

FORMATION AND DISPERSION OF EXPECTATIONS

An Empirical Analysis with Special Focus on Inflation

Inaugural-Dissertation
zur Erlangung des akademischen Grades eines Doktors
der Wirtschafts- und Sozialwissenschaften
der Wirtschafts- und Sozialwissenschaftlichen Fakultät der
Christian-Albrechts-Universität zu Kiel

vorgelegt von

Jonas Dovern
aus Aachen

Kiel, 2009

Gedruckt mit Genehmigung der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

Dekan: Prof. Dr. Thomas Lux
Erstberichterstattender: Prof. Dr. Helmut Herwartz
Zweitberichterstattender: Prof. Dr. Christian Merkl

Tag der Abgabe der Arbeit: 04. Mai 2009
Tag der Disputation: 23. Juli 2009

Contents

1	Introduction	1
1.1	Survey-Data	2
1.2	Inflation Dynamics	4
1.3	Structure of the Thesis	6
2	The <i>Consensus</i> Survey Data Set	9
2.1	Background of the Survey	9
2.2	Sample Issues	10
2.3	Rationality of Forecasts	12
2.3.1	Model	14
2.3.2	Results	22
2.3.3	Discussion of Results	31
2.4	Concluding Remarks	32
3	Inflation (Expectation) Dynamics	33
3.1	Estimating Expectation Formation of Households	33
3.1.1	Model	35
3.1.2	Data Issues	35
3.1.3	Results	38
3.1.4	Discussion of Results	48
3.2	Estimating the Sticky Information Phillips Curve	49
3.2.1	Model	50
3.2.2	Data Issues	50
3.2.3	Results	52
3.2.4	Discussion of Results	55
3.3	Concluding Remarks	55
4	Dispersion of Forecasts	57
4.1	Measuring Dispersion from Fixed Event Forecasts	58
4.1.1	Approaches	60
4.1.2	Data Issues	64

4.1.3	Empirical Results	65
4.2	Dispersion of Macroeconomic Forecasts in the G7	69
4.2.1	Introduction	69
4.2.2	Data Issues	72
4.2.3	Drivers of Disagreement	77
4.2.4	Disagreement Across Countries	85
4.2.5	Robustness Check Using Non-Transformed Raw Data	87
4.3	Concluding Remarks	91
5	Summary	93
	Bibliography	95
	Appendix	107
A	Carlson-Parking-Method for Quantifying Survey Data	107
B	Detailed Results about Determinants of Disagreement	108
C	Graphical Appendix	130
	Eidesstattliche Erklärung	145

List of Tables

2.1	Tests of Alternative Models for Idiosyncratic Forecast Errors for GDP	18
2.2	Bias of Individual Forecasts	25
2.3	Efficiency of Individual Forecasts	27
2.4	Bias of Consensus Forecasts	28
2.5	Sequential Test of Forecast Unbiasedness for Consensus Forecasts	29
2.6	Efficiency of Consensus Forecasts (Without Horizon Effects)	30
2.7	Efficiency of Consensus Forecasts (With Horizon Effects)	31
3.1	Tests for Granger Non-causality	40
3.2	Restricted Cross-Country Results	41
3.3	Country-by-Country Unrestricted Results	44
3.4	Tests for Cointegration between Expectations	47
3.5	Sticky Expectations in the VEC Form	48
3.6	SIPC: Equation-by-Equation Regression Results	53
3.7	SIPC: Seemingly Unrelated Regressions	54
4.1	Correlation of dispersion measures calculated from different types of forecasts	60
4.2	Labeling of Different Methods	66
4.3	Correlation Results	68
4.4	Results of Forecast Combination Regressions	69
4.5	Correlation between the Two Disagreement Measures	74
4.6	Summary Statistics on Individual Forecasts and Targeted Variables	76
4.7	Average Disagreement across Countries and Variables	78
4.8	Descriptive Statistics about Inflation, GDP Growth and Interest Rates	79
4.9	Central Bank Independence in G7 Countries, 1989–2006	83
B-1	Disagreement and Business Cycle – Panel Results for Inflation	108

B-2	Disagreement and Business Cycle – Panel Results for the Interest Rate	109
B-3	Disagreement and Business Cycle – Panel Results for GDP	110
B-4	Disagreement and Business Cycle – Panel Results for Consumption	111
B-5	Disagreement and Business Cycle – Panel Results for Investment	112
B-6	Disagreement and Business Cycle – Panel Results for Unemployment	113
B-7	Determinants of Disagreement – Country-by-Country Results for Inflation	114
B-8	Determinants of Disagreement – Country-by-Country Results for the Interest Rate	115
B-9	Determinants of Disagreement – Country-by-Country Results for GDP	116
B-10	Determinants of Disagreement – Country-by-Country Results for Consumption	117
B-11	Determinants of Disagreement – Country-by-Country Results for Investment	118
B-12	Determinants of Disagreement – Country-by-Country Results for Unemployment	119
B-13	Average Disagreement Across Countries and Variables	120
B-14	Correlation of Disagreement Across Countries	121
B-15	Correlation of Disagreement Within Countries	122
B-16	Horizon Effect for Disagreement across Forecasts for the Next Year	123
B-17	Horizon Effect for Disagreement across Forecasts for the Current Year	123
B-18	Rel. Change of Disagreement per Month for Forecasts for the Next Year	124
B-19	Rel. Change of Disagreement per Month for Forecasts for the Current Year	124
B-20	Quadratic Trend Effect per Month for Disagreement across Forecasts for the Next Year	124
B-21	Quadratic Trend Effect per Month for Disagreement across Forecasts for the Current Year	125
B-22	Time Trends in Disagreement across Forecasts for the Next Year	125
B-23	Time Trends in Disagreement across Forecasts for the Current Year	126
B-24	Horizon Effects for Multivariate Measure of Disagreement	126
B-25	Business Cycle Effect on Disagreement across Forecasts for the Next Year	127

B-26 Business Cycle Effect on Disagreement across Forecasts for the Current Year	127
B-27 Effect of Recession in Next Year on Disagreement of Forecasts for the Next Year	128
B-28 Effect of Recession in Current Year on Disagreement of Forecasts for the Current Year	128
B-29 Business Cycle Effect on Multivariate Measure of Disagreement	129

List of Figures

2.1	Structure of Survey Data	14
2.2	Autocorrelation Functions of Idiosyncratic Forecast Errors for GDP	17
3.1	Household and Expert Expectations and Actual Inflation . . .	37
4.1	Cross Sectional Standard Deviation of Forecasts in the SPF .	65
C-1	Expected and Actual Inflation Rate	130
C-2	Expected and Actual Short-term Interest Rates	131
C-3	Expected and Actual GDP Growth	132
C-4	Expected and Actual Consumption Growth	133
C-5	Expected and Actual Investment Growth	134
C-6	Expected and Actual Unemployment Rate	135
C-7	Disagreement and MSE of Expectations vs. Level and Variance of the Underlying Variable	136
C-8	Disagreement about Inflation Rate	137
C-9	Disagreement about Short-term Interest Rate	138
C-10	Disagreement about GDP Growth	139
C-11	Disagreement about Consumption Growth	140
C-12	Disagreement about Investment Growth	141
C-13	Disagreement about Unemployment Rate	142
C-14	Typical Shape of Horizon Effect	143
C-15	Multivariate Disagreement Measure	144

Chapter 1

Introduction¹

This thesis is a collection of papers that use survey data to analyze expectations about macroeconomic variables and the way these expectations are formed. Since inflation expectations are probably the most relevant expectations for understanding macroeconomic dynamics and certainly for analyzing and conducting monetary policy, some parts of the thesis focus exclusively on the formation and dynamics of inflation expectations while most of the parts are covering a broader range of macroeconomic expectations.

The papers should be seen as contributions to applied econometrics although in some cases they do also add to the literature on the methodological side. Foremost, the papers provide ample empirical evidence on various issues related to survey data on macroeconomic forecasts from an international point of view. They contribute to a literature that has been very much US-centric in the past. Not only do all papers look at various countries, but they also use – for the most part – disaggregate data on individual forecasts that have not been used in the literature before. In addition, we go beyond the existing literature with respect to the analysis of the determinants of dispersion of forecasts. Besides, in some cases the papers innovate in the way survey data are used in empirical analyses.

The work of this thesis is mainly connected to three streams of the economic literature. First, the part of the thesis that analyzes the properties of the survey data, which is used for the further analysis, connects to the literature that deals with the evaluation of rationality of observed forecasts (Chapter 2). Second, the part of the thesis, which tests empirically the validity of a certain type of model for the formation of inflation expectations

¹I would like to thank Pia Pinger very much for valuable help in putting together the final version of this thesis. In addition, I would like to thank all colleagues and the audience of various presentations for their valuable comments that helped to improve the papers collected in this thesis.

and inflation dynamics, connects to the literature on inflation determination in the New Keynesian Framework (Chapter 3). Finally, the part of the thesis that analyzes the behavior of the dispersion of forecasts relates to the emerging literature that deals with higher moments of expectations and issues of uncertainty (Chapter 4).

In the next two sections of this introduction, we briefly review important issues of the first and second streams of the literature.² We collect these reviews in the first chapter of the thesis to avoid repetitive references in the later sections.³

1.1 Survey-Data

The main purpose survey data have been used for is the assessment of the performance of forecasters that provide publicly available forecasts. The foci of papers on this evolve around the following issues: Are professional forecasts efficient and unbiased? Are forecasts obtained by averaging a set of independent forecasts superior to single forecasts? Are predictions obtained from survey data sets better than those obtained from simple benchmark models?

Regarding the first question, influential contributions assessing the efficiency of *fixed horizon* forecasts were made among others by Keane and Runkle (1990); Zarnowitz and Lambros (1987), and Zarnowitz and Braun (1993). Test on efficiency using *fixed event* forecasts usually follow the approach introduced by Nordhaus (1987). Modifications to this approach taking into account the special covariance structure of the panel survey data are presented among others in Davies and Lahiri (1995); Harvey et al. (2001); Loungani (2001), and Isiklar et al. (2006). The study presented in Section 2.3 falls into this category. The overall conclusion from existing studies is that most forecasters are not rational in the strict sense. While in the majority of cases the forecasts are found to be unbiased, they mostly do not seem to be efficient, i.e. forecasters tend to not use all information that is available when making their predictions.

Second, studies assessing *consensus* forecasts, i.e. pooled forecasts, usually find that they outperform single forecasts in terms of forecast precision. Batchelor (2001) for example finds that the *consensus* forecast provided by *Consensus Economics* outperforms the forecasts published by the Interna-

²The Volume of literature on dispersion of forecasts is still relatively small and we relate our research to this literature in Chapter 4 below.

³Though we try to minimize those repetitions to a minimum, in some cases they cannot be avoided completely to enable readers to understand the individual sections of this thesis.

tional Monetary Fund (IMF) and the World Bank for most of the variables, countries, and samples considered in their paper. A good overview of the reasons and the theoretical background of the superiority of *consensus* forecasts is given by Hendry and Clements (2004).

Finally, most studies show that forecasts published by professional forecasters are often outperformed by very simple or even naive forecasting methods - especially for longer forecasting horizons. A nice recent contribution on this issue is the paper by Isiklar et al. (2006), in which the authors show that even the forecast performance of the *consensus* forecast decreases rapidly in most cases when the forecast horizon increases and is outperformed by naive forecasts already for very moderate forecast horizons.

Another field of research in which survey data sets offer a valuable input to take economic theory to the real world is the literature that is concerned with questions about why forecasters disagree on the future at all. The studies presented below in Sections 4.1 and 4.2 add to this stream of literature.

The most prominent potential explanations for the dispersion of forecasts are the limited information processing capacity of agents (see e.g. Sims, 2003), the behavioral approach (Kahneman and Tversky, 1979; Kahneman, 2003), or the sticky information approach (Mankiw and Reis, 2002). Branch (2004) develops a model in which agents have to choose a predictor function from a set of costly alternatives since they are assumed not to have the capacity to grasp the true structure of the economy.⁴ Using a panel household survey data set the author finds that the data support a model in which agents use a specific predictor function and switch to another prediction method only infrequently. The model is at odds, however, with the fact that agents seem to have a 'genetic' bias towards specific predictor functions that is not justified by the functions' performance. Ehrbeck and Waldmann (1996) use a survey data set on interest rate predictions to falsify a theoretical model of 'strategic forecast bias' in which agents are assumed to have additional factors affecting their loss function rather than simply aiming at minimizing the expected squared forecast error. Lamont (2002) shows on basis of a survey data set of forecasts on real GDP growth for the US that reputation issues are a likely factor that makes forecasters tend to more extreme predictions, which deviate from the respective rational forecast.

Finally, survey data have been used to estimate macroeconomic relationships that are explicitly formulated in terms of expectations of future variables as for instance the forward looking and hybrid New Keynesian Phillips

⁴Other papers that derive theoretical models of why forecasters might rationally choose to publish diverging forecasts include Brock and Hommes (1997); Laster et al. (1999); Branch and Evans (2006).

Curve (see e.g. Gali and Gertler, 1999) or models based on the sticky information approach by Mankiw and Reis (2002). By using directly observable expectations in the estimation of such models, it is possible to circumvent difficulties (and to some degree arbitrariness) resulting from the application of instrumental variables or GMM estimation techniques. Examples of this literature include among others Carroll (2003) and Paloviita (2005). The results of Sections 3.1 and 3.2 below add to this strand of literature.

1.2 Inflation Dynamics

The understanding of how inflation is determined is probably one of the most crucial issues of macroeconomics and surely the most important issue in monetary policy. The way how inflation dynamics evolve has major implications for the conduct of monetary policy as a stabilization instrument for the macroeconomy.

The current modeling standard for analyzing monetary issues is the New Keynesian Model (NKM) also known under the label New Neoclassical Synthesis Model.⁵ Its main features are that i) prices are not entirely flexible which causes nominal rigidities, ii) firms act in a monopolistically competitive environment, and iii) agents' behavior is usually forward looking (Gali, 2008). Especially the last feature brings about a prominent role for expectations to play in the NKM. Most importantly, in the modeling of inflation expectations about future inflation are usually the key factor.

In its most simple form, the New Keynesian Phillips Curve (NKPC) can be written as

$$\pi_t = \beta \mathbf{E}_t(\pi_{t+1}) + \gamma y^* + u_t , \quad (1.1)$$

where π_t denotes inflation, \mathbf{E}_t is the expectation operator, and y^* is a measure of excess demand, usually proxied by an output gap or unemployment gap measure. u_t is a cost push shock that affects inflation in an idiosyncratic fashion.

In the traditional Backward Looking Phillips Curve (BLPC) it was assumed that inflation expectations were a function of past inflation. Consequently, (1.1) could be written as

$$\pi_t = \beta \alpha(L) \pi_t + \gamma y^* + u_t , \quad (1.2)$$

where now $\alpha(L)$ is a lag polynomial. Following for instance Sims (1980) or Calvo (1983) the concept of rational expectations gained importance in

⁵See Woodford (2003) for a comprehensive overview about this type of model.

macroeconomics during the 1980s, and the BLPC was criticized for being not consistent with optimizing (i.e. consumption or profit maximizing) behavior of economic agents. Under the assumption of rational expectations, we can write inflation expectations as

$$\mathbf{E}_t(\pi_{t+1}) = \pi_{t+1} + e_t, \quad (1.3)$$

where e_t denotes the one-step-ahead expectation error made at time t which is independently distributed across time under that assumption. Consequently, (1.1) can be written

$$\pi_t = \beta\pi_{t+1} + \gamma y^* + u_t + \beta e_t. \quad (1.4)$$

This equation for the inflation dynamic is consistent with a framework, in which monopolistically competitive firms set their prices according to the well known staggered mechanics à la Calvo (1983), or alternatively with a setup where the introduction of price adjustment costs leads firms to adjust their prices only infrequently (Rotemberg, 1982).

Now, one of the key issues is how to model inflation dynamics in such a way that the model generates a degree of inflation persistence that is comparable to the degree which is observed empirically. The problem with the – from a theoretical point of view appealing – pure forward looking Phillips Curve (FLPC) in (1.4) is that it fails exactly along this dimension.

To bring in line the dynamics implied by the NKM and the inflation persistence observed in reality, various alternatives have been proposed in the literature. The hybrid Phillips Curve (Gali and Gertler, 1999) has gained a lot of attention in recent years in the theoretical and empirical work on inflation dynamics. Its functional form can be written as

$$\pi_t = \delta^f \pi_{t+1} + \delta^b \pi_{t-1} + \gamma y^* + u_t + \delta^f e_t, \quad (1.5)$$

where δ^f and δ^b are the relative weights attached to the forward looking and the backward looking component respectively. It is usually derived from the ad-hoc assumption that there is a certain fraction of firms that use a rule of thumb and set their future prices based on their most recent inflation observations (Gali and Gertler, 1999).⁶

Recently, Mankiw and Reis (2002) have proposed a different concept to model inertia in inflation dynamics. They abandon the idea of sticky prices

⁶There have been attempts recently to derive more persistence of nominal and real variables in the NKM in a more rigorous manner from optimizing behavior of economic agents by adding additional market frictions – especially to the labor market models (see among others Christoffel and Linzert, 2005; Walsh, 2005; Krause and Lubik, 2007; Lechthaler et al., 2008).

and introduce the notion of sticky information instead. In their model firms are able to adjust prices continuously whereas they update their information set only infrequently; this behavior can be explained by the existence of costs for acquiring information (Reis, 2006a,b). Mankiw and Reis (2002) assume that each period, only a fraction λ of firms gathers the up-to-date information about the current state of the economy and re-computes and adjusts the optimal path of future prices. The remaining $(1 - \lambda)$ firms continue using their previous plans and thus set prices based on outdated information. The firm's probability of information updating is exogenously determined and independent of its price adjustment history. Under this assumption Mankiw and Reis derive the following formulation of the SIPC:

$$\pi_t = \frac{\lambda\alpha}{1-\lambda}y_t^* + \lambda \sum_{j=0}^{\infty} (1-\lambda)^j \mathbf{E}_{t-1-j}(\pi_t + \alpha\Delta y_t^*) + \varepsilon_t, \quad (1.6)$$

where the parameter α measures the sensitivity of the optimal relative price to the current output gap and depends on the structure of the economy (e.g. on preferences, technology, and the market structure parameters).⁷ Note that in contrast to the FLPC where current expectations for future inflation matter and to the BLPC where past inflation matters, in the SIPC what matters for determination of inflation rates are past expectations of current inflation.

1.3 Structure of the Thesis

The thesis is structured in the following way. In Chapter 2, we present a comprehensive analysis of the properties of the *Consensus Economics* survey data set that is used in most parts of the thesis and that has been specifically compiled for the studies presented in this thesis. Specifically, we present some information about the survey's background and some basic information about the coverage of the data set in Section 2.1 and 2.2. In Section 2.3, we present a formal analysis of the rationality properties of the forecasts included in the data set.

In Chapter 3, we present two studies that use survey data to provide empirical evidence in favor of the sticky information concept for describing inflation (expectation) dynamics. In Section 3.1, we show that a sticky information model for the expectation formation process of households fits

⁷The parameter α can be interpreted as a measure of the degree of real rigidity, see e.g. Ball and Romer (1990).

European data well. In Section 3.2, we show that in a partial equilibrium setup the SIPC fits observed inflation dynamics well.

In Chapter 4, we present different studies that contribute to the analysis of dispersion of expectations. In Section 4.1, we compare different methods to extract a measure of dispersion from fixed event forecasts. In Section 4.2, we use the method that we find to perform best to extract dispersion measures for forecasts in the G7 countries and use these to analyze the determinants of dispersion. Chapter 5 summarizes the main findings of the thesis and discusses possible extensions.

Chapter 2

The *Consensus* Survey Data Set

This section is devoted to the description of the data set, the *Consensus* survey data set, that we use most during the empirical analysis presented in this thesis. First, we present some very basic information about the data set including some background about the way the data is collected. Second, we elaborate on the size of our sample. Finally, in Section 2.3, we investigate the rationality properties of the forecasts contained in the data set. We present both tests for unbiasedness and tests for efficiency of individual and average forecasts.

