Dynamics or diversity? An empirical appraisal of distinct means to measure inflation uncertainty

March 22, 2012

Matthias Hartmann* and Helmut Herwartz§

Abstract

The evolution of the sovereign debt crisis currently commands nonconventional arrangements of monetary and fiscal policy. As a result, concerns about accelerating and less predictable inflation are shifting high up on the agenda of consumers, investors and monetary authorities alike. The question of how to measure uncertainty about inflation, however, is an open issue. The theoretical literature suggests that a suitable inflation uncertainty measure should improve interest rates predictions. Therefore we rank several candidate measures by their relative out-of-sample predictive content in the framework of an accordingly augmented Fisher relation. We document that measuring uncertainty by means of the diversity of expectations offers higher predictive content than dynamic specifications. Moreover, investigating alternative theoretical arguments, it is found that the impact of inflation uncertainty on bond yields is best explained as an inflation risk premium.

JEL classification: C52, C53, E37, E52
Keywords: Inflation uncertainty, Disparity of expectations, Sovereign bond yields, Fisher equation.

*Corresponding author. Christian-Albrechts-Universität zu Kiel, Institute of Statistics and Econometrics, Olshausenstraße 40-60, 24118 Kiel, Germany. Tel.: +49 431 8802224; Fax: +49 431 8807605. email: m.hartmann@stat-econ.uni-kiel.de.

§Christian-Albrechts-University at Kiel, Institute of Statistics and Econometrics. email: h.herwartz@stat-econ.uni-kiel.de.

1 Introduction

Inflation expectations are a crucial component in some of the most well known economic models. For example, the ex-ante formulation of the Fisher relation (Fisher 1930) relates interest rates directly to expectations about future inflation. Moreover, inflation expectations play a major role in distinct recitations of the Phillips curve (Friedman 1968, Phelps 1968, Clarida, Galí and Gertler 1999, Gordon 2009). Schwab (1982) describes how inflation expectations may affect the demand for housing through the anticipated impact of future inflation on mortgage payment schemes. Other examples of decision or negotiation outcomes affected by inflation forecasts are consumption smoothing (Reinhard and Végh 1994, Hördahl 2008), firms' investment and price setting (Levi and Makin 1979, Taylor 2000) or the determination of wages (Akerlof, Dickens and Perry 2000). Deviations of expectations from realised inflation are usually costly for the decision taker. Therefore, the more crucial the role of inflation expectations is, the more meaningful it is to explicitly consider the uncertainty about future inflation as an additional determinant of the decision process. The question about the importance of inflation uncertainty (abbreviated as IU hencefourth) effects has been examined in a sizeable empirical literature. Empirical studies, however, face the problem that IU cannot be directly observed. Several methods have been proposed to measure this latent quantity. The contribution of this study is a comparative assessment of some widely employed methods to measure IU. In addition, we highlight the role of IU in the current sovereign debt crisis by investigating how IU influences government bond yields.

The two most prominent families of IU approximations are time series based approaches like conditionally autoregressive heteroscedastic (GARCH, Engle 1982, Bollerslev 1986) processes and its descendants, and dispersion measures of forecast surveys

(Lahiri and Sheng 2010). The former draw explanatory content from historical observations. The latter emphasise the heterogeneity of opinions about future inflation. A widely used representative of the latter approach is the standard deviation of expert forecasts. Either of the two approaches relies on distinct sources of information. They also process information in different ways (Batchelor and Dua 1996, Mankiw and Reis 2004). Therefore, time series and survey-based methods might provide diverging estimates of IU in many situations. Lahiri and Liu (2005) find that such distinct IU indications deviate most during turbulent times, e.g. in the case of the US after the first oil price shock. This implies that the problem of which IU measure to choose is particularly difficult in circumstances of highest relevance.

In the related literature, the problem to choose from a set of potential IU measures has been recognised since several years. At least three distinct avenues to single out an empirical IU measure have emerged. Firstly, compliance with an economic definition of uncertainty may be a necessary condition of a meaningful IU approximation. For example, Giordani and Söderlind (2003) or Rich and Tracy (2003) point out that GARCH models does not express IU from an ex-ante point of view. Therefore they argue that measures based on forecast surveys are preferable to GARCH-based measures. Other objections to the GARCH model class are made on the grounds that these specifications are merely derived to fit stylised facts like volatility clustering in inflation data. Goodness-of-fit, however, may not be a sufficient criterion to evaluate IU measures, especially if the aim is to formulate an economic interpretation of IU (Peng and Yang 2008). A second way to select IU measures is to rely on statistical arguments. For example, Lahiri and Liu (2005) show that widely used dispersion measures such as the standard deviation of expert forecasts can be biased measures of IU. Third, one may rank alternative IU metrics according to the strength of their relation to a certain benchmark.

Bomberger (1996) employs a GARCH specification as the benchmark and compares the implied IU approximations to those obtained from the disagreement of expert forecasts. In a similar way, Chua, Kim and Suardi (2011) assess several distinct expressions of IU by means of their relation to a survey dispersion measure.

In this study, we propose an alternative way to examine the suitability of distinct IU measures. In Levi and Makin (1979) or Blejer and Eden (1979), the Fisher equation is augmented by IU terms. They document that estimates of the influence of inflation expectations on interest rates are more in line with theoretical predictions if IU terms are incorporated in the relation. We draw upon an augmented Fisher equation to rank IU measures according to their marginal predictive content. Forecasting proceeds in the framework of a Bayesian model averaging (BMA) approach. In this way, distinct model specification choices are linearly combined by means of exact posterior probability weights. Such methods have been documented to yield higher forecast performance than model selection or other forecast combination approaches (Koop and Potter 2003, Wright 2009). Moreover, we investigate the scope of IU measures to predict yields on long term government bonds for a large cross section of 18 developed economies. We focus on these securities because of their prominent role in the current debate on the sustainability of sovereign debt. Moreover, the risk to incur losses due to inflation is largest for debt obligations with long maturities. Missale and Blanchard (1994) argue that short-term debt obligations might be regarded as broadly equivalent to inflation-indexed bonds.

In large parts of the related literature, the concept of IU refers to the risk of welfare losses from surprises in future inflation. Therefore, it appears sensible to compare IU measures by means of ex-ante forecasting. Barnea, Dotan and Lakonishok (1979) or Friedman (1977) assert that IU influences anticipated returns on both financial or tangible assets. For example, Brenner and Landskroner (1983) describe how IU affects bond returns in the form of an inflation risk premium. In general, investment and savings decisions require the consideration of intertemporal tradeoffs with regard to streams of nominal income. Hence, the ex-ante predictive content of IU measures might be important for consumers, investors and also for the conduct of monetary policy. Moreover, due to the adverse nature of the theoretical explanations for the impact of IU, it is not clear if excess IU increases or reduces interest rates (Lahiri et al. 1988). Therefore, in addition to the forecast evaluation, we examine the direction of the IU effect on interest rates.

To summarise some results, we find that dynamic and dispersion measures differ significantly from each other in terms of their explanatory content for ex-ante interest rates. In terms of predictive content, both groups of IU measures compare favourably to benchmark IU measures as employed in the related literature. One candidate IU measure from the group of disparity statistics outperforms all other IU quantifications. Namely, the average over individual models' uncertainties contributes most to predictive accuracy. During the recent period of turmoil until the year 2011, the dynamics of IU as expressed by the dispersion statistics differ markedly from those of the dynamic IU metrics. It is also found that the distinction of forecast content is most clear-cut during turbulent times. Regarding the interpretation of the relation between IU and interest rates, we find that the IU impact can be described as an inflation risk premium.

In Section 2, alternative ways to express IU are introduced. Section 3 recalls distinct channels by which IU may influence interest rates. Subsequently, the forecasting methodology to select the most sensible IU measure from the set of candidates is introduced. Section 4 describes the data set and some particularities of the modelling framework. Forecasting results are discussed in Section 5, along with an interpretation of the impact of IU on interest rates from an economic point of view. Section 6 summarises and concludes.

2 Measuring IU

There is no unique way to define or measure IU. Hence, as a first step in empirical investigations, it has to be decided which IU measure shall be used. For the comparative evaluation of distinct IU metrics, we consider specifications which mimic commonly used dynamic and disparity approaches. To obtain measures of disparity, we replace forecast survey data by model based predictions. In this way it is guaranteed that distinct IU measures are conditional on sample information with equal timing. This is typically not the case when both survey data and aggregate time series are used to determine distinct expressions of IU. Surveyed experts might have access to more timely or private information while time series measures are confined to publicly available data (Rich and Tracy 2003). Moreover, replacing expert data by model forecasts makes our analysis less dependent on the availability of survey data. Thus, it is possible to consider a larger cross section of economies for robustness analysis. Finally, the models we propose do not require large samples of historical observations. A focus on recent data might account for potential changes in inflation regimes (Evans and Wachtel 1993). As has been noted by Lahiri and Liu (2005), the incorporation of regime shifts in the framework of estimated GARCH models is difficult because the timing of eventual changes is unknown. In the following, eight distinct measures of IU are discussed. We firstly consider time series based methods, which measure ex-ante IU by drawing upon historical sample information. As an alternative to these dynamic methods, 4 further approaches are based on the dispersion of individual forecasts.

Two of the uncertainty measures introduced below are based on a specification widely used for inflation forecasting, the linear autoregressive (AR) model. The success of AR models or random walk schemes in predicting inflation is documented in several empirical studies, including Canova (2007), Stock and Watson (2007, 2008) or Hartmann and Herwartz (2012). Given the widely documented predictive success of random walk specifications and allowing for the possibility of local trends in the inflation series, the AR scheme is formulated as

$$\pi_{t+\ell} = \alpha_0 + \alpha_1 t + \alpha_2 \pi_t + \epsilon_{t+\ell}, \quad t = \tau - B + 1, ..., \tau, \tag{1}$$

where $\epsilon_{t+\ell} \overset{iid}{\sim} (0, \sigma_\epsilon^2)$, $\ell \in \{1, 2, 3, 4\}$, is the forecast horizon, and B denotes the length of a (rolling) estimation sample window[1]. Out-of-sample forecasts implied by (1) are denoted $\hat{\pi}_{\tau+\ell|\tau}$, where $\tau = T_0 - \ell, ..., T - \ell$ is the rolling forecast origin. The time instances T_0 and T delimit the evaluation sample which is employed in the comparative evaluation of IU measures. In the following, we discuss distinct IU measures, some of which are also discussed in Hartmann and Herwartz (2012). We begin with an introduction of time series methods, then proceed to IU quantifications based on a cross section of inflation expectations.

1. Dynamic specifications

1.1 Predictive standard deviation

At forecast origin τ, the estimated predictive error standard deviation obtained from (1) is

$$\hat{\sigma}_{\tau+\ell|\tau} = \tilde{\sigma}_\epsilon \sqrt{(1 + \boldsymbol{z}_\tau' (Z_\tau' Z_\tau)^{-1} \boldsymbol{z}_\tau)}, \tag{2}$$

[1]Extracting inflation expectations form higher-order AR specifications obtains qualitatively equivalent results which are available from the authors upon request.

where Z_τ is the autoregressive design matrix and $\boldsymbol{z}_\tau$ are the most recent observations employed to obtain out-of-sample forecasts. The statistic in (2) is composed of time-local expressions of the variance of inflation surprises $\tilde{\sigma}_\epsilon$ and estimation uncertainty $\tilde{\sigma}_\epsilon \sqrt{\boldsymbol{z}'_\tau (Z'_\tau Z_\tau)^{-1} \boldsymbol{z}_\tau}$.

