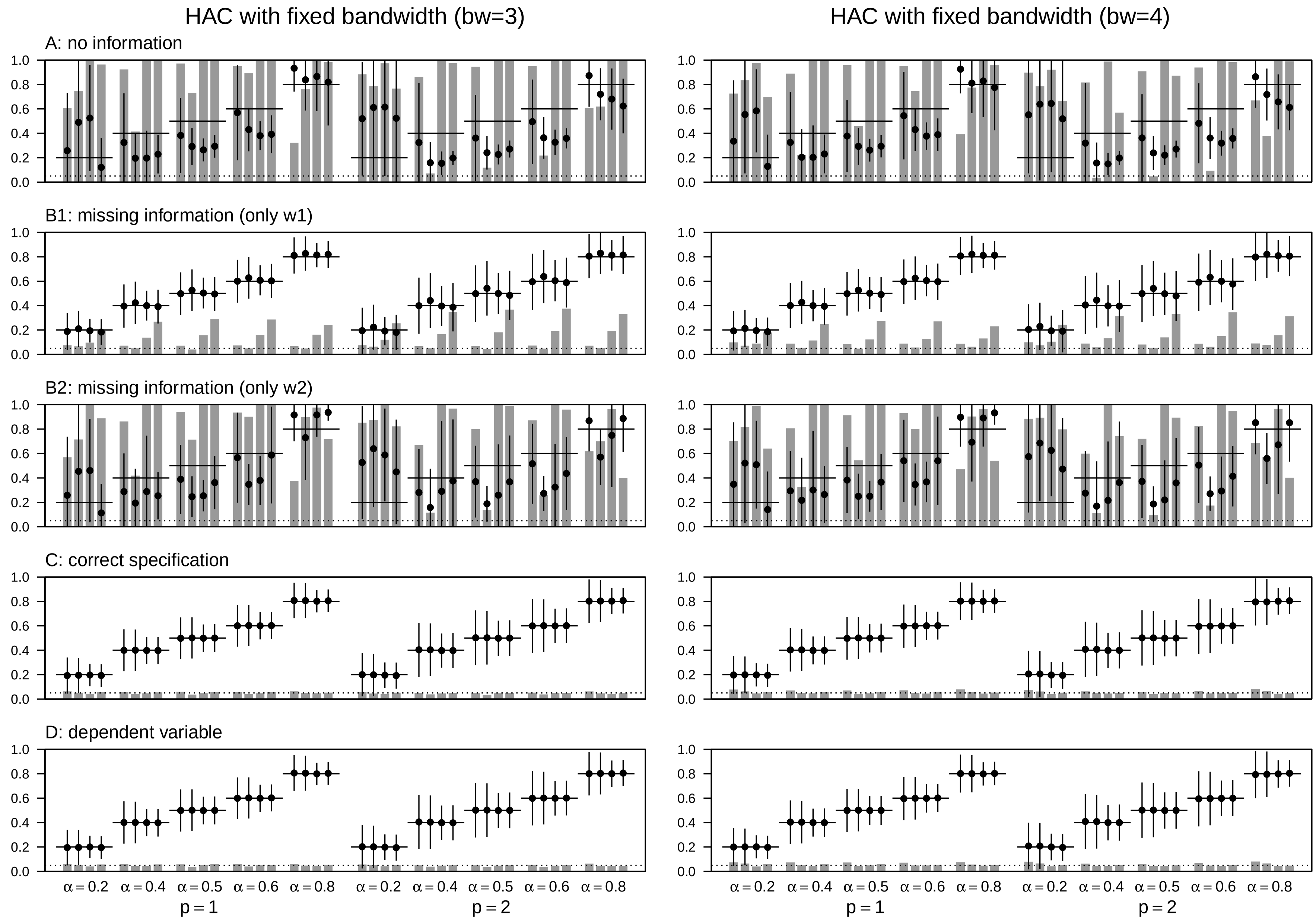


Figure B4: Fishbone plots - realistic scenario with outlier and fixed bandwidth I

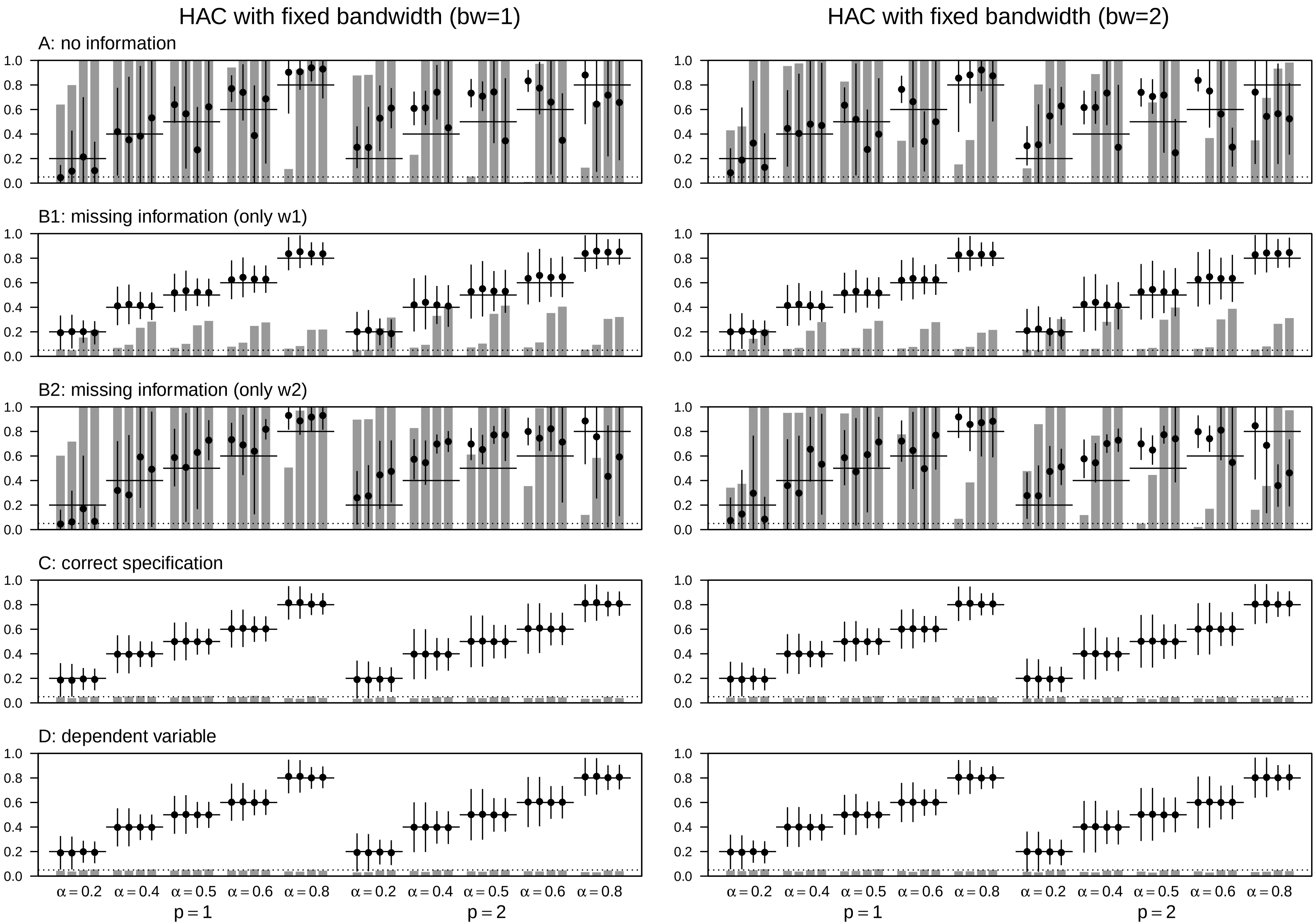
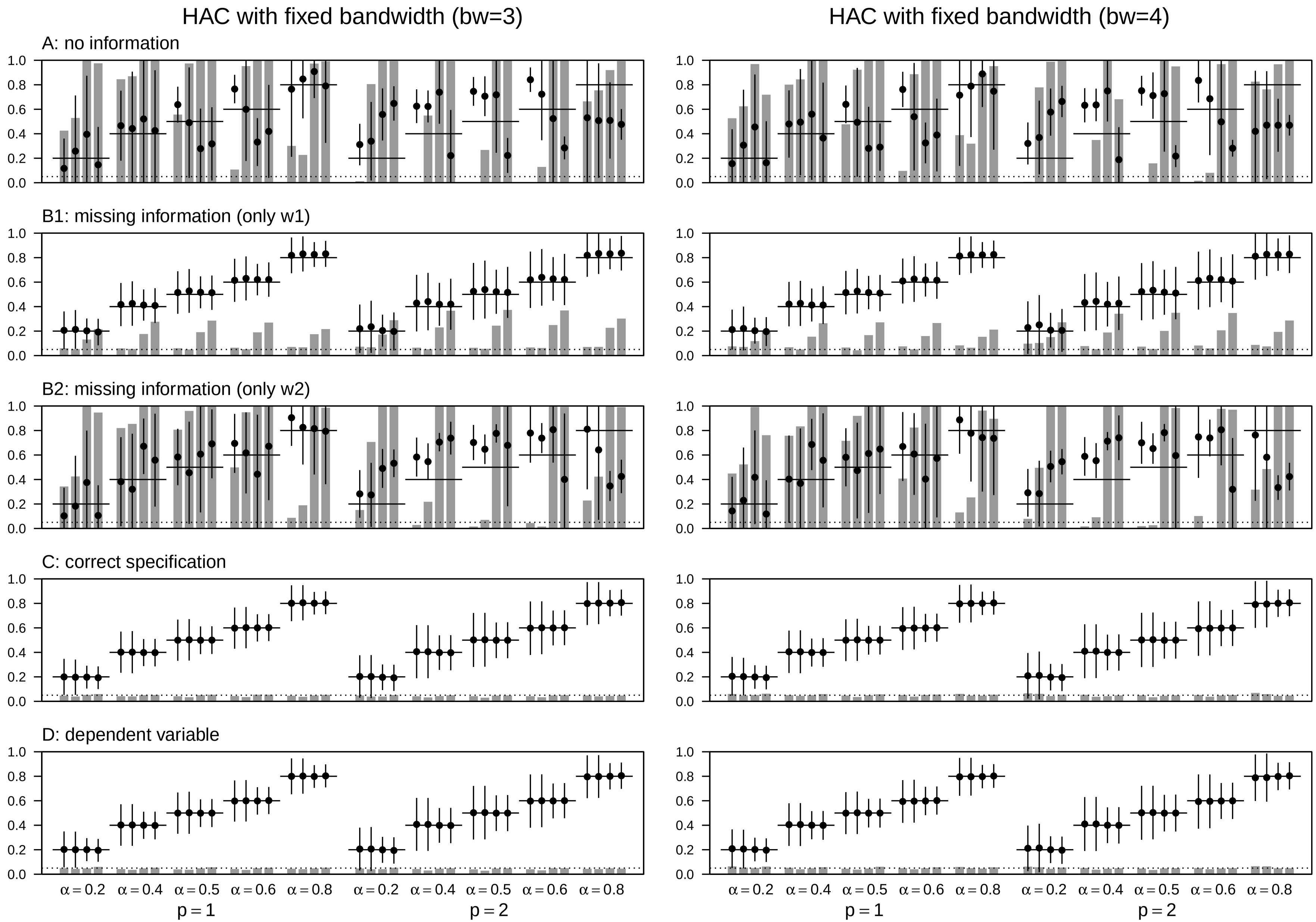