2.1 Background of the Survey

Most of empirical work presented in this thesis is based on survey data collected from a data set provided by *Consensus Economics*¹. Each month, starting in October 1989, the London-based firm *Consensus Economics* polls major economic organizations like important investment banks or research institutes about their predictions for the most common macroeconomic variables.² Since most of the panelists are located in the country they are forecasting upon, country-specific expertise is guaranteed. Initially conducting the surveys only for the G7 countries, *Consensus Economics* meanwhile provides forecasts for more than 70 countries from over 700 panelists. Data sets

¹We will describe what additional data we use in the respective subsections.

²Variables included in the survey are: Real growth of gross domestic product (GDP), real growth of private consumption, real growth of industrial production, consumer price inflation (CPI), short and long term interest rates, changes in the real effective exchange rate, real growth of fixed investment, the unemployment rate, the current account balance, and a few other macroeconomic variables.

are available for the categories G7 and Western Europe, Asia Pacific, Eastern Europe and Latin America.³

In line with the most widely communicated type of forecasts in the media, the participating institutions are asked to state their predictions for the annual averages of specific variables for the current and the subsequent calendar year rather than a forecast for the development over a fixed horizon of time. There are also such fixed horizon forecasts available from *Consensus Economics*,⁴ but the data summarized and analyzed here solely consist of monthly observations on the first mentioned type of forecast, the so called fixed event forecasts. A big advantage of the data set is, that estimates are comparable across countries as well as panelists. This is ensured by the survey procedure; *Consensus Economics* uses the same selection procedure across countries and publishes its survey for all countries simultaneously in the second week of each month.

Since the surveys by *Consensus Economics* have been conducted for a relatively short period of time only, few scientific work has utilized its data.⁵ In addition, there has been not much work presented in the literature that makes use of the full (disaggregated) information that is contained in the *Consensus* data set - presumably since coding the full data set is quite a cumbersome task.

2.2 Sample Issues

Our sample from the survey by *Consensus Economics* contains the forecasts for the G7 countries starting in October 1989 and ending in October 2006. This means we are able to work with a maximal time series dimension of 205 observations. During the sample period some breaks occurred with respect to the definition of variables which the panelists had to report on. In the United Kingdom inflation was traditionally measured by the growth rate of the retail

³In the studies presented in this thesis, we concentrate on data for the G7-countries or even only those data for the four largest European countries.

⁴Once a quarter *Consensus Economics* procures forecasts for a fixed time horizon of four quarters.

⁵Other prominent survey data sets on macroeconomic forecasts that are widely used in macroeconomic research include the Survey of Professional Forecasters collected by the Federal Reserve Bank of Philadelphia (see <http://www.phil.frb.org/econ/spf/> for more information) that is used in section 4.1 below, the corresponding Survey of Professional Forecasters conducted by the ECB (see Bowles et al. (2007)), the Livingston survey (see Curtin (1996) for details), the household survey of the European Commission (more information at http://ec.europa.eu/economy_finance/indicators/businessandconsumersurveys_en.htm), and the Michigan Survey of Consumer Expectations and Behavior (see Curtin (1996) for details).

price index (RPI). In May 1997 the relevant index changed to RPIX, which is the RPI excluding mortgage interest payments. A second change for this series is dated to January 2004 when the target changed to the CPI. Other target changes were those from GNP to GDP. For the US this took place in January 1992, for Japan in January 2000, and for Germany in January 1993. The shift from West-German data to data for re-unified Germany took place in June 1997.

As time passed, a number participating institutions suffered bankruptcy, others were acquired, and some merged with other participating panelists. Using synergy effects many newly formed companies closed one of their former research facilities. Other panelists joined the panel of forecasters later in the sample period. Therefore, the number of participating institutions varies across time. Some of the panelists are also listed under varying names in the publications of *Consensus Economics*. We connected series that refer to the same panelist. In addition, we tried to merge series of two panelists whenever they are related because one of them came to existence due to a merger or acquisition involving the other panelist. We have also checked the individual expectations, which substantially differ from others and made sure they correctly reflect the figures in the hard copy reports from *Consensus Economics*.⁶ Besides, for each respondent some observations were linearly interpolated when a single observation was missing and both adjacent observations were available and equal to each other. Typically, less than 10 percent of observations were gained by this procedure. Eventually, this leaves us with an average of 46 panelists per country. The highest numbers are counted for the United Kingdom (57) and the USA (56) whereas fewest forecasters took part in the survey for Canada (29). However, through the varying length of individual time series, the overall numbers of panelists in each country do not translate directly into the number of predictions observed in each month which means that we are facing an average cross-section dimension that is far smaller. It turns out that there are substantial differences between the G7 countries. While the average size of the monthly cross-section is fairly comparable between Germany (27.0), the UK (29.5) and the US (25.9) there is a huge gap to that of Italy, Canada and France, where the sample contains only an average of 13.6, 15.0 and 17.7 panelists respectively.

⁶For instance three observations for inflation forecasts for Japan in February and March 2002.

2.3 Rationality of Forecasts⁷

If we use the *Consensus* survey data set for the estimation of various economic models below, we should test whether the expectations expressed in the forecasts of the data set fulfil basic properties that we can expect macroeconomic expectations to have. If the properties of this specific data set on the contrary showed substantial deviation from what is usually found for other data sets, that would question the general interpretability of our results (for instance from Chapter 3 or Section 4.2). Therefore, we present an assessment of the basic properties of the forecasts in this section. As a side effect, by using this large amount of disaggregate data on individual macroeconomic forecasts we are able to provide much broader evidence on the properties of macroeconomic forecasts than has been available in the literature so far.

One weak point of the empirical literature on survey data is that there is only a limited number of non-US data sets, which provide information on forecasts. Consequentially, existing evidence is predominantly based on US data. Notable exceptions are Harvey et al. (2001), who analyze a set of selected individual forecasts for the UK from the *Consensus* data set, Gallo et al. (2002), who analyze the evolution of macroeconomic forecasts for the US, the UK, and Japan, Bowles et al. (2007), who analyze the performance of forecasts summarized in the Survey of Professional Forecasters conducted by the European Central Bank, Isiklar et al. (2006) or Ager et al. (2009), who use data from the *Consensus* data set on forecasts for a set of industrialized countries, Timmermann (2007), who analyzes the performance of IMF forecasts from the World Economic Outlook for various countries, and Batchelor (2001), who compares the forecasts made by the IMF and the OECD to private sector forecasts. However, all the existing international studies, with the exception of Harvey et al. (2001), make exclusive use of consensus forecasts rather than analyzing individual forecasts. This part of the thesis is written to fill this gap by covering forecasts for all G7 countries.

Our results are based on an approach that is commonly used in the literature to model the structure of macroeconomic forecasts. This literature dates back to early contributions by Ball (1962), Mincer and Zarnowitz (1969), Figlewski and Wachtel (1981), or Nordhaus (1987), who introduced the basic model framework for analyzing fixed event forecasts.⁸ Fixed event forecasts refer to the case that a forecaster constructs a sequence of forecasts over time

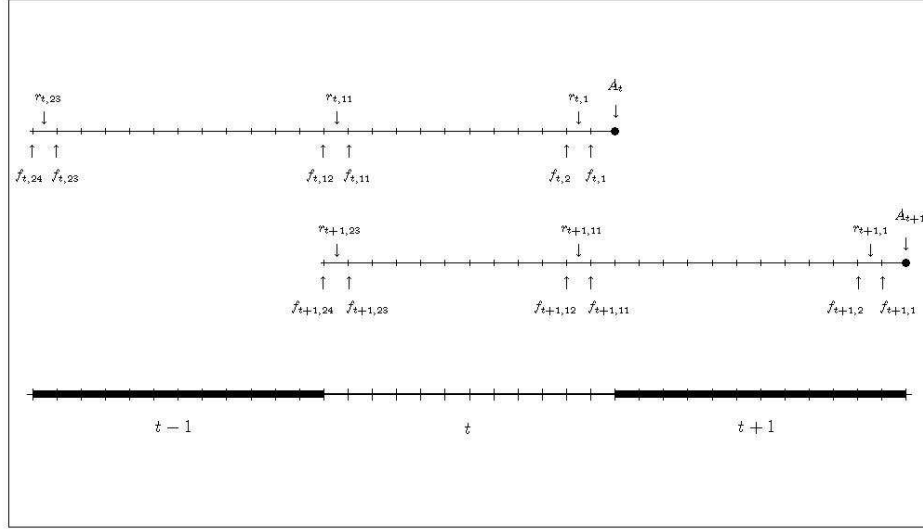
⁷Large parts of this section of the thesis are based on a paper that I wrote jointly with a co-author (Dovern and Weisser, 2008).

⁸Pesaran and Weale (2006) and Smant (2002) present nice summaries of the commonly used approaches. The latter contribution also provides an overview about the most prominent survey data sets that are used in empirical research on forecast efficiency.

for the same event (such as an annual figure for a macroeconomic variable). The data we are using below is of this type. A couple of more recent contributions have made proposals to improve the econometric approach for testing rationality of such large panels of fixed event forecasts. These include Keane and Runkle (1990) and Batchelor and Dua (1990), who introduce the analysis in a panel framework using the Generalized Methods of Moments (GMM) method, or Davies and Lahiri (1995), who develop the analytic framework for analyzing three dimensional panels of survey data which makes the use of information along all dimensions possible. To make our results comparable to existing studies, we follow the approach suggested by Davies and Lahiri (1995) and recently used by Clemens et al. (2007) and Ager et al. (2009) very closely and suggest only minor modifications.

Using this model framework we test whether the forecasts provided by the panelists of the survey are unbiased and efficient. Assuming that forecast accuracy is the only objective of a forecaster and that her loss function is symmetric and increasing in the forecast error, these two properties are necessarily a feature of a rational forecast. Since the work by Pigou (1927) or Keynes (1936), it is widely accepted that expectations and forecasts play a crucial role in all kinds of economic dynamics. Muth (1961) introduced the notion of rational expectation, which has since played a central role in economic thinking. So, although in a strict sense the concept of rational expectations does always refer to model-consistent expectations in economic theory, we believe that it is important for econometricians to analyze observed forecasts to check whether they show at least the basic features of “rational” expectations.

It should be noted, however, at this point that there are also arguments against the assumption that published forecasts reflect true expectations and should, thus, be rational if made by rational agents. Some of the cases against this assumption are the following. First, forecasters might seek to maximize public attention. If this is the case, an unbiased forecast is not optimal anymore, since the utility of the forecaster depends on more than one argument (Laster et al., 1999). Second, forecasters might produce so-called “intentional” forecast in some situations (Stege, 1989). A forecaster could, for example, predict a specific event to provoke a policy action that actually prevents the realization of the event. Third, Forecasters might have asymmetric loss functions (Capistran and Timmermann, 2006). They could, for example, have different weights concerning a possible over- or underestimation of an outcome. We believe, however, that these arguments are not particularly strong a priori. We, therefore, abstract from them and start this analysis from the hypothesis of rational forecasts, which are unaffected by these issues.

Figure 2.1: Structure of Survey Data

Notes: The figure shows a schematic representation of the timing of a sequence of forecasts made by one forecaster for two consecutive years.

2.3.1 Model

The *Consensus* data set is a panel data set and contains so-called fixed event forecasts. Since due to this structure it exhibits a special correlation structure, it is important to give a clear picture of the nature of the data before moving to the description of the tests that we are going to use.

The panel possesses a three dimensional structure of the kind introduced in Davies and Lahiri (1995). For each country and variable we have a NTH -vector of forecasts for T years made by N forecasters with forecast horizons ranging from one month to H months

$$F = [f_{1,1,H}, f_{1,1,H-1}, \dots, f_{1,1,1}, f_{1,2,H}, \dots, f_{1,T,1}, f_{2,1,H}, \dots, f_{N,T,1}]' \quad (2.1)$$

In other words, for each year we collect a sequence of H forecasts from each forecaster, starting H months before the year ends and ending in the last month of the respective year. Figure 2.1 shows a schematic diagram of the data structure. This structure will be of importance later on when we derive the correlation between different forecast errors.

PhD-Thesis written by: Jonas Dovern

Following Davies and Lahiri (1995), we assume that the forecast error for each forecast can be decomposed into three different parts⁹

$$e_{i,t,h} \equiv A_t - f_{i,t,h} = \phi_i + \lambda_{t,h} + \epsilon_{i,t,h} , \quad (2.2)$$

where A_t denotes the realization of a variable for year t . The first error term ϕ_i is the individual bias of the forecasts made by forecaster i . The second error term $\lambda_{t,h}$ is common to all forecasters and reflects the occurrence of macroeconomic shocks that hit an economy between the date at which the forecasts are made and the end of year t . Following the literature, we assume that these shocks are cumulated over the h months in an arithmetic way, so that the error term can be written $\lambda_{t,h} = \sum_{k=1}^h u_{t,k}$. We assume that $u_{t,h}$ is distributed with a zero mean and a variance of σ_u^2 . Since $u_{t,h}$ and $u_{t+1,h+12}$ occur at the same point in time, they will be correlated (Davies and Lahiri, 1995).

Modeling of Individual Errors¹⁰

The third error term $\epsilon_{i,t,h}$ refers to the forecaster specific error component of the forecast error. The literature proposes two alternative ways to model this error. On the one hand, it can be seen as an independently and identically distributed (*iid*) shock. This is the view taken for the estimation in Davies and Lahiri (1995). On the other hand, Davies and Lahiri mention that one could assume that over time each forecaster receives a flow of private information on the outcome. Under this assumption, one can model the individual error as $\epsilon_{i,t,h} = \sum_{k=1}^h \eta_{i,t,k}$, where the $\eta_{i,t,k}$ are distributed with mean 0 and variance σ_i^2 . Again, $\eta_{i,t,k}$ and $\eta_{i,t+1,k+12}$ have a non-zero correlation, since these information shocks occur at the same point in time.

It is clear that the two model variants for $\epsilon_{i,t,h}$ have very different implications. In the first case, the individual errors are assumed to be a white noise process, while in the second case they are assumed to follow a random walk for each target year t . In the first case, there would be no correlation between consecutive individual errors, while in the second case the autocorrelation would be very high and decaying only slowly for higher distances between two forecast errors for the same target year. Intuitively, the second model is much more appealing: Consider a forecaster whose forecast is above the consensus forecast in one month. Isn't it very likely that he is going to publish an above-average forecast also in the following month? That is, it is

⁹To enhance readability of the formulas we suppress an index referring to a specific variable throughout the following sections.

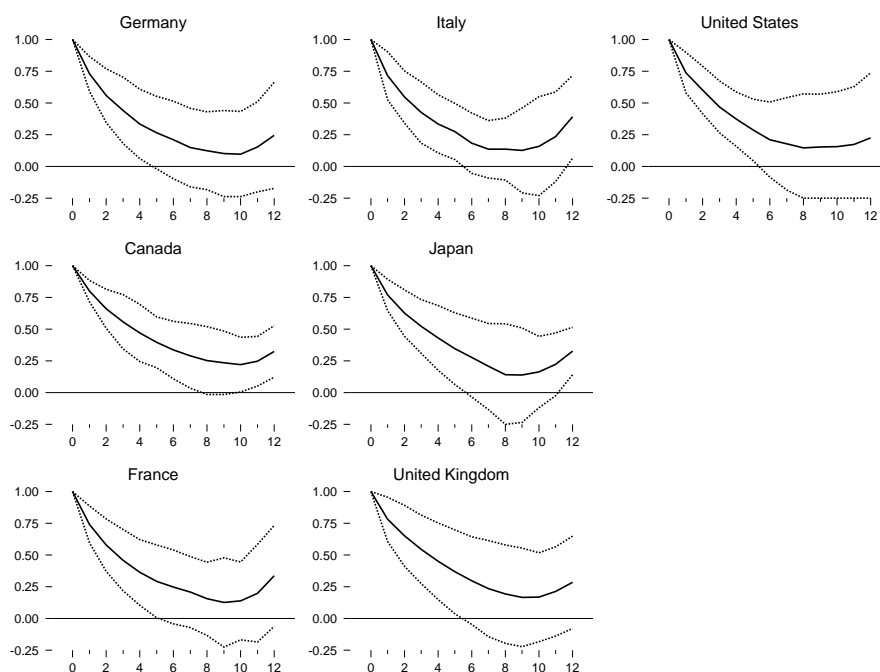
¹⁰This subsection is based on a note that I wrote (Dovern, 2009a).

very strange to believe that individual forecasts fluctuate randomly around the consensus forecast without any persistence. Rather a forecaster is likely to be persistently more optimistic or pessimistic than the average for some time. This behavior would be better captured by the second model that implies a high autocorrelation of the individual errors.

Eventually, choosing from the two alternatives is a matter of empirical facts. In our data set, the estimates of the individual errors, say $\hat{\epsilon}_{i,t,h}$, show a fairly high degree of autocorrelation. The empirical autocorrelation functions are usually declining slowly and approach zero only after twelve month or so. We have formally, investigated this issue using the GDP forecasts for the G7 countries. We propose to assess which model is more in line with the empirical evidence by looking at the empirical autocorrelation function (ACF) of $\hat{\epsilon}(i,t,h)$ and by testing implications of the two models. Under the first option the autocorrelation should be 0 for all lags. Under the second option the series of $\hat{\epsilon}(i,t,h)$ for each target year is a unit root process (effectively a random walk) and the autocorrelation function declines only slowly as the lag increases.

Figure 2.2 shows the estimated average ACFs for all countries together with a confidence band of \pm two standard deviations (computed across all forecasters for each country). It is evident for all cases that the idiosyncratic errors are highly persistent. This persistence is highest in Canada and lowest in Germany and Italy. Though this visual inspection suggests that the modeling of the idiosyncratic errors as accumulation of *iid* shocks is more appropriate than the alternative, it is better to also look at additional formal tests.

Table 2.1 shows the summary of results for ADF tests for unit roots in the data and the test for zero autocorrelation. Row 1 contains for each country the average estimated first order autocorrelation and its standard deviation; the values are quite high - ranging from 0.71 for Italy to 0.79 for Canada. In all of the individual cases we can reject the hypothesis that the series have zero autocorrelation which would be implied by the iid option. Row 3 contains for each country the average test statistic of the ADF tests and its standard deviation; the values range from -5.11 for UK to -6.40 for Italy. In all of the individual cases we can reject the hypothesis of a unit root based on a critical value of -2.87 on a 5% confidence level. Thus, none of the two models, which are proposed in the literature to model the idiosyncratic part of the forecast error of fixed event forecasts, are truly consistent with the data. Still, we can make use of Bayesian information criteria to judge which of the two models is preferred by the data. Row 5 shows the average across forecasters of the Bayes Factor obtained for the random walk model against the white noise model under flat priors, i.e. the prior belief that both

Figure 2.2: Autocorrelation Functions of Idiosyncratic Forecast Errors for GDP

Notes: The line in the middle is the average estimated autocorrelation function across forecasters. The two outer lines are upper and lower confidence bands calculated by using two times the standard deviation of the autocorrelations across all forecasters.

models are equally likely a priori. The values are all above 1 indicating that among the two alternatives one would, on average, clearly favor the random walk model against the white noise model to describe the idiosyncratic errors. Effectively, there is not one single case for which the white noise model would be preferred.

So, while we have to reject the null hypothesis of a unit root in the individual errors for virtually all cases, we usually prefer the second model against the first model based on Bayesian information criteria.¹¹ Against this background, we think there are good reasons to modify the model used by Davies and Lahiri (1995). Hence, we will model the individual error components according to the assumption that forecasters receive private information which are accumulated over time.

¹¹We expect results for other variables to follow the results for the GDP forecasts closely.

Table 2.1: Tests of Alternative Models for Idiosyncratic Forecast Errors for GDP

	Germany	Canada	France	Italy	Japan	UK	US
Avg. Autocorrelation	0.73 (0.08)	0.79 (0.05)	0.75 (0.06)	0.71 (0.1)	0.75 (0.06)	0.78 (0.08)	0.73 (0.1)
Avg. ADF-Statistic	-5.96 (1.58)	-5.3 (1.1)	-5.84 (1.43)	-6.4 (2.07)	-5.55 (1.31)	-5.11 (1.41)	-5.67 (1.29)
<i>BayesFactor_{RW}</i>	1.42 (0.22)	1.65 (0.2)	1.43 (0.24)	1.36 (0.19)	1.45 (0.19)	1.61 (0.44)	1.38 (0.27)
# Forecasters	29	14	17	11	13	30	22

Notes: Numbers are calculated across all forecasters for each country. Numbers in parenthesis denote the corresponding standard deviations across forecasters. The corresponding critical value for the individual ADF tests on a 5% confidence level is -2.87.