1.2 Exponential smoothing

Among the most prominent ways to measure IU are GARCH processes. To obtain ex-ante formulations of IU, however, this class of models is not uniformly recommendable (Hwang and Pereira 2006). As nonlinear specifications, estimated GARCH models are likely to suffer from inefficiencies if samples of moderate size are considered. We suggest an alternative which is designed to balance both the arrival of news and inertia in second-order inflation dynamics. Being related to the RiskMetrics exponential smoothing approach (Zangari 1996), this IU measure reads as

$$h^{(\lambda)}_{\tau+1|\tau} = \sqrt{\lambda(\Delta\pi_\tau)^2 + (1-\lambda)\overline{(\Delta\pi)^2}}. \tag{3}$$

In (3), $\Delta\pi_t = \pi_t - \pi_{t-1}$, and $\overline{(\Delta\pi)^2} = (1/(B-1)) \sum_{t=\tau-B+1}^{\tau-1} (\Delta\pi_t)^2$, where smoothing over past observations is restricted to express IU only by means of the most currently observed data[2]. The tradeoff between news response and past information is addressed by choosing $\lambda \in \{0.1, 0.2\}$ (Christoffersen and Diebold 2000). The prescription of parameter values for λ facilitates the quantification of the level of IU at the current end of sample information. We select $\lambda = 0.1, 0.2$ since such magnitudes are typically obtained as parameter estimates from GARCH models when quarterly inflation data is considered (Bollerslev 1986).

[2]Distinct forecasts based on alternative choices of the estimation window B obtain qualitatively equivalent results, which are available from the authors upon request.

1.3 Unanticipated volatility

The statistics in (2) and (3) are obtained as ex-ante quantifications of IU. An alternative IU indicator might be obtained as the realised prediction error

$$\hat{a}_{\tau+\ell} = |\hat{\pi}_{\tau+\ell|\tau} - \pi_{\tau+\ell}|, \tag{4}$$

from the AR model in (1). This measure expresses the common view that the ex-post track record of inflation forecasting success (or loss) may serve as an indicator of currently perceived inflation risk (Giordani and Söderlind 2003).

2. Measuring IU by means of expectation heterogeneity

A common way of measuring uncertainty is to exploit the variation across individual expectations. We model the dispersion of opinions by considering forecasts from $J = 5$ alternative linear forecasting specifications. Constituting models are the AR specification in (1) and four autoregressive distributed lag schemes which arise when lagged inflation in (1) is complemented with further predictors. A detailed description of these specifications along with further references is given in the Appendix.

2.1 Disagreement of expectations

Based on five rival predictions of inflation, the disagreement measure obtains as

$$\hat{s}_{\tau+\ell|\tau} = \sqrt{\frac{1}{J-1}\sum_{j=1}^{J}(\hat{\pi}_{j,\tau+\ell|\tau} - \overline{\pi}_{\tau+\ell|\tau})^2}, \tag{5}$$

with $\overline{\pi}_{\tau+\ell|\tau} = (1/J)\sum_{j=1}^{J}\hat{\pi}_{j,\tau+\ell|\tau}$. Variants of an analogous measure based on expert opinions are employed in numerous studies like e.g. Cukierman and Wachtel (1979) or Batchelor and Dua (1996). This measure has been critisised by Zarnowitz and Lambros (1987) to be biased in at least two situations. First, if the constituting forecasts are uncertain to a large extent but individuals mostly agree, the measure in (5) may provide a rather modest indication of IU. Conversely, there might be situations when individuals are very certain about the evolution of future inflation but their anticipations largely disagree. In both cases, IU as quantified by $\hat{s}_{\tau+\ell|\tau}$ may not be interpretable in a straightforward way.

2.2 Average uncertainty

In addition to (5), Zarnowitz and Lambros (1987) propose to average individual predictive standard deviations. Adapted to the forecasting models that we consider, this measure obtains as

$$\bar{\sigma}_{\tau+\ell|\tau} = \frac{1}{J}\sum_{j=1}^{J} \hat{\sigma}_{j,\tau+\ell|\tau}, \tag{6}$$

with $\hat{\sigma}_{j,\tau+\ell|\tau}$ denoting predictive standard deviations obtained according to (2) for the AR scheme in (1) and the inflation forecast models listed in the Appendix. Although it is based on distinct time-series models, we regard $\bar{\sigma}_{\tau+\ell|\tau}$ as a dispersion IU measure like $\hat{s}_{\tau+\ell|\tau}$. We classify IU metrics in this way because both entail characteristics which only arise as a matter of pooling. For example, $\bar{\sigma}_{\tau+\ell|\tau}$ is less likely to obtain 'eccentric' (Zarnowitz and Lambros 1987) quantifications of IU than the individual dynamic IU statistics on which the combination is based. As Zarnowitz and Lambros (1987) note, $\bar{\sigma}_{\tau+\ell|\tau}$ may be interpreted as a combination of IU forecasts. Forecast combinations have been found to provide superior predictions of various macroeconomic processes

(Bates and Granger 1969, Timmerman 2006). In the theoretical literature, no optimality conditions for combinations of interval forecasts are provided (Wallis 2005). Forecast combination strategies for predictions of conditional second moments have been evaluated by Becker and Clements (2008) or by Patton and Sheppard (2009), who investigate volatility forecasts for the S&P500 index and IBM stock returns, respectively. In both cases, averages of single model based volatility forecasts as in (6) cannot be outperformed by any competing prediction scheme.

2.3 Augmenting the disagreement measure

As noted by Lahiri and Liu (2005), estimates of IU like $\hat{\sigma}_{\tau+\ell|\tau}$ in (2) might be characterised by individual biases. They suggest a combination of (5) and (6), given by[3]

$$\varsigma_{\tau+\ell|\tau} = 0.5(\hat{s}_{\tau+\ell|\tau} + \bar{\sigma}_{\tau+\ell|\tau}). \tag{7}$$

In cases when individual biases in $\hat{\sigma}_{j,\tau+\ell|\tau}$ are not symmetrically distributed around $\bar{\sigma}_{\tau+\ell|\tau}$, the resulting bias in (6) might be balanced by the disagreement term $\hat{s}_{\tau+\ell|\tau}$ in (7). For situations when surveyed experts report individual forecasts of density functions, Lahiri and Liu (2005) or Wallis (2005) point out the equivalence of this measure to the variance of combined forecast density functions (cf. Diebold et al. 1999, Giordani and Söderlind 2005).

Finally, Lahiri and Sheng (2010) propose a combination of disagreement measures and IU quantifications from GARCH models as a further ex-ante approximation of IU.

[3] In Lahiri and Liu (2005), the scaling factor of 0.5 is not applied. We specify $\varsigma_{\tau+\ell|\tau}$ in this way to align our analysis with the literature on forecast combinations, where the sum of combination weights is typically constrained to unity.

We determine a similar combination measure as

$$\zeta_{\tau+\ell|\tau} = 0.5(\hat{s}_{\tau+\ell|\tau} + h^{(0.1)}_{\tau+1|\tau}), \tag{8}$$

where the exponential smoothing measure $h^{(0.1)}_{\tau+1|\tau}$ is regarded as a substitute to GARCH quantifications.

3 Evaluation of empirical IU measures

In this section, the methodology to evaluate alternative ex-ante IU measures is described. We start by recalling theoretical assertions on how IU might be economically relevant. In several studies, it is suggested that IU matters for both macroeconomic policy and individual investment decisions. Based on such theoretical arguments, we extend the Fisher equation by incorporating candidate IU measures one after the other. Then, the predictive performance of the respective model reformulations yields a ranking of IU measures based on their predictive content.

3.1 Theoretical hypotheses on the impact of IU

Numerous theoretical contributions like Barnea et al. (1979) or Grauer and Litzenberger (1980) discuss the influence of IU on interest rates. However, distinct theories predict opposite signs of the impact of IU on interest rates. Blejer and Eden (1979), for example, argue that IU should have a negative effect on interest rates as a result of a diminishing demand for loanable funds. This might occur, for example, if firms delay previously intended investment projects due to the uncertainty about real payoffs. Conversely, Barnea et al. (1979) or Brenner and Landskroner (1983) describe how a positive influence of IU may arise in the form of an inflation risk premium. Recent empirical studies, where the presence of inflation risk premia is documented for distinct interest rates include Buraschi and Jiltsov (2005) or Gürkaynak et al. (2010).

3.2 The modelling framework

In the following, we describe the empirical models employed to assess the strength and the direction of the IU impact on interest rates. We determine ℓ-periods-ahead predictions regarding interest rates, denoted $\hat{R}_{\tau+\ell|\tau}$, by means of an autoregressive distributed lag (ADL) model. The ADL scheme reads as

$$R_{\tau+\ell} = \gamma_{10} + \gamma_{11}\tau + \sum_{p=1}^{P}\gamma_{12,p}\pi_{\tau-p+1} + \sum_{p=1}^{P}\gamma_{13,p}R_{\tau-p+1} + \sum_{p=1}^{P}\gamma_{14,p}IU_{\tau-p+\ell+1|\tau} + e_{\tau+\ell} \tag{9}$$

with $\tau = T_0 - \ell, ..., T - \ell$, the term $IU_{\tau+\ell|\tau}$ representing a particular inflation uncertainty measure and $e_{\tau+\ell} \overset{iid}{\sim} (0, \sigma_e^2)$. The formulation in (9) corresponds to the 'augmented Fisher relation' in Blejer and Eden (1979) or Levi and Makin (1979). Based on this model, the value of distinct IU measures is assessed by means of their potential to improve predictions of $R_{\tau+\ell}$. Particularly, the overall impact of IU on $R_{\tau+\ell}$, $\bar{\gamma}^{(IU)} = \sum_{p=1}^{P}\gamma_{14,p}$, indicates if it might be interpreted as a risk premium or as an impediment to aggregate investment.

4 Data and implementation details

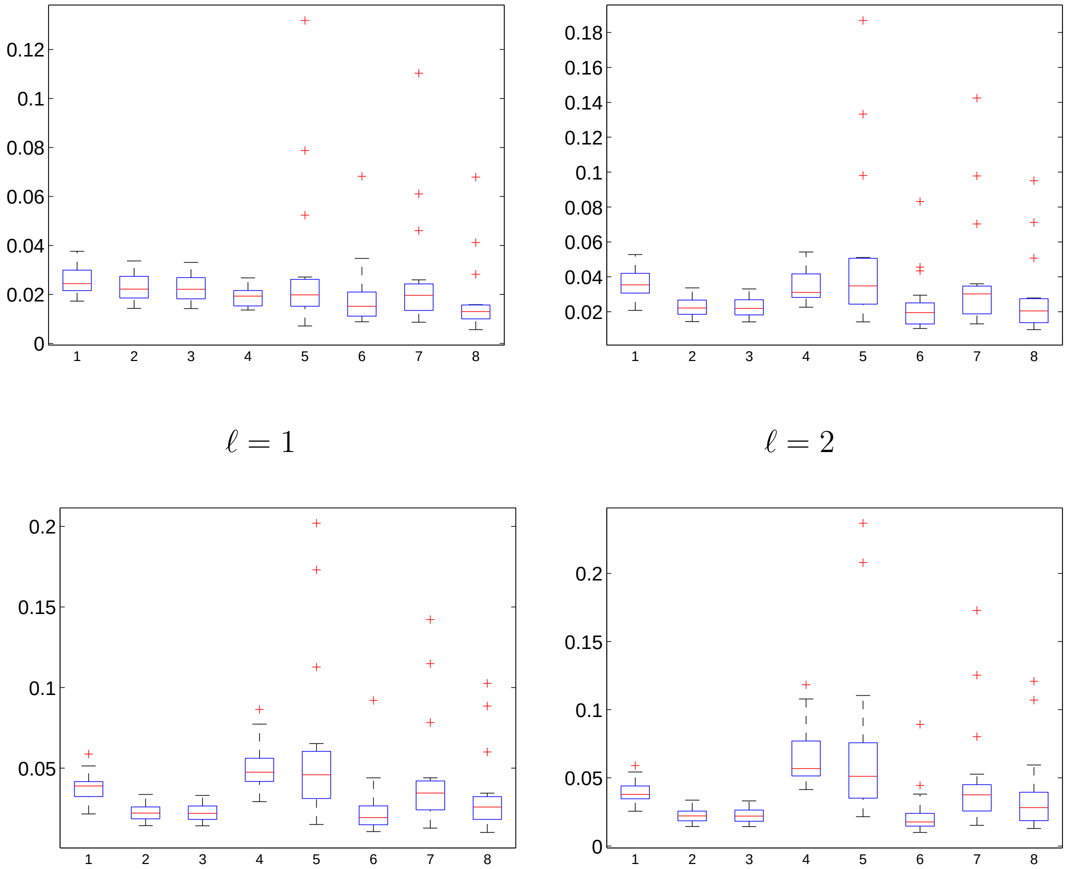
In this Section, the comparative evaluation of distinct IU measures is described and respective results are discussed. We firstly introduce the data set and discuss statistical properties of distinct IU measures. Next, certain particularities of the forecasting design are described. The outcomes of the forecasting study are summarised and interpreted subsequently.