Figure B5: Fishbone plots - realistic scenario with outlier and fixed bandwidth II



Appendix C

Table C1: Mean and median forecast errors

Forecaster	GDP growth Median Real	GDP growth Median Revised	GDP growth *MAE* Real-time	GDP growth *MAE* Revised	HICP inflation Median Real	HICP inflation Median Revised	HICP inflation *MAE* Real-time	HICP inflation *MAE* Revised	Unemployment rate Median Real	Unemployment rate Median Revised	Unemployment rate *MAE* Real-time	Unemployment rate *MAE* Revised
aggregate	-0.430	-0.303	0.969	1.093	0.451	0.406	0.733	0.721	0.030	0.224	0.533	0.565
id 1	–	–	–	–	0.486	0.363	0.784	0.761	–	–	–	–
id 4	-0.565	-0.358	0.883	0.954	0.442	0.456	0.708	0.691	0.005	0.202	0.551	0.568
id 5	-0.528	-0.232	1.106	1.229	0.476	0.468	0.745	0.721	-0.002	0.156	0.648	0.638
id 7	-0.396	-0.272	1.016	1.107	0.723	0.684	0.901	0.888	0.047	0.002	0.588	0.576
id 14	-0.343	-0.232	0.938	1.070	0.566	0.549	0.858	0.848	-0.212	0.013	0.520	0.526
id 15	-0.585	-0.601	1.158	1.269	0.373	0.369	0.737	0.718	0.062	0.228	0.576	0.604
id 16	-0.202	-0.066	0.980	1.188	0.421	0.424	0.754	0.744	-0.012	0.203	0.609	0.610
id 20	-0.463	-0.260	0.985	1.104	0.436	0.424	0.735	0.723	-0.032	0.186	0.588	0.581
id 22	-0.326	-0.198	0.758	0.874	0.387	0.456	0.757	0.745	-0.244	-0.067	0.519	0.502
id 23	-0.315	-0.061	0.933	1.043	0.484	0.452	0.874	0.858	0.106	0.188	0.602	0.635
id 24	-0.399	-0.287	1.061	1.208	0.568	0.482	0.811	0.797	0.050	0.180	0.546	0.536
id 26	-0.402	-0.266	0.918	1.018	0.446	0.452	0.790	0.774	0.062	0.239	0.609	0.654
id 29	-0.734	-0.581	1.300	1.419	0.584	0.595	0.898	0.892	–	–	–	–
id 31	-0.534	-0.360	1.062	1.174	0.534	0.476	0.765	0.750	-0.044	0.210	0.621	0.623
id 33	-0.740	-0.486	1.135	1.273	0.113	0.130	0.613	0.605	0.195	0.406	0.651	0.710
id 37	-0.434	-0.142	0.883	0.980	0.342	0.286	0.657	0.647	0.064	0.249	0.474	0.538
id 38	-0.153	0.091	0.928	1.117	0.407	0.386	0.807	0.794	-0.072	0.057	0.645	0.644
id 39	-0.243	-0.142	0.847	0.971	0.284	0.269	0.684	0.662	-0.110	0.109	0.471	0.493
id 41	-0.448	-0.223	1.115	1.254	0.421	0.368	0.688	0.675	0.064	0.258	0.607	0.628
id 42	-0.371	-0.381	0.958	1.094	0.424	0.458	0.760	0.747	0.213	0.415	0.627	0.723
id 47	-0.020	0.204	0.961	1.222	0.264	0.272	0.732	0.735	-0.064	0.192	0.618	0.598
id 48	-0.529	-0.431	1.173	1.262	0.460	0.486	0.930	0.907	0.144	0.284	0.558	0.609
id 52	-0.460	-0.263	0.979	1.134	0.480	0.468	0.852	0.851	-0.212	0.172	0.635	0.652
id 54	-0.510	-0.352	1.071	1.154	0.423	0.440	0.748	0.735	0.023	0.081	0.636	0.612
id 56	-0.608	-0.386	1.143	1.282	0.576	0.558	0.753	0.728	0.162	0.302	0.596	0.594
id 82	0.012	0.014	0.896	1.140	0.459	0.458	0.867	0.863	–	–	–	–
id 85	-0.508	-0.295	0.789	0.842	0.407	0.407	0.663	0.643	–	–	–	–
id 89	-0.530	-0.329	1.080	1.199	0.428	0.391	0.750	0.734	0.285	0.544	0.590	0.715
id 90	-0.608	-0.283	1.181	1.310	0.429	0.423	0.740	0.724	0.061	0.280	0.557	0.616
id 93	–	–	–	–	0.266	0.249	0.715	0.699	–	–	–	–
id 94	-0.229	-0.169	0.948	1.062	0.236	0.210	0.812	0.796	-0.114	0.110	0.574	0.614
id 95	-0.353	-0.187	0.932	1.076	0.375	0.356	0.742	0.723	0.133	0.154	0.551	0.544
id 96	-0.459	-0.166	0.916	1.018	0.111	0.119	0.681	0.660	-0.016	0.157	0.502	0.534

Table C2: Estimates of α under Linex loss - GDP growth

	Real-time										Revised									
Forecaster	**Base**	**Equal Weights**			**Expon. Weights**			**Disagreement**			**Base**	**Equal Weights**			**Expon. Weights**			**Disagreement**		
		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4
aggregate	2.175	0.301	0.592	0.245	0.381	0.775	0.356	1.657	1.687	1.589	1.119	0.142	0.220	0.148	0.223	0.315	0.237	1.036	0.969	1.058
average	1.454	0.218	0.314	0.220	0.281	0.377	0.268	1.132	1.152	1.112	0.654	0.120	0.157	0.124	0.152	0.184	0.152	0.568	0.569	0.576
id 4	3.587	0.189	0.473	0.206	0.290	0.419	0.294	2.209	2.284	2.145	1.590	0.152	0.174	0.164	0.222	0.269	0.229	1.192	1.207	1.190
id 5	1.983	0.342	0.595	0.348	0.357	0.571	0.352	1.510	1.485	1.416	1.054	0.221	0.318	0.239	0.287	0.319	0.278	0.849	0.868	0.852
id 7	1.119	0.130	0.250	0.141	0.171	0.310	0.181	0.963	0.974	0.960	0.538	0.143	0.227	0.152	0.137	0.179	0.142	0.450	0.456	0.445
id 14	1.083	0.185	0.312	0.196	0.224	0.305	0.224	0.897	0.905	0.874	0.630	0.120	0.137	0.128	0.152	0.185	0.153	0.483	0.483	0.470
id 15	1.425	0.313	0.421	0.305	0.406	0.649	0.375	1.159	1.236	1.188	0.862	0.233	0.367	0.226	0.273	0.392	0.272	0.834	0.832	0.826
id 16	0.942	0.110	0.120	0.114	0.217	0.246	0.217	0.783	0.776	0.791	0.452	-0.020	-0.013	-0.014	-0.053	-0.057	-0.040	0.318	0.309	0.363
id 20	1.274	0.259	0.358	0.274	0.334	0.424	0.331	0.992	1.012	0.980	0.688	0.139	0.162	0.150	0.168	0.183	0.166	0.583	0.581	0.581
id 22	1.519	0.169	0.194	0.180	0.236	0.287	0.241	1.182	1.197	1.150	0.337	0.042	0.046	0.045	0.042	0.046	0.045	0.222	0.224	0.204
id 23	0.815	0.082	0.099	0.090	0.139	0.174	0.139	0.670	0.682	0.658	0.559	0.013	0.045	0.014	-0.001	-0.002	-0.002	0.486	0.471	0.506
id 24	1.493	0.166	0.217	0.177	0.299	0.398	0.301	1.253	1.297	1.213	0.821	0.106	0.115	0.112	0.145	0.246	0.153	0.707	0.718	0.707
id 26	1.137	0.236	0.339	0.254	0.310	0.322	0.307	0.991	0.997	0.985	0.399	0.125	0.144	0.133	0.154	0.167	0.158	0.355	0.356	0.356
id 29	1.605	0.297	0.454	0.302	0.306	0.759	0.323	1.218	1.237	1.216	0.744	0.170	0.190	0.177	0.266	0.314	0.267	0.651	0.654	0.652
id 31	1.663	0.388	0.575	0.380	0.475	0.888	0.410	1.470	1.481	1.468	0.846	0.236	0.321	0.249	0.296	0.326	0.286	0.759	0.762	0.756
id 33	1.653	0.391	0.458	0.346	0.432	0.491	0.391	1.382	1.401	1.369	0.886	0.264	0.308	0.264	0.307	0.338	0.289	0.774	0.761	0.776
id 37	2.178	0.166	0.199	0.176	0.280	0.337	0.293	1.578	1.604	1.472	1.152	0.152	0.150	0.134	0.204	0.230	0.210	0.991	0.999	0.992
id 38	0.617	0.003	0.006	0.005	0.069	0.044	0.034	0.464	0.447	0.464	-0.013	-0.047	-0.045	-0.045	-0.062	-0.058	-0.058	-0.006	-0.009	-0.199
id 39	1.445	0.100	0.154	0.099	0.198	0.290	0.197	0.995	1.149	0.971	0.202	0.025	0.029	0.028	0.049	0.049	0.049	0.280	0.246	0.264
id 41	1.467	0.405	0.364	0.386	0.516	0.466	0.372	1.305	1.218	1.274	0.571	0.235	0.234	0.255	0.239	0.247	0.232	0.677	0.690	0.673
id 42	1.967	0.311	0.763	0.283	0.380	0.432	0.401	1.486	1.533	1.445	0.841	0.153	0.340	0.158	0.187	0.242	0.196	0.689	0.705	0.684
id 47	0.454	0.045	0.050	0.048	0.047	0.054	0.054	0.330	0.330	0.330	0.005	-0.026	-0.031	-0.032	-0.034	-0.040	-0.042	0.005	0.005	0.005
id 48	1.039	0.173	0.230	0.180	0.187	0.275	0.191	0.871	0.880	0.871	0.623	0.174	0.363	0.185	0.255	0.335	0.251	0.565	0.568	0.566
id 52	1.285	0.197	0.221	0.210	0.268	0.326	0.274	1.047	1.035	0.994	0.450	0.065	0.073	0.070	0.110	0.109	0.111	0.490	0.491	0.498
id 54	1.891	0.293	0.449	0.304	0.360	0.431	0.356	1.478	1.501	1.457	0.758	0.155	0.172	0.157	0.228	0.244	0.224	0.662	0.658	0.661
id 56	2.117	0.460	0.493	0.432	0.572	0.668	0.514	1.689	1.710	1.664	1.045	0.304	0.322	0.303	0.346	0.427	0.326	0.920	0.921	0.916
id 82	0.004	-0.029	-0.026	-0.027	-0.053	-0.049	-0.048	-0.102	-0.031	-0.086	-0.244	-0.086	-0.082	-0.082	-0.154	-0.226	-0.126	-0.349	-0.171	-0.189
id 85	1.180	0.150	0.334	0.164	0.229	0.312	0.238	1.019	0.994	1.074	0.582	0.121	0.124	0.136	0.179	0.257	0.192	0.558	0.550	0.527
id 89	1.421	0.269	0.396	0.256	0.310	0.542	0.282	1.260	1.258	1.246	0.786	0.177	0.208	0.172	0.202	0.309	0.212	0.753	0.662	1.159
id 90	1.281	0.463	0.509	0.482	0.365	0.329	0.274	1.011	1.105	0.993	0.454	0.148	0.141	0.138	0.136	0.156	0.122	0.396	0.395	0.393
id 94	1.258	0.008	0.039	0.030	0.122	0.131	0.130	0.859	0.838	0.847	0.586	0.061	0.132	0.063	0.106	0.157	0.102	0.450	0.452	0.445
id 95	2.167	0.194	0.282	0.191	0.289	0.398	0.296	1.528	1.586	1.454	1.019	0.093	0.097	0.099	0.156	0.195	0.172	0.848	0.801	0.820
id 96	1.827	0.280	0.412	0.286	0.371	0.466	0.362	1.520	1.527	1.501	0.911	0.076	0.086	0.081	0.160	0.154	0.147	0.914	0.864	0.860

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets.