Test of Unbiasedness

Testing the unbiasedness of forecaster i is equivalent to testing whether $\phi_i = 0$ in (2.2). We can examine this hypothesis by testing the zero restriction on the elements of $\Phi = [\phi_1, \dots, \phi_N]'$ in

$$e = A - F = \Phi \otimes i_{TH} + \underbrace{\lambda + \epsilon}_{=\nu}, \quad (2.3)$$

where e is the vector of stacked forecast errors, A is given by $i_N \otimes (A^+ \otimes i_H)$ with $A^+ = (A_1, A_2, \dots, A_T)'$ and i_{TH} , i_N and i_H are vectors of ones of dimension TH , N and H respectively.¹² λ and ϵ are vectors of length NTH in which we stack the appropriate $\lambda_{t,h}$ and $\epsilon_{i,t,h}$ respectively.

Now, while a simple OLS regression gives consistent point estimates for the bias, we cannot base our inference on the OLS standard errors, since the elements of ν are clearly not *iid* due to the special correlation structure caused by the structure of the panel data set. Davies and Lahiri (1995) show that it is neither diagonal nor homoscedastic. Recalling that due to our assumption about the individual errors our specification differs from their model, we formally have the following elements of $\Sigma = E[\nu\nu']$ for two forecasters, say i and j :

$$Cov(\nu_{i,t_1,h_1}, \nu_{j,t_2,h_2}) = Cov\left(\sum_{k=1}^{h_1} u_{t_1,k} + \sum_{k=1}^{h_1} \eta_{i,t_1,k}, \sum_{k=1}^{h_2} u_{t_2,k} + \sum_{k=1}^{h_2} \eta_{j,t_2,k}\right) \quad (2.4)$$

$$Cov(\nu_{i,t_1,h_1}, \nu_{j,t_2,h_2}) = \begin{cases} \min\{h_1, h_2\} [\sigma_u^2 + \sigma_i^2] & \text{if } i = j, t_1 = t_2, h_1 = h_2 \\ \min\{h_1, h_2 - 12\} [\sigma_u^2 + \sigma_i^2] & \text{if } i = j, t_1 = t_2 - 1, h_2 \geq 12 \\ \min\{h_1, h_2\} \sigma_u^2 & \text{if } i \neq j, t_1 = t_2, h_1 = h_2 \\ \min\{h_1, h_2 - 12\} \sigma_u^2 & \text{if } i \neq j, t_1 = t_2 - 1, h_2 \geq 12 \\ 0 & \text{else} \end{cases}$$

¹²The operator \otimes denotes the Kronecker Product.

Clearly, the different non-zero cases deserve some more explanation. The forecast errors ν are correlated across several dimensions. First, they are correlated within the maximum forecast horizon H since $\lambda_{t,h}$ and $\epsilon_{i,t,h}$ are the accumulation of period-specific shocks; this refers to the first case shown in (2.4). Second, the forecast errors are correlated between subsequent years since the forecast horizons are of overlapping nature; this refers to the second case shown in (2.4). Finally, the forecast errors are correlated across different forecasters, since forecast errors are produced at the same time and are all subject to the same subsequent aggregate shocks summarized by $\lambda_{t,h}$; this refers to the third and fourth case shown in (2.4).

Given Σ , the covariance matrix of the Generalized Methods of Moments (GMM) estimator is given by

$$\text{Var}(\hat{\Phi}) = [(I_N \otimes i_{TH})'(I_N \otimes i_{TH})]^{-1} [(I_N \otimes i_{TH})'\Sigma(I_N \otimes i_{TH})][(I_N \otimes i_{TH})'(I_N \otimes i_{TH})]^{-1} \quad (2.5)$$

and can be used to derive valid t-statistics for testing $\phi_i = 0$. Naturally, Σ is not observed and has to be replaced by a consistent estimate, say $\hat{\Sigma}$, before computation of the test statistics is possible.

Though Σ has a complicated pattern, it depends only on $N+1$ parameters, namely $\sigma_1^2, \dots, \sigma_N^2$ and σ_u^2 . Davies and Lahiri (1995) propose to obtain a consistent estimate by first estimating these $N+1$ parameters and then replacing the parameters in Σ by the corresponding estimates. We will follow this approach. Note that an estimator of ϕ_i is simply given by the average forecast error of forecaster i and that we can estimate the two other parts of the forecasts errors by

$$\hat{\lambda}_{t,h} = \frac{1}{N} \sum_{i=1}^N (A_t - f_{i,t,h} - \hat{\phi}_i) \quad (2.6)$$

and

$$\hat{\epsilon}_{i,t,h} = A_t - f_{i,t,h} - \hat{\phi}_i - \hat{\lambda}_{t,h} . \quad (2.7)$$

We can obtain estimates for the unknown parameters as the estimated coefficients from the following regressions:

$$\hat{\lambda} \odot \hat{\lambda} = \sigma_u^2 \kappa_H + \omega_\lambda \quad (2.8)$$

$$\hat{\epsilon} \odot \hat{\epsilon} = (I_N \otimes \kappa_H) \sigma^2 + \omega_\epsilon \quad , \quad (2.9)$$

where $\kappa_H = i_T \otimes [H, H-1, \dots, 1]'$ and $\sigma^2 = [\sigma_1^2, \dots, \sigma_N^2]'$. Clemens et al. (2007) and Ager et al. (2009) show how the problem simplifies when $N = 1$. Essentially, all terms that refer to the individual component of the forecasting

errors vanish if one assumes that there simply is no “private” shocks in this setup with one agent only (Ager et al., 2009) or if one models these “private” shocks like we do – so that they become observationally indistinguishable from the aggregate shocks (Clemens et al., 2007).¹³

Test of (Weak) Efficiency

For testing the efficiency of the forecasts, we use the concept of weak-form efficiency that has been originally proposed by Nordhaus (1987). The concept starts from the notion of strong efficiency of forecasts which requires that all information, which has been revealed at the time a forecast is made, is taken into account during the forecasting process. In other words: If a series of forecasts is strongly efficient, it would have not been possible to improve the forecast performance by using any information available also to the forecaster. Since the amount of potentially relevant information is immense and any selection for an empirical analysis would be ad-hoc,¹⁴ Nordhaus (1987) proposes to restrict the relevant information set to lagged values of the forecasts themselves. He shows that under weak form efficiency the revisions of forecasts should be uncorrelated under certain assumptions. It should be intuitively clear that for efficient forecasts the current forecast should not reveal any information on future revisions – or as Nordhaus states (p. 673):

If I could look at your most recent forecasts and accurately say, “Your next forecast will be 2% lower than today’s”, then you can surely improve your forecasts.

Against this background, weak-form efficiency of a sequence of forecasts can be formally tested using an equation of the form

$$r_{i,t,h} = \beta_i r_{i,t,h+k} + \xi_{i,t,h} , \quad (2.10)$$

where $r_{i,t,h}$ is defined as $f_{i,t,h} - f_{i,t,h+1}$, $k \geq 1$, and $\xi_{i,t,h}$ is the error term. The hypothesis of weak-form efficiency implies $\beta_i = 0$; a consistent estimate of β_i can be obtained by the OLS estimator treating $\xi_{i,t,h}$ as white noise. But again – due to the special structure of the fixed event forecasts – the covariance

¹³Ager et al. (2009) analyze consensus forecasts with this $N = 1$ setup by noting that the individual errors should cancel out in the aggregate. This is true, however, only when the consensus forecast is based on a large sample of panelists. For small panels, one gets components of the covariance matrix of the aggregate error that reflect the aggregated variance of the individual errors. Hence, when the consensus forecast is based on a small number of panelists, treating the former as an $N = 1$ case does bias the estimated variances and covariances downward leading to too many rejections of unbiasedness (or efficiency).

¹⁴Not to mention the problem of constructing large data sets with real-time vintages.

matrix of $\xi = [\xi_{1,1,H-(k+1)}, \dots, \xi_{N,T,1}]'$, say $\Xi = E[\xi\xi']$, is non-diagonal and heteroscedastic.

To derive the exact form of Ξ , we first note that, using (2.2), we can re-write the forecast revisions as

$$r_{i,t,h} = f_{i,t,h} - f_{i,t,h+1} = \lambda_{t,h+1} - \lambda_{t,h} + \epsilon_{i,t,h+1} - \epsilon_{i,t,h} = u_{t,h+1} + \eta_{i,t,h+1}. \quad (2.11)$$

Now, it is evident that under the Null hypothesis $\beta_i = 0$ we obtain the following expressions for the elements of Ξ :¹⁵

$$\begin{aligned} Cov(\xi_{i,t_1,h_1}, \xi_{j,t_2,h_2}) &= Cov(u_{t_1,h_1+1} + \eta_{i,t_1,h_1+1}, u_{t_2,h_2+1} + \eta_{j,t_2,h_2+1}) \quad (2.12) \\ Cov(\xi_{i,t_1,h_1}, \xi_{j,t_2,h_2}) &= \begin{cases} \sigma_u^2 + \sigma_i^2 & \text{if } i = j, t_1 = t_2, h_1 = h_2 \\ \sigma_u^2 + \sigma_i^2 & \text{if } i = j, t_1 = t_2 - 1, h_1 = h_2 - 1 \\ \sigma_u^2 & \text{if } i \neq j, t_1 = t_2, h_1 = h_2 \\ \sigma_u^2 & \text{if } i \neq j, t_1 = t_2 - 1, h_1 = h_2 - 1 \\ 0 & \text{else} \end{cases} \quad (2.13) \end{aligned}$$

Given Ξ , the covariance matrix for the GMM estimator of β can be written as

$$Var(\hat{\beta}) = (r'_{+k} r_{+k})^{-1} r'_{+k} \Xi r_{+k} (r'_{+k} r_{+k})^{-1}, \quad (2.14)$$

where $r_{+k} = [r_{1,1,H-1}, \dots, r_{1,1,(k+1)}, r_{1,2,H-1}, \dots, r_{N,T,(k+1)}]'$ and $\beta = [\beta_1, \dots, \beta_N]'$. $Var(\hat{\beta})$ can be used to derive valid t-statistics for testing $\beta_i = 0$. Naturally, Ξ is not observed and has to be replaced by a consistent estimate, say $\hat{\Xi}$, before computation of the test statistics is possible.

To obtain $\hat{\Xi}$ we can use the same method that we used to derive $\hat{\Sigma}$. First, we derive estimates for the single elements of Ξ and replace these elements in a second step by their estimates to consistently estimate Ξ . Note that the structure of Ξ is much more simple than that of Σ so that its elements are simply given by

$$\hat{\sigma}_u^2 = \frac{1}{T(H - (k + 1))} \sum_{t=1}^T \sum_{h=1}^{H-(k+1)} (\hat{u}_{t,h+1}^2) \quad (2.15)$$

and

$$\hat{\gamma}_i^2 = \frac{1}{T(H - (k + 1))} \sum_{t=1}^T \sum_{h=1}^{H-(k+1)} (\hat{\eta}_{i,t,h+1}^2), \quad (2.16)$$

¹⁵Note that at this point the assumption of private information for $\epsilon_{i,t,h}$ is crucial for the result that under weak-form efficiency $\beta_i = Cov(r_{i,t,h}, r_{i,t,h+1}) = Cov(u_{t,h+1} + \eta_{i,t,h+1}, u_{t,h+2} + \eta_{i,t,h+2}) = 0$. Under the assumption that the $\epsilon_{i,t,h}$ represent ordinary iid shocks we would get $\beta_i = Cov(u_{t,h+1} + \epsilon_{i,t,h+1} - \epsilon_{i,t,h}, u_{t,h+2} + \epsilon_{i,t,h+2} - \epsilon_{i,t,h+1}) = -\sigma_i^2 \neq 0$.

where $\hat{u}_{t,h}$ and $\hat{\eta}_{i,t,h}$ are consistently estimated by

$$\hat{u}_{t,h} = \frac{1}{N} \sum_{i=1}^N r_{i,t,h-1} \quad (2.17)$$

and

$$\hat{\eta}_{i,t,h} = r_{i,t,h-1} - \hat{u}_{t,h} . \quad (2.18)$$

Given this formal framework, we will now move to the empirical analysis of the macroeconomic forecasts in the G7 countries.

2.3.2 Results

Individual Forecasts

In this section we present all results concerning the properties of the individual forecasts. For both – the tests of unbiasedness as well as the tests of weak efficiency – we include those panelists in the sample who made a forecast at more than 50% of the possible dates. Thereby, we avoid the influence of small sample problems which could arise from those panelists that reported only a few forecasts.¹⁶

An additional feature of the data that we have to deal with is given by the fact that the record of most of the forecasters includes a bunch of missing values, i.e. the panel is heavily unbalanced. There are two reasons for that. First, the set of panelists who take part in the *Consensus* survey changes continuously. Hence, there are some forecasters that enter the panel at a later stage, while other forecasters leave the panel after the first part of the time period covered by our data set. Second, some forecasters do not submit their forecast on a regular basis, i.e. some of them do not provide their current forecasts for some of the months. To minimize the reduction of our data base due to the second issue, we interpolate a missing value in all those cases, in which a forecast is missing only for one month in a row *and* the two adjacent forecasts are equal to each other. Formally, if $f_{i,t,h}$ is missing *and* $f_{i,t,h+1} = f_{i,t,h-1}$, we set the missing forecast equal to $f_{i,t,h+1}$.

For the estimation, we follow Davies and Lahiri (1995) and deal with missing values by simply deleting the appropriate elements in the vectors of forecast errors or revisions and the corresponding rows and columns in the covariance matrices respectively. Those compressed matrices can be directly used in the GMM estimation procedure (Blundell et al., 1992).

¹⁶The threshold of 50% is of course arbitrary. Results for the included panelists are, however, robust to the inclusion of more forecasters in the used sample.

Unbiasedness

The analysis of the biases present in the individual forecasts reveal some interesting differences across countries as well as variables. The results are summarized in Table 2.2.¹⁷ In general, most of the individual forecasts are unbiased. The overall performance is best for the inflation forecasts; only for France, the UK and the US there are one or two forecasters respectively who produce biased inflation forecasts. Especially surprising is the good performance of inflation forecasts for Italy that underwent a significant transition from a high inflation regime towards a low inflation regime during the early sample period. One could imagine that forecasters adjusted only slowly to this new environment causing forecasts to be biased upwards.

This behavior is actually what can be observed for the inflation forecasts in the UK, where inflation was also very high at the beginning of our sample period and declined considerably to low levels in the mid of the 1990s. All but three panelist, which entered the sample rather late, have overestimated inflation on average. After all, only 2 out of 30 did so significantly on a 95% confidence level.

A similar argument applies to the bias that is found for most of the forecasts for GDP growth in the European countries. Here, the wide majority of forecasters overestimate growth on average. This phenomenon is most pronounced in Germany and Italy but applies to a lesser extend also to France. The same is also true for the forecast for growth of private consumption in Germany. Batchelor (2007) shows that this kind of bias can be inevitable in an environment of declining trend growth rates since forecasters have to gradually learn about the new trend.

A very special picture is given by the combination of forecasts for GDP growth and for growth of industrial production in the UK. While forecasts for the former are generally unbiased, the results show strong evidence for rejecting the hypothesis of unbiased forecasts for the latter forecasts; most panelist on average overestimate growth of industrial production by about 1 to 1.5 percentage points. This might reflect the fact that although the trend growth of overall output remained relatively constant over the sample, there has been a shift of the structural composition of the economy in the UK from production oriented sectors towards services – especially towards the financial sector – which had to be learned by the forecasters. A similar phenomenon can be observed when comparing forecasts on GDP growth for the US, which are generally unbiased, to forecasts for growth of private consumption in the US, which tend to underestimate consumption growth. Again it seems that it has been hard for a large number of panelists to anticipate the gradual

¹⁷Detailed results on basis of individual panelists are available on request.

decline in the saving rate of private households as well as to properly estimate additional consumption effects of huge increases in household wealth that was caused by the stock market boom of the late 1990s and the real estate booms during the time from 2002 until the end of our sample.

In general, we can conclude that biased forecasts seem to be produced in times of structural changes or gradual developments that have to be learned by the forecasters; this source for bias in macroeconomic forecasts is also supported by the results in Andolfatto et al. (2008) who analyze the properties of artificial forecasts generated within a standard dynamic stochastic general equilibrium model. On the contrary, forecasts seem to be unbiased in general for stable economies without large structural shocks. One example is Canada where the structure of the economy and the medium term growth trend have not fundamentally changed since the introduction of inflation targeting in 1991. As a consequence, there is not one single case among all forecasts for the Canadian economy in which the panelist produced biased forecasts.

Weak Efficiency

For testing weak efficiency of individual forecasts we followed the literature (Clemens, 1995; Harvey et al., 2001; Isiklar et al., 2006) by setting k in (2.10) equal to 1. This makes indeed sense, since by the time a new revision is made each forecaster knows about his most recent previous forecast revision. The results are summarized in Table 2.3.¹⁸

The analysis of the individual forecasts' properties in terms of weak efficiency reveals an interesting contrast between the forecasts made for GDP growth and those for the other variables under investigation. For the majority of forecasts for growth of industrial production and private consumption as well as for the inflation rate we cannot reject the hypothesis of weakly efficient forecasts; only few series of forecasts show a significant correlation between proceeding forecast revisions. In those cases, the estimated coefficient is mostly negative which means that the corresponding forecasters tend to overreact to incoming news, i.e. they initially revise their forecasts by an amount that is too large and undo part of this revision during the next forecasting round.

In contrast, we find more evidence for deviations from weak efficiency for forecasts of GDP growth in all countries but Italy and Japan.¹⁹ The main dif-

¹⁸Again, detailed results for all individual panelists are available on request.

¹⁹The fact that we find weakly efficient forecasts for GDP growth for Japan is in contrast to the results of Ashiya (2003), who analyzes the reaction to news of forecasters for GDP growth in a slightly different modeling framework and based on a different set of private sector forecasts; he concludes that the individual forecasters tend to significantly overreact to new information.

Table 2.2: Bias of Individual Forecasts

Gross Domestic Product							
	# obs	# bias	# bias pos	# bias neg	mean(bias)	var(bias)	cons biased?
Germany	29	7	0	7	-0.49	0.021	no
Canada	14	0	0	0	-0.27	0.020	no
France	17	4	0	4	-0.40	0.016	no
Italy	11	10	0	10	-0.61	0.002	yes (-)
Japan	13	0	0	0	-0.22	0.042	no
UK	30	0	0	0	-0.25	0.025	no
USA	22	0	0	0	0.23	0.023	no
Inflation							
	# obs	# bias	# bias pos	# bias neg	mean(bias)	var(bias)	cons biased?
Germany	30	0	0	0	-0.06	0.011	no
Canada	14	0	0	0	0.25	0.028	no
France	17	2	0	2	-0.14	0.021	no
Italy	11	0	0	0	0.20	0.010	no
Japan	13	0	0	0	-0.13	0.011	no
UK	30	2	0	2	-0.20	0.025	no
USA	22	1	0	1	-0.14	0.027	no
Industrial Production							
	# obs	# bias	# bias pos	# bias neg	mean(bias)	var(bias)	cons biased?
Germany	28	0	0	0	-1.00	0.078	no
Canada	4	0	0	0	-0.93	0.120	no
France	8	2	0	2	-0.98	0.093	yes (-)
Italy	8	4	0	4	-1.53	0.048	no
Japan	13	0	0	0	-1.26	0.089	no
UK	28	22	0	22	-1.40	0.056	yes (-)
USA	21	0	0	0	-0.46	0.077	no
Private Consumption							
	# obs	# bias	# bias pos	# bias neg	mean(bias)	var(bias)	cons biased?
Germany	29	18	0	18	-0.52	0.039	yes (-)
Canada	14	0	0	0	0.03	0.051	no
France	17	0	0	0	-0.18	0.008	no
Italy	11	0	0	0	-0.40	0.005	no
Japan	13	0	0	0	-0.31	0.012	no
UK	29	0	0	0	0.15	0.033	no
USA	22	10	10	0	0.51	0.019	no

Notes: *#obs* indicates the number of individual panelists, *#bias* the number of them which provides significantly biased forecasts and *#bias pos/neg* the direction of the biases. *Mean(bias)* and *Var(bias)* indicate the mean and the variance of the individual biases and *cons biased?* provides comparison to the corresponding consensus forecast.

ference, however, is that the estimated coefficients are positive in all but one of the significant cases. Thus, those forecasts for GDP growth, which deviate from weak efficiency, show a strong tendency towards forecast smoothing in

general. This means that forecasters tend to process new information only slowly which results in positively autocorrelated revisions.²⁰ Also Gallo et al. (2002) find that forecasters tend to stick to their past forecast even when the authors control in their study for the most recently observed average forecast and the dispersion of forecasts. Batchelor and Dua (1992) rationalize such a forecasting behavior by noting that in reality forecasters might not have a single objective which is minimizing the expected squared errors. They are likely to take into account as well that their clients might “mistrust forecasters who make frequent [erratic] revisions to forecasts” (p. 179). The fact that usually the forecast for GDP growth is the part of a comprehensive macroeconomic forecast published by a forecaster, which is most widely anticipated and discussed by clients or the media, might bring about that it is exactly this forecast, for which incentive and reputation considerations make forecasters deviate most from their true expectations. This would explain why we find the strong tendency for forecast smoothing only for forecasts on GDP growth.