4.1 Data set

The data set comprises quarterly observations for 18 developed economies, namely Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, UK and the US for the time period 1988Q1 to 2011Q1, with T_0 referring to the initial observation. A focus on more recent time periods helps to avoid the incorporation of observations from structurally distinct regimes. The target variable $R_{\tau+\ell}$ represents annual constant maturity yields on government bonds with a contract length of about 10 years issued in quarter $\tau+\ell$. Data on consumer prices, denoted CPI_τ, are used to express inflation as annualised quarterly price changes, i.e. $\pi_t = \ln(CPI_\tau/CPI_{\tau-4})$. Further series in the data set are industrial production, oil prices and a monetary aggregate. All series are seasonally adjusted and drawn from Datastream. The use of annual yield data avoids difficulties associated with the matching of anticipation horizons inherent in the series $R_{\tau+\ell}$ and $IU_{\tau+\ell|\tau}$ which are encountered if yields to distinct maturities are considered (Batchelor and Dua 1996). In this data set, both $R_{\tau+\ell}$ and the inflation rate $\pi_{\tau+\ell}$ on which IU series are based refer to annual rates.

4.2 Descriptive analysis of IU measures

Before turning to the assessment of IU measures' predictive ability, some features of the alternative IU statistics like their relative magnitudes or mutual correlations are discussed in the following. The Box-and-Whisker plots in figure 1 show the magnitudes and variation of distinct IU approximations for the 18 sample economies. The plots depict the median of $IU_{\tau+\ell|\tau}$ over the period between 1988Q1 and 2011Q1. The magnitudes of distinct IU quantifications are also compared in studies like Batchelor and Dua (1996) or Bomberger (1996). Batchelor and Dua (1996) find that the size of the standard deviation of individual forecasts exceeds the aggregate over individual forecasters' uncertainties. In their study, the former IU statistic is equivalent to the disagreement measure in (5) and the latter is obtained in a similar way to (6). Lahiri and Sheng (2010) document that the disagreement of individual predictions is typically larger than the average over idiosyncratic uncertainties. However, this ordering is reversed during certain time periods. In contrast, Zarnowitz and Lambros (1987) and Lahiri et al. (1988) find that the average of individual uncertainty estimates is typically larger than the disagreement statistic. This might be regarded as an indication of potential biases in the disagreement statistic. Zarnowitz and Lambros (1987) concede that such findings might be partly the result of a rather small sample size. The relative magnitudes of $\hat{s}_{\tau+\ell|\tau}$ and $\bar{\sigma}_{\tau+\ell|\tau}$ are, for our data, in line with the findings of Bomberger (1996), Batchelor and Dua (1996) and Lahiri and Sheng (2010). The disagreement statistics indicate, on average over economies, a higher level of IU than the average over individual uncertainties as expressed by $\bar{\sigma}_{\tau+\ell|\tau}$. The distinction between dynamic and disparity measures is also clearly reflected in their respective average magnitudes. All dispersion measures indicate considerably higher levels of IU than the dynamic IU measures. The median of all dynamic IU quantifications is scaled up by a factor of 5 to facilitate comparisons in one graph. The relation between the magnitude of time-series based IU measures and the disagreement statistics stands in contrast to findings of Bomberger (1996). In the study of Bomberger (1996), the level of IU as measured by means of ARCH specifications typically exceeds the magnitude of the disagreement of inflation forecasts collected in the *NBER-ASA Quarterly Economic Outlook Survey* data set. The level of all IU quantifications increases with the forecast horizon ℓ. The $h^{(\lambda)}_{\tau+1|\tau}$ measures are only determined for $\ell = 1$, and are of smaller magnitude than the other IU measures obtained for higher horizons. In addition to the comparison of relative magnitudes of IU measures based on dynamic and disagreement measures, Bomberger (1996) and Batchelor and Dua (1993) document that the latter approaches yield higher variation in anticipated IU. The variation of our IU measures reinstates these findings. This can be seen from the cross-sectional variation depicted in the Box-and-Whisker plots. The dispersion measures of IU, notably the disagreement metric, are characterised by higher variability across economies than the dynamic IU statistics.



$\ell = 1$ $\ell = 2$

$\ell = 3$ $\ell = 4$

Figure 1: Box-and-Whisker plots for th median IU over 1988Q1 to 2011Q1, in 18 economies. The numbers on the abscissa refer to the following IU measures: 1.) $\hat{\sigma}_{\tau+\ell|\tau}$, 2.) $h^{(0.1)}_{\tau+1|\tau}$, 3.) $h^{(0.2)}_{\tau+1|\tau}$, 4.) $\hat{a}_{\tau}$, 5.) $\hat{s}_{\tau+\ell|\tau}$, 6.) $\bar{\sigma}_{\tau+\ell|\tau}$, 7.) $\varsigma_{\tau+\ell|\tau}$, 8.) $\zeta_{\tau+\ell|\tau}$.

The number of distinct IU approximations discussed in the related literature is an indication of how difficult it can be to select a suitable IU measure. This ambiguity can be best addressed by means of comparing a set of IU metrics which provide, to some extent, idiosyncratic information. Hence, we report correlations of IU measures to ascertain the degree to which alternative IU approximations deliver overlapping or mutually exclusive information. Respective correlation coefficients are reported in table 1. The a priori classification of IU measures into dynamic and dispersion metrics is empirically confirmed by the correlation numbers. The relation between dynamic measures on the one hand and dispersion statistics on the other hand does not appear to be very strong. Correlations are particularly high between the dispersion measures $\hat{s}_{\tau+\ell|\tau}$, $\varsigma_{\tau+\ell|\tau}$ and $\zeta_{\tau+\ell|\tau}$. The relations between time series based measures $\hat{\sigma}_{\tau+\ell|\tau}$, $h^{(\lambda)}_{\tau+1|\tau}$ and $\hat{a}_\tau$ are smaller but still markedly positive.

Table 1: Mutual correlations of IU measures for $\ell = 1$

	$\hat{\sigma}_{\tau+\ell\|\tau}$	$h^{(0.1)}_{\tau+1\|\tau}$	$h^{(0.2)}_{\tau+1\|\tau}$	$\hat{a}_\tau$	$\hat{s}_{\tau+\ell\|\tau}$	$\bar{\sigma}_{\tau+\ell\|\tau}$	$\varsigma_{\tau+\ell\|\tau}$
$h^{(0.1)}_{\tau+1\|\tau}$	0.70	·	·	·	·	·	·
$h^{(0.2)}_{\tau+1\|\tau}$	0.69	0.98	·	·	·	·	·
$\hat{a}_\tau$	0.44	0.45	0.56	·	·	·	·
$\hat{s}_{\tau+\ell\|\tau}$	0.39	0.13	0.12	0.11	·	·	·
$\bar{\sigma}_{\tau+\ell\|\tau}$	0.45	0.19	0.18	0.09	0.72	·	·
$\varsigma_{\tau+\ell\|\tau}$	0.45	0.16	0.15	0.11	0.97	0.85	·
$\zeta_{\tau+\ell\|\tau}$	0.43	0.18	0.18	0.14	0.99	0.72	0.97

Cell entries represent average correlation coefficients across 18 economies.

In addition, we compare the model based IU metrics and two well known reference IU measures which are frequently employed in the related literature. As a first benchmark, we obtain IU estimates based on estimated GARCH(1,1) models (Bollerslev 1986). Similar to the $h^{(\lambda)}_{\tau+1|\tau}$ measures, the GARCH estimates are only obtained for $\ell = 1$. For most of the 18 economies we find rather high and positive correlations between IU measures and GARCH. In particular, the dynamic measures $\hat{\sigma}_{\tau+\ell|\tau}$ and $h^{(\lambda)}_{\tau+1|\tau}$ seem to be strongly related to the GARCH measure.

	$\hat{\sigma}_{\tau+1\|\tau}$	$h^{(0.1)}_{\tau+1\|\tau}$	$h^{(0.2)}_{\tau+1\|\tau}$	$\hat{a}_{\tau}$	$\hat{s}_{\tau+1\|\tau}$	$\bar{\sigma}_{\tau+1\|\tau}$	$\varsigma_{\tau+1\|\tau}$	$\zeta_{\tau+1\|\tau}$
Austria	-0.01	-0.20	-0.19	0.01	0.45	**0.58**	0.50	0.45
Belgium	0.07	-0.11	-0.09	0.01	0.20	**0.26**	0.23	0.18
Canada	-0.37	-0.17	-0.17	-0.18	0.17	**0.34**	0.21	0.16
Denmark	**0.17**	0.08	0.09	0.13	0.08	0.03	0.06	0.08
Finland	**0.42**	0.33	0.34	0.16	-0.14	0.17	-0.01	-0.11
France	-0.41	-0.18	-0.18	-0.29	-0.21	**0.02**	-0.16	-0.22
Germany	0.39	-0.09	-0.09	0.00	0.75	**0.83**	0.80	0.75
Ireland	**0.42**	0.11	0.10	0.02	0.26	0.04	0.21	0.26
Italy	**0.74**	0.28	0.31	0.48	0.31	0.38	0.36	0.31
Japan	0.49	-0.07	-0.08	0.00	0.68	0.65	**0.69**	0.68
Netherlands	**0.57**	0.26	0.27	0.11	0.30	0.43	0.36	0.31
Norway	0.15	0.35	0.36	0.23	0.46	**0.50**	0.48	0.46
Portugal	0.78	0.46	0.45	0.14	0.34	**0.83**	0.72	0.42
Spain	-0.72	-0.49	-0.46	-0.10	**0.03**	-0.06	0.01	0.01
Sweden	0.14	0.02	0.02	0.01	0.53	0.53	**0.54**	0.53
Switzerland	-0.01	-0.21	-0.21	-0.08	0.39	**0.44**	0.41	0.37
UK	0.45	0.04	0.03	0.01	0.59	**0.68**	0.62	0.59
US	-0.10	-0.24	-0.24	-0.15	0.08	**0.34**	0.15	0.05

The table reports correlations between distinct IU measures introduced in (2) to (8) for $\ell = 1$ and IU as implied by a GARCH(1,1) model for the time between 1988Q1 and 2011Q1. For each economy, the highest correlation between $u_{\tau+4|\tau}$ and one of the model based IU statistics appears in boldface. Respective results for $\ell > 1$ are qualitatively similar and available from the authors on request.

Furthermore, we examine how closely model based approaches are linked to survey based IU quantifications. The corresponding benchmark IU estimate is based on the inflation expectation survey provided by the Center for European Economic Research (ZEW) for the time period between 1992Q1 until 2011Q1. This data set reports percentages out of 350 respondents who expect inflation either to rise or to remain at most equal during the year after each wave of the survey. We denote by $\hat{\mathcal{P}}$ the percentage of respondents who expect a rising inflation rate. Then, the standard deviation of $\hat{\mathcal{P}}$ which is given by

$$u_{\tau+4|\tau} = \sqrt{\frac{\hat{\mathcal{P}}(1-\hat{\mathcal{P}})}{350}} \tag{10}$$

is considered as a survey-based measure of IU. In Table 2, cell entries denote correlation coefficients between $u_{\tau+4|\tau}$ and the IU measures (2) to (8).