Table C3: Estimates of α under Linex-Linex loss - GDP growth

	Real-time										Revised									
Forecaster	**Base**	**Equal Weights**			**Expon. Weights**			**Disagreement**			**Base**	**Equal Weights**			**Expon. Weights**			**Disagreement**		
		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4
aggregate	0.730	0.699	0.723	0.693	0.694	0.709	0.687	0.710	0.710	0.710	0.630	0.642	0.634	0.652	0.628	0.650	0.649	0.629	0.629	0.629
average	0.705	0.644	0.716	0.649	0.642	0.669	0.644	0.694	0.696	0.695	0.626	0.594	0.636	0.609	0.595	0.615	0.606	0.622	0.623	0.624
id 4	0.784	0.679	0.780	0.675	0.682	0.711	0.680	0.772	0.772	0.772	0.694	0.629	0.736	0.637	0.624	0.638	0.624	0.684	0.685	0.685
id 5	0.749	0.675	0.780	0.669	0.682	0.703	0.663	0.728	0.730	0.729	0.716	0.636	0.689	0.643	0.627	0.667	0.633	0.702	0.702	0.702
id 7	0.718	0.616	0.674	0.620	0.618	0.657	0.616	0.686	0.709	0.687	0.610	0.578	0.632	0.590	0.554	0.604	0.581	0.608	0.608	0.608
id 14	0.694	0.644	0.708	0.644	0.640	0.664	0.629	0.663	0.663	0.662	0.617	0.573	0.607	0.580	0.549	0.610	0.559	0.615	0.616	0.616
id 15	0.716	0.670	0.815	0.671	0.659	0.692	0.660	0.709	0.709	0.708	0.661	0.611	0.711	0.636	0.602	0.630	0.624	0.646	0.647	0.647
id 16	0.614	0.623	0.723	0.624	0.626	0.636	0.614	0.610	0.610	0.610	0.582	0.561	0.597	0.583	0.549	0.576	0.575	0.580	0.580	0.580
id 20	0.743	0.675	0.747	0.675	0.666	0.682	0.660	0.731	0.732	0.731	0.657	0.614	0.656	0.623	0.603	0.623	0.609	0.652	0.652	0.652
id 22	0.699	0.657	0.693	0.664	0.649	0.668	0.653	0.692	0.692	0.692	0.618	0.584	0.606	0.600	0.583	0.609	0.607	0.616	0.616	0.616
id 23	0.639	0.570	0.690	0.602	0.562	0.651	0.612	0.633	0.634	0.633	0.561	0.558	0.583	0.571	0.552	0.575	0.570	0.562	0.562	0.563
id 24	0.681	0.631	0.660	0.642	0.644	0.672	0.649	0.671	0.672	0.671	0.598	0.582	0.610	0.614	0.610	0.628	0.631	0.596	0.596	0.596
id 26	0.715	0.671	0.738	0.674	0.667	0.680	0.668	0.704	0.705	0.704	0.615	0.594	0.631	0.608	0.581	0.615	0.604	0.614	0.614	0.614
id 29	0.738	0.663	0.772	0.659	0.653	0.712	0.653	0.724	0.724	0.724	0.629	0.565	0.613	0.594	0.613	0.621	0.639	0.596	0.624	0.624
id 31	0.777	0.715	0.816	0.711	0.710	0.722	0.698	0.773	0.773	0.773	0.682	0.683	0.708	0.685	0.681	0.691	0.677	0.677	0.677	0.677
id 33	0.761	0.712	0.790	0.710	0.673	0.713	0.693	0.745	0.746	0.745	0.690	0.642	0.725	0.656	0.633	0.649	0.644	0.685	0.686	0.685
id 37	0.767	0.640	0.728	0.641	0.646	0.683	0.639	0.747	0.748	0.747	0.672	0.584	0.642	0.587	0.602	0.626	0.593	0.668	0.668	0.657
id 38	0.607	0.572	0.685	0.593	0.571	0.621	0.597	0.603	0.604	0.604	0.602	0.545	0.591	0.564	0.554	0.582	0.570	0.596	0.596	0.596
id 39	0.698	0.639	0.716	0.639	0.648	0.667	0.646	0.693	0.693	0.693	0.602	0.597	0.618	0.612	0.598	0.609	0.607	0.606	0.606	0.616
id 41	0.714	0.684	0.746	0.689	0.691	0.697	0.675	0.709	0.710	0.709	0.591	0.630	0.616	0.637	0.625	0.629	0.630	0.584	0.584	0.610
id 42	0.748	0.711	0.774	0.709	0.689	0.712	0.688	0.741	0.741	0.741	0.641	0.618	0.641	0.638	0.628	0.624	0.653	0.650	0.650	0.650
id 47	0.543	0.519	0.555	0.541	0.525	0.552	0.547	0.546	0.546	0.546	0.496	0.490	0.528	0.516	0.492	0.519	0.515	0.500	0.501	0.501
id 48	0.756	0.617	0.766	0.627	0.594	0.618	0.602	0.750	0.750	0.750	0.697	0.652	0.835	0.634	0.624	0.642	0.614	0.692	0.693	0.692
id 52	0.654	0.635	0.661	0.639	0.632	0.655	0.635	0.646	0.647	0.646	0.608	0.582	0.594	0.604	0.586	0.588	0.591	0.604	0.605	0.605
id 54	0.725	0.661	0.696	0.661	0.672	0.694	0.670	0.717	0.718	0.717	0.592	0.590	0.603	0.600	0.587	0.607	0.608	0.592	0.592	0.592
id 56	0.745	0.676	0.694	0.679	0.682	0.694	0.685	0.739	0.739	0.739	0.665	0.626	0.646	0.632	0.629	0.646	0.635	0.664	0.664	0.664
id 82	0.530	0.510	0.570	0.553	0.534	0.550	0.576	0.533	0.533	0.533	0.498	0.509	0.562	0.559	0.544	0.539	0.551	0.506	0.507	0.507
id 85	0.688	0.636	0.673	0.644	0.634	0.658	0.638	0.690	0.690	0.690	0.603	0.595	0.599	0.598	0.603	0.609	0.609	0.607	0.608	0.608
id 89	0.707	0.634	0.787	0.639	0.633	0.659	0.633	0.702	0.702	0.702	0.658	0.625	0.666	0.632	0.619	0.638	0.622	0.658	0.658	0.658
id 90	0.750	0.684	0.722	0.671	0.664	0.680	0.647	0.734	0.735	0.734	0.644	0.604	0.642	0.620	0.597	0.617	0.599	0.643	0.643	0.643
id 94	0.705	0.606	0.666	0.615	0.618	0.662	0.626	0.684	0.685	0.684	0.606	0.561	0.590	0.584	0.563	0.594	0.579	0.606	0.606	0.606
id 95	0.739	0.662	0.680	0.665	0.669	0.696	0.672	0.728	0.729	0.728	0.663	0.591	0.613	0.603	0.621	0.636	0.635	0.656	0.657	0.657
id 96	0.738	0.662	0.692	0.667	0.655	0.684	0.656	0.724	0.726	0.725	0.626	0.623	0.624	0.632	0.601	0.613	0.610	0.610	0.610	0.610

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets.

Table C4: Percentage of rationality rejections - GDP growth

IV set	Base	Equal Weights			Expon. Weights			Disagreement			Base	Equal Weights			Expon. Weights			Disagreement		
		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4
Linex loss																				
	Real-time										**Revised**									
1	29	61	71	61	42	55	48	19	26	19	23	55	61	55	61	65	61	29	32	29
2	45	23	48	19	23	42	13	45	52	48	65	45	52	42	23	45	19	68	68	68
3	52	52	68	55	48	58	48	52	48	55	45	65	58	65	71	71	74	52	58	52
4	6	0	0	0	0	0	0	6	6	6	0	0	0	0	0	0	0	0	0	0
5	26	6	16	6	6	26	6	29	35	23	55	3	3	3	0	13	0	42	42	45
6	39	26	13	26	13	6	10	42	42	45	26	29	26	29	23	19	23	29	29	29
7	65	81	74	81	74	74	74	71	68	71	71	68	65	68	71	74	74	61	61	61
8	68	84	77	84	81	81	81	71	68	68	71	71	71	68	81	81	81	68	65	61
9	68	74	71	71	68	68	68	65	68	61	81	90	87	94	94	87	94	84	81	77
10	0	6	6	6	6	6	6	0	0	0	0	0	0	0	0	0	0	0	0	0
11	35	29	42	29	52	23	48	35	42	32	55	39	42	42	58	45	55	65	58	65
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	87	84	84	84	71	87	74	81	84	81	68	87	81	81	74	71	71	81	77	71
14	48	29	23	29	35	39	32	58	58	58	45	19	19	23	29	26	29	48	48	48
15	52	42	39	42	39	45	42	61	58	61	65	58	52	58	55	55	55	58	58	58
Linex-Linex loss																				
	Real-time										**Revised**									
1	16	19	52	19	19	26	13	16	16	16	19	45	55	55	39	45	45	19	19	19
2	6	3	23	3	3	3	3	3	3	3	10	16	26	19	3	16	6	6	6	6
3	16	26	35	26	19	26	29	10	10	10	13	29	26	45	23	32	29	13	13	10
4	0	0	0	0	0	0	0	0	0	0	0	3	0	6	3	0	3	0	0	0
5	0	3	6	6	3	0	3	0	0	0	0	0	19	6	0	0	3	3	3	3
6	16	10	3	3	6	3	13	13	13	13	3	13	23	13	3	13	6	0	0	0
7	45	29	32	26	19	19	16	42	42	42	13	35	19	32	32	26	26	13	13	13
8	32	39	16	26	39	23	26	32	32	32	6	16	16	10	3	0	0	6	6	6
9	48	26	77	26	29	29	23	52	48	48	23	23	58	39	32	42	32	19	19	19
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	35	39	45	42	42	48	52	35	35	35	23	45	35	65	48	39	61	16	16	16
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	35	29	58	32	39	45	42	32	35	32	10	26	42	26	19	26	32	10	13	13
14	23	16	23	19	16	29	26	32	29	29	26	13	3	13	29	29	26	29	29	29
15	19	23	10	13	39	23	19	16	16	16	6	3	6	0	10	0	0	10	10	10

Notes: The table provides the percentage of rationality rejections for each instrument set and across individual forecaster.