Consensus Forecast

In this section, we present the results concerning the properties of the average forecast, the so-called *consensus forecast*.²¹ Average forecasts have been frequently used in empirical research although results based on them should be treated with caution because of inconsistency problems due to the aggregation bias (Bonham and Cohen, 2001) caused for instance by not-accounting for private information (Figlewski and Wachtel, 1981) or the fact that the aggregation might cancel out deviations from unbiasedness of individual forecasters (Keane and Runkle, 1990). We, therefore, present the results based on average forecasts to compare their properties to those of the individual forecasts.

Unbiasedness

We can simply use the framework presented in the last section with $N = 1$ to analyze the bias of the consensus forecasts. In order to be able to compare the results for consensus and private forecasts, we start by discussing the results under the assumption of a homogenous bias across horizons. They are given in Table 2.4. The estimation outcomes show that for all countries we cannot

²⁰This phenomenon is also known as *conservatism* in psychology (Phillips and Edwards, 1966; Edwards, 1968).

²¹Note that all panelists are included in the computation of the average forecast. Hence, unlike in the analysis of individual forecasts we do not exclude those panelists who reported less than 50% of all possible forecasts over the entire sample.

Table 2.3: Efficiency of Individual Forecasts

Gross Domestic Product							
	# obs	# ineff	# ineff pos	# ineff neg	mean(beta)	var(beta)	cons eff?
Germany	24	9	9	0	0.10	0.007	no (+)
Canada	12	5	5	0	0.13	0.011	no (+)
France	15	9	9	0	0.15	0.014	no (+)
Italy	10	2	2	0	0.05	0.010	no (+)
Japan	10	0	0	0	0.02	0.003	no (+)
UK	26	7	6	1	0.07	0.012	no (+)
USA	21	8	8	0	0.10	0.009	no (+)
Inflation							
	# obs	# ineff	# ineff pos	# ineff neg	mean(beta)	var(beta)	cons eff?
Germany	26	4	0	4	-0.03	0.009	no (+)
Canada	13	1	0	1	-0.03	0.008	no (+)
France	15	2	1	1	-0.01	0.008	no (+)
Italy	10	1	0	1	-0.01	0.004	yes
Japan	11	1	0	1	-0.07	0.004	no (+)
UK	25	7	0	7	-0.08	0.023	yes
USA	20	4	1	3	-0.05	0.013	no (+)
Industrial Production							
	# obs	# ineff	# ineff pos	# ineff neg	mean(beta)	var(beta)	cons eff?
Germany	23	6	3	3	-0.03	0.019	no (+)
Canada	2	0	0	0	0.02	0.014	yes
France	7	2	1	1	0.01	0.014	no (+)
Italy	7	3	1	2	-0.01	0.020	no (+)
Japan	10	0	0	0	0.02	0.003	no (+)
UK	24	4	2	2	0.02	0.011	no (+)
USA	20	4	1	3	-0.02	0.013	no (+)
Private Consumption							
	# obs	# ineff	# ineff pos	# ineff neg	mean(beta)	var(beta)	cons eff?
Germany	27	9	0	9	-0.10	0.009	no (+)
Canada	12	1	0	1	-0.02	0.006	no (+)
France	15	3	2	1	0.02	0.007	no (+)
Italy	10	1	0	1	-0.03	0.006	no (+)
Japan	10	0	0	0	-0.03	0.002	no (+)
UK	24	3	1	2	-0.01	0.010	no (+)
USA	20	2	1	1	0.02	0.007	no (+)

Note: *#obs* indicates the number of individual panelists, *# ineff* the number of them which provides significantly weakly inefficient forecasts and *# ineff pos/neg* the sign of the estimated coefficient. *Mean(beta)* and *Var(beta)* indicate the mean and the variance of the estimated individual coefficients and *cons eff?* provides comparison to the corresponding consensus forecast.

reject the hypothesis that the consensus forecasts for inflation are unbiased. For all other variables the picture is mixed. First, the average forecasts for

growth of private consumption are biased upwards in Germany while they are significantly, but only weakly, too pessimistic in the US. For the other five countries the corresponding forecasts are unbiased. Second, consensus forecasts for GDP growth are unbiased in all countries but Germany, France and Italy where they tend to be too optimistic on average. And finally, the average forecasts for growth of industrial production are biased upwards in France, Italy, and the UK while being unbiased for the remaining four countries. These figures broadly reflect the picture given by the individual forecasts, i.e. in cases where a sizable fraction of individual forecasts was found to be biased also the consensus forecast is biased.

Table 2.4: Bias of Consensus Forecasts

		Germany	Canada	France	Italy	Japan	UK	USA
GDP	phi	-0.52*	-0.28	-0.44*	-0.63***	-0.28	-0.24	0.18
	t-stat	-1.79	-0.90	-1.66	-2.64	-0.67	-0.83	0.62
Inflation	phi	-0.03	0.23	-0.15	0.18	-0.13	-0.20	-0.13
	t-stat	-0.19	0.54	-1.09	0.61	-0.99	-1.31	-0.88
Ind. Prod.	phi	-1.00	-0.91	-1.06**	-1.56**	-1.35	-1.43**	-0.48
	t-stat	-1.56	-1.52	-1.96	-2.25	-1.28	-2.85	-0.83
Priv. Cons.	phi	-0.56**	0.04	-0.19	-0.44	-0.36	0.18	0.48*
	t-stat	-2.24	0.14	-1.07	-1.50	-1.27	0.60	1.87

Notes: *, **, *** denotes rejection of the null at the 10%, 5% and 1% significance level, respectively.

So far we have assumed that bias does not vary with respect to the forecast horizon. Since it is a reasonable hypothesis that this might be wrong, we relax the restriction of a constant bias for the average forecasts.²² To do this robustness analysis, we write the aggregated forecast errors as

$$e_{t,h} = A_t - f_{t,h} = \phi_h + \sum_{k=1}^h u_{t,k}^*,$$

where now ϕ_h is a horizon-specific bias and $u_{t,k}^* = u_{t,k} + \eta_{t,k}$. A consistent estimator for ϕ_h is given by $\hat{\phi}_h = 1/T \sum_{t=1}^T e_{t,h}$. Note that under the joint hypothesis that the forecasts are unbiased for all forecast horizons ($H_0 : \phi_H = \dots = \phi_1 = 0$), the structure for the covariance matrix of the residuals is equal to the one in the model with biases across forecast horizons restricted to be equal to each other.

²²Note that we cannot do the same in the analysis of individual forecasts since for the wide majority of panelists the data sets includes so many missing values that the estimate for each horizon-specific bias would be based on 10 or even less observations.

As it is not the central focus of this study, we refrain from tabulating the results for horizon-specific biases in detail. Instead we give a broad picture by referring to the sequential test of forecast unbiasedness as introduced by Ager et al. (2009). The null-hypothesis of this test is that all horizon-specific biases up to some horizon h are jointly equal to 0 ($H_0 : \phi_1 = \dots = \phi_h = 0$). The Wald statistic is given by $W = (\phi_h)'(Var(\phi_h))^{-1}(\phi_h)$ and χ^2 -distributed with h degrees of freedom. The corresponding F-statistic, which possesses better small sample properties is obtained by dividing the Wald statistic by h .

The results are summarized in Table 2.5.²³ They show that for most variables and countries the null hypothesis of joint insignificance of the horizon specific biases can be rejected for horizons greater than 14 to 16. Noteworthy exceptions are the forecasts on GDP for Canada, on inflation for Germany and Italy, as well as the ones on industrial production for France and Japan, where joint insignificance can not be rejected for any horizon. Furthermore the test rejects the null hypothesis also for low horizons in some cases. These are the forecasts on inflation in Canada and France and those for private consumption for Japan.

Table 2.5: Sequential Test of Forecast Unbiasedness for Consensus Forecasts

	Germany	Canada	France	Italy	Japan	UK	US
GDP	15-24		14-24	14-24	15-24	15-24	16-24
Inflation		1-4; 24	1; 14; 15; 19-24		15-24	14-24	14-24
Ind. Prod.	16-24	15-24		24		14-24	16-24
Priv. Cons.	15-24	14-24	15-24	15-24	1-3; 14-24	24	15-24

Notes: Each cell shows the horizons h for which the null-hypothesis of joint insignificance of biases can be rejected at the 5% level.

Weak Efficiency

The results from the tests of weak efficiency for the consensus forecasts demonstrate very well that caution is required when working with average forecast data. In contrast to the setup for the analysis of individual forecasts we set $k = 2$ for the analysis of the consensus forecasts (Isiklar, 2005). The reason is that it is not clear whether each forecaster knows already about the most recent consensus forecast when a new forecast is produced, since in the extreme case the forecasts have to be reported two weeks before a new consensus forecast is published and additionally the production process for each forecast might last more than a week depending on the institutional

²³More detailed results are available on request.

framework of a specific forecaster. In any case, each forecaster should know about the average forecast published two months ago.

Table 2.6 shows the results for implementations of the test based on (2.10). It is obvious that the results taken at face value would lead to completely different conclusions than those seen in the previous section. Clearly, all average forecasts except in two cases²⁴ show evidence for forecast smoothing, i.e. incoming information gets reflected in the average forecast in a very sluggish way. The effect is indeed most strongly visible for forecasts for GDP growth also here, but even for variables for which the individual forecasters tend to overreact to news we find the opposite deviation from weak efficiency in the consensus forecast (e.g. growth of private consumption in Germany).

Table 2.6: Efficiency of Consensus Forecasts (Without Horizon Effects)

		Germany	Canada	France	Italy	Japan	UK	USA
GDP	beta	0.49***	0.27***	0.34***	0.34***	0.25***	0.37***	0.23***
	t-stat	7.79	4.31	4.92	5.23	3.94	6.49	3.67
Inflation	beta	0.19***	0.17***	0.14**	0.06	0.20***	0.09*	0.23***
	t-stat	2.99	3.09	2.23	1.02	3.27	1.66	3.65
Ind. Prod.	beta	0.46***	0.00	0.33***	0.38***	0.33***	0.49***	0.28***
	t-stat	6.94	0.06	5.02	5.46	4.91	8.10	4.67
Priv. Cons.	beta	0.32***	0.17***	0.24***	0.19***	0.18***	0.29***	0.16**
	t-stat	4.94	2.71	3.79	3.13	2.82	4.92	2.53

Notes: *, **, *** denotes rejection of the null at the 10%, 5% and 1% significance level, respectively.

Like mentioned already in the previous section it is possible to estimate horizon specific biases for the consensus forecasts. If we relax the assumption of one single bias for all forecast horizons, this has implications for the construction of the test on weak efficiency. Namely, the unconditional expectation for a revision is no longer equal to zero under the null hypothesis in that case, since $r_{t,h} = f_{t,h} - f_{t,h+1} = \phi_{h+1} - \phi_h + u_{t,h+1}^*$. We, therefore, expand (2.10) by including constant terms for each forecast horizon. Dropping the index i , the new equation on which we base the robustness check for our results is

$$r_{t,h} = \beta r_{t,h+2} + \phi_{h+1} - \phi_h + \xi_{t,h}. \quad (2.19)$$

The coefficient β can be consistently estimated by regressing the consensus forecasts' revisions on its second lags and a set horizon-specific dummy-variables using OLS. An estimate for the variance of $u_{t,h+1}^*$ can be obtained

²⁴Those are the inflation forecasts in Italy and the forecasts for growth of industrial production in Canada.

by regressing the revisions on the set of horizon-specific dummy-variables and taking the mean of the squared residuals from that regression.

The results that are given in Table 2.7 confirm the evidence presented in the previous paragraph. The point estimates for the correlations between two subsequent revisions do not change qualitatively. The maximum difference between two corresponding point estimates is 0.12 in absolute values.

Table 2.7: Efficiency of Consensus Forecasts (With Horizon Effects)

		Germany	Canada	France	Italy	Japan	UK	USA
GDP	beta	0.43***	0.26***	0.29***	0.24***	0.26***	0.36***	0.22***
	t-stat	6.94	4.11	4.27	3.78	4.01	6.30	3.55
Inflation	beta	0.20***	0.16***	0.14**	0.10*	0.19***	0.06	0.27***
	t-stat	3.07	2.98	2.25	1.68	3.22	1.13	4.27
Ind. Prod.	beta	0.44***	-0.02	0.30***	0.28***	0.35***	0.37***	0.27***
	t-stat	6.67	-0.37	4.44	4.06	5.28	6.21	4.44
Priv. Cons.	beta	0.23***	0.17***	0.21***	0.12**	0.16***	0.30***	0.12**
	t-stat	3.69	2.71	3.37	1.99	2.55	4.97	1.96

Notes: *, **, *** denotes rejection of the null at the 10%, 5% and 1% significance level, respectively.

2.3.3 Discussion of Results

In general, our findings confirm, based on a broad data base, previous results about rationality properties of macroeconomic forecasts. In addition, we have already pointed to some interesting particularities across countries and variables in the previous section. In what follows, we discuss briefly the relation between the results regarding individual and consensus forecasts and summarize the findings of the study.

To facilitate comparison, Tables 2.2 and 2.3 also provide information about the results for consensus forecasts in their last columns. As demonstrated above, almost all consensus forecasts exhibit characteristics of forecast smoothing and are, thus, not weakly efficient. It is evident that this holds even if the wide majority of individual forecasters is weakly efficient. This phenomenon can be explained by the fact that new information is processed by some forecasters slower than by others. It results in positive autocorrelation of the revisions of the consensus forecasts. Similar results hold for the analysis of biases. Naturally, the consensus is unbiased if there are only a few individual panelists who produce biased forecasts. As soon as there is a significant fraction of forecasters who report biased forecasts, it depends on the correlation of their biases whether the consensus will be biased or not. If

forecasters deviate into both directions from unbiasedness, the biases might cancel out in the aggregate. This, however, is not the case in our sample. As already mentioned, forecasters tend to be biased into the same direction for a specific target variable. Therefore, the consensus is biased if there is a sizable fraction of biased individual forecasts.

To recapitulate the findings of this section: First, our results confirm that data on average forecasts should be used with caution since even in a situation where all individual forecasts are rational the hypothesis of rationality is often rejected based on aggregate data. Second, we find large difference in the performance of forecasters across countries and different macroeconomic variables. Third, among the four kinds of forecasts that we analyze, inflation forecasts perform best in terms of unbiasedness. Fourth, forecasts tend to be biased in situations where forecasters have to learn about large structural shocks or gradual changes in the trend of a variable. Nevertheless, our results show that forecasts from the *Consensus* data set have desirable properties by and large.²⁵

2.4 Concluding Remarks

In this chapter of the thesis, we have presented a characterization of the survey data set that will be most widely used in the empirical analysis below. We have elaborated on the background and geographical coverage of the survey as well as about the available data sample size. In the foregoing section we have demonstrated that the forecasts included in the *Consensus* data set have fairly good properties in terms of unbiasedness and efficiency which makes us quite confident that they represent true expectations of the panelists and are distorted by institutional incentive issues only to a marginal extend.

²⁵There are several dimensions along which the study could be expanded in the future. For simplicity, we have assumed that the variance of the macroeconomic shocks ($\lambda_{t,h}$) as well as the variance of the idiosyncratic forecast error ($\epsilon_{i,t,h}$) decay linearly if h goes to 1. One could also imagine other (more general) functional forms. As soon as enough longer time series become available for individual forecasters one could implement the assumption of a horizon specific bias also in the analysis of individual forecasts. Taking into account correlations across countries – like Isiklar et al. (2006) do in their analysis of consensus forecasts – would clearly be desirable. Currently however, this would require too much computational power for the estimation of the covariance matrices. Finally, it would be promising to investigate whether the results concerning the inefficiency of the consensus forecasts could be used to construct some kind of “adjusted consensus forecasts” that are superior in terms of forecast accuracy.

Chapter 3

Inflation (Expectation) Dynamics

As discussed in Section 1.2 one of the most crucial questions in macroeconomic theory – especially monetary theory – is how inflation is determined. Since all competing models describing inflation dynamics attribute a fairly large role to inflation expectations, it is of utmost importance to understand how certain economic agents form expectations about the future path of inflation. In the next section, we use the *Consensus* data set to present empirical evidence for a model of expectation formation of households that has been proposed by Carroll (2003). Subsequently, in section 3.2, we use similar data to estimate the Sticky Information Phillips Curve that has been proposed by Mankiw and Reis (2002) to describe the aggregate price setting mechanism in the economy.

3.1 Estimating Expectation Formation of Households¹

Several recent papers (including Mankiw and Reis, 2002, 2006) argue that sticky information models, in which agents update their information occasionally rather than instantaneously, resolve some puzzles in the output–inflation dynamics. For example, sticky information models are able to account for considerable inflation persistence and substantial sacrifice ratios (recessionary disinflations) typically observed in the data.

¹Large Parts of this section of the thesis are based on a paper that I published jointly with three co-authors (Döpke et al., 2008a).

Microeconomic foundations for the sticky information paradigm were elaborated in Carroll's (2003) work on the "epidemiological model of expectations." Carroll argues that US survey data on inflation expectations are consistent with a model in which in each period only a fraction of households adopts inflation forecasts of rational experts. The remaining households find it costly to update their information and continue using their past expectations rather than form fully rational predictions. In related work Sims (2003), Branch (2004) and others provide alternative justifications for models with agents that do not instantaneously incorporate all available information as implied by most standard macro models.

While the sticky information approach seems useful for modeling US data, corresponding evidence for European countries is, to the best of our knowledge, still lacking.² The analysis presented in this section attempts to fill this gap by investigating inflation expectation data from four major EU economies (France, Germany, Italy, and the United Kingdom).

In general, our findings support the usefulness of the Carroll's sticky information model for the description of inflation dynamics in European countries. We find that household inflation expectations adjust sluggishly to the more precise predictions of professional forecasters. The average speed of this adjustment varies little across the four countries in our sample and is somewhat lower than that in the US: a typical household updates its inflation expectations roughly once in eighteen months (compared to once a year previously found in the US). While this result is quite robust across the two estimation methods, we find that the frequency of information updating in Europe is somewhat higher for the vector error-correction specification, amounting to about once a year. Similarly to the US, European households are not backward-looking: they tend to update their expectations from experts' rational forecasts rather than actual past inflation rates. Finally, the estimates are stable over time: we cannot distinguish any statistically significant differences between various institutional settings (e.g., inflation-targeters and non-targeters).

For policy-makers these results imply that inflation expectations of (European) consumers are sluggish. Consequently, the channel from household expectations to actual inflation rates is likely to remain an important source of inflation persistence even when central banks gain (even) more credibility (in fighting inflation) and even if expert expectations are rational (in that they adjust instantaneously to new information).

²The only work testing sticky information models on international data is Khan and Zhu (2002) and Khan and Zhu (2006). However, these two papers have to adopt restrictive assumptions to circumvent data limitations and approximate agents' expectations with forecasts from a VAR model.