	$\hat{\sigma}_{\tau+\ell\|\tau}$	$h^{(0.1)}_{\tau+1\|\tau}$	$h^{(0.2)}_{\tau+1\|\tau}$	$\hat{a}_{\tau}$	$\hat{s}_{\tau+\ell\|\tau}$	$\bar{\sigma}_{\tau+\ell\|\tau}$	$\varsigma_{\tau+\ell\|\tau}$	$\zeta_{\tau+\ell\|\tau}$
$\ell = 1$								
Germany	0.09	0.39	**0.40**	0.26	-0.28	-0.20	-0.30	-0.26
France	0.17	0.06	0.10	0.21	0.35	**0.37**	0.37	0.37
Italy	-0.23	-0.24	-0.20	0.08	0.11	**0.14**	0.12	0.09
Japan	0.17	0.11	0.15	0.20	0.13	**0.20**	0.15	0.13
UK	0.04	0.18	0.22	**0.29**	-0.13	0.20	0.06	-0.12
US	0.12	0.21	0.24	**0.32**	-0.01	-0.00	-0.01	0.03
$\ell = 2$								
Germany	-0.11	0.36	**0.38**	0.18	-0.33	-0.36	-0.44	-0.31
France	0.08	-0.02	-0.00	-0.02	0.35	0.33	**0.36**	0.35
Italy	-0.12	-0.31	-0.29	0.08	0.23	0.23	**0.24**	0.23
Japan	0.15	0.02	0.07	0.12	0.31	0.32	**0.32**	0.31
UK	-0.01	0.05	0.09	0.07	0.02	**0.17**	0.09	0.02
US	-0.02	0.12	**0.15**	0.08	-0.12	-0.08	-0.12	-0.10
$\ell = 3$								
Germany	-0.23	0.34	**0.34**	0.07	-0.35	-0.48	-0.52	-0.34
France	0.18	-0.07	-0.06	0.08	**0.43**	0.23	0.41	0.42
Italy	-0.01	-0.37	-0.36	**0.21**	0.12	0.19	0.14	0.12
Japan	0.13	-0.06	-0.03	0.03	0.37	0.30	0.37	**0.37**
UK	-0.04	-0.08	-0.03	-0.10	0.07	**0.08**	0.07	0.06
US	-0.05	0.07	**0.08**	-0.02	-0.13	-0.13	-0.14	-0.12
$\ell = 4$								
Germany	-0.30	0.34	**0.35**	-0.06	-0.31	-0.52	-0.52	-0.29
France	0.31	-0.09	-0.09	0.29	**0.36**	0.14	0.29	0.36
Italy	-0.02	-0.41	-0.41	**0.26**	0.00	0.12	0.02	-0.01
Japan	0.12	-0.04	-0.02	-0.00	0.29	0.15	0.28	**0.29**
UK	-0.03	-0.21	-0.17	-0.11	-0.01	**0.03**	0.00	-0.01
US	-0.06	**0.04**	0.04	-0.14	-0.17	-0.15	-0.18	-0.16

The table reports correlations between distinct IU measures introduced in (2) to (8) for $\ell = 1$ and IU as implied by a disagreement of experts surveyed by the ZEW. For each economy, the highest correlation between $u_{\tau+4|\tau}$ and one of the model based IU statistics appears in boldface.

The table entries indicate that in most economies, model- and survey based expressions of IU are positively associated. In the case of Germany and the US, however, correlations between the model based and survey measures are predominantly negative. The dynamic and the dispersion based IU statistics seem to provide rather distinct sorts of information. Respective linkages to the benchmark do not show a strong common pattern. The correlations of distinct disparity measures with the survey based IU are largely similar in magnitude. The same holds for the dynamic IU metrics $\hat{\sigma}_{\tau+\ell|\tau}$ and $h^{(\lambda)}_{\tau+1|\tau}$. The $\hat{a}_\tau$ statistic, however, exhibits a correlation pattern which is rather distinct from the other IU measures. A slightly stronger association is found between the survey measure and the model based dispersion statistics. Hence we find some evidence that the dispersion measures are related to the benchmark approaches of IU measurement in a more pronounced way than the dynamic IU metrics. Lahiri and Sheng (2010) or Chua, Kim and Suardi (2011) report correlation statistics for the US, where long samples of survey expectation data are available. They document high correlations between IU quantifications based on survey data and those determined e.g. from GARCH models, which range from small negative up to magnitudes of to 0.9, depending on the sample period and the anticipation horizon. Similarly, Zarnowitz and Lambros (1987) report correlations between survey-based representations of (5) and (6), which are varying between -0.29 and 0.74 for distinct forecast horizons.

Furthermore, we complement the correlation analysis by providing a graphical impression about the evolution of distinct IU approximations during the sample period. Distinct IU series are plotted as the median across economies. For all IU measures, a large reduction of the average level of IU is indicated before the beginning of the financial crisis in its most severe form in 2008. This development parallels the reduction and stabilisation of international inflation dynamics known as the Great Moderation (Giannone and Surico 2006, Benati 2008). During the crisis period, however, the signal delivered by dynamic IU metrics markedly differs from the one provided by the disparity statistics. The uprise of IU after 2008 as indicated by the dynamic IU measures exceeds the initially high level of IU indicated before the year 1990. The IU quantifications based on disparity statistics, however, remain at the average level prior to 2008. This is also reflected in the trajectory of the benchmark IU measure $u_{\tau+4|\tau}$ which is based on the ZEW survey data. The statistic $u_{\tau+4|\tau}$ shows a temporary uprise in IU during the years 2008 and 2009, however, the overall level of IU as indicated by this measure has been relatively high before the financial crisis. This underscores the assertion of Lahiri and Liu (2005) that it might be especially difficult to determine the appropriate IU measure during turbulent times. The findings from this preliminary data analysis suggest that the problem of selecting an IU measure might amount to the choice between a dynamic and a dispersion statistic. Differences among the candidate measures from one of these groups appear to be less pronounced. In the following discussion, the merits of distinct IU measures are evaluated by means of out-of-sample forecasting.

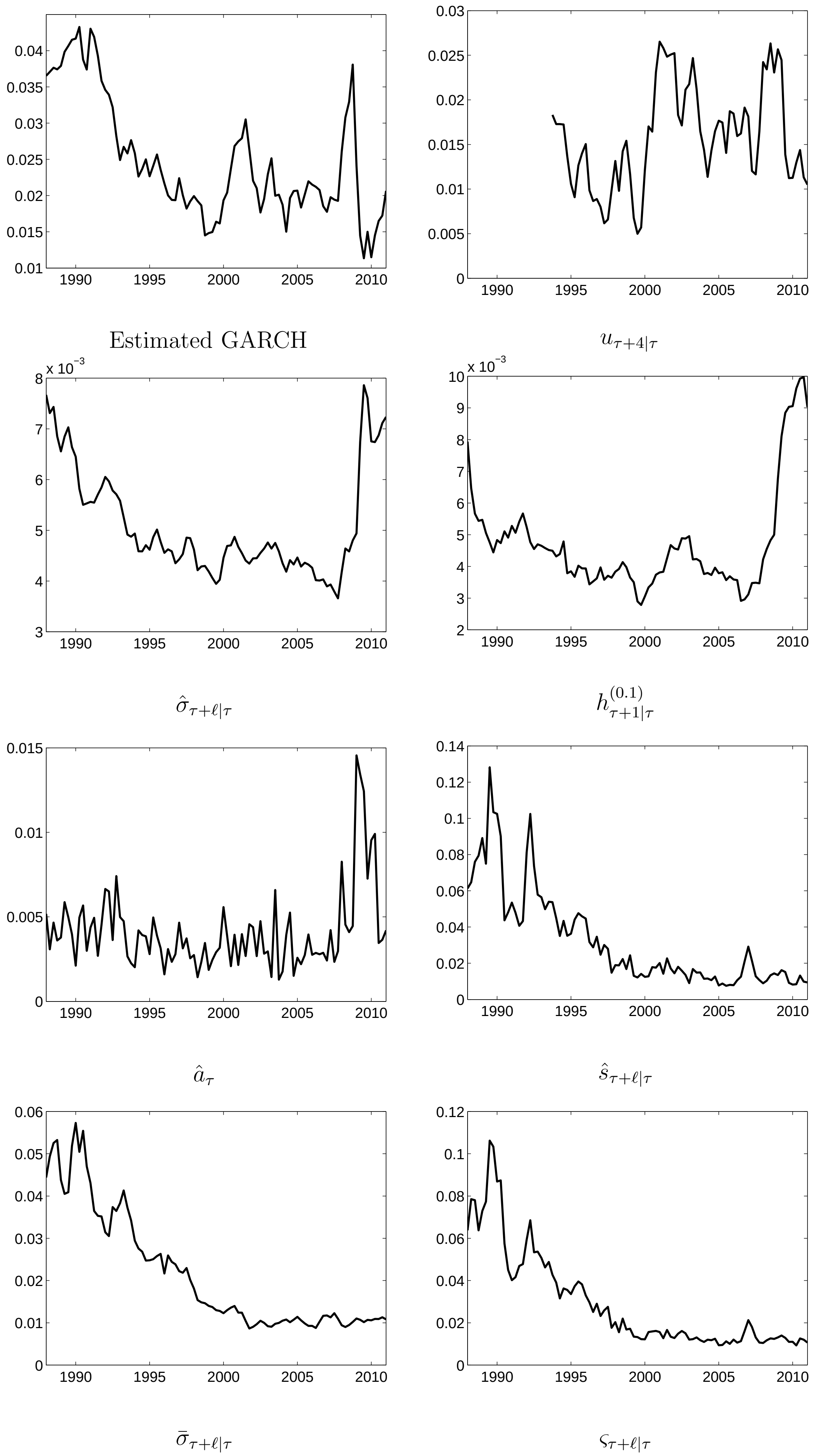


Figure 2: The plots show the median across economies of IU estimates based on distinct models, for a forecast horizon of $\ell = 1$. The IU measure $u_{\tau+4|\tau}$ is based on survey data on inflation expectations provided by the ZEW. Trajectories for $\ell > 1$ are qualitatively equivalent and available from the authors upon request.

4.3 The forecasting design

Based on the ex-ante formulation in (9), we measure the ability of modelling strategies dependent on the IU measure in question to forecast interest rates. Forecasting performance is assessed in the framework of pseudo out-of-sample cross validation. Similar to cross-validation (CV) techniques, each observation $R_{\tau+\ell}$ from the considered period $\tau = T_0 - \ell, ..., T - \ell$ is predicted ℓ-steps ahead by means of a respective leave-one-out estimate. The computation of ℓ-steps-ahead predictions is straightforward due to the linear relation between $R_{\tau+\ell}$ and the explanatory variables which are conditional on information up to period τ. This way of forecasting amounts to the so-called direct multistep prediction method (Chevillon 2005). With a maximum lag order of $P = 4$, distinct predictor selection procedures regarding the predetermined variables π_τ, R_τ and $IU_{\tau+\ell|\tau}$ gives rise to a total of $M = 2^{12}$ distinct specifications of (9). One way to address the uncertainty about the best forecasting model for interest rates is to select a most suitable candidate out of the multitude of distinct model specifications. Alternatively, the literature on forecast combinations suggests that the incorporation of inferior models is often beneficial in terms of predictive accuracy (Timmermann 2006). A sizeable literature has documented that linear combinations of forecasts outperform predictions based on the selection of one particular model (Avramov 2002, Cremers 2002). Relying on the scope of forecast combinations to improve forecast accuracy, we consider all M subset models of the augmented Fisher relation and subsequently combine the corresponding forecasts by means of BMA. The following exposition of a feasible BMA procedure which relies on exact posterior probabilities follows Wasserman (2000). The combined predictions obtain as

$$\hat{R}_{\tau+\ell|\tau} = \sum_{m=1}^{M} w_m^* \hat{R}_{\tau+\ell|\tau}^{(m)}, \tag{11}$$

with

$$w_m^* = \frac{w_m}{\sum_m w_m} \text{ and } w_m = \int L_m(\boldsymbol{\gamma}^{(m)}) p_m(\boldsymbol{\gamma}^{(m)}) d\boldsymbol{\gamma}^{(m)}. \tag{12}$$

In (12), $L_m(\boldsymbol{\gamma}^{(m)})$ and $p_m(\boldsymbol{\gamma}^{(m)})$ represent the likelihood and the a-priori distribution regarding the parameters $\boldsymbol{\gamma}^{(m)}$ from $m = 1, ..., M$ reformulations of (9), respectively. Based on the log-likelihood function $l(\boldsymbol{\gamma}^{(m)}) = \ln L(\boldsymbol{\gamma}^{(m)})$, exact posterior probabilities w_m in (12) can be approximated as

$$\ln \hat{w}_m = l(\hat{\boldsymbol{\gamma}}^{(m)}) - \frac{n_m}{2} \ln(T - T_0), \tag{13}$$

where $\hat{\boldsymbol{\gamma}}^{(m)}$ denotes the (Q)ML estimator of $\boldsymbol{\gamma}^{(m)}$ and n_m stands for the number of right hand side variables in model m. A feasible rule to compute forecast combination weights is to replace w_m in (12) by $\exp\left(l(\hat{\boldsymbol{\gamma}}^{(m)}) - \frac{n_m}{2}\ln(T - T_0)\right)$. For horizons $\ell > 1$, the weights w_m^* in (12) might not be strictly suitable because of potential serial correlation in the forecast errors, and thus, misspecification of the likelihood function (Wright 2009). For this reason, we provide forecasting results for combined ℓ-step-ahead predictions which are drawn from likelihood estimates determined under $\ell = 1$ in addition to the ones obtained for $\ell > 1$. This alternative approach reveals the extent to which our findings are affected by potentially biased weights for higher forecast horizons.