Table C5: Estimates of α under Linex loss - HICP inflation

	Real-time										Revised									
Forecaster	**Base**	**Equal Weights**			**Expon. Weights**			**Disagreement**			**Base**	**Equal Weights**			**Expon. Weights**			**Disagreement**		
		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4
aggregate	-1.079	-0.887	-0.838	-0.801	-0.777	-0.819	-0.763	-1.006	-1.005	-0.999	-1.085	-0.908	-0.866	-0.832	-0.792	-0.835	-0.782	-1.017	-1.016	-1.010
average	-1.122	-0.784	-0.745	-0.710	-0.728	-0.772	-0.713	-0.987	-0.986	-0.972	-1.095	-0.787	-0.756	-0.716	-0.722	-0.758	-0.702	-0.975	-0.971	-0.959
id 1	-0.761	-0.716	-0.683	-0.672	-0.622	-0.670	-0.638	-0.770	-0.764	-0.764	-0.745	-0.728	-0.696	-0.688	-0.629	-0.682	-0.648	-0.767	-0.757	-0.760
id 4	-1.229	-0.906	-0.832	-0.781	-0.836	-0.897	-0.810	-1.103	-1.098	-1.085	-1.270	-1.033	-0.955	-0.905	-0.885	-0.989	-0.894	-1.192	-1.187	-1.174
id 5	-0.932	-0.741	-0.720	-0.706	-0.691	-0.743	-0.716	-0.896	-0.898	-0.896	-0.905	-0.754	-0.734	-0.719	-0.693	-0.598	-0.549	-0.890	-0.892	-0.892
id 7	-1.422	-1.243	-1.102	-1.024	-1.164	-1.064	-1.006	-1.302	-1.303	-1.287	-1.427	-1.243	-1.162	-1.041	-1.164	-1.069	-1.025	-1.311	-1.312	-1.297
id 14	0.082	-0.064	-0.091	-0.067	0.219	0.258	0.159	0.054	0.053	0.053	0.168	0.006	-0.028	-0.025	0.294	0.285	0.274	0.136	0.143	0.131
id 15	-1.026	-0.851	-0.805	-0.759	-0.766	-0.801	-0.753	-0.976	-0.974	-0.970	-1.020	-0.832	-0.787	-0.741	-0.757	-0.792	-0.743	-0.974	-0.972	-0.969
id 16	-0.983	-0.700	-0.655	-0.630	-0.684	-0.728	-0.659	-0.899	-0.898	-0.890	-0.978	-0.730	-0.693	-0.637	-0.691	-0.735	-0.682	-0.898	-0.897	-0.889
id 20	-1.569	-1.010	-1.009	-0.933	-1.043	-0.998	-0.914	-1.359	-1.360	-1.302	-1.567	-1.054	-1.006	-0.919	-1.027	-1.064	-0.916	-1.358	-1.358	-1.295
id 22	-1.267	-0.950	-0.869	-0.792	-0.891	-0.939	-0.829	-1.155	-1.151	-1.138	-1.259	-0.970	-0.892	-0.825	-0.906	-0.963	-0.862	-1.154	-1.150	-1.138
id 23	-0.932	-0.710	-0.630	-0.606	-0.680	-0.705	-0.665	-0.865	-0.862	-0.856	-0.939	-0.715	-0.707	-0.655	-0.689	-0.716	-0.672	-0.876	-0.873	-0.867
id 24	-1.423	-0.965	-0.912	-0.857	-0.992	-1.028	-0.932	-1.265	-1.238	-1.222	-1.414	-0.962	-0.913	-0.861	-0.984	-1.011	-0.930	-1.258	-1.234	-1.220
id 26	-1.138	-0.953	-0.908	-0.879	-0.879	-1.236	-0.850	-1.053	-1.051	-1.044	-1.159	-0.987	-0.946	-0.918	-0.909	-1.264	-0.882	-1.077	-1.075	-1.068
id 29	-0.220	-0.344	-0.342	-0.349	-0.168	-0.214	-0.205	-0.333	-0.334	-0.334	-0.190	-0.324	-0.324	-0.331	-0.179	-0.126	-0.143	-0.290	-0.290	-0.291
id 31	-1.176	-0.922	-0.926	-0.891	-0.824	-0.886	-0.851	-1.093	-1.095	-1.089	-1.192	-0.942	-0.911	-0.882	-0.854	-0.925	-0.892	-1.114	-1.116	-1.112
id 33	-0.108	-0.165	-0.158	-0.155	-0.008	0.085	0.060	-0.111	-0.111	-0.111	-0.090	-0.105	-0.103	-0.103	0.090	0.152	0.141	-0.093	-0.093	-0.093
id 37	-0.899	-0.713	-0.689	-0.671	-0.680	-0.713	-0.703	-0.843	-0.843	-0.841	-0.925	-0.714	-0.695	-0.656	-0.700	-0.729	-0.716	-0.871	-0.871	-0.869
id 38	-0.808	-0.659	-0.641	-0.626	-0.596	-0.630	-0.602	-0.746	-0.746	-0.741	-0.846	-0.693	-0.683	-0.671	-0.610	-0.636	-0.634	-0.783	-0.783	-0.779
id 39	-0.796	-0.544	-0.509	-0.496	-0.431	-0.494	-0.457	-0.716	-0.719	-0.714	-0.802	-0.561	-0.531	-0.521	-0.446	-0.512	-0.475	-0.728	-0.731	-0.726
id 41	-1.294	-1.083	-1.038	-0.975	-0.922	-0.986	-0.976	-1.194	-1.197	-1.194	-1.310	-1.136	-1.075	-1.011	-0.962	-1.021	-1.008	-1.213	-1.215	-1.212
id 42	-1.325	-0.920	-0.878	-0.845	-0.836	-0.869	-0.876	-1.261	-1.246	-1.231	-1.327	-0.927	-0.891	-0.865	-0.878	-0.880	-0.852	-1.254	-1.243	-1.228
id 47	-0.974	-0.720	-0.703	-0.677	-0.698	-0.716	-0.699	-0.874	-0.874	-0.871	-0.977	-0.729	-0.710	-0.685	-0.702	-0.720	-0.703	-0.877	-0.877	-0.875
id 48	-0.976	-0.789	-0.758	-0.735	-0.747	-0.779	-0.738	-0.908	-0.908	-0.902	-0.967	-0.798	-0.769	-0.750	-0.750	-0.784	-0.743	-0.902	-0.902	-0.897
id 52	-1.194	-0.856	-0.816	-0.719	-0.785	-0.880	-0.811	-1.100	-1.101	-1.084	-1.130	-0.826	-0.763	-0.691	-0.588	-0.649	-0.726	-1.051	-1.052	-1.031
id 54	-0.948	-0.768	-0.735	-0.715	-0.684	-0.725	-0.694	-0.911	-0.911	-0.909	-0.931	-0.774	-0.741	-0.722	-0.684	-0.728	-0.698	-0.901	-0.901	-0.900
id 56	-2.057	-1.381	-1.334	-1.218	-1.295	-1.323	-1.245	-1.749	-1.756	-1.718	-2.071	-1.415	-1.367	-1.244	-1.323	-1.358	-1.252	-1.778	-1.786	-1.750
id 82	-0.992	-0.585	-0.551	-0.516	-0.596	-0.621	-0.574	-0.915	-0.914	-0.907	-0.984	-0.590	-0.580	-0.516	-0.589	-0.615	-0.575	-0.914	-0.912	-0.906
id 85	-4.408	-1.206	-1.117	-1.001	-1.642	-1.898	-1.456	-2.622	-2.612	-2.470	-3.663	-1.079	-1.008	-0.916	-1.559	-1.709	-1.371	-2.339	-2.326	-2.220
id 89	-0.736	-0.616	-0.598	-0.591	-0.534	-0.566	-0.544	-0.682	-0.681	-0.679	-0.737	-0.621	-0.608	-0.605	-0.549	-0.578	-0.562	-0.687	-0.686	-0.684
id 90	-1.812	-1.101	-1.062	-1.037	-1.079	-1.108	-1.048	-1.549	-1.555	-1.524	-1.809	-1.042	-1.048	-1.026	-1.048	-1.063	-1.014	-1.536	-1.478	-1.449
id 93	-0.866	-0.692	-0.668	-0.628	-0.616	-0.662	-0.637	-0.825	-0.825	-0.822	-0.854	-0.685	-0.665	-0.592	-0.634	-0.668	-0.643	-0.815	-0.814	-0.812
id 94	-0.900	-0.610	-0.493	-0.585	-0.512	-0.573	-0.556	-0.797	-0.798	-0.793	-0.900	-0.604	-0.598	-0.581	-0.488	-0.496	-0.459	-0.683	-0.684	-0.680
id 95	-1.055	-0.826	-0.798	-0.763	-0.734	-0.767	-0.721	-0.967	-0.966	-0.959	-1.037	-0.829	-0.809	-0.771	-0.746	-0.779	-0.733	-0.956	-0.954	-0.948
id 96	-0.902	-0.573	-0.556	-0.545	-0.604	-0.633	-0.614	-0.797	-0.797	-0.793	-0.885	-0.562	-0.550	-0.540	-0.596	-0.624	-0.607	-0.784	-0.784	-0.780

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets.