3.1.1 Model

Carroll (2003) proposes the following model of the transmission of inflation expectations. The economy consists of two types of agents: experts (professional forecasters) and households. Experts collect in every period relevant information on future inflation, make rational inflation forecasts and publish them in newspapers. Because reading newspapers (or making informed inflation forecasts) is costly, households—in contrast to the standard frictionless rational framework—choose to update their expectations occasionally rather than instantaneously. As a result, new information about inflation spreads slowly across population in the following “epidemiological” way. In each period only a randomly chosen fraction λ_1 of households follows the latest inflation stories and updates its inflation expectations. The remaining $1 - \lambda_1$ households stick to their forecasts from the previous period. The evolution of the (average) household (denoted *HH*) inflation (π) expectation (\mathbf{E}) follows:

$$\mathbf{E}_t^{HH} \pi_{t,t+1} = \lambda_1 \mathbf{E}_t^{EX} \pi_{t,t+1} + (1 - \lambda_1) \mathbf{E}_{t-1}^{HH} \pi_{t,t+1},$$

where $\mathbf{E}_t^{HH} \pi_{t,t+1}$ and $\mathbf{E}_t^{EX} \pi_{t,t+1}$ denote one-period-ahead inflation expectations of households and experts (*EX*), respectively.

Thus, news about inflation can be thought of as a disease that spreads slowly across the population, infecting a fraction λ_1 of all households in each period. The calculation outlined in Carroll (2003) leads to a similar equation formulated for expected one-year-ahead inflation rates collected at quarterly frequency, which are typically reported in surveys of inflation expectations:

$$\mathbf{E}_t^{HH} \pi_{t,t+4} = \lambda_1 \mathbf{E}_t^{EX} \pi_{t,t+4} + (1 - \lambda_1) \mathbf{E}_{t-1}^{HH} \pi_{t-1,t+3}. \quad (3.1)$$

(3.1) holds if (i) inflation follows a random walk process (or households believe that inflation is a random walk) or (ii) $\mathbf{E}_{t-1}^{HH} \pi_{t-1,t+3} \approx \mathbf{E}_{t-1}^{HH} \pi_{t,t+4}$. Both of these assumptions are likely to be satisfied in our dataset. As discussed below, the underlying CPI inflation process in the core European economies has, indeed, been very persistent recently, warranting the random walk approximation. In addition, given the high persistence of the inflation process, there is not much difference between household expectations as of time $t - 1$ of inflation rates at $t + 3$ and $t + 4$, which, in turn, implies that condition (ii) is also likely to be met.

3.1.2 Data Issues

To test the model of information diffusion, we need two inflation expectations series: inflation forecasts of households and professional forecasters.

PhD-Thesis written by: Jonas Dovern

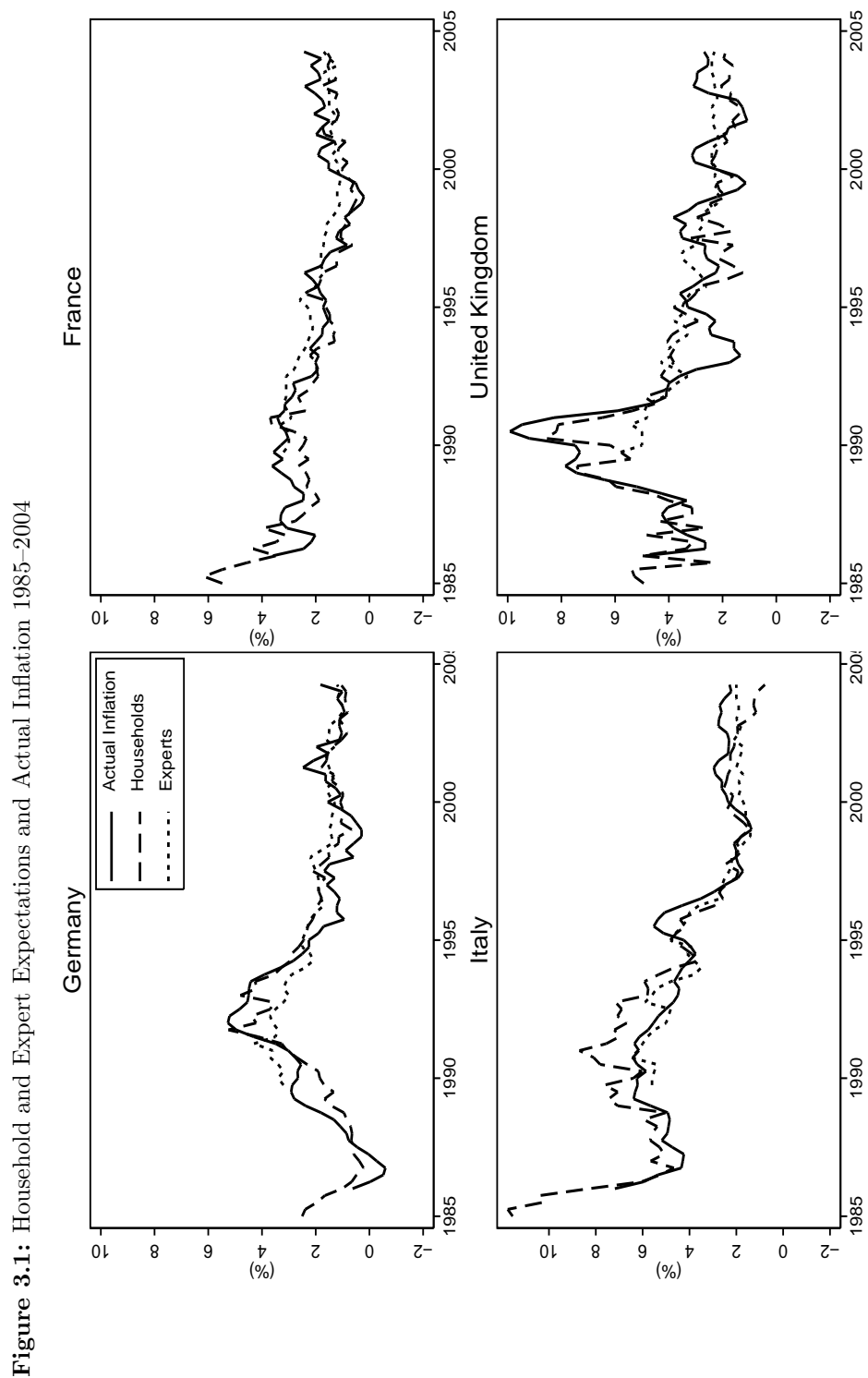
The forecasts of households were obtained from the European Commission's (EC) consumer survey and those of professional forecasters from *Consensus Economics*, a London-based macroeconomic survey firm.³

Household expectations were constructed using the EC survey's question 6, which asks how, by comparison with the last 12 months, the respondents expect that consumer prices will develop in the next 12 months.⁴ Unfortunately, the answers are qualitative rather than quantitative (unlike, for example, the question on expected inflation in the US Michigan Survey of Consumer Sentiment). This means that the respondents are asked about the direction of the expected movement of consumer prices (increase/fall), not about the exact value of this movement. Consequently, care needs to be taken when transforming these data into quantitative measures of household expectations, required to test (3.1). We follow much of the existing literature (including Gerberding (2001) and Mankiw et al., 2003) in adopting the Carlson and Parkin (1975) method, which is briefly explained in Appendix A.

To appropriately rescale the expectations it is necessary to specify a variable that captures the perceived current level of the inflation rate. We experimented with three choices of $\tilde{\mu}_t$: (i) recursive Hodrick–Prescott (HP) filter, (ii) past inflation, (iii) inflation extracted as fitted values from the regression of inflation on the past balance statistics from the survey. The recursive HP filter was calculated using the following quasi-real-time procedure to minimize the well-known end-of-sample problems. For each period, t , we first forecast the underlying inflation process for the next twelve quarters with an ARMA model, selected with the Akaike criterion (with the maximum number of four lags on both AR and MA terms). We then apply the filter on this artificially extended series (with the HP filter with the usual quarterly

³We only investigate Germany, France, Italy and the United Kingdom because expectations of professional forecasters from other European countries, such as the Netherlands and Spain, are not available before 1996. Note that the data used in this section differs from the data set that is described in Section 2.1 in that the forecasts used here are only reported as averages across forecasters, they are only collected on a quarterly basis, and they enable us to directly infer the one-year-ahead forecasts rather than expectations for the annual averages.

⁴The exact wording of question 6 of the Consumer Survey of the Joint Harmonised EU Programme of Business and Consumer Surveys is: “By comparison with the past 12 months, how do you expect that consumer prices will develop in the next 12 months?” For more information on the survey, see the Commission's web page, http://ec.europa.eu/economy_finance/indicators/businessandconsumersurveys_en.htm.



penalty parameter $\lambda_{HP} = 1600$). Finally, we set $\tilde{\mu}_t$ equal to the value of the HP filtered inflation as of time t .⁵

Figure 3.1 compares expert and household inflation expectations with actual inflation rates. Most of the time both expert and household predictions are close to actual inflation. However, sometimes there are rather persistent differences between expectations and actual inflation. More importantly, because household and expert expectations occasionally differ quite considerably (such as in the early 1990s in the UK), a closer examination of the dynamic interaction of both variables is warranted.

3.1.3 Results

The choice of the appropriate empirical strategy depends on the time series properties of the underlying expectations. If the series are stationary, (3.1) can be estimated directly using OLS (as in Carroll, 2003). If they are non-stationary (I(1)) and cointegrated, the model should be transformed into an vector error-correction (VEC) form prior to estimation.

A number of recent papers investigate the degree of persistence of various measures of inflation in Europe.⁶ Although these studies agree that inflation is a very persistent process, the evidence on its order of integration is less conclusive. Most papers cited by ECB (2005) reject the null hypothesis that inflation in large European countries has recently been non-stationary. In contrast, the recent work of O'Reilly and Whelan (2005) as well as our preliminary tests (investigating inflation and inflation expectations in the EU-4 countries) in general do not reject the null hypothesis of a unit root in inflation.

A potential criticism of our results is that the sample is too short to allow reliable inference. The fact that we are unable to reject the null may well result from the notoriously low power of the unit root tests under such circumstances, rather than the existence of the unit root.

Since the main focus of this paper is not on providing the definitive answer on the order of integration of inflation (or inflation expectations), we directly move on to estimating Carroll's theoretical model and investigating how sensitive its implications are with respect to the stationarity assumptions. We

⁵Most results reported in the paper hold for all normalizations considered. In particular, we find that the alternative rescaling procedures typically imply that λ_1 is positive and significant and the summing-up restriction $\lambda_1 + \lambda_2 = 1$ is met. One result that does not hold for normalizations (ii) and (iii) is that household expectations are insensitive to past inflation. This is not surprising because for these two normalizations the past inflation rate directly enters $\tilde{\mu}_t$.

⁶See ECB (2005), Table 3.4, for a summary of the literature on European countries.

first estimate the Carroll (2003) model in the stationary framework. We then briefly consider how the results are affected if the non-stationary (VEC) setup is adopted.

The Stationary Case: The Carroll Model

Before estimating (3.1), we will examine some preliminary evidence on the relationship between expert and household expectations. Given the interest in the interaction between the expectations of both professional forecasters and households, a natural starting point is to ask: Which of the two group forecasts is (on average) better? And what is the causality between the two expectations? We provide the answers in the next subsection.

Relationship Between Expert and Household Expectations

First, expert expectations are substantially more precise than household expectations. The root mean squared errors of expert forecasts are between 15% to 35% lower in Germany, Italy and the UK than for household expectations. The two expectations are comparably precise in France.⁷ This result does not, of course, come as a surprise since households may know expert forecasts when forming their own expectations. According to the epidemiology model, at least some households update their own expectations by following experts. In addition, in an environment with stable inflation like that in France in the post-1980s, neither forecast varies much.

Second, we can examine whether expert forecasts Granger-cause household forecasts. Results are given in Table 3.1; columns 3 and 4 indicate that lags of expert expectations are typically significant predictors of household expectations. Household expectations, on the other hand, tend not to Granger-cause experts. Thus, in all countries, except for Italy we conclude that the direction of causality goes from experts toward households.⁸

What Determines Household Expectations?

Having found supportive preliminary evidence for the epidemiological model of expectation formation, we now turn to direct estimation of the speed of

⁷This is in line with the findings of Gerberding (2001).

⁸Admittedly, the p-values on lagged household expectations in expert equations in the upper right corner of each country-specific panel, which range between 0.08 and 0.15, are quite low, which may suggest the existence of some feedback from households to experts. Note, however, that an additional piece of evidence supporting the causality from experts to households comes from the VEC setup of (3.4) below: While the loading coefficients in the household equations are, as reported in Table 3.5 below, significant, in the expert equations they are statistically indistinguishable from zero or have the wrong (positive) sign.

Table 3.1: Tests for Granger Non-causality

$$\mathbf{E}_t^i \pi_{t,t+4} = \beta_0 + \sum_{j=1}^p \beta_j \mathbf{E}_{t-j}^{EX} \pi_{t-j,t+4-j} + \sum_{k=1}^p \gamma_k \mathbf{E}_{t-k}^{HH} \pi_{t-k,t+4-k} + \varepsilon_t$$

Country	Dependent Variable: Expectations of ...	$\beta_j = 0, \forall j$ p-value	$\gamma_k = 0, \forall k$ p-value
Germany	Experts	0.000	0.125
	Households	0.000	0.000
France	Experts	0.000	0.076
	Households	0.000	0.000
Italy	Experts	0.000	0.010
	Households	0.620	0.000
United Kingdom	Experts	0.000	0.149
	Households	0.009	0.000

Notes: Sample runs from 1989Q4 to 2004Q2. The tests were computed with $p = 2$ lags of independent variables.

information updating, λ_1 . Table 3.2 summarizes the estimation results of various constrained versions of the following regression:

$$\mathbf{E}_t^{HH} \pi_{t,t+4} = \lambda_0 + \lambda_1 \mathbf{E}_t^{EX} \pi_{t,t+4} + \lambda_2 \mathbf{E}_{t-1}^{HH} \pi_{t-1,t+3} + \lambda_3 \pi_{t-5,t-1} + \varepsilon_t. \quad (3.2)$$

All models are estimated with seemingly unrelated regressions (SUR) with coefficients restricted to be equal across the four countries.⁹

The format of Table 3.2 follows that of Carroll (2003), Table III. The left panel (the first four columns) displays the point estimates of λ_1 s together with standard errors; the right panel shows specification tests (p-values of various tests of coefficients). The last column tests whether the coefficients are the same across countries.¹⁰ The models are labeled M1–M6. The first

⁹Analogous results hold when the models are estimated with equation-by-equation OLS. However, since the cross-correlation between residuals in our dataset is up to 0.3, SUR improve efficiency of our estimates. In addition, SUR also make it possible to test (and impose) cross-equation restrictions and answer questions such as: “Does the speed of information updating vary across countries?” We report some results unrestricted across countries below in Table 3.3.

¹⁰To conserve space we do not report measures of fit, which of course differ slightly for each country (and model). \bar{R}^2 s vary between 0.75 and 0.96. For more information of how well selected models explain household inflation expectations see Table 3.3 below. Similarly to Carroll’s findings, the Durbin–Watson statistics detect virtually no autocorrelation in residuals for models M1–M6: Since it is well-known that OLS estimates in a setup with lagged dependent variables and autocorrelated residuals are inconsistent, we have reestimated the models allowing for first-order autocorrelation in disturbances and found no statistically significant evidence (at the 5 % level) of autocorrelated errors.

Table 3.2: Restricted Cross-Country Results

$$\mathbf{E}_t^{HH} \pi_{t,t+4} = \lambda_0 + \lambda_1 \mathbf{E}_t^{EX} \pi_{t,t+4} + \lambda_2 \mathbf{E}_{t-1}^{HH} \pi_{t-1,t+3} + \lambda_3 \pi_{t-5,t-1} + \varepsilon_t$$

Model	λ_0	λ_1	λ_2	λ_3	Test	Cross eqn
					p-value	p-value
M1		0.17*** (0.04)	0.83*** (0.03)		$\lambda_1 + \lambda_2 = 1$ 0.912	0.04
M2		0.17*** (0.03)	0.83*** (0.03)		$\lambda_1 = 0.25$ 0.016	0.62
M3	-0.22*** (0.07)	0.29*** (0.05)	0.78*** (0.03)		$\lambda_0 = 0$ 0.003	0.15
M4		0.31*** (0.05)	0.77*** (0.05)	0.00 (0.04)	$\sum_{i=1}^3 \lambda_i = 1$ 0.003	0.03
M5	-0.22*** (0.07)	0.29*** (0.05)	0.78*** (0.05)	-0.01 (0.04)	$\lambda_3 = 0$ 0.900	0.13
M6			0.92*** (0.04)	0.05 (0.04)	$\lambda_2 + \lambda_3 = 1$ 0.015	0.34

Notes: Sample runs from 1989Q4 to 2004Q2. Seemingly unrelated regressions. Standard errors in brackets. *, **, *** denotes rejection of the null at the 10%, 5% and 1% significance level, respectively.

model, M1, estimates the following version of (3.2):

$$\mathbf{E}_t^{HH} \pi_{t,t+4} = \lambda_1 \mathbf{E}_t^{EX} \pi_{t,t+4} + \lambda_2 \mathbf{E}_{t-1}^{HH} \pi_{t-1,t+3} + \varepsilon_t, \quad (3.3)$$

in which coefficients λ_1 and λ_2 are estimated without any restriction. The estimates of λ_1 and λ_2 are 0.17 and 0.83, respectively; both are statistically significant. The summing-up restriction implied by the Carroll model, $\lambda_1 + \lambda_2 = 1$, is clearly satisfied.

Model M2 is estimated for the restricted version with the summing-up restriction imposed. Given how close the restriction is to being met in M1, it does not come as a surprise that the point estimates of λ_1 barely change.¹¹ Our baseline $\lambda_1 = 0.17$ suggests that on average, a European household reads economic updates or consults economic experts roughly once in 18 months.¹² In addition, the estimate implies about 47% of households use information outdated more than one year and about 23% more than two years. Thus

¹¹That the summing up restriction holds is not particularly surprising because it also holds outside the epidemiology framework. For example, this restriction is met in any model which implies that the household and expert expectations are cointegrated with a cointegrating vector $(1, -1)^\top$.

¹²This frequency is calculated as $1/\lambda_1 \approx 6$ quarters.

household expectations seem to react very sluggishly to the available macroeconomic news (that has already been processed by forecasters). In fact, our estimate λ_1 is much lower than one, which suggests the costs of acquiring new macroeconomic information are substantial (relative to the benefits).¹³

The speed of adjustment $\lambda_1 = 0.17$ is lower than Carroll's baseline coefficient of 0.27. Because the standard error on λ_1 is small, the difference is statistically significant. However, much of the difference between λ_1 in Europe and the US can be accounted for with different time ranges: Carroll's sample (1981Q3–2000Q2) differs from ours (1989Q4–2004Q2). Re-estimating model M2 with the US data and our sample range gives $\lambda_1 = 0.16$. This matches Carroll's evidence that updating is faster when inflation is in the news, including the early period in his sample. In contrast, in the 1990s λ_1 has fallen because inflation (in the US) received less press coverage than before 1985. In addition, the recent monetary policy leading to low and stable inflation has reduced uncertainty and, together with smaller exogenous shocks hitting the economy, has presumably further decreased households' incentives to update.

Models M3–M6 investigate four alternative structures of household expectations. First, we add a constant to (3.3). This parameter turns out to be negative and significantly different from zero. As we will see below (in Table 3.3), this result, which contradicts the simple epidemiology model (3.1), is driven mainly by the early sample in the UK where (around 1991) the actual inflation rate was falling considerably. In such an environment the epidemiology model (3.1) may not be adequate in that the non-updating $(1 - \lambda_1)$ households decide to adjust for the falling trend in inflation by adding a negative number to their past forecasts $\mathbf{E}_{t-1}^{HH} \pi_{t-1,t+3}$, which may then cause the negative estimate of λ_0 in Table 3.2 (and for the UK in Table 3.3 below).¹⁴

Models M4–M6 allow for the possibility that consumers are at least in part backward-looking (adaptive) by adding past inflation to the right-hand side of (3.2). Similarly to the US, there is very little of the backward-looking element in household inflation expectations: the coefficient λ_3 is small both in terms of its size and significance. Thus, households recognize that when

¹³The finding that uncertainty about aggregate variables matters little compared with idiosyncratic uncertainty is in line with the independent evidence on the small cost of eliminating business cycles, reported, e.g., in Krusell and Smith (1999), and in the sticky information setup with consumption dynamics by Carroll and Slacalek (2008).