4.4 Forecasting strategies and performance measurement

In (9), alternative IU measures are employed interchangeably as conditioning variables, which yields a set of distinct forecasts. From the set of disparity statistics, the IU measures (6) to (8) are determined as averages of other IU statistics. The construction of these IU metrics is motivated in section 2 and in the studies of Wallis (2005) or Lahiri and Sheng (2010). Additionally, we consider several IU measures which are determined as the maximum, the minimum and the median of the IU metrics (2) to (8). Moreover, evidence regarding the comparison between dynamic and dispersion IU measures is provided by incorporating the average over the time-series IU metrics on the one hand and the disparity IU measures on the other hand. These statistics are denoted as $\overline{(TS)}$ and $\overline{(DS)}$, respectively. Finally, the predictive content of dynamic and dispersion IU measures is compared with the GARCH(1,1) and the $u_{\tau+4|\tau}$ benchmark measures introduced in section 4.2 for anticipation horizons $\ell = 1$ and $\ell = 4$, respectively. This provides insight on the relative performance of the model based IU measures to widely used of quantifying IU. In case of $u_{\tau+4|\tau}$, the comparison is limited to a cross section of six economies as listed in section 4.2 and the time period between 1992Q2 and 2011Q1. A ranking of IU measures is constructed by means of mutual comparisons of absolute forecast errors (AE) as given by

$$|e^{\bullet}_{i,\tau+\ell|\tau}| = |\hat{R}^{\bullet}_{i,\tau+\ell|\tau} - R_{\tau+\ell}| \tag{14}$$

for economies $i = 1, ..., 18$, where '$\bullet$' indicates that forecasts are obtained for distinct IU measures, $\bullet \in \{\hat{\sigma}_{\tau+\ell|\tau}, h^{(\lambda)}_{\tau+1|\tau}$,
$\hat{a}_{\tau}, \hat{s}_{\tau+\ell|\tau}, \bar{\sigma}_{\tau+\ell|\tau}, \varsigma_{\tau+\ell|\tau}, \zeta_{\tau+\ell|\tau}, \max(\text{IU}), \min(\text{IU}), \text{median}(\text{IU}), \overline{(TS)}, \overline{(DS)}\}$. The percentage of cases when an IU measure obtains predictions which are among the q best is then

expressed as

$$\text{TOP}q^{\bullet} = (1/((T - T_0 + 1) \times 18)) \sum_{\tau=T_0-\ell}^{T-\ell} \sum_{i=1}^{18} \mathrm{I}(|e^{\bullet}_{i,\tau+\ell}| \leq |e^{(q)}_{i,\tau+\ell}|), \tag{15}$$

where $\mathrm{I}(\cdot)$ stands for the indicator function and $|e^{(q)}_{i,\tau+\ell}|$ denotes the q-th smallest absolute prediction error (Stock and Watson 1999).

It seems unlikely that the predictive content provided by individual IU measures remains largely equivalent for distinct times or economies. For example, Lahiri and Liu (2005) find that distinct IU approximations for US data differ by the largest degree during periods like the oil crisis or shortly after changes in the conduct of monetary policy. Under such circumstances, economic decisions might be most strongly affected by the uncertainty about future inflation. To evaluate the relative predictive content of IU measures under distinct scenarios, the TOP$q^{\bullet}$-criterion in (15) is firstly considered conditional on whether predicted observations are drawn from either calm or turbulent subperiods of the evaluation sample. The examination of median IU trajectories in section 4.2 shows that distinct IU measures differ considerably in their indication of IU during the turbulent years after 2008. Hence we distinguish between turbulent and calm periods by means of the standard deviation over the IU metrics in (2) to (8) at each forecasting step $\tau = T_0 - \ell, ..., T - \ell$, denoted as $SD_{\tau+\ell|\tau}$. Conditional forecast rankings are determined by computing the average TOP$q^{\bullet}$ measure separately for all sample observations above and below the median of $SD_{\tau+\ell|\tau}$, respectively. Moreover, we distinguish performance rankings between earlier and later periods of the estimation sample by splitting the available time instances into two periods of equal length. The former period starts in 1988Q1 and ends in 1998Q3, the latter comprises observations until 2011Q1. This may reveal if certain IU statistics may have become more relevant for interest rate forecasting in the course of the two recent decades.

Furthermore, the usefulness of distinct IU measures might depend on historical experiences of distinct economies with respect to inflation rates. For example, prolonged periods of high inflation rates are frequently conjectured to affect the degree of inflation aversion of monetary authorities like in the case of the German Bundesbank (Clarida and Gertler 1997). Moreover, IU is likely strongly associated with the inflation level. This relation has been theoretically described by Friedman (1977) and Ball (1992) and empirically documented, e.g. in a recent study of Hartmann and Herwartz (2012). Hence, the formulation of recommendations for investors or monetary authorities based on indications about IU might depend on the economy-specific average level of inflation experienced over a longer time period. Therefore, we examine the relative merits of distinct IU measures for predicting interest rates separately in nine high-inflation and nine low-inflation economies. Finally, the overall contribution to predictive accuracy from the set of IU metrics is evaluated by comparing predictive performance provided by the respective reformulations of the Fisher equation to a specification which does not incorporate an IU term.

5 Empirical findings

This section summarises and interprets the results of the forecast comparison study. Subsequently, we discuss the role of IU for the determination of interest rates.

5.1 Related literature

The literature on forecasting interest rates and other macroeconomic variables documents that uncertainty measures similar to the IU metrics considered in our investigation can be useful predictors in many situations. Li and Zhao (2004) provide evidence for a significant impact of GARCH-implied uncertainty terms on density forecasts of 1-month US T-bill rates in an out-of-sample study. Höhrdahl et al. (2006) compare the out-of-sample forecast accuracy of a term structure model which incorporates distinct inflation risk terms to ad-hoc models for interest rate forecasting. They find that term structure models including macroeconomic variables outperform simpler models like the random walk in terms of predictive performance. Park (2005) documents that the disagreement among experts' earnings forecasts has predictive content for stock returns. Moreover, Kurz and Motolese (2011) find that the disagreement of market analysts' predictions is positively associated with the time variation in the risk premia of stock returns. In contrast, Elliot and Ito (1999) find that the disagreement as measured by the variance across expert forecasts of the Yen/Dollar exchange rate is not significantly related to corresponding profits from several trading rules.

5.2 Forecasting interest rates

In Table 4, the outcomes of the forecasting study based on the TOP$q^\bullet$ criterion are summarised. In the left and the right part of the table, respectively, the frequencies in which forecasting models based on particular IU measures show either the best performance ($q = 1$) or are among the three most accurate ($q = 3$) models are reported. In both cases, the disparity measure $\bar{\sigma}_{\tau+\ell|\tau}$ is the msot informative predictor. This ranking is particularly clear-cut for anticipation horizons $\ell > 1$. Other IU statistics from the dispersion class contribute less to the forecasting accuracy of the augmented Fisher relation. From the set of time-series statistics, the measure $\hat{\sigma}_{\tau+\ell|\tau}$ is relatively frequently among the best IU metrics. Summary measures of IU like, for example, the median over all IU metrics or the mean of the dynamic IU statistics also take up high ranks in some cases, e.g. for $\ell = 1$. However, they are less successful than $\bar{\sigma}_{\tau+\ell|\tau}$. The top ranking frequencies of $\overline{TS}$ show that forecast precision can be improved in several cases if we combine the informative content of individual dynamics IU statistics. This does not apply to a similar extent for the dispersion measures. Thus, choosing a candidate from the group of time-series IU metrics appears to be more difficult than selection among the disparity statistics.

Table 4: TOP$q^{\bullet}$

	$q=1$				$q=3$			
	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$
$\hat{\sigma}_{\tau+\ell\|\tau}$	8.91	11.63	11.63	12.21	21.45	24.16	25.32	25.19
$h^{(0.1)}_{\tau+1\|\tau}$	8.27	7.04	6.40	6.72	23.32	22.03	21.77	21.77
$h^{(0.2)}_{\tau+1\|\tau}$	4.13	4.26	3.17	3.23	23.26	21.77	17.57	18.09
$\hat{a}_{\tau}$	12.02	7.49	7.82	9.37	26.94	23.06	22.93	23.13
$\hat{s}_{\tau+\ell\|\tau}$	8.98	7.82	7.62	7.49	21.51	20.09	20.54	20.74
$\bar{\sigma}_{\tau+\ell\|\tau}$	11.69	15.89	16.73	17.25	22.35	27.65	28.62	27.97
$\varsigma_{\tau+\ell\|\tau}$	2.45	3.17	5.56	4.26	15.96	18.15	20.80	21.90
$\zeta_{\tau+\ell\|\tau}$	5.10	4.01	7.36	7.95	19.44	20.22	22.22	20.93
max(IU)	4.72	4.07	5.68	3.88	15.70	16.86	20.80	20.09
min(IU)	5.36	7.75	7.30	5.17	19.51	21.83	20.16	17.12
median(IU)	8.53	9.88	8.59	8.72	20.74	23.00	22.48	23.39
$\overline{TS}$	9.30	5.88	5.49	5.68	28.81	24.22	21.38	20.99
$\overline{DS}$	1.10	2.58	2.20	3.36	19.06	18.28	18.99	21.12
$\circ$	9.63	9.37	6.72	5.88	22.16	19.51	18.22	19.32

Cell entries represent the frequencies in which distinct IU measures lead to forecasts which are best ($q = 1$ case) or among the 3 most accurate ($q = 3$). The row labelled as '$\circ$' reports respective ranking frequencies for a forecasting model without an IU term.

Table 5 reports how frequently distinct IU measures lead to forecasts which outperform the Fisher equation without IU terms (as indicated by '$\circ$'). The left panel shows that it is generally beneficial to predict interest rates by incorporating an IU term, since frequencies are in basically all cases above 50%. In the right panel, the benchmark absolute errors are downward scaled by a factor of $c = 0.8$ to enable more clear-cut distinctions among the IU measures. This reveals that the $\bar{\sigma}_{\tau+\ell|\tau}$ IU measure is the best performing candidate IU statistic also in this respect. The comparisons of dynamic and dispersion IU measures to the respective frequencies of benchmark IU approximations GARCH and the survey-based standard deviation $u_{\tau+4|\tau}$ from (10) are shown in table 6. The results demonstrate that both groups of IU metrics compare favourably to their respective benchmark in terms of the frequency to outperform the Fisher equation without IU terms. Though the GARCH(1,1) measure provides relatively high percentages for $\ell = 1$, the overall indication is that the model based IU metrics are suitable to forecast interest rates also when compared with prominent measures from the related literature.