Table C6: Estimates of α under Linex-Linex loss - HICP inflation

	Real-time										Revised									
Forecaster	**Base**	**Equal Weights**			**Expon. Weights**			**Disagreement**			**Base**	**Equal Weights**			**Expon. Weights**			**Disagreement**		
		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4
aggregate	0.283	0.312	0.313	0.316	0.317	0.316	0.319	0.290	0.290	0.290	0.292	0.314	0.316	0.318	0.320	0.319	0.322	0.299	0.299	0.299
average	0.303	0.326	0.327	0.330	0.329	0.329	0.334	0.308	0.309	0.309	0.306	0.330	0.331	0.334	0.334	0.335	0.338	0.312	0.312	0.312
id 1	0.305	0.326	0.328	0.330	0.331	0.329	0.332	0.311	0.311	0.311	0.306	0.328	0.329	0.331	0.333	0.331	0.334	0.313	0.312	0.313
id 4	0.263	0.298	0.291	0.294	0.293	0.292	0.296	0.270	0.270	0.270	0.267	0.292	0.294	0.297	0.295	0.294	0.298	0.274	0.274	0.274
id 5	0.302	0.328	0.328	0.330	0.347	0.347	0.348	0.319	0.319	0.319	0.305	0.343	0.344	0.345	0.349	0.349	0.372	0.310	0.310	0.311
id 7	0.204	0.223	0.224	0.226	0.228	0.228	0.230	0.210	0.210	0.210	0.206	0.224	0.226	0.228	0.229	0.229	0.231	0.211	0.211	0.211
id 14	0.297	0.364	0.367	0.371	0.404	0.403	0.455	0.306	0.306	0.307	0.327	0.396	0.399	0.401	0.455	0.477	0.472	0.336	0.336	0.336
id 15	0.271	0.301	0.303	0.307	0.313	0.311	0.314	0.278	0.278	0.279	0.274	0.308	0.310	0.313	0.314	0.313	0.316	0.281	0.281	0.281
id 16	0.306	0.326	0.328	0.330	0.332	0.331	0.333	0.312	0.312	0.313	0.309	0.328	0.330	0.333	0.334	0.333	0.335	0.315	0.314	0.315
id 20	0.261	0.288	0.291	0.293	0.289	0.316	0.318	0.267	0.267	0.267	0.264	0.292	0.294	0.296	0.319	0.318	0.341	0.269	0.269	0.270
id 22	0.236	0.258	0.259	0.262	0.270	0.268	0.271	0.243	0.243	0.244	0.240	0.262	0.263	0.266	0.271	0.269	0.273	0.248	0.248	0.248
id 23	0.287	0.312	0.314	0.317	0.317	0.315	0.318	0.293	0.293	0.293	0.288	0.312	0.314	0.317	0.312	0.311	0.313	0.292	0.292	0.292
id 24	0.291	0.312	0.314	0.317	0.315	0.313	0.315	0.295	0.295	0.296	0.294	0.312	0.314	0.316	0.314	0.313	0.315	0.298	0.297	0.298
id 26	0.317	0.326	0.326	0.327	0.336	0.335	0.336	0.321	0.321	0.321	0.320	0.329	0.329	0.330	0.338	0.338	0.340	0.325	0.325	0.325
id 29	0.358	0.366	0.368	0.370	0.391	0.392	0.394	0.362	0.362	0.362	0.362	0.374	0.386	0.378	0.394	0.419	0.416	0.366	0.366	0.366
id 31	0.288	0.303	0.304	0.306	0.308	0.307	0.309	0.292	0.292	0.292	0.291	0.305	0.307	0.308	0.311	0.310	0.312	0.295	0.295	0.295
id 33	0.463	0.467	0.469	0.469	0.470	0.470	0.470	0.465	0.465	0.465	0.469	0.474	0.475	0.476	0.476	0.476	0.477	0.471	0.471	0.471
id 37	0.314	0.314	0.332	0.333	0.306	0.306	0.307	0.298	0.298	0.298	0.300	0.321	0.323	0.338	0.312	0.312	0.313	0.305	0.305	0.305
id 38	0.365	0.375	0.374	0.376	0.378	0.377	0.379	0.368	0.368	0.368	0.367	0.379	0.379	0.380	0.378	0.369	0.371	0.370	0.370	0.370
id 39	0.372	0.382	0.384	0.386	0.389	0.388	0.390	0.376	0.376	0.376	0.375	0.382	0.385	0.386	0.391	0.388	0.390	0.379	0.379	0.379
id 41	0.219	0.259	0.262	0.266	0.244	0.245	0.248	0.203	0.229	0.229	0.221	0.261	0.264	0.268	0.246	0.247	0.250	0.231	0.231	0.231
id 42	0.290	0.310	0.311	0.312	0.313	0.312	0.314	0.296	0.296	0.296	0.290	0.311	0.312	0.314	0.315	0.314	0.316	0.297	0.297	0.297
id 47	0.371	0.386	0.387	0.388	0.386	0.385	0.386	0.374	0.374	0.374	0.370	0.383	0.384	0.385	0.385	0.384	0.385	0.373	0.373	0.373
id 48	0.275	0.296	0.298	0.300	0.303	0.302	0.304	0.282	0.282	0.282	0.279	0.302	0.303	0.306	0.306	0.305	0.308	0.286	0.286	0.286
id 52	0.270	0.293	0.296	0.299	0.294	0.292	0.295	0.276	0.276	0.276	0.271	0.294	0.297	0.300	0.294	0.315	0.318	0.276	0.276	0.276
id 54	0.293	0.322	0.324	0.327	0.322	0.321	0.323	0.298	0.298	0.298	0.296	0.317	0.319	0.322	0.317	0.316	0.319	0.302	0.302	0.302
id 56	0.256	0.275	0.275	0.277	0.276	0.275	0.277	0.261	0.261	0.261	0.257	0.280	0.282	0.284	0.282	0.280	0.283	0.264	0.264	0.264
id 82	0.256	0.278	0.280	0.282	0.289	0.282	0.290	0.258	0.258	0.258	0.256	0.279	0.280	0.283	0.289	0.284	0.291	0.263	0.262	0.262
id 85	0.268	0.302	0.305	0.307	0.281	0.284	0.283	0.287	0.286	0.287	0.263	0.303	0.298	0.301	0.281	0.281	0.283	0.276	0.276	0.276
id 89	0.369	0.386	0.387	0.389	0.386	0.385	0.387	0.374	0.374	0.374	0.373	0.389	0.391	0.392	0.390	0.389	0.390	0.378	0.378	0.378
id 90	0.213	0.252	0.253	0.254	0.238	0.236	0.258	0.227	0.227	0.227	0.218	0.265	0.266	0.267	0.259	0.257	0.259	0.228	0.228	0.228
id 93	0.374	0.388	0.388	0.390	0.392	0.391	0.393	0.378	0.378	0.378	0.375	0.393	0.394	0.396	0.394	0.401	0.401	0.380	0.380	0.380
id 94	0.368	0.387	0.389	0.390	0.400	0.400	0.401	0.373	0.373	0.373	0.372	0.391	0.393	0.408	0.403	0.402	0.403	0.376	0.376	0.376
id 95	0.327	0.354	0.355	0.356	0.356	0.355	0.356	0.332	0.332	0.332	0.329	0.355	0.356	0.358	0.358	0.356	0.358	0.334	0.334	0.334
id 96	0.359	0.387	0.388	0.390	0.375	0.374	0.376	0.364	0.364	0.364	0.363	0.388	0.390	0.392	0.380	0.378	0.380	0.368	0.367	0.368

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets.

Table C7: Percentage of rationality rejections - HICP inflation

IV set	Base	Equal Weights			Expon. Weights			Disagreement			Base	Equal Weights			Expon. Weights			Disagreement		
		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4
Linex loss																				
	Real-time										**Revised**									
1	15	6	6	6	15	15	15	12	12	12	15	6	6	6	15	12	15	15	15	12
2	15	9	9	6	12	12	12	12	12	12	15	9	9	9	12	12	12	12	12	12
3	30	24	30	24	30	30	30	27	27	24	30	24	21	18	30	33	27	21	21	21
4	76	70	70	70	79	70	79	70	70	70	70	61	67	64	73	64	67	67	67	67
5	70	67	76	85	52	55	64	70	67	70	70	67	76	82	61	52	70	73	73	70
6	76	88	76	79	76	73	67	79	73	79	76	79	82	79	70	76	64	76	76	76
7	97	85	88	79	97	97	97	97	97	97	94	85	85	79	97	97	97	97	97	97
8	88	70	70	67	88	88	85	82	82	82	85	64	67	67	85	76	82	82	79	79
9	97	79	73	67	97	97	94	94	94	91	94	76	70	61	97	97	97	94	94	94
10	0	3	3	3	0	0	0	0	0	0	0	0	3	3	0	0	0	0	0	0
11	94	88	91	94	91	88	88	91	91	91	94	91	91	91	88	88	88	91	91	91
12	15	12	12	12	12	12	12	18	15	15	12	12	12	12	12	12	12	15	15	12
13	91	91	82	82	94	91	91	88	91	91	94	85	82	70	88	91	91	91	94	94
14	91	79	85	85	82	82	85	79	79	79	88	76	85	82	85	79	85	79	79	79
15	85	79	85	85	88	85	88	88	88	91	85	76	82	79	88	82	88	85	85	85
Linex-Linex loss																				
	Real-time										**Revised**									
1	3	3	3	3	3	0	0	3	3	3	3	3	3	3	0	0	0	3	3	3
2	0	3	3	3	0	0	0	0	0	0	3	0	0	0	0	0	0	3	3	3
3	3	12	12	12	9	9	9	3	3	0	3	15	15	15	9	9	9	3	3	3
4	3	9	9	9	6	6	6	6	6	6	6	9	9	9	9	9	9	6	6	6
5	6	12	12	12	9	9	9	6	6	6	6	12	9	9	9	9	9	9	9	9
6	3	6	6	6	3	3	3	3	3	3	3	6	6	9	3	3	3	3	3	3
7	9	21	24	24	9	9	6	9	9	9	9	12	21	21	6	6	6	9	9	9
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	3	0	0	0	3	3	3	0	0	0	3	3	0	0	6	9	9	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	6	12	12	12	9	9	9	6	6	6	6	12	12	12	9	9	9	6	6	6
12	3	0	0	0	0	0	0	3	3	3	3	0	0	0	0	0	0	3	3	3
13	6	0	3	3	3	3	6	3	3	3	3	0	0	3	6	6	6	3	3	3
14	6	6	3	3	15	15	15	6	6	6	6	9	9	9	9	15	15	9	9	9
15	3	3	0	6	0	3	6	6	6	6	3	3	3	3	9	12	9	6	6	6

Notes: The table provides the percentage of rationality rejections for each instrument set and across individual forecaster.