¹⁴One piece of evidence that is consistent with this hypothesis is that re-estimating model M3 for the post-1992 sample gives an insignificant $\lambda_0 = -0.05$ (with the standard error of 0.07). (Estimating the same “unrestricted” model of Table 3.3 below for the same sample also gives for the UK an insignificant constant $\lambda_0 = -0.14$ with the standard error of 0.31.)

reading newspapers it makes more sense to learn from the forward-looking experts than to naively extrapolate past inflation rates.

Other, more general specifications of the epidemiology model could be considered. For example, how does augmenting the model (3.3) with more lags of expert and/or household expectations affect the estimates of λ_1 ? The results, which we do not reported here, suggest that adding an additional lag of household expectations does not affect the results much (because the coefficient is small and statistically indistinguishable from zero). Adding a lag of expert expectations ($\mathbf{E}_t^{EX} \pi_{t-1,t+3}$) causes the standard error on λ_1 to increase considerably and coefficients on expert expectations to become individually insignificant; jointly, however, they remain very significant and positive (and sum close to the number implied by the epidemiology model).

Generally, there appears to be a lot of homogeneity across countries. As indicated in the last column, in four of the six (M1–M6) models considered the null of equal coefficients in the four countries is satisfied; two models (M1 and M4) yield borderline rejections of homogeneity (at the 5% significance level).

Cross-Country Differences

Having found supportive evidence for Carroll’s sticky information model in European data, let us now investigate in more detail how the findings vary across countries. Table 3.3 summarizes estimation results obtained from seemingly unrelated regressions, unrestricted across countries, for models M1–M3.

The findings parallel those in the previous section: First, the speed of updating λ_1 varies between 0.11 and 0.32 (as estimated with models M1 and M2) and is, for all countries except Italy, highly statistically significant. Second, for all countries, except France, the summing-up restriction, $\lambda_1 + \lambda_2 = 1$, implied by the Carroll’s sticky information model, is met. Even for France, the two coefficients sum up to 0.91. Third, the intercept term λ_0 is insignificant for all countries except the UK.

The estimated speed of updating is only about half as large in Italy as in the other countries. This finding seems counterintuitive because the benefits from acquiring information in an environment with a relatively high level and uncertainty about inflation should actually be larger. We believe the econometric result is driven by the persistent differences between household and professionals’ expectations at the beginning of the sample, before 1993 (apparent in Figure 3.1). These deviations then may translate into a lower estimate of λ_1 . To investigate this hypothesis, we have re-estimated the regressions for the post-1992 period and found the Italian λ_1 substantially higher—about 0.25—and more significant.

Table 3.3: Country-by-Country Unrestricted Results

$$\mathbf{E}_t^{HH} \pi_{t,t+4} = \lambda_0 + \lambda_1 \mathbf{E}_t^{EX} \pi_{t,t+4} + \lambda_2 \mathbf{E}_{t-1}^{HH} \pi_{t-1,t+3} + \varepsilon_t$$

Model	λ_0	λ_1	λ_2	\bar{R}^2	Test
					p-value
Germany					
M1		0.18*** (0.06)	0.82*** (0.06)	0.91	$\lambda_1 + \lambda_2 = 1$ 0.764
M2		0.20*** (0.06)	0.80*** (0.06)	0.91	$\lambda_1 = 0.25$ 0.368
M3	-0.21* (0.12)	0.29*** (0.08)	0.80*** (0.06)	0.91	
France					
M1		0.32*** (0.08)	0.59*** (0.09)	0.85	$\lambda_1 + \lambda_2 = 1$ 0.002
M2		0.18*** (0.07)	0.82*** (0.07)	0.83	$\lambda_1 = 0.25$ 0.322
M3	-0.04 (0.12)	0.33*** (0.10)	0.61*** (0.09)	0.80	
Italy					
M1		0.14 (0.11)	0.86*** (0.09)	0.96	$\lambda_1 + \lambda_2 = 1$ 0.991
M2		0.11* (0.06)	0.89*** (0.06)	0.96	$\lambda_1 = 0.25$ 0.022
M3	-0.18 (0.15)	0.25* (0.13)	0.81*** (0.09)	0.95	
United Kingdom					
M1		0.23*** (0.08)	0.77*** (0.08)	0.89	$\lambda_1 + \lambda_2 = 1$ 0.763
M2		0.23*** (0.08)	0.77*** (0.08)	0.89	$\lambda_1 = 0.25$ 0.781
M3	-0.67** (0.30)	0.53*** (0.16)	0.66*** (0.09)	0.89	

Notes: Sample runs from 1989Q4 to 2004Q2. Seemingly unrelated regressions. Standard errors in brackets. *, **, *** denotes rejection of the null at the 10%, 5% and 1% significance level, respectively.

It is a bit surprising that the speed of updating does not seem to vary more across countries, which differ very much from each other in many respects.¹⁵ Distinct differences may show up more clearly once more data are available.

We could now in principle similarly test how stable λ_1 has been over time. However, due to the limited number of observations the tests of structural stability have weak power. We looked for differences in the speed on expectation updating by including dummy variables for, e.g., members of the euro area and inflation-targeters. Unfortunately, these investigations did not lead to any statistically clear-cut results about the determinants of λ_1 .

Most of our findings suggest that the epidemiology model of information diffusion performs similarly well, qualitatively as well as quantitatively, for the core European countries as it does for the US. Expert inflation expectations are typically more precise than household expectations. Econometric tests indicate that the Carroll model is adequate along several dimensions.¹⁶ Most models imply that European households update somewhat more slowly than US households, on average once in 18 months compared with once a year. Finally, there is strong evidence that, as suggested by the epidemiology model, European households update information from the professional forecasters rather than based on the past inflation rate.¹⁷

Econometric Issues: Measurement Error and Endogeneity

One concern about our SUR estimates is that they are affected by measurement errors in household inflation expectations (caused because household

¹⁵At least two kinds of institutional differences are potentially relevant for the size of λ_1 . First, the monetary policy setup and recent inflation dynamics in various EMU countries, the UK and the US are quite varied. For example, whereas Germany, under the Bundesbank regime, has always had moderate and stable inflation rates, Italy faced considerably higher inflation rates in the early 1990s and has witnessed pronounced declines in price level increases over the past decade in the run-up to and after the introduction of the euro. In addition, different communication strategies of central banks might affect how information spreads across households. (See for example Ehrmann and Fratzscher, 2005 and the literature cited therein.) Second, both the size and structure of the forecasting industry are dissimilar. This profession is in the US and the UK dominated by private forecasters, while in much of continental Europe, in particular in Germany, public forecasters play a more prominent role.

¹⁶For example, the speed of updating is positive and statistically significant, the summing-up restriction holds fairly well and household inflation expectations are not sensitive with respect to the past inflation.

¹⁷Consideration might be given to the possibility that households update their expectations by referring directly to other publicly available information, such as foreign prices. However, in the epidemiology framework this information is already captured and processed by professional forecasters, who are assumed to be rational. Moreover, obtaining such information is presumably much more costly than simply referring to the published professional forecasts.

expectations have to be rescaled). In presence of such classical measurement error coefficients λ_2 for model M1 in Table 3.3 are biased toward zero. To get a feel for how serious this bias is we have reestimated the models of Table 3.3 using past household and expert expectations as instruments for (the potentially mismeasured) $\mathbf{E}_t^{HH}\pi_{t-1,t+3}$. Compared to the SUR procedure the instrumental variables estimates of λ_2 fall roughly from 0.8 to 0.6. Together with this falling coefficient on lagged household expectations, the estimated speed of consumers' information updating λ_1 rises from about 0.2 to 0.4.

Another potentially relevant econometric issue with equations like (3.1) is endogeneity: Because experts presumably take into account households' expectations when making forecasts, they are affected by shocks to households' beliefs, which may induce correlation between $\mathbf{E}_t^{EX}\pi_{t,t+4}$ and the error term ε_t . We have attempted to address this issue by estimating the regressions instrumenting current expert expectations with their lags and found that the estimates were similar to those shown in Table 3.3.

The Non-Stationary Case: The Vector Error-Correction Form

Having estimated the epidemiology model in the stationary framework, let us now examine how the implications change when we assume that the expectation series are I(1) instead. Suppose we collect expert and household expectations in a vector $x_t = (\mathbf{E}_t^{HH}\pi_{t,t+4}, \mathbf{E}_t^{EX}\pi_{t,t+4})^\top$. If the two series are cointegrated with the cointegrating vector $\alpha = (1, -\alpha_1)^\top$, the system has the following vector error correction (VEC) representation:

$$\Delta x_t = \lambda_1 \alpha^\top x_{t-1} + \beta(L)\Delta x_t + \varepsilon_t, \quad (3.4)$$

where $\lambda_1 = (\lambda_1^{HH}, \lambda_1^{EX})^\top$ denotes the vector of loading coefficients and $\beta(L)$ is a matrix lag polynomial. Similarly to the stationary model (3.1), λ_1 determines the speed of adjustment toward the (long-run) equilibrium.¹⁸

We are particularly interested in λ_1^{HH} , which corresponds to the speed of adjustment observed for households. Furthermore, note that the theoretical derivation of the epidemiology model predicts a cointegrating vector $\alpha = (1, -1)^\top$. This is due to the fact that in the long-run households completely adapt to professional forecasts.

Before estimating the VEC representation (3.4) and its “ α -restricted” counterpart some preliminary specification tests need to be performed. First, we test whether there exists a valid cointegrating relationship between the

¹⁸The adjustment pattern in the partial adjustment version of the model (3.1), however, differs from the VEC analysis in two ways: First, the adjustment in the VEC is analyzed in an interdependent system and feedback effects are considered and second, the short-run dynamics in the VEC might influence the dynamic adjustment path.

Table 3.4: Tests for Cointegration between Household and Expert Expectations

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	5% Critical Value	p-value [†]
Germany				
None	0.20	16.31	12.32	0.01
At most 1	0.06	3.48	4.13	0.07
France				
None	0.22	15.92	12.32	0.01
At most 1	0.03	1.96	4.13	0.19
Italy				
None	0.36	21.69	12.32	0.00
At most 1	0.09	3.65	4.13	0.07
United Kingdom				
None	0.25	18.90	12.32	0.00
At most 1	0.05	2.60	4.13	0.13

Notes: [†] MacKinnon–Haug–Michelis (1999) p-values. Samples run from 1989Q4 to 2004Q2 (Italy: 1992Q4–2002Q4).

expert and household expectations as shown in Table 3.4. The findings show that, for all four countries, the two series are cointegrated (at the 5% significance level). In addition, we checked whether the theoretical restriction on $\alpha = (1, -1)^\top$ is supported in data. The values for α_1 are close to -1 (the value predicted by the model) and range from -1.21 for the UK to -1.00 for Germany. As illustrated by the likelihood ratio statistics presented in Table 3.5, we find that α is not significantly different from $(1, -1)^\top$ (except in the UK).

The VEC findings are summarized in Table 3.5.¹⁹ All estimates of λ_1^{HH} are significant (although for the restricted case in France and Italy only at the 10% level) and lie except for the restricted case in Italy in the neighborhood of 0.25, typically somewhat higher than implied by the “stationary” results above. We again find a lot of homogeneity among the four countries with French and Italian households updating presumably somewhat more slowly than British and German ones. The estimates in Table 3.5 imply updating frequencies between three and seven quarters.

Carroll’s sticky information model is also supported by how the deviations from the long-term equilibrium are corrected. The error-correction process is

¹⁹The models were estimated based on the sample from 1989Q4 to 2004Q2, except for Italy, where a valid cointegrating relationship was found between 1992Q4 and 2002Q4 only.

Table 3.5: Sticky Expectations in the VEC Form

Model		Germany	France	Italy	UK
Unrestricted	$\hat{\lambda}_{HH}$	-0.30***	-0.26***	-0.18**	-0.20***
	std. error	(0.09)	(0.09)	(0.07)	(0.05)
Restricted	$\hat{\lambda}_{HH}$	-0.30***	-0.14*	-0.23*	-0.27***
	std. error	(0.09)	(0.08)	(0.12)	(0.08)
Test for restriction (1, -1) on α	LR stat.	0.00	2.29*	2.97*	3.86**
	p value	0.988	0.070	0.085	0.049

Notes: Samples run from 1989Q4 to 2004Q2 (Italy: 1992Q4–2002Q4). “Unrestricted” refers to the unrestricted VEC model. “Restricted” refers to the VEC estimation results under the restriction $\alpha = (1, -1)^\top$. *, **, *** denotes rejection of the null at the 10%, 5% and 1% significance level, respectively.

primarily driven by the adjustment in household rather than expert expectations. This is implied by the estimates of λ_1^{EX} , which are not significantly different from zero or have the wrong (positive) sign.

Our findings thus imply that the epidemiology model of Carroll (2003) can be easily extended to the “non-stationary world.” The derived VEC epidemiology model of information diffusion performs similarly well as compared to the stationary model. This result is especially useful for the analysis of European countries, which plausibly have highly persistent inflation rates (see O’Reilly and Whelan, 2005 and references in ECB, 2005). Thus, even though it seems to be difficult to draw clear conclusions about the stationarity properties of the series with the small sample size at hand, the VEC representation might be preferable once more data are available.

3.1.4 Discussion of Results

Inflation expectations are crucial determinants of future inflation dynamics. The model estimated here attempts to analyze how these expectations are formed and how information is transmitted from professional forecasters to households. Our estimates of the speed of information updating have important implications for the persistence of inflation and inflation expectations. We document that the qualitative and quantitative findings previously reported for the US generalize to major European countries. Most European households adjust rather sluggishly to new information; they update their information on average once in twelve to eighteen months. In addition, it turns out that households are forward-looking in that they use forecasts provided by experts rather than just past information to update their beliefs about

future inflation. These findings are robust to the two estimation methods (suited for data with various stochastic properties) we consider.

The analysis of this section can be extended through a number of avenues. Survey data can be used to directly estimate the sticky information Phillips curve in addition to its epidemiological micro-foundations (as in the following section or Döpke et al. (2008b)). Alternatively, it would be possible, in the spirit of Mankiw et al. (2003), to analyze the evolution of the cross-sectional distribution of inflation expectations in Europe rather than just their mean values (as in Section 4.2 or Dovern et al. (2009)). Finally, the epidemiology model could, in principle, be estimated for additional countries, using cross-sectional dependence among countries to alleviate problems related to short samples.

3.2 Estimating the Sticky Information Phillips Curve²⁰

Several recent papers, including Mankiw and Reis (2002, 2003, 2006), argue that models in which agents update their information occasionally rather than instantaneously resolve some stylized business cycle puzzles.²¹ These puzzles include the fact that inflation is considerably persistent and empirically disinflations are found to be costly.²² Theoretical foundations for the new sticky information paradigm were elaborated in the work by Carroll (2003) on the “epidemiological model of expectations” that has been discussed in the previous section. Furthermore, Reis (2006a,b) and Mankiw and Reis (2006) discuss the microfoundations of the sticky information approach and propose to replace the widely used backward-looking and New Keynesian Phillips curves with the Sticky Information Phillips curve (SIPC).

Interestingly, there has been relatively little research on estimation the key parameters of the SIPC empirically. Carroll (2003) and Döpke et al. (2008a) estimate the epidemiological model of transmission of information between households and forecasters using US and European data, respectively. Among the few papers we are aware of that estimate the SIPC directly are Khan and Zhu (2002, 2006). However, due to data limitations Khan and Zhu have to use inflation and output forecasts generated using a VAR model as a proxy for the actual forecasts. Similarly, the estimation of the SIPC of Kiley (2007)

²⁰Large Parts of this section of the thesis are based on a paper that I published jointly with three co-authors (Döpke et al., 2008b).

²¹Compare also the overview in section 1.2 for this point.

²²The sticky-information models are related to models of rational inattention (Sims, 2003) and learning (Branch, 2004).

also proxies for inflation expectations. In contrast to these papers, we use survey-based inflation expectations directly.

3.2.1 Model

As discussed in Section 1.2, Mankiw and Reis (2002) derive the following version of the SIPC:

$$\pi_t = \frac{\lambda\alpha}{1-\lambda}\tilde{y}_t + \lambda \sum_{j=0}^{\infty} (1-\lambda)^j \mathbf{E}_{t-1-j}(\pi_t + \alpha\Delta\tilde{y}_t) + \varepsilon_t, \quad (3.5)$$

where π_t is the inflation rate and \tilde{y}_t the output gap. $\mathbf{E}_t(\cdot)$ denotes the rational (mathematical) expectation as of time t . Parameter α measures the sensitivity of the optimal relative price to the current output gap and depends on the structure of the economy (e.g. the preferences, technology, and the market structure parameters).²³ To increase the precision of the estimates of λ , on which we primarily focus, we impose that α lies between 0.10 and 0.20, a range considered plausible in the literature²⁴.

Note that in contrast to the standard (forward-looking) sticky-price model, in which current expectations of future state of the economy play an important role, what matters in the sticky-information model (3.5) are the past expectations of present events.

3.2.2 Data Issues

We use quarterly data between 1993Q2 and 2004Q4 for Germany, France, Italy and the United Kingdom. The actual GDP and inflation series were obtained from OECD's Main Economic Indicators database.

The experts' inflation and output forecasts we use were collected by *Consensus Economics*, a major London-based macroeconomic survey firm. Each quarter since 1989 Consensus Economics publishes the consensus forecasts constructed as the median of 20–30 individual predictions of major banks and research institutes (in each country). The consensus forecasts are available up to six quarters ahead, i.e. for quarters t through $t + 6$.

We use the GDP growth forecasts to extract estimates of future output gap, \tilde{y}_t , as follows. First, we take into account that the expectations reported

²³The parameter α can be interpreted as a measure of the degree of real rigidity, see e.g. Ball and Romer (1990).

²⁴We also estimated both parameters jointly. While the estimates of λ remain about the same as in Tables 3.6 and 3.7, α is estimated imprecisely. Therefore we impose α as suggested by, e.g., Mankiw and Reis (2002) and Khan and Zhu (2006).

in the survey refer to year-on-year changes rather than annualized quarterly changes as implied by the SIPC model. Second, we proxy the expected output gap based on expected GDP growth as follows. Denote y_t and y_t^* the log of output and the log of potential output, respectively. We first recursively construct a prolonged GDP series, $y_S^*(t)$ ($S = t_0, \dots, t, \dots, t+6$), for each sample point t by setting $y_S^*(t) = y_t$ for $S \leq t$ and $y_{t+1}^*(t) = y_{t-3}^*(t) + \mathbf{E}_t \Delta y_{t-3,t+1}$, $y_{t+2}^*(t) = y_{t-2}^*(t) + \mathbf{E}_t \Delta y_{t-2,t+2}$, \dots , $y_{t+6}^*(t) = y_{t+2}^*(t) + \mathbf{E}_t \Delta y_{t+2,t+6}$, where $\mathbf{E}_t \Delta y_{i,j}$ denotes the expectation for GDP growth between time i and j formed at time t . We then apply the Christiano and Fitzgerald (2003) full sample asymmetric band-pass filter to filter out the cyclical component of $y_S^*(t)$, say $\tilde{y}_S^*(t)$, as a proxy for the output gap. We interpret the last six observations of this series as the expectation of the output gap for periods $t+1$ through $t+6$ as of time t .