Table 5: Percentage of cases where $|e^{\bullet}_{\tau+\ell|\tau}| < c \times |e^{(\circ)}_{\tau+\ell|\tau}|$

	$c = 1$				$c = 0.8$			
	$\ell = 1$	$\ell = 2$	$\ell = 3$	$\ell = 4$	$\ell = 1$	$\ell = 2$	$\ell = 3$	$\ell = 4$
$\hat{\sigma}_{\tau+\ell\|\tau}$	51.03	53.29	55.49	54.84	22.22	29.07	30.75	31.20
$h^{(0.1)}_{\tau+1\|\tau}$	51.87	54.20	52.71	52.00	18.09	21.25	21.25	19.77
$h^{(0.2)}_{\tau+1\|\tau}$	51.74	53.94	52.84	51.74	15.50	17.64	17.70	15.31
$\hat{a}_{\tau}$	49.55	51.16	52.78	52.78	26.94	29.13	28.94	28.10
$\hat{s}_{\tau+\ell\|\tau}$	51.42	53.55	54.97	54.97	26.49	31.65	35.79	34.82
$\bar{\sigma}_{\tau+\ell\|\tau}$	49.68	53.04	53.29	55.10	22.87	34.04	35.34	36.82
$\varsigma_{\tau+\ell\|\tau}$	50.19	53.10	56.07	55.56	25.32	31.65	35.47	35.92
$\zeta_{\tau+\ell\|\tau}$	50.45	52.71	54.91	54.72	27.45	33.01	37.34	35.21
max(IU)	50.45	53.10	56.20	55.62	25.26	32.11	35.47	35.59
min(IU)	49.94	53.62	52.97	50.97	18.80	22.42	22.80	22.87
median(IU)	51.55	55.88	52.26	55.62	21.45	30.62	30.30	34.30
$\overline{TS}$	51.16	51.36	52.39	53.23	26.94	28.42	28.55	27.78
$\overline{DS}$	50.65	53.23	56.14	55.62	25.97	31.91	35.85	35.79

The symbol '$\circ$' represents forecast errors obtained for the model (9) which does not include an IU term. In the right part of the table, $|e^{(\circ)}_{\tau+\ell|\tau}|$ is scaled downwards to obtain more pronounced distinctions among alternative IU measures.

Table 6: Percentages of $|e^{\bullet}_{\tau+\ell|\tau}| < c \times |e^{(\circ)}_{\tau+\ell|\tau}|$, with benchmark measures

	$\hat{\sigma}_{\tau+1\|\tau}$	$h^{(0.1)}_{\tau+1\|\tau}$	$h^{(0.2)}_{\tau+1\|\tau}$	$\hat{a}_{\tau}$	$\hat{s}_{\tau+1\|\tau}$	$\bar{\sigma}_{\tau+1\|\tau}$	$\varsigma_{\tau+1\|\tau}$	$\zeta_{\tau+1\|\tau}$	GARCH(1,1)
$c = 1$	51.03	51.87	51.74	49.55	51.42	49.68	50.19	50.45	51.42
$c = 0.8$	22.22	18.09	15.50	26.94	26.49	22.87	25.32	27.45	29.01
	$\hat{\sigma}_{\tau+4\|\tau}$	$h^{(0.1)}_{\tau+1\|\tau}$	$h^{(0.2)}_{\tau+1\|\tau}$	$\hat{a}_{\tau}$	$\hat{s}_{\tau+4\|\tau}$	$\bar{\sigma}_{\tau+4\|\tau}$	$\varsigma_{\tau+4\|\tau}$	$\zeta_{\tau+4\|\tau}$	$u_{\tau+4\|\tau}$
$c = 1$	54.84	52.00	51.74	52.78	54.97	55.10	55.56	54.72	48.94
$c = 0.8$	31.20	19.77	15.31	28.10	34.82	36.82	35.92	35.21	17.72

The table reports frequencies where the Fisher equation with distinct IU terms. and GARCH(1,1), outperforms the relation without IU terms. The upper part of the table reports percentages for $\ell = 1$ for the comparison of results to those obtained by means of a GARCH(1,1) model. The lower part of the table shows outperformance frequencies for $\ell = 4$ and with IU as quantified by $u_{\tau+4|\tau}$. Predictions based on the IU measures max(IU), min(IU), median(IU), mean(TS) and mean(DS) and without IU terms are omitted from the table to economise on space. Respective numbers are reported in table 5.

Subsample-specific rankings are reported in table 7. To economise on space, we report only comparisons between the 3 most successful candidate IU measures in this table. In general, the lead of $\bar{\sigma}_{\tau+\ell|\tau}$ over other IU measures is strongest for higher forecast horizons. In particular, a high predictive contribution of this IU measure is found during turbulent periods, when selecting an IU statistic might be especially important. The results for calm periods, however, do only show that the $\bar{\sigma}_{\tau+\ell|\tau}$ IU measure outperforms its rivals in a slightly less pronounced way. The split of the sample into early and recent observations further reinstates the findings regarding the robust performance of $\bar{\sigma}_{\tau+\ell|\tau}$. We find that predictive content of this metric is more pronounced for the latter period, that is, the importance of $\bar{\sigma}_{\tau+\ell|\tau}$ has been increasing during the recent two decades. While the performance numbers of $\bar{\sigma}_{\tau+\ell|\tau}$ are not overly different from those of $\hat{\sigma}_{\tau+\ell|\tau}$ until 1998, the distinction becomes rather strong during subsequent years until 2011. Furthermore, from the forecast comparisons for low- and high inflation economies, it is apparent that the ranking of IU measures is more clear-cut for economies with higher average inflation rates. This might be due to the well-documented positive association between the level and the uncertainty of inflation. The explicit consideration of uncertain periods appears to be a suitable means to obtain most pronounced distinctions among measures of IU. This is in line with the findings of Lahiri and Liu (2005). Our results suggest that the $\bar{\sigma}_{\tau+\ell|\tau}$ IU measure is particularly useful during such situations, when the measurement of IU might be of highest relevance for economics decision takers.

Table 7: TOP1$^{\bullet}$, results for subsamples

	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$
	Turbulent periods				Calm periods			
$\hat{\sigma}_{\tau+\ell\|\tau}$	33.46	33.46	30.62	29.07	27.78	31.40	30.75	31.52
$\hat{a}_{\tau}$	30.23	24.94	24.68	25.45	37.98	29.72	30.10	28.42
$\bar{\sigma}_{\tau+\ell\|\tau}$	36.30	41.60	44.70	45.48	34.24	38.89	39.15	40.05
	Sample period 1988Q1-1998Q3				Sample period 1998Q4-2011Q1			
$\hat{\sigma}_{\tau+\ell\|\tau}$	27.91	32.30	32.69	33.33	33.33	32.56	28.68	27.26
$\hat{a}_{\tau}$	38.63	29.33	27.13	25.84	29.59	25.32	27.65	28.04
$\bar{\sigma}_{\tau+\ell\|\tau}$	33.46	38.37	40.18	40.83	37.08	42.12	43.67	44.70
	High inflation economies				Low inflation economies			
$\hat{\sigma}_{\tau+\ell\|\tau}$	25.97	29.46	28.81	29.72	35.27	35.40	32.56	30.88
$\hat{a}_{\tau}$	34.50	26.23	27.13	23.90	33.72	28.42	27.65	29.97
$\bar{\sigma}_{\tau+\ell\|\tau}$	39.53	44.32	44.06	46.38	31.01	36.18	39.79	39.15

Turbulent and calm periods are distinguished according to whether the standard deviation over the IU metrics in (2) to (8) exceeds its median value. Similarly, the cross section of 18 economies is split into 2 groups labelled 'High-' and 'Low inflation' according to their average inflation rate over the sample period. For further descriptions see table 4.

The high performance of the dispersion IU statistic $\bar{\sigma}_{\tau+\ell|\tau}$ still holds if the weights for the BMA scheme are only obtained for $\ell = 1$, though the advantage over other IU measures is reduced to a certain extent. This is documented in table 8 and, furthermore, in table 9 for the subsamples as discussed above. The IU metrics $\hat{\sigma}_{\tau+\ell|\tau}$ and $\hat{a}_{\tau}$ provide higher predictive content than $\bar{\sigma}_{\tau+\ell|\tau}$ in some cases, but neither of them outperforms the latter uniformly over distinct horizons. Moreover, conditional predictive performance comparisons, particularly for turbulent periods, reveal the reliably good performance of $\bar{\sigma}_{\tau+\ell|\tau}$ also for the alternative BMA weighting scheme. Hence it appears that the risk from potential biases in BMA weights to affect the qualitative conclusions of the forecast comparison among IU measures is rather limited.

Table 8: TOP$q^{\bullet}$: BMA based on $\ell = 1$-weights

	$q=1$				$q=3$			
	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$
$\hat{\sigma}_{\tau+\ell\|\tau}$	8.91	12.14	10.85	12.34	21.45	26.42	23.06	22.67
$h^{(0.1)}_{\tau+1\|\tau}$	8.27	8.85	9.11	8.20	23.32	22.80	23.71	23.26
$h^{(0.2)}_{\tau+1\|\tau}$	4.13	3.62	3.23	3.17	23.26	23.64	20.67	20.41
$\hat{a}_{\tau}$	12.02	8.53	8.66	7.75	26.94	25.39	23.32	22.35
$\hat{s}_{\tau+\ell\|\tau}$	8.98	8.01	7.88	7.82	21.51	19.57	21.19	19.25
$\bar{\sigma}_{\tau+\ell\|\tau}$	11.69	10.72	10.34	10.79	22.35	21.71	22.42	24.29
$\varsigma_{\tau+\ell\|\tau}$	2.45	3.94	5.36	5.23	15.96	17.57	19.32	22.74
$\zeta_{\tau+\ell\|\tau}$	5.10	3.68	6.14	6.40	19.44	18.48	19.32	17.44
max(IU)	4.72	5.17	4.78	5.88	15.70	16.34	19.19	20.48
min(IU)	5.36	6.01	5.94	5.81	19.51	19.38	18.28	17.05
median(IU)	8.53	9.63	8.91	12.14	20.74	21.58	23.51	24.61
$\overline{TS}$	9.30	8.98	9.04	5.94	28.81	27.45	25.45	22.87
$\overline{DS}$	1.10	2.20	3.10	2.20	19.06	20.67	21.64	23.19
$\circ$	9.63	8.79	6.91	7.11	22.16	19.25	19.38	19.70

The results in this table are obtained for the alternative BMA weighting scheme where combination weights in (12) are throughout taken from the likelihood of (9) given $\ell = 1$. For further descriptions see table 4.

Table 9: TOP1$^\bullet$: BMA based on $\ell = 1$-weights, results for subsamples

	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$
	Turbulent periods				Calm periods			
$\hat{\sigma}_{\tau+\ell\vert\tau}$	33.46	35.27	35.40	35.01	27.78	31.65	26.87	25.45
$\hat{a}_\tau$	30.23	27.65	26.61	24.81	37.98	31.91	34.11	31.52
$\bar{\sigma}_{\tau+\ell\vert\tau}$	36.30	37.08	37.98	40.18	34.24	36.43	39.02	43.02
	Sample period 1988Q1-1998Q3				Sample period 1998Q4-2011Q1			
$\hat{\sigma}_{\tau+\ell\vert\tau}$	27.91	33.59	32.43	32.56	33.33	33.33	29.84	27.91
$\hat{a}_\tau$	38.63	32.04	29.97	25.84	29.59	27.52	30.75	30.49
$\bar{\sigma}_{\tau+\ell\vert\tau}$	33.46	34.37	37.60	41.60	37.08	39.15	39.41	41.60
	High inflation economies				Low inflation economies			
$\hat{\sigma}_{\tau+\ell\vert\tau}$	25.97	28.42	28.42	28.94	35.27	38.50	33.85	31.52
$\hat{a}_\tau$	34.50	30.49	27.91	26.61	33.72	29.07	32.82	29.72
$\bar{\sigma}_{\tau+\ell\vert\tau}$	39.53	41.09	43.67	44.44	31.01	32.43	33.33	38.76

For further descriptions see tables 4 and 7.