Table C8: Estimates of α under Linex loss - Unemployment rate

	Real-time										Revised									
Forecaster	**Base**	**Equal Weights**			**Expon. Weights**			**Disagreement**			**Base**	**Equal Weights**			**Expon. Weights**			**Disagreement**		
		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4
aggregate	-0.865	-0.210	-0.210	-0.209	-0.387	-0.314	-0.319	-0.627	-0.633	-0.618	-2.367	-0.739	-0.722	-0.683	-1.226	-1.193	-1.107	-1.767	-1.734	-1.676
average	-0.620	-0.131	-0.130	-0.125	-0.257	-0.246	-0.222	-0.439	-0.441	-0.427	-2.086	-0.703	-0.688	-0.633	-1.072	-1.075	-0.951	-1.534	-1.518	-1.452
id 4	-0.349	-0.072	-0.070	-0.070	-0.149	-0.150	-0.144	-0.264	-0.264	-0.263	-2.401	-0.646	-0.646	-0.618	-1.205	-1.306	-1.155	-1.692	-1.720	-1.660
id 5	-1.149	-0.287	-0.288	-0.286	-0.664	-0.690	-0.650	-0.941	-0.927	-0.931	-2.712	-0.939	-0.941	-0.888	-1.407	-1.436	-1.249	-2.051	-2.005	-1.960
id 7	0.312	0.137	0.145	0.140	0.251	0.247	0.238	0.245	0.246	0.245	-0.669	-0.362	-0.351	-0.342	-0.421	-0.397	-0.386	-0.514	-0.502	-0.491
id 14	0.533	0.473	0.483	0.475	0.551	0.609	0.664	0.377	0.383	0.389	-0.352	-0.068	-0.067	-0.067	-0.083	-0.101	-0.068	-0.274	-0.273	-0.272
id 15	-1.124	-0.288	-0.286	-0.284	-0.616	-0.575	-0.555	-0.926	-0.932	-0.906	-2.971	-0.743	-0.682	-0.640	-1.290	-1.261	-1.110	-1.969	-1.979	-1.900
id 16	0.054	0.026	0.022	0.021	0.090	0.046	0.069	0.047	0.021	0.014	-1.691	-0.584	-0.568	-0.546	-0.898	-0.897	-0.861	-1.307	-1.317	-1.309
id 20	-0.647	-0.142	-0.139	-0.138	-0.277	-0.266	-0.249	-0.475	-0.444	-0.469	-1.411	-0.428	-0.429	-0.430	-0.723	-0.696	-0.672	-1.085	-1.082	-1.078
id 22	1.542	0.555	0.536	0.522	1.093	1.008	0.949	1.087	1.073	1.047	-0.183	-0.074	-0.074	-0.074	-0.120	-0.129	-0.126	-0.144	-0.148	-0.147
id 23	-0.706	-0.183	-0.183	-0.182	-0.411	-0.403	-0.385	-0.580	-0.585	-0.585	-1.682	-0.693	-0.666	-0.631	-1.005	-0.932	-0.892	-1.383	-1.371	-1.342
id 24	-0.505	-0.111	-0.110	-0.109	-0.202	-0.194	-0.189	-0.372	-0.375	-0.377	-1.884	-0.463	-0.457	-0.444	-0.914	-0.857	-0.810	-1.458	-1.460	-1.433
id 26	-1.418	-0.380	-0.378	-0.369	-0.817	-0.791	-0.750	-1.111	-1.119	-1.114	-2.180	-0.745	-0.709	-0.609	-1.081	-0.993	-0.935	-1.682	-1.668	-1.657
id 31	-0.721	-0.156	-0.157	-0.156	-0.408	-0.402	-0.383	-0.569	-0.587	-0.587	-2.012	-0.731	-0.736	-0.695	-1.178	-1.123	-1.058	-1.638	-1.629	-1.608
id 33	-0.480	-0.305	-0.308	-0.299	-0.337	-0.325	-0.314	-0.425	-0.422	-0.418	-1.303	-0.726	-0.722	-0.692	-0.819	-0.823	-0.793	-1.125	-1.112	-1.099
id 37	-1.862	-0.437	-0.438	-0.425	-0.805	-0.762	-0.706	-1.232	-1.220	-1.165	-4.138	-0.940	-0.955	-0.868	-1.537	-1.319	-1.250	-2.606	-2.484	-2.275
id 38	-0.430	-0.102	-0.077	-0.086	-0.133	-0.222	-0.121	-0.338	-0.337	-0.343	-1.382	-0.608	-0.588	-0.564	-0.787	-0.774	-0.739	-1.100	-1.091	-1.073
id 39	-0.320	-0.032	-0.031	-0.031	-0.106	-0.116	-0.111	-0.240	-0.234	-0.244	-1.951	-0.499	-0.492	-0.476	-0.988	-0.998	-0.915	-1.399	-1.415	-1.385
id 41	-2.129	-0.321	-0.353	-0.332	-1.046	-1.069	-0.868	-1.272	-1.422	-1.278	-4.444	-1.600	-1.487	-1.273	-2.311	-3.429	-2.020	-3.221	-3.122	-2.872
id 42	-1.581	-0.656	-0.664	-0.634	-0.918	-0.965	-0.857	-1.245	-1.221	-1.189	-2.076	-1.072	-1.068	-0.992	-1.409	-1.312	-1.239	-1.674	-1.584	-1.511
id 47	0.112	0.044	0.051	0.054	0.134	0.250	0.157	0.225	0.226	0.255	-1.629	-0.623	-0.633	-0.602	-1.053	-0.967	-0.956	-1.331	-1.390	-1.277
id 48	-1.349	-0.238	-0.236	-0.233	-0.533	-0.498	-0.470	-0.955	-0.949	-0.916	-2.906	-0.627	-0.634	-0.600	-1.447	-1.184	-1.086	-2.013	-1.892	-1.754
id 52	0.596	0.238	0.244	0.238	0.468	0.431	0.446	0.486	0.529	0.521	-0.920	-0.564	-0.503	-0.461	-0.536	-0.552	-0.529	-0.728	-0.725	-0.691
id 54	0.214	0.090	0.092	0.088	0.194	0.337	0.184	0.177	0.179	0.168	-0.524	-0.331	-0.335	-0.332	-0.406	-0.396	-0.396	-0.466	-0.465	-0.464
id 56	-1.829	-0.431	-0.414	-0.396	-0.805	-0.837	-0.710	-1.147	-1.160	-1.132	-3.866	-1.315	-1.337	-1.000	-1.843	-1.742	-1.553	-2.807	-2.747	-2.645
id 89	-2.384	-0.734	-0.719	-0.648	-1.106	-0.990	-0.889	-1.762	-1.713	-1.624	-3.943	-1.171	-1.132	-1.007	-1.743	-1.489	-1.324	-2.592	-2.631	-2.276
id 90	-1.415	-0.270	-0.271	-0.269	-0.544	-0.550	-0.522	-0.734	-0.744	-0.721	-3.022	-1.164	-1.155	-1.062	-1.656	-1.614	-1.629	-2.284	-2.236	-2.104
id 94	-0.006	0.011	0.010	0.010	0.014	0.010	0.010	-0.026	-0.028	-0.032	-1.477	-0.532	-0.536	-0.515	-0.822	-0.794	-0.764	-1.132	-1.126	-1.096
id 95	-0.509	-0.178	-0.182	-0.176	-0.272	-0.171	-0.220	-0.510	-0.511	-0.495	-2.835	-0.875	-0.782	-0.731	-1.340	-1.673	-1.231	-1.889	-1.870	-1.856
id 96	0.233	0.071	0.072	0.071	0.158	0.130	0.126	0.188	0.194	0.192	-1.759	-0.573	-0.582	-0.555	-0.945	-1.057	-0.849	-1.319	-1.382	-1.349

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets

Table C9: Estimates of α under Linex-Linex loss - Unemployment rate

	Real-time										Revised									
Forecaster	**Base**	**Equal Weights**			**Expon. Weights**			**Disagreement**			**Base**	**Equal Weights**			**Expon. Weights**			**Disagreement**		
		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4
aggregate	0.425	0.444	0.444	0.444	0.442	0.445	0.446	0.427	0.427	0.427	0.323	0.360	0.360	0.358	0.352	0.353	0.355	0.329	0.330	0.330
average	0.444	0.462	0.461	0.462	0.458	0.461	0.462	0.443	0.443	0.443	0.320	0.357	0.357	0.360	0.350	0.354	0.356	0.328	0.326	0.328
id 4	0.468	0.471	0.470	0.470	0.477	0.479	0.480	0.468	0.467	0.467	0.347	0.371	0.371	0.372	0.365	0.366	0.367	0.352	0.352	0.352
id 5	0.335	0.398	0.401	0.402	0.391	0.391	0.393	0.340	0.340	0.340	0.241	0.286	0.285	0.288	0.273	0.286	0.289	0.247	0.247	0.247
id 7	0.535	0.515	0.516	0.515	0.535	0.537	0.536	0.530	0.530	0.530	0.403	0.418	0.420	0.422	0.424	0.426	0.427	0.410	0.410	0.410
id 14	0.645	0.623	0.625	0.623	0.634	0.635	0.635	0.636	0.636	0.635	0.484	0.485	0.485	0.485	0.489	0.492	0.493	0.481	0.481	0.481
id 15	0.388	0.424	0.417	0.418	0.404	0.406	0.407	0.393	0.392	0.392	0.272	0.337	0.336	0.338	0.319	0.322	0.324	0.295	0.296	0.296
id 16	0.496	0.527	0.527	0.526	0.511	0.514	0.514	0.495	0.494	0.494	0.375	0.381	0.382	0.383	0.387	0.389	0.390	0.375	0.374	0.374
id 20	0.434	0.458	0.458	0.459	0.454	0.458	0.459	0.435	0.435	0.435	0.364	0.415	0.416	0.417	0.396	0.400	0.412	0.369	0.369	0.369
id 22	0.652	0.629	0.621	0.625	0.649	0.650	0.650	0.636	0.636	0.636	0.488	0.502	0.503	0.503	0.503	0.503	0.503	0.488	0.488	0.488
id 23	0.417	0.433	0.433	0.434	0.429	0.432	0.433	0.416	0.416	0.416	0.311	0.348	0.349	0.351	0.341	0.343	0.344	0.317	0.317	0.317
id 24	0.463	0.460	0.461	0.461	0.469	0.474	0.474	0.460	0.460	0.460	0.351	0.385	0.386	0.388	0.377	0.379	0.381	0.354	0.354	0.354
id 26	0.347	0.381	0.381	0.382	0.368	0.369	0.370	0.352	0.352	0.353	0.239	0.290	0.291	0.294	0.283	0.303	0.305	0.251	0.252	0.252
id 31	0.433	0.464	0.461	0.462	0.442	0.445	0.446	0.433	0.432	0.432	0.305	0.333	0.334	0.336	0.324	0.326	0.328	0.308	0.308	0.308
id 33	0.418	0.430	0.431	0.431	0.445	0.448	0.448	0.418	0.418	0.418	0.314	0.331	0.333	0.334	0.337	0.338	0.339	0.318	0.318	0.318
id 37	0.356	0.382	0.381	0.383	0.385	0.389	0.391	0.364	0.365	0.365	0.236	0.292	0.292	0.296	0.284	0.289	0.292	0.249	0.250	0.251
id 38	0.463	0.477	0.476	0.477	0.472	0.476	0.478	0.460	0.460	0.460	0.356	0.378	0.379	0.380	0.378	0.381	0.382	0.360	0.360	0.360
id 39	0.480	0.503	0.502	0.502	0.489	0.492	0.492	0.480	0.479	0.479	0.350	0.401	0.401	0.403	0.378	0.378	0.380	0.358	0.358	0.358
id 41	0.374	0.390	0.387	0.387	0.380	0.385	0.386	0.368	0.367	0.367	0.217	0.246	0.247	0.248	0.248	0.250	0.251	0.224	0.224	0.224
id 42	0.326	0.367	0.366	0.368	0.360	0.362	0.364	0.334	0.334	0.334	0.227	0.280	0.282	0.291	0.281	0.282	0.285	0.235	0.236	0.236
id 47	0.504	0.506	0.507	0.507	0.507	0.510	0.510	0.499	0.499	0.499	0.372	0.387	0.389	0.390	0.391	0.394	0.394	0.374	0.374	0.374
id 48	0.355	0.428	0.429	0.431	0.381	0.385	0.386	0.361	0.361	0.362	0.262	0.346	0.346	0.349	0.315	0.327	0.330	0.277	0.277	0.277
id 52	0.559	0.544	0.545	0.544	0.558	0.559	0.560	0.550	0.551	0.552	0.409	0.418	0.421	0.424	0.422	0.431	0.427	0.412	0.412	0.412
id 54	0.489	0.487	0.487	0.487	0.496	0.498	0.498	0.486	0.486	0.486	0.364	0.385	0.385	0.386	0.384	0.386	0.388	0.366	0.367	0.367
id 56	0.371	0.402	0.400	0.401	0.382	0.387	0.388	0.376	0.375	0.375	0.212	0.260	0.244	0.260	0.254	0.256	0.259	0.217	0.218	0.218
id 89	0.261	0.318	0.321	0.325	0.302	0.304	0.307	0.272	0.272	0.272	0.198	0.261	0.263	0.267	0.237	0.242	0.245	0.209	0.210	0.210
id 90	0.423	0.429	0.429	0.429	0.434	0.435	0.436	0.426	0.426	0.424	0.253	0.329	0.330	0.331	0.318	0.318	0.318	0.298	0.245	0.299
id 94	0.498	0.503	0.502	0.502	0.500	0.503	0.503	0.486	0.485	0.485	0.356	0.396	0.396	0.398	0.386	0.388	0.389	0.363	0.364	0.364
id 95	0.430	0.465	0.456	0.460	0.443	0.455	0.458	0.430	0.430	0.430	0.325	0.347	0.347	0.348	0.348	0.347	0.348	0.333	0.333	0.333
id 96	0.514	0.514	0.513	0.513	0.518	0.519	0.520	0.510	0.510	0.510	0.330	0.396	0.396	0.398	0.361	0.361	0.363	0.339	0.339	0.339