We use the GDP growth forecasts to extract expectations as of time s for the future output gap, $\mathbf{E}_s \tilde{y}_t$, as follows. First, we have to bear in mind that the expectations reported in the survey refer to year-on-year changes rather than annualized quarterly changes as implied by the SIPC model. Second, we proxy the expected output gap based on the expectations for GDP growth as follows. Denote y_t and y_t^* the log of output and the log of potential output, respectively. For each time period, s , in our sample, we construct a prolonged GDP time series, say $\hat{y}_s(t)$, by setting $\hat{y}_s(t) = y_t$ for $t \leq s$ and recursively computing $\hat{y}_s(t+1) = y_{t-3} + \mathbf{E}_s \Delta y_{t-3,t+1}$, $\hat{y}_s(t+2) = y_{t-2} + \mathbf{E}_s \Delta y_{t-2,t+2}$, \dots , $\hat{y}_s(t+5) = \hat{y}_s(t+1) + \mathbf{E}_s \Delta y_{t+1,t+5}$, and $\hat{y}_s(t+6) = \hat{y}_s(t+2) + \mathbf{E}_s \Delta y_{t+2,t+6}$, where $\mathbf{E}_s \Delta y_{i,j}$ denotes the expectation for GDP growth between time i and j formed at time s . We then apply the Christiano and Fitzgerald (2003) full sample asymmetric band-pass filter on this prolonged time series to filter out $y_s^*(t)$ and the cyclical component of $\hat{y}_s(t)$, $\tilde{y}_s(t) = \hat{y}_s(t) - y_s^*(t)$. We see the cyclical component as a proxy for the output gap. More specifically, we interpret the last six observations of this series as the expectation for the output gap for periods $s+1$ through $s+6$ as of time s .

For the expert expectations of the inflation rate we also face the first problem mentioned above that the expectations reported in the survey refer to year-on-year changes rather than annualized quarterly changes. Analogously to the previous paragraph, we compute annualized expected quarterly inflation rates by prolonging the actual consumer price index time series based on the expected year-to-year inflation rates and transforming this prolonged series into expected quarterly inflation rates.

3.2.3 Results

We assume that the updating firms each period simply adopt professional forecasts to form rational expectations of inflation and output gap up to six quarters ahead. Consequently, the infinite sum in (3.5) is truncated alternatively at four and six lags.²⁵ We estimate (3.5) first individually for each country using non-linear least squares (in Table 3.6) and then jointly using seemingly unrelated regressions (SUR) (in Table 3.7).

Equation-by-Equation Estimation

Table 3.6 summarizes the results of estimating (3.5) with truncation lags $n = 4$ and 6 for values of α between 0.1 and 0.2 for Germany, France, Italy, and the United Kingdom. As the theoretical model (3.5) does not have a constant we exclude it in the empirical estimation and report the uncentered R^2 .²⁶

We find the following five key results. First, all estimates of λ are highly significant for all parameterizations of the model. Second, for France, Germany, and the UK their values lie around 0.20 to 0.30. This is about the size one would expect and in line with findings in Khan and Zhu (2002) or Döpke et al. (2008a). Third, there is a lot of homogeneity across the latter three countries. Given the same parametrization, $\hat{\lambda}$ s do not differ by more than 0.02. Fourth, the results for Italy deviate quite substantially from the outcomes for the other countries. λ is estimated around 0.5 to 0.6 which implies about twice as high frequency of information updating compared to the findings for other countries. Furthermore, the estimates for Italy are more sensitive with respect to the values chosen for α . This is not the case for the other countries. Finally, the models including up to 6 lags of the sequence of expectation terms generally show a better fit to the data and smaller $\hat{\lambda}$ (this latter result is also evident from the results in Khan and Zhu, 2006).

Our estimates of λ are typically a bit smaller than the estimates for the US presented in Carroll (2003). This indicates that the information transmission process is somewhat slower in the three European countries considered here in this study. This is in line with the evidence of Döpke et al. (2008a), who estimate the Carroll (2003) model for European countries, and find the information updating process of households to be also somewhat slower than for the US economy.

²⁵The results with 5 lags do not differ considerably.

²⁶If the constant is included it is insignificant.

Table 3.6: SIPC: Equation-by-Equation Regression Results

Truncation at lag:		4		6	
		λ	uncent. R^2	λ	uncent. R^2
France	$\alpha = .10$	0.271	0.56	0.188	0.58
		5.4		5.4	
	$\alpha = .15$	0.268	0.56	0.189	0.59
		5.5		5.3	
	$\alpha = .20$	0.271	0.56	0.191	0.59
		5.4		5.3	
Germany	$\alpha = .10$	0.257	0.64	0.181	0.63
		5.7		5.4	
	$\alpha = .15$	0.258	0.64	0.181	0.63
		5.7		5.4	
	$\alpha = .20$	0.257	0.64	0.181	0.63
		5.7		5.4	
Italy	$\alpha = .10$	0.608	0.90	0.456	0.89
		6.1		3.4	
	$\alpha = .15$	0.577	0.89	0.492	0.90
		4.6		3.7	
	$\alpha = .20$	0.608	0.90	0.540	0.90
		6.1		5.0	
UK	$\alpha = .10$	0.271	0.64	0.201	0.72
		6.2		6.2	
	$\alpha = .15$	0.270	0.64	0.202	0.72
		6.2		6.1	
	$\alpha = .20$	0.271	0.64	0.202	0.72
		6.2		6.1	

Notes: Numbers below the estimates are t-statistics.

Seemingly Unrelated Regression (SUR) Estimation

As the residuals of the individual equations are substantially cross-correlated, we investigate in Table 3.7 how using the SUR affects our baseline results approach to improve the efficiency of the estimation.

Again, all coefficients were found to be highly significant and to be (with the exception of Italy) in the range of 0.14 to 0.18 for truncation at lag 6 and in the range of 0.20 to 0.29 for truncation at lag 4. In addition, likelihood-ratio tests confirm that we cannot reject the null hypothesis that the λ are equal for France, Germany, and the UK. We only present the test statistic for one particular value of α as for other specifications the outcomes are very similar. For $\alpha = 0.15$ and truncation at lag 4, the LR-statistic is 1.19 (p-val: 0.55). For $\alpha = 0.15$ and truncation at lag 6, the LR-statistic is 0.84

PhD-Thesis written by: Jonas Dovern

Table 3.7: SIPC: Estimation by Seemingly Unrelated Regressions

Truncation at lag:		4		6	
	α	λ	uncent. R^2	λ	uncent. R^2
France	$\alpha = .10$	0.213	0.58	0.146	0.57
		5.4		5.1	
	$\alpha = .15$	0.216	0.58	0.146	0.57
		5.5		5.1	
$\alpha = .20$	0.219	0.58	0.144	0.57	
	5.5		5.1		
Germany	$\alpha = .10$	0.296	0.67	0.158	0.63
		5.8		5.6	
	$\alpha = .15$	0.294	0.66	0.160	0.63
		5.8		5.6	
$\alpha = .20$	0.292	0.66	0.160	0.63	
	5.9		5.6		
Italy	$\alpha = .10$	0.451	0.77	0.525	0.69
		8.2		5.6	
	$\alpha = .15$	0.471	0.77	0.568	0.70
		8.0		6.5	
$\alpha = .20$	0.494	0.78	0.571	0.71	
	7.9		7.2		
UK	$\alpha = .10$	0.190	0.57	0.177	0.72
		5.1		5.6	
	$\alpha = .15$	0.193	0.57	0.177	0.72
		5.1		5.6	
$\alpha = .20$	0.195	0.58	0.176	0.72	
	5.2		5.6		

Notes: Numbers below the estimates are t-statistics.

(p-val: 0.66). Obviously, the hypothesis that λ for Italy is also equal to the parameters in the other three countries is rejected at any sensible significance levels. A possible explanation for this finding of a somewhat higher λ in Italy is a higher level of and uncertainty about inflation in the estimation sample in Italy compared to the other three countries.

Estimating λ jointly for France, Germany, and the UK with λ being constrained to be equal across countries (assuming that the price setting mechanism is the same in all three countries) yields no big surprises. For all parameterizations $\hat{\lambda}$ is highly significant and lies in between the individual country estimates. For truncation at lag 4 we find $\hat{\lambda} = 0.3$ and for truncation at lag 6 we find $\hat{\lambda} = 0.16$. The estimates again seem to be very robust to the particular value chosen for α .

PhD-Thesis written by: Jonas Dovern

3.2.4 Discussion of Results

In this paper, we make the first attempt to estimate the main parameter of the SIPC developed in Mankiw and Reis (2002) for four large European countries using survey-based expectations. We find that λ ranges between 0.15 and 0.3 for Germany, France and the United Kingdom and between 0.5 and 0.6 for Italy.

To recapitulate the results of this section: Using recent data from four major European economies we find that producers in France, Germany and the United Kingdom update their information sets about once a year, those in Italy about once every six months. These findings are quite robust to the two estimation methods we use (equation-by-equation estimation and seemingly unrelated regressions) and to the number of lags of right-hand side variables included. The estimates of λ are consistent with those of Döpke et al. (2008a) except for Italy whose λ they pin down to be comparable to the other countries.²⁷

3.3 Concluding Remarks

In this part of the thesis we have used survey data on (inflation) expectations of households and firms to assess the empirical validity of two different models that are based on the sticky information concept. First, we have demonstrated that the expectation formation process of households in four major European economies can be characterized by the “epidemiological model of expectations” proposed by Carroll (2003). Second, we have shown that the aggregate inflation dynamics in those four countries can be explained – at least in a partial equilibrium framework – by means of the SIPC proposed by Mankiw and Reis (2002). All in all, the results support the sticky information concept for modeling inflation dynamics. We should note, however, that recent research has shown that the sticky information framework does not outperform the conventional sticky price framework when analyzed in a general equilibrium model (see e.g. Trabandt, 2007; Korenok, 2008).

²⁷Possible extensions of this work include investigating how the frequency of updating varies across other countries and time periods or more generally what other factors determine the size of λ .

Chapter 4

Dispersion of Forecasts

Macroeconomic models often impose homogeneity. Agents have the same preferences, beliefs, information sets, are hit by the same shocks or process information in the same way. Such assumptions are convenient because they make models simple and tractable while keeping them useful for aggregate policy analysis. However, evidence from micro data and casual observations show that people differ from each other, and economists have recently put much effort into constructing and studying models that can account for some of the differences.¹

The extent of disagreement about the future paths of macroeconomic variables is for instance remarkably high – even among professional forecasters (Zarnowitz and Lambros, 1987; Gallo et al., 2002). The degree of disagreement is sometimes seen as a reflection of the uncertainty of the macroeconomic environment that economic agents face. Viewed in this way, it of high relevance to understand the underlying drivers of dispersion, i.e uncertainty, since several theories suggest that increased uncertainty leads to costs in terms of welfare. Friedman (1977) and Ball (1992) demonstrate for instance how increasing inflation uncertainty leads to higher losses in aggregate output. Consequently, it is desirable for economic agents as well as for

¹For example, models in which some households are more impatient than others (or are subject to liquidity constraints) are useful in studying the monetary policy transmission mechanism (Iacoviello, 2005). Models with heterogenous beliefs/expectations are becoming popular in asset pricing literature (see Scheinkman and Xiong, 2003 for a survey). Carroll (1997) and Krusell and Smith (1998) model reaction of agents' consumption–saving behavior to idiosyncratic (and aggregate) income shocks. Morris and Shin (2005b) investigate the value of providing of public information to agents depending on the amount of private information they have.

researchers to have at hand a good proxy for the uncertainty attached to a given forecast of a variable.²

Giordani and Söderlind (2003) claim that using cross-sectional dispersion measures from survey data on forecasts constitutes a valuable approach for estimating uncertainty that is superior or at least complementary to time series approaches like e.g. conditional volatilities derived from GARCH models. This approach accentuates the heterogeneity of expectations and is rooted in the idea that agents have different ideas and models of the functioning of the economy (Sims, 2003; Branch, 2004; Mankiw and Reis, 2002, 2003). The assumption is that the differences in predictions derived from these models are higher the higher the true uncertainty in the economy is.³ Following Giordani and Söderlind (2003), measures on the dispersion of predictions are therefore frequently used to proxy the degree of uncertainty surrounding point forecasts for macroeconomic variables.

In what follows we analyze the dispersion of forecasts and how it can be measured using fixed event forecasts. However, we only punctually relate our findings to the notion of uncertainty. Still, from the perspective of e.g. a central bank that has a mandate to target a certain inflation rate the issue of diverging inflation expectation is of interest in itself.

4.1 Measuring Dispersion from Fixed Event Forecasts⁴

Practitioners and applied researchers are usually interested in a measure of uncertainty that is unaffected by changing institutional factors and – more

²See also the recent paper by Bloom et al. (2008) who claim that uncertainty can be a major cause for recessionary developments.

³It should be mentioned that this view is disputed frequently (Bomberger, 1996; Rich and Tracy, 2003; Döpke and Fritsche, 2006), as a bunch of other theories exist about the emergence of disagreement, i.e. forecast dispersion. First, models of information transmission stress the role of time lags in the transmission of “news” to different agents in the macroeconomy (Carroll, 2003; Mankiw et al., 2003). Second, disagreement can be explained to some extent by the usage of differing models and the existence of ideological beliefs (Fuller and Geide-Stevenson, 2003); Unfortunately, there is – with the exception of Batchelor and Dua (1990) – very little empirical evidence on this issue. Third, it is sometimes argued, that different forecasters face different incentives to cheat, to seek rents or to influence the public debate. Hence, forecast accuracy might not be the only aim of the forecasters (Laster et al., 1999; Ehrbeck and Waldmann, 1996). Finally, another source of forecasters’ disagreement might be due to forecasting being part of the policy advising process (Stege, 1989; Kirchgässner, 1999).

⁴Large parts of this section of the thesis are based on a paper that I wrote jointly with a co-author (Dovern and Fritsche, 2008).

importantly – time varying forecast horizons. This is why exclusively survey data on fixed horizon forecasts have been used when forecast dispersion has been used as a measure of uncertainty (or for other purposes).

Unfortunately, a good number of data sets provide only fixed event forecasts rather than fixed horizon forecasts. And given a scarce data situation for a particular country or variable, one would like to use these fixed event forecasts to measure uncertainty. The problem is that every measure of fundamental dispersion is distorted by the fact that the forecast horizon is time varying, when it has been derived from fixed event forecasts. It is well documented, that forecast dispersion in fixed event panels has a remarkable proportion which is “non-fundamental” in a sense that this part of cross-section dispersion is driven by the time varying forecast horizons rather than by macroeconomic uncertainty (Gallo et al., 2002; Patton and Timmermann, 2007). Theoretical justifications can be found in early works by Lucas (1973) or Townsend (1983).

For empirical researchers and policy makers it would be desirable to be able to extract the “fundamental” component of dispersion from such data by controlling for “non-fundamental” factors such as the influence of the forecast horizon. In this section of the thesis, we suggest some empirical approaches for this task and assesses their relative performance.

To illustrate the issue treated in this section, Table 4.1 shows the correlation coefficients between the cross-sectional standard deviation for four US variables derived from fixed event forecasts and those derived from fixed horizon forecasts for the US economy.⁵ Two observations are worth pointing out. First, the correlation between the dispersions derived from fixed horizon forecasts and those based on predictions for the current year’s annual figures are in all cases smaller than those with the dispersion based on predictions for the subsequent year’s annual figures. This is not surprising as one would expect that the dispersion across panelists is mostly affected by the shrinking forecast horizon when the latter is close to zero, because this implies that some of the relevant data for the forecast has already been in the information set of the forecasters. Patton and Timmermann (2007) indeed show that much of the reduction of cross-sectional dispersion is observed when the forecast horizon becomes smaller than one year. Second, the correlations are especially low for forecasts for the growth rate of real GDP.

⁵We use the well-known U.S. data set of quarterly macroeconomic forecasts from the *Survey of Professional Forecasters* (see below).

Table 4.1: Correlation of dispersion measures calculated from different types of forecasts

	CY	NY	12M	CY	NY	12M
	GDP Growth			Unemployment Rate		
Current year	1.00			1.00		
Next year	0.65	1.00		0.75	1.00	
12-Month	0.58	0.61	1.00	0.63	0.85	1.00
	CPI Inflation			T-Bill Rate		
Current year	1.00			1.00		
Next year	0.82	1.00		0.81	1.00	
12-Month	0.82	0.94	1.00	0.74	0.85	1.00

Notes: Displayed are linear correlation coefficients. CY refers to the forecasts made for the current calendar year, NY refers to the forecasts made for the next calendar year, and 12M refers to the 12 months ahead forecasts.

4.1.1 Approaches

In this part, we present different modeling frameworks that are potentially adequate to estimate the “fundamental” component of cross-sectional dispersion derived from fixed event forecasts. Throughout the remainder of this section we adopt the following notation: Let $\mathbf{E}_{t,i}^0(\tilde{y})$ denote the forecast for a variable for the current calendar year made by forecaster i at time t . Analogously, $\mathbf{E}_{t,i}^1(\tilde{y})$ denotes her forecast for next year’s annual figure. In case of growth rates being forecasted, the forecast for the quarter-to-quarter growth rate at time s made by the same forecaster at time t is given by $\mathbf{E}_{t,i}^s(y)$. We compute the twelve-months-ahead growth forecast of each panelist as $\mathbf{E}_{t,i}^{12}(\hat{y}) = [\prod_{k=0}^3 (\mathbf{E}_{t,i}^{t+k}(y)/100 + 1) - 1] * 100$.

The different cross-sectional dispersions at each sample point are calculated as the standard deviation across all N forecasts or their interquartile range. We denote them as $D_t^{\mathbf{E}^0(\tilde{y})}$, $D_t^{\mathbf{E}^1(\tilde{y})}$, and $D_t^{\mathbf{E}^{12}(\hat{y})}$ respectively. Since $\mathbf{E}_{t,i}^{12}(\hat{y})$ is unaffected by seasonal influences and the forecast horizon is fixed over time, we do not expect $D_t^{\mathbf{E}^{12}(\hat{y})}$ to show any seasonal patterns. Rather, it should only reflect disagreement due to the prevailing macroeconomic uncertainty.

From the six candidate approaches which we will consider in this paper, one differs fundamentally from the remaining five approaches. This approach is non-parametric, intuitive, and simple; it involves the approximation of the twelve-months-ahead forecasts in a first step, and the calculation of a dispersion measure across those approximative fixed horizon forecasts in a second step. In contrast, all other methods take the dispersion measured over fixed event forecasts as input and use different parametric time-series

approaches to decompose this dispersion into different components one of which represents the fundamental degree of dispersion we are interested in.

Estimation via Approximation of Fixed Horizon Forecasts

If someone is interested in the dispersion across (unobserved) fixed horizon forecasts, a natural way of calculation is based on an approximation of those unknown forecasts. To this end, we construct simple proxies for the twelve-months-ahead forecasts by taking a weighted moving average of fixed event forecasts (Heppke-Falk and Hüfner, 2004; Smant, 2002), namely the forecasts for two subsequent calendar years.

$$\hat{\mathbf{E}}_{t,i}^{12}(\hat{y}) = \frac{4-q+1}{4}\mathbf{E}_{t,i}^0(\tilde{y}) + \frac{q-1}{4}\mathbf{E}_{t,i}^1(\tilde{y}), \quad (4.1)$$

where q is equal to one in each first quarter of the year, equal to two in each second quarter of the year, and so on. As an example, consider the situation in the second quarter of 2007. We would compute a proxy for the twelve-months-ahead forecast with target date 2008Q1 by taking 3/4 of $\mathbf{E}_{07Q2,i}^{07}(\tilde{y})$ and adding 1/4 of $\mathbf{E}_{07Q2,i}^{08}(\tilde{y})$.

In a second step, we compute a measure of dispersion, let's say $D_t^{\hat{\mathbf{E}}^{12}}(\hat{y})$, like the standard deviation or the interquartile range, across all individual forecasters at each point in time.

Estimation via Time-Series Decompositions

The other methods take a different route and start from the dispersion calculated across fixed event forecasts. Formally, we assume that we can write $D_t^{\mathbf{E}^k(\hat{y})}$, $k \in \{0, 1\}$, as the sum of two components and a residual term

$$D_t^{\mathbf{E}^k(\hat{y})} = D_t^{\mathbf{E}^k(\hat{y})f} + D_t^{\mathbf{E}^k(\hat{y})h} + \epsilon_t. \quad (4.2)$$

Here $D_t^{\mathbf{E}^k(\hat{y})h}$ denotes the component that is driven by the time varying forecast horizon and contains no valuable information about the fundamental disagreement among forecasters. On the other hand, $D_t^{\mathbf{E}^k(\hat{y})f}$ is the fundamental component. It is driven by the same underlying factors as $D_t^{\mathbf{E}^{12}(\hat{y})}$ and should follow a sample path with similar dynamic properties. It is this component that we want to use as a proxy for the (in case of survey data on fixed event forecasts) unobserved process $D_t^{\mathbf{E}^{12}(\hat{y})}$.