5.3 The effect of IU on interest rates

Theoretical explanations of the IU influence assert that both the demand for loanable funds from investors and the supply of savings tend to be discouraged by higher IU (Lahiri et al. 1988). This means that, ceteris paribus, the sign of the respective IU impact on interest rates will be negative if the former effect dominates and positive in the contrary case. Hence, our estimates provide evidence on which influence is the prevailing one. To highlight the evolution of the relation over time, we provide a graphical display of the IU effect. Denoting estimated overall IU effects at prediction steps $\tau = T_0, ..., T$ in economies $i = 1, ..., 18$ as $\bar{\hat{\gamma}}_{i\tau}^{(IU)}$, we partition the sequence of estimates into equally sized subwindows W_k, comprising 5 quarters each[4]. Window-specific average coefficient estimates are obtained as

$$\bar{\bar{\gamma}}_{ik}^{(IU)} = (1/5) \sum_{\tau \in W_k} \bar{\hat{\gamma}}_{i\tau}^{(IU)}. \tag{16}$$

[4]The sample period covers 86 observations in total, hence we let the first subperiod comprise 6 quarters.

Figures 3 to 5 show economy-specific trajectories of $\bar{\bar{\gamma}}_{ik}^{(IU)}$ for $k = 1, ..., 17$, covering the sample period between 1988Q1 and 2011Q1. In the forecasting study, the average ranking of distinct IU measures across time instances and economies is considered as an indication of their respective predictive content. For a discussion of the IU effect during a certain subperiod k, the weights of each subset model specification $m = 1, ..., M$ of (9) based on the corresponding likelihood provide a means to determine their relative importance. Therefore, we consider confidence bands around the time paths of IU effects which are based on the variation in estimates of $\bar{\bar{\gamma}}_{ik}^{(IU)}$ over $m = 1, ..., M$. We find that for all economies, the estimated impact of IU on sovereign bond yields is positive. This suggests the interpretation of $\bar{\bar{\gamma}}_{ik}^{(IU)}$ as an inflation risk premium is the most meaningful way to explain the influence of IU on interest rates. These results do only in parts agree with the findings of Berument et al. (2007). They document that IU as measured by a GARCH model and short-term interest rates from the G7 and emerging economies are in most cases significantly related. However, Berument et al. (2007) find that the IU effect varies in sign across economies. Duffee (2012) also reports negative estimates of inflation risk premia in US Treasury bonds of varying maturity. The results of Duffee (2012) are obtained in the framework of term structure models with multiple factors, one of which is the inflation rate. Moreover, the plots in the figures 3 to 5 show that for several economies, the price of inflation risk appears to have increased during the most recent years.

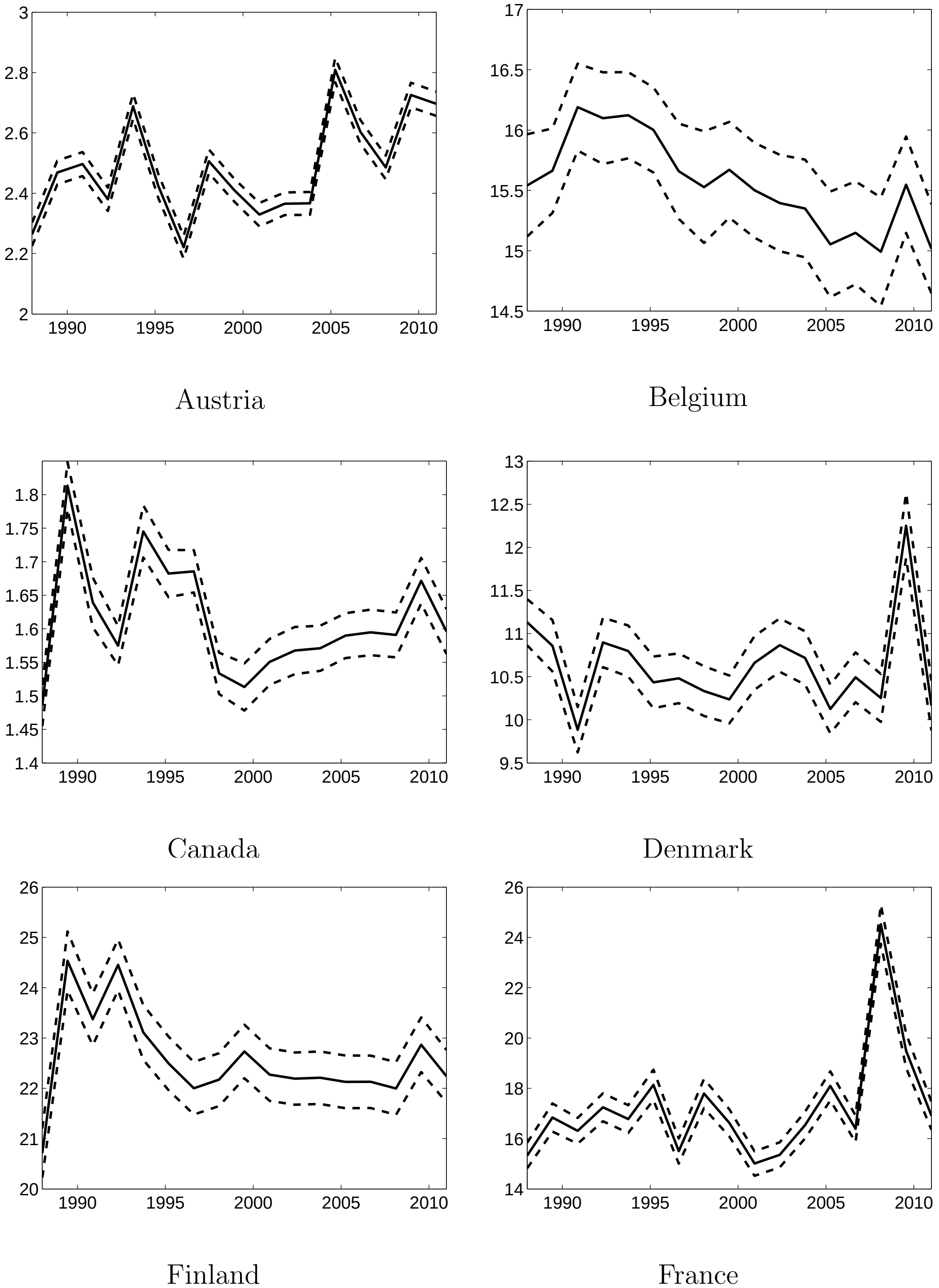


Figure 3: Accumulated inflation uncertainty effects $\bar{\gamma}_{ik}^{(IU)}$ for distinct economies and the forecasting horizon $\ell = 1$. Dashed lines are approximate ± 2 standard error confidence bands indicating likelihood-weighted subset model variation with respect to distinct specifications of the Fisher equation in (9). Results for $\ell > 1$ are qualitatively similar and available from the authors upon request.

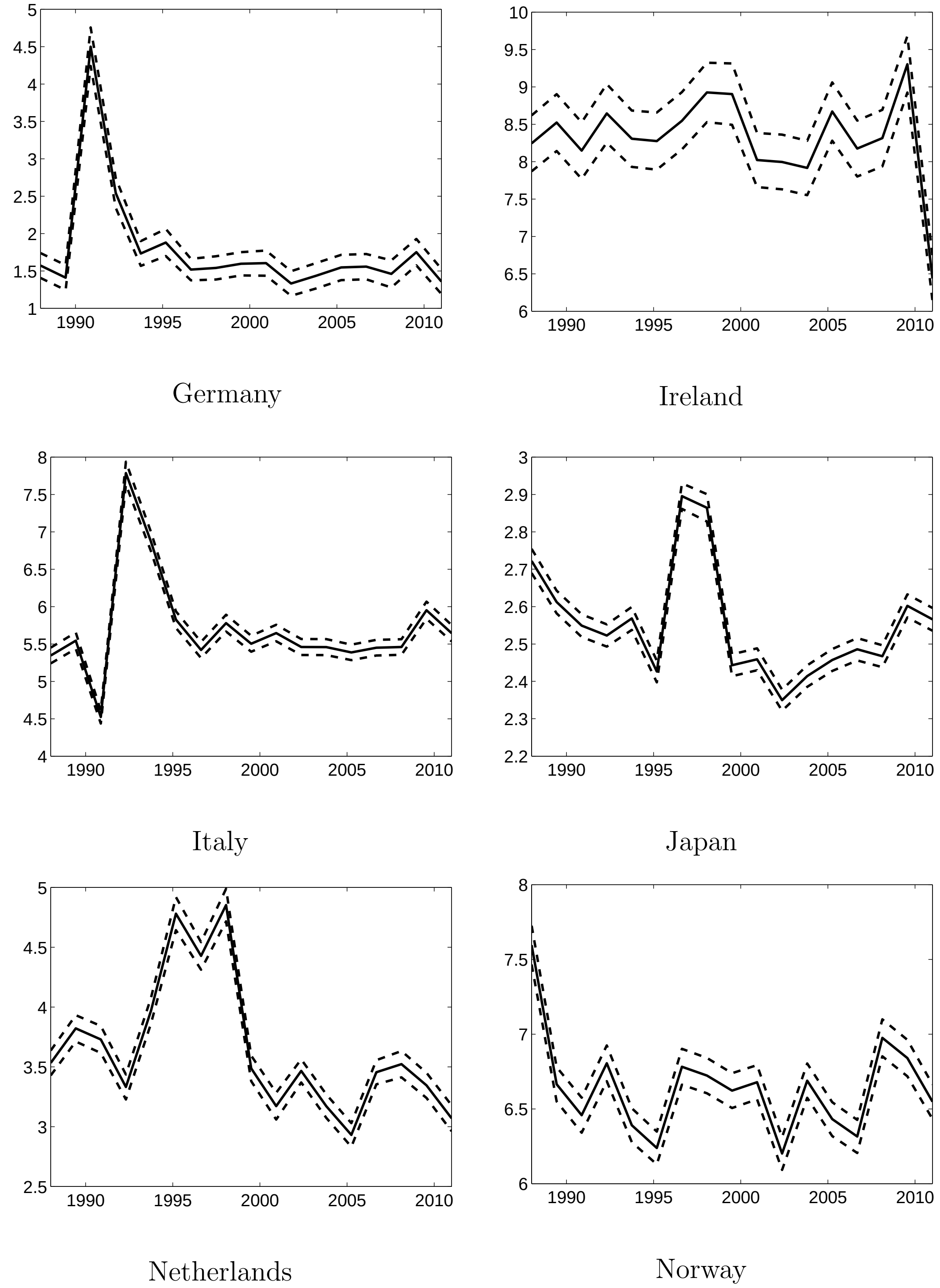
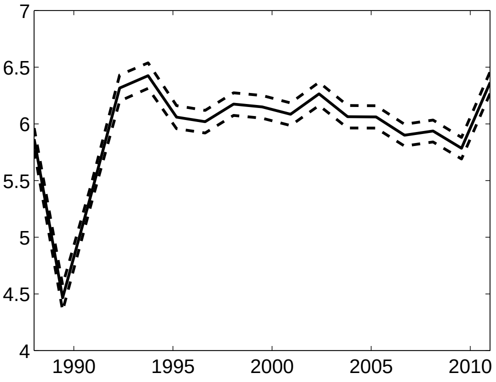
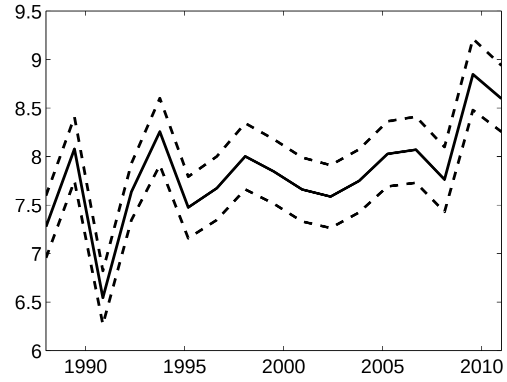


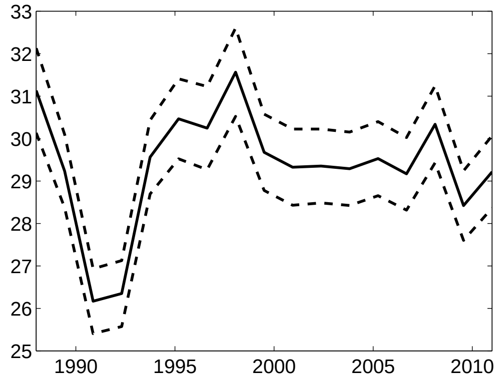
Figure 4: For a description see figure 3.