Notes: $\hat{\alpha}$ estimates are the mean results across the instrument sets

Table C10: Percentage of rationality rejections - Unemployment rate

IV set	Base	Equal Weights			Expon. Weights			Disagreement			Base	Equal Weights			Expon. Weights			Disagreement		
		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4		σ^2	s^3	k^4	σ^2	s^3	k^4	σ^2	s^3	k^4
Linex loss																				
	Real-time										**Revised**									
1	7	7	7	7	7	7	7	7	7	7	36	11	14	11	18	36	25	29	29	29
2	36	36	36	36	32	32	36	39	39	39	29	14	14	14	21	25	18	21	25	18
3	7	14	14	14	11	14	11	7	7	4	36	14	14	14	14	21	21	25	25	21
4	36	54	54	54	43	46	50	36	36	36	11	18	14	14	14	14	14	14	14	11
5	64	64	64	64	64	61	61	57	68	64	36	7	11	11	39	25	21	32	32	29
6	25	50	50	50	36	36	39	29	29	25	11	14	14	14	18	18	14	14	14	11
7	4	4	4	4	4	4	4	4	4	4	18	0	0	0	14	25	21	18	18	18
8	36	29	29	29	36	36	36	36	36	32	11	4	4	4	4	4	4	4	4	4
9	21	43	43	43	29	39	32	18	18	18	14	4	4	4	0	4	4	7	7	7
10	4	4	4	4	0	4	4	4	0	0	0	0	0	0	0	4	4	0	0	0
11	89	86	82	82	89	89	86	89	89	93	71	75	75	75	82	75	82	75	75	75
12	0	4	7	7	4	4	4	0	0	0	0	0	4	4	4	4	4	0	0	0
13	46	57	54	54	54	61	64	50	50	50	46	57	68	54	50	57	57	43	43	39
14	54	46	50	50	54	57	57	64	64	64	36	32	29	29	39	32	36	36	32	32
15	32	29	29	29	21	21	21	29	29	29	11	14	11	11	11	11	18	7	7	7
Linex-Linex loss																				
	Real-time										**Revised**									
1	4	4	7	4	4	4	4	4	4	4	7	7	7	7	4	4	4	7	7	7
2	29	21	25	25	25	25	25	29	29	29	14	14	11	11	14	14	14	18	18	18
3	4	4	4	4	7	7	7	4	4	4	0	4	4	4	0	0	0	0	0	0
4	43	54	54	54	54	54	54	50	50	50	18	21	21	21	14	21	21	18	18	18
5	54	71	68	71	54	50	50	54	54	54	25	32	32	32	29	21	21	25	25	25
6	21	54	57	57	25	25	25	21	21	21	14	18	18	18	18	14	18	14	14	14
7	4	4	4	4	4	4	4	4	4	4	0	0	0	0	4	4	4	0	0	0
8	25	29	29	29	21	25	25	21	21	21	4	0	0	0	4	4	4	4	4	4
9	29	36	36	36	32	36	36	25	25	25	7	4	4	4	4	4	4	7	7	7
10	7	4	4	4	7	4	4	7	7	7	4	4	4	4	0	0	0	4	4	4
11	68	75	75	79	68	68	68	68	68	68	68	75	75	75	68	68	71	64	64	64
12	7	4	4	4	4	0	0	4	4	4	4	4	4	4	4	4	4	4	4	4
13	32	46	46	46	43	43	43	32	32	32	39	46	46	46	50	50	50	36	36	36
14	29	46	46	50	46	43	43	36	36	36	29	29	25	21	29	29	29	25	25	25
15	18	18	21	21	21	21	21	14	14	14	7	7	7	7	7	7	7	7	7	7

Notes: The table provides the percentage of rationality rejections for each instrument set and across individual forecaster.

Table C11: Results for instrument sets and Linex loss - GDP growth

IV set	$\hat{\alpha}$ estimates				$H_0: \alpha \leq 0$				$H_0: J = 0$			
	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4
Real-time												
0	1.044	0.798	0.791	0.779	46	50	55	55	–	–	–	–
1	1.619	1.010	0.981	0.944	96	86	86	86	36	36	32	32
2	2.799	1.446	1.437	1.336	100	96	96	96	55	45	45	50
3	1.599	0.955	0.974	0.881	86	82	86	82	59	59	55	50
4	1.195	0.889	0.879	0.858	73	73	73	77	0	0	0	0
5	3.270	2.010	1.951	1.779	100	100	100	100	32	45	45	45
6	1.018	0.826	0.823	0.807	73	77	82	82	50	32	32	32
7	0.856	0.542	0.521	0.503	36	32	27	32	64	59	59	64
8	0.679	0.421	0.414	0.407	23	27	27	27	64	55	55	55
9	1.931	1.107	1.137	0.982	91	73	68	68	64	59	59	55
10	1.120	0.860	0.850	0.830	64	77	77	82	0	0	0	0
11	2.368	1.657	1.513	1.396	100	100	100	100	36	64	59	55
12	1.068	0.761	0.751	0.729	59	68	73	77	0	0	0	0
13	1.399	0.979	1.030	0.870	91	77	73	73	91	82	77	82
14	1.236	0.906	0.885	0.864	82	77	77	77	59	45	45	45
15	1.060	0.621	0.601	0.579	55	46	46	46	55	55	55	55
Revised												
0	0.433	0.357	0.356	0.356	0	14	14	14	–	–	–	–
1	1.085	0.669	0.661	0.653	96	77	77	77	27	27	27	27
2	1.547	0.775	0.765	0.748	96	77	77	77	73	45	45	45
3	1.054	0.642	0.634	0.625	77	68	68	68	45	59	59	64
4	0.702	0.530	0.527	0.524	77	73	73	73	0	0	0	0
5	1.594	1.042	1.024	0.987	96	86	86	86	64	32	32	32
6	0.572	0.357	0.353	0.361	77	50	55	55	32	36	36	36
7	0.026	0.013	-0.029	-0.033	5	9	9	9	77	59	55	55
8	-0.209	-0.066	-0.063	-0.060	5	0	0	0	77	73	73	73
9	0.177	0.228	0.205	0.164	18	23	18	18	82	73	73	77
10	0.516	0.378	0.378	0.377	9	27	27	27	0	0	0	0
11	1.715	1.129	1.108	1.015	100	96	96	96	55	55	45	50
12	0.433	0.273	0.292	0.294	9	32	32	32	0	9	5	5
13	0.402	0.370	0.240	0.241	68	55	41	41	64	73	64	64
14	0.558	0.400	0.387	0.379	73	64	64	64	45	45	45	45
15	0.332	0.137	0.140	0.142	41	36	36	36	68	59	59	64

Notes: $\hat{\alpha}$ estimates are the mean results across the individual forecasters. $H_0: \alpha \leq 0$ and $H_0: J = 0$ show the rejection frequencies across forecasters on a 5 percent level.

Table C12: Results for instrument sets and Linex-Linex loss - GDP growth

IV set	**$\hat{\alpha}$ estimates**				$H_0: \alpha \leq 0.5$				$H_0: J = 0$			
	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4
					Real-time							
0	0.673	0.666	0.666	0.665	96	91	96	96	–	–	–	–
1	0.754	0.721	0.721	0.720	100	100	100	100	18	18	18	18
2	0.817	0.777	0.777	0.775	100	100	100	100	5	9	5	0
3	0.772	0.740	0.740	0.738	96	96	96	96	14	32	27	27
4	0.696	0.684	0.684	0.683	91	91	91	91	0	0	0	0
5	0.831	0.806	0.805	0.802	100	100	100	100	0	5	5	5
6	0.663	0.652	0.652	0.652	86	82	82	82	18	18	14	14
7	0.637	0.625	0.629	0.629	73	73	73	73	55	45	45	45
8	0.617	0.628	0.628	0.628	68	59	55	55	32	27	27	27
9	0.688	0.709	0.708	0.707	73	73	73	73	45	36	36	36
10	0.671	0.667	0.666	0.666	96	91	96	96	0	0	0	0
11	0.827	0.791	0.778	0.775	96	96	96	96	45	32	32	32
12	0.669	0.664	0.664	0.663	91	91	91	91	0	5	5	5
13	0.733	0.727	0.716	0.715	82	77	73	73	32	27	27	27
14	0.704	0.690	0.690	0.688	96	82	82	82	23	23	23	23
15	0.650	0.649	0.650	0.649	77	77	77	77	23	23	23	23
					Revised							
0	0.629	0.625	0.625	0.625	68	68	68	68	–	–	–	–
1	0.704	0.684	0.683	0.683	96	91	91	91	23	23	18	18
2	0.681	0.686	0.668	0.668	64	68	64	68	9	14	18	23
3	0.721	0.655	0.661	0.660	86	77	77	77	9	14	23	23
4	0.666	0.651	0.651	0.650	96	82	82	82	0	0	0	0
5	0.716	0.693	0.705	0.702	91	68	73	73	0	5	9	9
6	0.585	0.570	0.570	0.570	50	36	41	36	5	9	9	9
7	0.579	0.584	0.584	0.584	18	27	27	27	14	18	18	18
8	0.520	0.535	0.535	0.535	9	5	5	5	9	9	9	9
9	0.558	0.578	0.578	0.578	14	27	27	27	18	18	23	23
10	0.618	0.614	0.614	0.614	68	68	73	73	0	0	0	0
11	0.750	0.700	0.699	0.697	96	77	77	77	23	14	18	18
12	0.612	0.616	0.616	0.615	64	68	73	73	0	0	0	0
13	0.564	0.565	0.566	0.566	32	32	32	32	14	18	18	18
14	0.661	0.660	0.665	0.663	86	82	82	82	32	23	23	23
15	0.584	0.586	0.587	0.587	59	46	46	46	5	5	5	5

Notes: $\hat{\alpha}$ estimates are the mean results across the individual forecasters. $H_0: \alpha \leq 0.5$ and $H_0: J = 0$ show the rejection frequencies across forecasters on a 5 percent level.

Table C13: Results for instrument sets and Linex loss - HICP inflation

IV set	$\hat{\alpha}$ estimates				$H_0: \alpha \geq 0$				$H_0: J = 0$			
	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4
Real-time												
0	-0.544	-0.461	-0.464	-0.460	0	9	13	13	–	–	–	–
1	-0.670	-0.556	-0.561	-0.559	22	22	22	22	4	13	13	9
2	-0.671	-0.587	-0.594	-0.592	30	39	39	39	4	13	13	9
3	-0.686	-0.511	-0.512	-0.507	22	30	30	30	13	13	13	13
4	-0.964	-0.842	-0.860	-0.837	57	61	61	61	61	48	43	35
5	-1.609	-1.251	-1.332	-1.223	96	87	87	87	52	35	35	39
6	-1.538	-1.244	-1.270	-1.221	96	87	87	87	65	65	61	65
7	-1.469	-1.357	-1.360	-1.324	83	87	87	91	91	74	74	65
8	-1.435	-1.442	-1.451	-1.403	83	87	87	87	78	48	57	48
9	-1.440	-1.398	-1.451	-1.416	87	96	96	96	91	83	83	78
10	-0.630	-0.603	-0.607	-0.597	17	26	30	30	0	0	0	0
11	-1.252	-1.225	-1.212	-1.230	91	91	91	91	87	70	70	61
12	-0.751	-0.702	-0.713	-0.699	35	52	52	52	9	4	4	4
13	-1.497	-1.404	-1.449	-1.402	91	96	96	96	96	78	78	70
14	-1.672	-1.471	-1.519	-1.437	91	91	91	91	78	48	48	48
15	-1.742	-1.587	-1.605	-1.552	91	96	96	96	87	57	57	57
Revised												
0	-0.533	-0.450	-0.454	-0.450	0	9	9	13	–	–	–	–
1	-0.697	-0.580	-0.585	-0.581	22	22	22	22	4	13	13	9
2	-0.674	-0.609	-0.617	-0.614	35	35	35	39	9	17	17	17
3	-0.730	-0.577	-0.545	-0.538	22	35	35	35	17	13	13	13
4	-0.927	-0.822	-0.839	-0.818	48	57	57	61	57	48	39	39
5	-1.544	-1.245	-1.321	-1.222	91	91	91	91	61	39	39	39
6	-1.511	-1.259	-1.284	-1.240	91	87	87	87	61	61	57	61
7	-1.431	-1.289	-1.376	-1.305	83	87	87	87	91	70	70	61
8	-1.392	-1.442	-1.448	-1.397	78	87	87	87	78	48	52	39
9	-1.420	-1.399	-1.453	-1.421	87	96	96	96	91	74	78	70
10	-0.621	-0.595	-0.599	-0.589	17	26	26	26	0	0	0	0
11	-1.277	-1.287	-1.285	-1.314	91	91	91	91	87	70	70	61
12	-0.710	-0.687	-0.700	-0.685	35	52	52	52	9	4	4	4
13	-1.501	-1.424	-1.469	-1.427	91	96	96	96	96	78	74	70
14	-1.689	-1.485	-1.532	-1.454	91	91	91	91	78	48	43	43
15	-1.751	-1.573	-1.601	-1.549	91	96	96	96	83	52	57	57

Notes: $\hat{\alpha}$ estimates are the mean results across the individual forecasters. $H_0: \alpha \geq 0$ and $H_0: J = 0$ show the rejection frequencies across forecasters on a 5 percent level.

Table C14: Results for instrument sets and Linex-Linex loss - HICP inflation

IV set	$\hat{\alpha}$ **estimates**				$H_0: \alpha \geq 0.5$				$H_0: J = 0$			
	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4
					Real-time							
0	0.440	0.441	0.441	0.441	4	13	13	13	–	–	–	–
1	0.394	0.396	0.396	0.397	26	26	26	26	0	0	0	0
2	0.370	0.380	0.380	0.380	52	39	39	39	0	0	0	0
3	0.360	0.380	0.381	0.381	39	39	39	39	9	9	9	9
4	0.332	0.338	0.338	0.339	74	78	74	74	4	4	4	4
5	0.268	0.286	0.286	0.287	96	96	96	96	0	0	0	0
6	0.229	0.249	0.249	0.250	96	100	100	100	0	0	0	0
7	0.251	0.264	0.263	0.264	96	100	100	100	0	0	0	0
8	0.251	0.260	0.260	0.261	96	100	100	100	0	0	0	0
9	0.198	0.213	0.213	0.214	100	100	100	100	0	0	0	0
10	0.421	0.414	0.414	0.414	22	30	30	30	0	0	0	0
11	0.238	0.254	0.254	0.255	96	96	96	96	0	4	9	9
12	0.388	0.387	0.387	0.387	44	48	48	48	0	0	0	0
13	0.186	0.209	0.209	0.210	100	100	100	100	0	0	0	0
14	0.250	0.257	0.257	0.258	96	96	96	96	4	4	4	4
15	0.234	0.236	0.235	0.236	100	100	100	100	0	0	0	0
					Revised							
0	0.442	0.443	0.443	0.443	4	13	13	13	–	–	–	–
1	0.391	0.393	0.393	0.394	26	30	30	30	0	0	0	0
2	0.366	0.372	0.372	0.373	48	44	44	44	0	0	0	0
3	0.369	0.374	0.375	0.375	26	44	44	44	4	4	4	4
4	0.337	0.342	0.342	0.343	74	70	70	70	4	4	4	4
5	0.270	0.288	0.288	0.288	96	96	96	96	0	0	0	0
6	0.237	0.256	0.256	0.257	96	96	96	96	0	0	0	0
7	0.255	0.267	0.267	0.268	96	100	100	100	0	0	0	0
8	0.254	0.264	0.263	0.264	96	100	100	100	0	0	0	0
9	0.202	0.217	0.216	0.218	100	100	100	100	0	0	0	0
10	0.423	0.416	0.416	0.416	22	26	26	26	0	0	0	0
11	0.241	0.257	0.257	0.258	96	96	96	96	0	0	4	4
12	0.393	0.395	0.395	0.395	39	39	39	39	0	0	0	0
13	0.191	0.212	0.213	0.214	100	100	100	100	0	0	0	0
14	0.254	0.261	0.261	0.261	96	96	96	96	4	4	4	4
15	0.237	0.239	0.238	0.238	100	100	100	100	0	0	0	0

Notes: $\hat{\alpha}$ estimates are the mean results across the individual forecasters. $H_0: \alpha \geq 0.5$ and $H_0: J = 0$ show the rejection frequencies across forecasters on a 5 percent level.

Table C15: Results for instrument sets and Linex loss - Unemployment rate

IV set	$\hat{\alpha}$ estimates				$H_0: \alpha \geq 0$				$H_0: J = 0$			
	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4
Real-time												
0	-0.507	-0.357	-0.364	-0.357	9	5	5	5	–	–	–	–
1	-0.502	-0.393	-0.404	-0.381	9	5	9	5	0	5	5	5
2	-0.933	-0.185	-0.338	-0.269	27	9	14	9	41	50	50	50
3	-0.519	-0.390	-0.386	-0.377	9	9	9	9	0	9	9	9
4	-1.110	-0.741	-0.748	-0.718	32	36	41	41	36	32	32	32
5	-1.539	-0.968	-0.924	-0.933	68	55	41	55	91	68	59	68
6	-1.156	-0.766	-0.784	-0.755	59	50	50	50	18	14	23	23
7	-0.339	-0.276	-0.286	-0.279	5	9	9	9	0	0	0	0
8	0.175	0.019	-0.011	-0.017	5	5	5	9	32	50	50	50
9	-0.327	-0.229	-0.230	-0.213	5	9	9	9	9	5	5	5
10	-0.580	-0.438	-0.454	-0.430	18	18	18	18	9	9	9	9
11	-1.192	-0.752	-0.750	-0.780	64	64	59	64	95	86	86	86
12	-0.630	-0.520	-0.487	-0.482	23	23	23	23	5	5	5	5
13	-1.095	-0.516	-0.492	-0.468	41	32	27	32	68	55	50	50
14	-1.084	-0.876	-0.790	-0.826	68	55	55	55	41	32	36	32
15	-0.763	-0.518	-0.545	-0.520	36	32	32	32	14	14	14	14
Revised												
0	-1.243	-0.880	-0.908	-0.845	27	27	32	36	–	–	–	–
1	-2.056	-1.324	-1.299	-1.232	68	68	68	68	18	18	14	14
2	-2.621	-1.487	-1.456	-1.385	86	91	86	91	41	23	23	23
3	-2.139	-1.825	-1.602	-1.650	68	68	73	73	32	14	18	14
4	-1.644	-1.090	-1.173	-1.027	68	68	55	73	27	32	32	32
5	-2.376	-1.525	-1.616	-1.457	91	77	77	82	59	55	59	55
6	-1.837	-1.165	-1.205	-1.065	82	77	68	77	9	23	23	27
7	-1.185	-0.776	-0.820	-0.748	36	27	32	32	5	14	14	14
8	-1.193	-0.780	-0.815	-0.747	46	50	46	50	18	27	27	27
9	-1.284	-0.816	-0.876	-0.771	41	36	36	41	9	5	5	5
10	-1.354	-0.909	-0.946	-0.882	55	50	50	59	5	0	0	0
11	-2.451	-1.454	-1.479	-1.325	86	86	86	86	86	82	82	82
12	-1.302	-0.912	-1.003	-0.920	59	64	68	68	5	0	5	5
13	-2.131	-1.255	-1.312	-1.082	82	73	64	68	73	77	73	77
14	-1.972	-1.282	-1.430	-1.106	82	86	82	82	50	45	55	45
15	-1.483	-1.060	-1.019	-0.914	59	55	55	55	14	14	14	14

Notes: $\hat{\alpha}$ estimates are the mean results across the individual forecasters. $H_0: \alpha \geq 0$ and $H_0: J = 0$ show the rejection frequencies across forecasters on a 5 percent level.

Table C16: Results for instrument sets and Linex-Linex loss - Unemployment rate

IV set	$\hat{\alpha}$ **estimates**				$H_0: \alpha \geq 0.5$				$H_0: J = 0$			
	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4	-	σ^2	s^3	k^4
Real-time												
0	0.434	0.432	0.430	0.429	14	14	14	14	–	–	–	–
1	0.428	0.420	0.418	0.417	23	27	27	27	0	0	0	0
2	0.418	0.420	0.419	0.418	46	36	41	41	27	32	32	23
3	0.427	0.421	0.419	0.419	23	27	27	27	0	0	0	0
4	0.372	0.387	0.386	0.386	59	55	55	50	55	36	36	36
5	0.319	0.336	0.334	0.335	73	68	68	68	45	36	36	36
6	0.370	0.395	0.390	0.389	64	55	55	55	14	23	18	18
7	0.439	0.432	0.431	0.430	14	14	14	14	0	0	0	0
8	0.559	0.534	0.532	0.530	9	14	14	14	5	5	5	5
9	0.406	0.405	0.404	0.404	36	27	27	27	14	5	5	5
10	0.422	0.415	0.414	0.414	32	27	27	27	5	5	5	5
11	0.363	0.381	0.380	0.381	73	59	59	59	82	64	59	59
12	0.434	0.404	0.402	0.402	27	41	41	41	5	5	0	0
13	0.367	0.386	0.388	0.389	68	68	68	64	36	32	32	32
14	0.392	0.384	0.379	0.378	64	59	59	59	27	18	23	23
15	0.416	0.407	0.406	0.406	46	50	55	55	5	5	5	5
Revised												
0	0.374	0.374	0.373	0.373	55	64	59	59	–	–	–	–
1	0.300	0.316	0.316	0.316	82	82	82	82	5	5	5	5
2	0.268	0.286	0.286	0.287	96	91	91	91	9	5	0	0
3	0.294	0.289	0.290	0.291	82	86	86	86	5	9	9	9
4	0.316	0.315	0.314	0.314	86	91	91	91	45	36	36	36
5	0.209	0.255	0.255	0.260	91	82	82	82	45	36	36	36
6	0.292	0.312	0.310	0.310	82	91	91	91	14	9	9	9
7	0.373	0.370	0.368	0.369	55	64	64	64	5	0	0	0
8	0.379	0.374	0.372	0.372	59	59	64	64	5	18	18	18
9	0.342	0.336	0.336	0.337	68	68	73	73	0	0	0	0
10	0.365	0.365	0.364	0.365	77	86	82	82	5	5	5	5
11	0.252	0.259	0.256	0.257	96	96	96	96	82	59	59	59
12	0.358	0.357	0.355	0.358	82	86	86	86	0	5	5	5
13	0.274	0.287	0.287	0.287	96	91	91	91	36	45	45	45
14	0.279	0.306	0.304	0.304	86	82	77	77	32	27	32	32
15	0.329	0.349	0.336	0.340	73	73	82	77	9	5	5	5

Notes: $\hat{\alpha}$ estimates are the mean results across the individual forecasters. $H_0: \alpha \geq 0.5$ and $H_0: J = 0$ show the rejection frequencies across forecasters on a 5 percent level.

Affirmation

I hereby declare that the dissertation entitled

“Macroeconomic Forecast Evaluation Under Asymmetric Loss”

is my own work. I have only used the sources indicated and have not made unauthorized use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given. I have not presented this thesis or parts thereof to an university as part of an examination or degree.

....	....
Place, Date	Signature