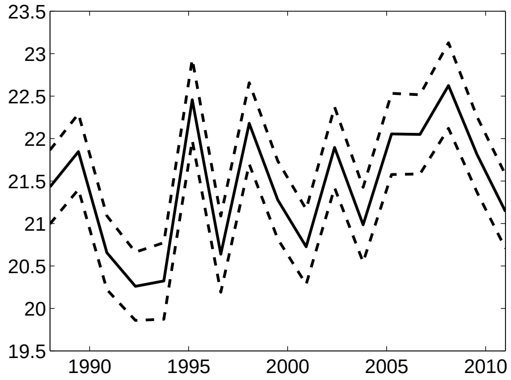
Portugal



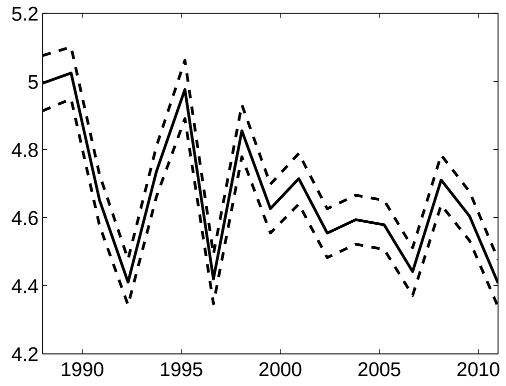
Spain



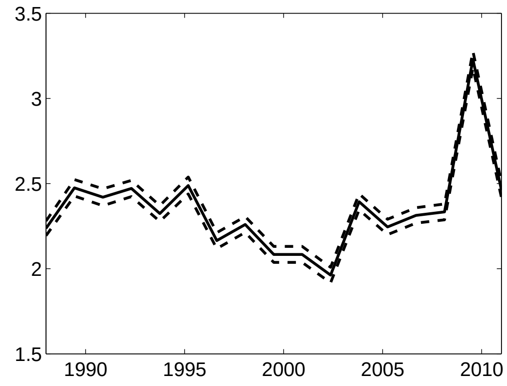
Sweden



Switzerland



UK



US

Figure 5: For a description see figure 3.

6 Summary and concluding remarks

This study is a comparative evaluation of several alternative measures of inflation uncertainty. The assessment is carried out by means of forecasting government bond interest rates, employing alternative inflation uncertainty measures as predictor variables. We predict interest rates for 18 mature economies at anticipation horizons of one month up to one year. In the related literature, two categories of inflation uncertainty quantification are distinguished. The first group are time-series measures, the other one is based on the heterogeneity of individual inflation forecasts. The forecast competition shows that the average over individual uncertainties as a representative of the dispersion family is the most viable predictor variable for interest rates. We further note that all uncertainty measures uniformly indicate a decrease in inflation uncertainty during the years of the so-called Great Moderation. However, the two groups of uncertainty metrics differ markedly in their indication of inflation uncertainty during the recent crisis period after the year 2008. While time-series measures indicate a considerable uprise of inflation risk after 2008, the dispersion measures remain at their average level of the pre-crisis years. Moreover, our estimates of the relation between inflation uncertainty and interest rates is uniformly positive across economies. This suggests the interpretation of the IU effect as a risk premium as the most likely explanation. Although the IU level as indicated by the most informative IU measure does not rise significantly during recent years, investors seem to demand for a higher inflation risk premium than during the previous years of the Great Moderation in several economies.

7 Appendix

7.1 Further inflation forecasting schemes

Extending the baseline AR in (1) with a lagged output gap term, $\tilde{y}_t = y_t - \bar{y}_t$, yields the backward looking Phillips curve following e.g. Stock and Watson (2007), i.e.

$$\pi_{t+\ell} = \alpha_{10} + \alpha_{11} t + \alpha_{12} \pi_{t-1} + \alpha_{13} \tilde{y}_{t-1} + \epsilon_{t+\ell}, \quad t = \tau - B + 1, ..., \tau. \tag{17}$$

In (17), estimates of the long run trend on which $\tilde{y}_t$ is based are computed recursively within the estimation window by means of the Hodrick-Prescott filter with smoothing parameter 129600 (Ravn and Uhlig 2002). To examine the predictive content of monetary variables, Stock and Watson (2008) predict inflation changes with the money augmented Phillips curve, initially proposed by Gerlach (2004). Similarly, the growth rate of core money, denoted $\overline{m}_t$, is typically interpreted as a proxy for inflation expectations. Consequently, this specification reads as

$$\pi_{t+\ell} = \alpha_{20} + \alpha_{21} t + \alpha_{22} \pi_{t-1} + \alpha_{23} \tilde{y}_{t-1} + \alpha_{24} \bar{m}_{t-1} + \epsilon_{t+\ell}. \tag{18}$$

Neumann and Greiber (2004) propose to augment (18) with an indicator of energy prices obtaining

$$\pi_{t+\ell} = \alpha_{30} + \alpha_{31} t + \alpha_{32} \pi_{t-1} + \alpha_{33} \tilde{y}_{t-1} + \alpha_{34} \bar{m}_{t-1} + \alpha_{35} \Delta^2 oil_{t-1} + \epsilon_{t+\ell}. \tag{19}$$

In (19), $\Delta^2 oil_t$ denotes second differences of the log oil price in terms of domestic currency. Note that (19) implicitly comprises log foreign exchange rate changes as predictors of inflation. An alternative model in the spirit of Cogley (2002) incorporates the deviation of inflation from its long run trend, denoted $\tilde{\pi}_t = \pi_t - \bar{\pi}_t$. This model is given by

$$\pi_{t+\ell} = \alpha_{40} + \alpha_{41}\tilde{\pi}_{t-1} + \epsilon_{t+\ell}. \tag{20}$$

This specification expresses the view that that in states deviating markedly from the long run inflation trend, additional adjustment dynamics might impact on $\pi_{t+\ell}$.

8 Acknowledgements

We thank the participants of the 2nd IWH/INFER Workshop on Applied Economics and Economic Policy in Halle for helpful suggestions. Financial support by Deutsche Forschungsgemeinschaft (HE 2188/3-1) is gratefully acknowledged.

References

Ball L (1992) Why does high inflation raise inflation uncertainty?. Journal of Monetary Economics 29: 371-388

Barnea A, Dotan A, Lakonishok J (1979) The effect of price level uncertainty on the determination of nominal interest rates: Some empirical evidence Southern Economic Journal 46: 609-614

Batchelor R, Dua P (1996) Empirical measures of inflation uncertainty: A cautionary note. Applied Economics 28: 333-341. doi: 10.1080/000368496328704

Black F (1987) Business cycles and equilibrium. Basil Blackwell, New York

Blanchard, OJ, Dell'Ariccia G, Mauro P (2010) Rethinking macroeconomic policy. Journal of Money, Credit and Banking 42: 199-215. DOI: 10.1111/j.1538-4616.2010.00334.x

Blejer MI, Eden B (1979) A note on the specification of the Fisher equation under inflation uncertainty. Economics Letters 3: 249-255

Bomberger WA (1996) Disagreement as a measure of uncertainty. Journal of Money, Credit and Banking 28: 381-392

Bomberger WA, Frazer W (1981) Interest rates, uncertainty and the Livingston data. Journal of Finance 36: 661-675

Buraschi A, Jiltsov A (2005) Inflation risk premia and the expectations hypothesis. Journal of Financial Economics 75: 429-490. doi: 10.1016/j.jfineco.2004.07.003

Chevillon G (2005) Direct multistep estimation and forecasting. Journal of Economic Surveys 21: 746-785. doi: 10.1111/j.1467-6419.2007.00518.x

Chua CL, Kim D, Suardi S (2011) Are all empirical measures of uncertainty alike? Journal of Economic Surveys 25: 801-827. doi: 10.1111/j.1467-6419.2010.00662.x

Clarida R, Gertler M (1997) How the Bundesbank conducts monetary policy. In: Romer C and Romer D (eds) Reducing inflation: Motivation and Strategy. NBER, Chicago, pp 363-412

Clarida R, Galí J, Gertler M (1999) The science of monetary policy: A New Keynesian perspective. NBER Working Paper 7147

Fountas S, Karanasos M, Kim J (2002) Inflation and output growth uncertainty and their relation with inflation and output growth. Economics Letters 68: 293-301. doi: 10.1016/S0165-1765(02)00009-5

Fountas S, Karanasos M, Kim J (2006) Inflation uncertainty, output growth uncertainty, and macroeconomic performance. Oxford Bulletin of Economics and Statistics 68: 319-343. doi: 10.1111/j.1468-0084.2006.00164.x

Friedman M (1968) The role of monetary policy. American Economic Review 58: 1-17

Friedman M (1977) Inflation and unemployment. Journal of Political Economy 85: 451-472

Giordani P, Söderlind P (2003) Inflation forecast uncertainty. European Economic Review 47: 1037-1059. doi: 10.1016/S0014-2921(02)00236-2

Gordon RJ (2009) The history of the Phillips Curve: Consensus and bifurcation. Economica 78: 10-50. doi: 10.1111/j.1468-0335.2009.00815.x

Gordon RJ, Brainard W, Juster T (2009) Inflation in recession and recovery. Brookings Papers on Economic Activity 1: 105-166

Grier KB, Perry MJ (2000) The effects of real and nominal uncertainty on inflation and output growth: Some GARCH-M evidence. Journal of Applied Econometrics 15: 45-58

Hartmann M, Herwartz, H (2012) Causal relations between inflation and inflation uncertainty - Cross sectional evidence in favour of the Friedman-Ball hypothesis. Economics Letters 115: 144-147. doi: 10.1016/j.econlet.2011.12.036

Kandel S, Ofer AR, SARIG O (1996) Real interest rates and inflation: An ex-ante empirical analysis. Journal of Finance 51: 205-225

Lahiri K, Liu F (2005) ARCH models for multi-period forecast uncertainty: A reality check using a panel of density forecasts. In: Terrell D, Fomby T (eds) Econometrics analysis of financial and economic time series part A. Elsevier, Amsterdam, pp 312-364. DOI: 10.1016/S0731-9053(05)20012-9

Lahiri K, Sheng X (2010) Measuring forecast uncertainty by disagreement: The missing link. Journal of Applied Econometrics 25: 514-538. doi: 10.1002/jae.1167

Levi MD, Makin, JH (1978) Anticipated inflation and interest rates: Further interpretation of findings on the Fisher equation. American Economic Review 68: 801-812

Levi MD, Makin JH (1979) Fisher, Phillips, Friedman and the measured impact of inflation on interest. Journal of Finance 34: 35-52

Mankiw N, Reis R (2004) Disagreement in inflation expectations. NBER Macroeconomics Annual 18: 147-209

Missale A, Blanchard, OJ (1994) The debt burden and debt maturity. American Economic Review 84: 309-319

Peng A, Yang L (2008) Modelling uncertainty: A recursive VAR bootstrapping approach. Economics Letters 99: 478-481. doi: 10.1016/j.econlet.2007.09.026

Phelps ES (1968) Money wage dynamics and labor market equilibrium. Journal of Political Economy 76: 687-711

Rich R, Tracy J (2003) Modeling uncertainty: Predictive accuracy as a proxy for predictive confidence. Federal Reserve Bank of New York Staff Reports 253

Schwab RM (1982) Inflation expectations and the demand for housing. American Economic Review 72: 143-153

Stock, JH, Watson MW (1999) A Comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. In: Engle RF, White H (eds) Cointegration, causality, and forecasting: A festschrift in honor of Clive W. J. Granger. Oxford University Press

Williams JC (2009) Heeding daedalus: Optimal inflation and the zero lower bound. Brookings Papers on Economic Activity Fall, 1-37. doi: 10.1353/eca.0.0066