In the remainder of this section, we present different time series models that serve to extract the fundamental component, $D_t^{\mathbf{E}^k(\hat{y})f}$, from the observed

time series, $D_t^{\mathbf{E}^k(\tilde{y})}$. The basic idea behind all five approaches is to determine the seasonal (forecast horizon dependent) component. The methods differ most crucially in the way residual terms are treated, i.e. whether they are assumed to be part of the fundamental component or not.

Seasonal Adjustment by X12-ARIMA

One natural approach to filter out the component, $D_t^{\mathbf{E}^k(\tilde{y})h}$, which moves over the year in a repetitive way due to the varying forecast horizon, is the application of a standard seasonal adjustment method. We have chosen the widely used X12-ARIMA procedure (US-Census-Bureau, 2007) for this purpose.

Constant Forecast-Horizon-Effects

Another very simple approach is to assume that the reduction of dispersion caused by a shrinkage of the forecast horizon is constant over time, i.e. there is one $D_t^{\mathbf{E}^k(\tilde{y})h}$ for all first quarters of the years, one for all second quarters of the years, and so on. We can estimate those fixed forecast-horizon-effects by regressing $D_t^{\mathbf{E}^k(\tilde{y})}$ on a set of quarter-dummies each of them being equal to one only in one specific quarter of the year. The regression equation takes on the form

$$D_t^{\mathbf{E}^k(\tilde{y})} = \sum_{i=1}^4 \beta_i Dum_i + v_t . \quad (4.3)$$

One can argue that the residuals, \hat{v}_t , of this kind of regression should be a good approximation to $D_t^{\mathbf{E}^k(\tilde{y})f}$. Note that the first two time series approaches simply filter out a deterministic seasonal component; all residual shocks are attributed to the fundamental component. This will be different for the following two approaches.

Univariate Unobserved Components Model

Yet another approach is to specify an unobserved components model (Harvey, 1989; Durbin and Koopman, 2001) for $D_t^{\mathbf{E}^k(\tilde{y})}$. This requires some assumptions about the processes behind $D_t^{\mathbf{E}^k(\tilde{y})f}$ and $D_t^{\mathbf{E}^k(\tilde{y})h}$. Since it is not unreasonable to assume that $D_t^{\mathbf{E}^k(\tilde{y})f}$ exhibits some degree of persistence, we assume here that it follows a random walk process:

$$D_t^{\mathbf{E}^k(\tilde{y})f} = D_{t-1}^{\mathbf{E}^k(\tilde{y})f} + v_t , \quad (4.4)$$

where we assume that $v_t \sim NID(0, \sigma_v^2)$ is independently distributed from ϵ_t above.

PhD-Thesis written by: Jonas Dovern

For $D_t^{\mathbf{E}^k(\tilde{y})h}$, we assume that it follows a stochastic seasonal pattern. More specifically, we specify it in such a way that it has a trigonometric form:

$$D_t^{\mathbf{E}^k(\tilde{y})h} = \sum_{i=1}^{s/2} \gamma_{j,t}, \quad (4.5)$$

where s is the number of seasonal frequencies in a given year (e.g. 4 for quarterly data) and each of the $\gamma_{j,t}$ follows:

$$\begin{bmatrix} \gamma_{j,t} \\ \gamma_{j,t}^* \end{bmatrix} = \begin{bmatrix} \cos \lambda_j & \sin \lambda_j \\ -\sin \lambda_j & \cos \lambda_j \end{bmatrix} \begin{bmatrix} \gamma_{j,t-1} \\ \gamma_{j,t-1}^* \end{bmatrix} + \begin{bmatrix} \omega_{j,t} \\ \omega_{j,t}^* \end{bmatrix}.$$

Here $\lambda = 2\pi j/s$ is the frequency and the disturbances $\omega_{j,t}$ and $\omega_{j,t}^*$ are mutually uncorrelated and $NID(0, \sigma_\omega^2)$. The filtered state estimates (conditional on past information only) of $D_t^{\mathbf{E}^k(\tilde{y})f}$ constitute a proxy for the fundamental dispersion.

Bivariate Unobserved Components Model

Whereas we used data on $D_t^{\mathbf{E}^0(\tilde{y})}$ and $D_t^{\mathbf{E}^1(\tilde{y})}$ only separably in the approaches so far, it might be worth specifying a bivariate model to use a richer information set to extract one fundamental component from data on both of the dispersion time series. Such an approach is proposed in this paragraph. More specifically, we assume that $D_t^{\mathbf{E}^0(\tilde{y})f}$ and $D_t^{\mathbf{E}^1(\tilde{y})f}$ are equal at each point in time. Hence, we require the fundamental component of disagreement among forecasters to be identical for both the disagreement on the current calendar year's annual growth rate and the disagreement on next year's growth rate. Given that these two kinds of forecasts are made by forecasters at the same point in time and facing the same information set about the stance of the economy this is a natural assumption. We denote this common fundamental component by $D_t^{\mathbf{E}(\tilde{y})f}$. The appropriate specification of the data generating process for this fundamental component is data driven and has to be specified for each set of forecasts analyzed. Some restrictions have to be made, however, to limit the number of possible models. We assume that it follows a stationary autoregressive process with a maximum lag order of q .

We capture the changes in dispersion induced by changing forecast horizons by including dummies for the number of quarters of each forecast horizon. Formally, the model is given by

$$\begin{bmatrix} D_t^{\mathbf{E}^0(\tilde{y})} \\ D_t^{\mathbf{E}^1(\tilde{y})} \end{bmatrix} = D_t^{\mathbf{E}(\tilde{y})f} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \sum_{i=0}^4 \left(\begin{bmatrix} \beta_i \\ \beta_{i+4} \end{bmatrix} Dum_i \right) + \begin{bmatrix} \epsilon_t^1 \\ \epsilon_t^2 \end{bmatrix}, \quad (4.6)$$

where $D_t^{\mathbf{E}(\tilde{y})f}$ evolves according to

$$D_t^{\mathbf{E}(\tilde{y})f} = L(\beta)D_t^{\mathbf{E}(\tilde{y})f} + \zeta_t .$$

$L(\beta)$ is a lag polynomial of order q and the three error terms ϵ_t^1 , ϵ_t^2 , and ζ_t are assumed to be uncorrelated and independently identically normally distributed with different but fixed variances. Again, the filtered state estimates for $D_t^{\mathbf{E}(\tilde{y})f}$ will serve as a proxy for the fundamental degree of dispersion. Note that in both the univariate and the bivariate unobserved components approach we do not add the residual terms to the fundamental component; this is a conceptual difference to the first two time series approaches above.

4.1.2 Data Issues

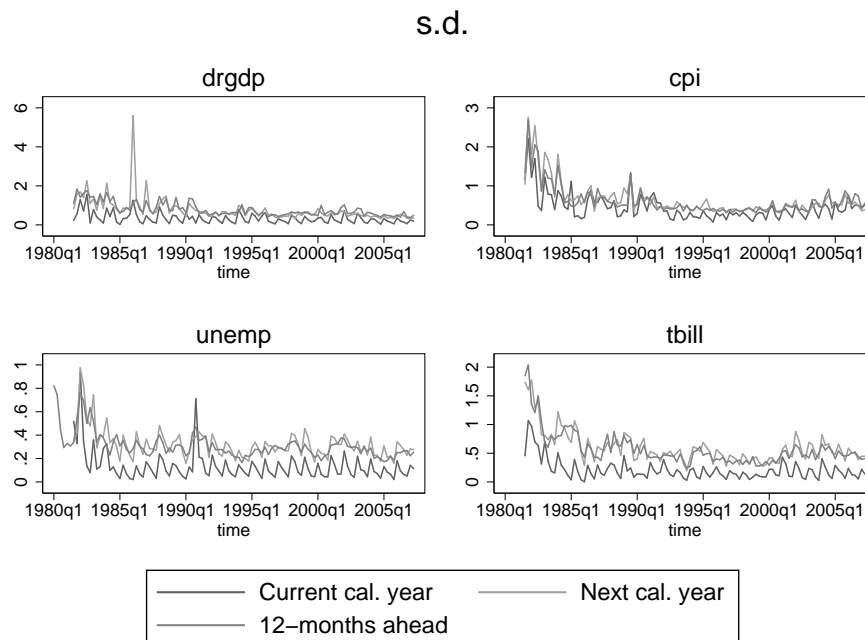
As the *Consensus* data set provides only fixed event forecasts, we cannot use it to evaluate the performance of the various proposed methods. Instead, we use data from the Survey of Professional Forecasters (SPF)⁶ for this purpose, since it has the big advantage for the purpose of this section that it simultaneously provides fixed event and fixed horizon forecasts. The forecasters are asked to report not only their predictions for the quarterly development over the next five quarters (from which e.g. four-quarter-ahead forecasts can be deduced) but also their predictions for the annual figures of the current calendar year and those for the next calendar year.

The data set reports forecasts on macroeconomic variables from professional forecasters collected through surveys among the panelists. The SPF is the oldest survey in the US that reports forecasts on macroeconomic variables at a quarterly frequency. Its beginning dates back to 1968, although the set of variables has been continuously extended in later years, such that the samples do not reach back to 1968 for all variables.⁷ Forecasters are anonymous which should minimize the problem of distorted forecasts due to incentive issues (Kahneman and Tversky, 1979; Ehrbeck and Waldmann, 1996; Batchelor, 2007).

We concentrate in this section on the most prominent macroeconomic variables of the data set, namely the growth rate of the real gross domestic product (*drgdp*), the inflation rate (*cpi*), the unemployment rate (*unemp*), and the treasury-bill rate (*tbill*). We use a sample that spans the period from 1981Q1 to 2007Q2.

⁶The data can be downloaded at <http://www.phil.frb.org/econ/spf/>. See Croushore and Stark (2001) and <http://www.philadelphiafed.org/econ/forecast/> for details on the survey.

⁷The cross-section dimension, i.e. the number of forecasters who take part in the survey, is currently around 30. For more details on the survey see e.g. Croushore (1993).

Figure 4.1: Cross Sectional Standard Deviation of Forecasts in the SPF

Notes: The figures show the dispersion of forecasts for the current and next calendar year respectively together with the dispersion of the 12-months ahead forecasts from the SPF.

To get an impression about the data and the problem one is facing when estimating cross-sectional dispersion from fixed event forecasts, we plot the cross-sectional standard deviations over time in Figure 4.1. The plots show the dispersion of the forecasts for the current and next calendar year respectively together with the dispersion of the 12-months ahead forecasts. It is evident that those dispersion measures based on the fixed event forecasts inherit seasonal patterns. These are naturally more pronounced for the results based on the forecasts for the current calendar year. Another observation is that the seasonal effect seems to be weakest for the dispersion of interest rate forecasts.

4.1.3 Empirical Results

In this part, we show the empirical results obtained using the cross-sectional standard deviation (s.d.) and the interquartile range (iqr) respectively. We use the linear correlation between the different potential proxies and the dispersion derived from the fixed horizon forecasts as performance criterion.

PhD-Thesis written by: Jonas Dovern

Table 4.2: Labeling of Different Methods

Label	Method
M1	Approximation of fixed horizon FCs
M2	Bivariate unobs. components model
M3	Univ. unobs. comp. model based on FCs for current cal. year
M4	Univ. unobs. comp. model based on FCs for next cal. year
M5	Extr. of sais. comp. with dummies based on FCs for current cal. year
M6	Extr. of sais. comp. with dummies based on FCs for next cal. year
M7	Saisonal adjustment of FCs for current cal. year by X12-ARIMA
M8	Saisonal adjustment of FCs for next cal. year by X12-ARIMA
M9	Unprocessed dispersion across FCs for current cal. year
M10	Unprocessed dispersion across FCs for next cal. year

Notes: FCs refers to forecasts.

To ease references to the different methods in tables and in the description of results, we introduce the labeling scheme presented in Table 4.2. Given its prominence against the other methods – not to mention its intuitiveness –, we will treat M1 as the reference method.

Table 4.3 contains a bunch of information that describes the empirical results. In what follows, we discuss the different aspects represented in the table. The most important information is given by the first number in each column. Those numbers are the linear correlation coefficient of the proxy obtained by the different methods and the dispersion measure derived directly from the fixed horizon forecasts given in the SPF. Ultimately, we would like to know whether the differences in performance according to the correlation coefficient of the different methods are statistically significant. To this end, we use a test based on Fisher’s z-transformation (Fisher, 1925) to infer whether we can reject the Null hypothesis that two correlation coefficients are statistically different from each other.⁸ Information on the test outcomes are given in the table by the numbers in parenthesis. They refer to the p-value of testing the Null hypothesis that the correlation of the corresponding method is equal to the correlation of M1. Values above 5% indicate that we cannot reject the hypothesis of equal correlation coefficients and, hence, equal performance of the two methods.

The test results indicate that the correlation coefficient in the overwhelming number of cases is statistically different. However, only in a minority of

⁸The test takes into account that we are dealing with dependent correlation coefficients in the sense that for three random variables x_1 , x_2 , and y we test whether $\text{corr}(x_1, y) - \text{corr}(x_2, y) = 0$, i.e. both correlations are computed against the same random variable.

cases sophisticated methods outperform the simple reference method M1. This result holds for the standard deviation as a common measure of dispersion; it does not hold when using the interquartile range to measure dispersion of forecasts. Considering that the interquartile range is a more appropriate measure of dispersion in those cases where the distribution of forecasts is not symmetric, this could indicate weaknesses of the sophisticated methods relative to M1 when the distribution is non-normal (for instance skewed).

The methods which indicate a higher correlation with the cross-section dispersion of fixed horizon forecasts compared to M1 are: method M2 in two out of eight cases, method M6 in four out of eight cases, and method M8 in two out of eight cases. In general this does not indicate a clear gain when using quite sophisticated methods. In fact, only M6 looks like a method which is a serious competitor to M1. In case of the interquartile range measure, there is no method which beats the moving-average transformation M1 in terms of a significantly higher correlation coefficient for three out of the four variables.

To test for the best-performing model more rigorously, we made use of the idea outlined by Granger and others for estimating the optimal weights in forecast combination exercises (Bates and Granger, 1969; Granger and Ramanathan, 1984). To that end, we used a panel regression (SUR) of the following form:

$$D_{it}^{\hat{y}^{12}} = \alpha_i + \sum_{j=1}^8 \beta_j D_{it}^{\tilde{y}^j f} + \varepsilon_{it}, \quad \sum_{j=1}^8 \beta_j = 1 \quad (4.7)$$

where $D_{it}^{\hat{y}^{12}}$ denotes the dispersion from the fixed horizon forecasts and $D_{it}^{\tilde{y}^j f}$, $j = 1, \dots, 8$ are the different dispersion approximations based on the competing models (except the unprocessed dispersion measures); the subscript i refers to the different variables analyzed in this section.

We estimated the regressions in levels and first differences of the series and for both dispersion measures. The results are summarized in Table 4.4. Once more it becomes clear, that M1 is by far the most promising method. Although, the results suggest that combining the proxy derived by M1 with other proxies (especially from M4 and M8) can improve the quality, we conclude that for practical work M1 constitutes a fairly good approach to proxy the fundamental dispersion from panels of fixed event forecasts.⁹

⁹While the dynamics are well captured by the well-performing proxies, the results show sizable deviations in the level of dispersion relative to the dispersion derived from fixed horizon forecasts. We leave the discussion on re-scaling the level of the proxies for future research. In most applications, practitioners and policy makers are likely to be interested primarily in the variation of dispersion measures.

Table 4.3: Correlation Results

	GDP-Growth				Inflation			
	s.d.		iqr		s.d.		iqr	
M1	0.82		0.76		0.92		0.60	
M2	0.80	(0.74)	0.74	(0.72)	0.94*	(0.05)	0.59	(0.88)
M3	0.78	(0.32)	0.64	(0.06)	0.84*	(0.00)	0.34*	(0.00)
M4	0.74	(0.08)	0.63*	(0.05)	0.86*	(0.01)	0.63	(0.78)
M5	0.67*	(0.00)	0.53*	(0.00)	0.84*	(0.00)	0.24*	(0.00)
M6	0.62*	(0.00)	0.83*	(0.05)	0.94*	(0.04)	0.68	(0.20)
M7	0.54*	(0.00)	0.54*	(0.00)	0.82*	(0.00)	0.28*	(0.00)
M8	0.70*	(0.00)	0.82	(0.10)	0.94	(0.08)	0.70	(0.10)
M9	0.58*	(0.00)	0.40*	(0.00)	0.82*	(0.00)	0.21*	(0.00)
M10	0.61*	(0.00)	0.83*	(0.05)	0.94	(0.07)	0.68	(0.25)
	Unempl. Rate				T-Bill Rate			
	s.d.		iqr		s.d.		iqr	
M1	0.83		0.65		0.89		0.88	
M2	0.90*	(0.00)	0.50*	(0.02)	0.85*	(0.04)	0.75*	(0.00)
M3	0.80	(0.38)	0.26*	(0.00)	0.75*	(0.00)	0.67*	(0.00)
M4	0.82	(0.91)	0.44*	(0.01)	0.90	(0.61)	0.76*	(0.00)
M5	0.24*	(0.00)	0.66*	(0.00)	0.37*	(0.00)	0.79*	(0.00)
M6	0.89*	(0.00)	0.67	(0.64)	0.93*	(0.00)	0.88	(0.75)
M7	0.56*	(0.00)	0.31*	(0.00)	0.80*	(0.00)	0.73*	(0.00)
M8	0.91*	(0.00)	0.66	(0.82)	0.94*	(0.00)	0.88	(0.92)
M9	0.62*	(0.00)	0.27*	(0.00)	0.69*	(0.00)	0.54*	(0.00)
M10	0.83	(0.94)	0.63	(0.68)	0.90	(0.60)	0.86	(0.45)

Notes: Numbers refer to the correlation to the dispersion of the actual twelve-months-ahead predictions. Number in parenthesis show the p-values corresponding to H_0 : Correlation of corresponding method is equal to correlation of M1. An * indicates rejection of H_0 at a 95% confidence level. We marked those cases with bold numbers for which an alternative method delivers a significantly higher correlation coefficient than M1 rather than a significantly lower one.

To recapitulate the results of this section: We can conclude that a moving-average transformation of the fixed event predictions on the level of individual forecasters (M1) performs extremely well for interquartile range and standard deviations measures. There are some other methods that perform comparably well in the case of the standard deviation measure, namely a bivariate unobserved components model (M2), the seasonal dummy method (M5, M6), and seasonal adjustment using a standard procedure like X12-ARIMA (M8). Also the forecast combination exercise reveals that the moving-average method seems to outperform all other candidates. It has by far the largest

Table 4.4: Results of Forecast Combination Regressions

Measure	β coefficients			
	s.d		iqr	
	Difference	Level	Difference	Level
M1	0.468***	0.506***	0.532***	0.741***
M2	0.216*	0.177	-0.058	-0.338***
M3	0.128	-0.120	0.303**	0.111
M4	0.120	0.273***	-0.006	0.242***
M5	-0.305***	-0.206*	-0.085*	-0.143***
M6	0.106	-0.061	0.156***	0.198***
M7	0.175***	0.139*	0.035	0.095*
M8	0.092 (NA)	0.293 (NA)	0.123 (NA)	0.094 (NA)

Notes: “NA” refers to non-availability of an estimated coefficient value due to the fact that an adding-up constraint was imposed on the sum of the coefficients. *, **, *** denotes significance at 10, 5, 1 % levels.

weight associated with any method, which clearly speaks in favor of this method. All in all, our results are quite useful for practitioners and researchers as a tested benchmark to calculate dispersion measures from panels of survey data on fixed event forecasts.

4.2 Dispersion of Macroeconomic Forecasts in the G7¹⁰

4.2.1 Introduction

Expectations are known to be a crucial determinant of economic dynamics.¹¹ Although the available micro data sets make it possible to measure and test many aspects of heterogeneity (e.g., differences in income, portfolios, demographics, shocks or labor force status), they typically contain little information about expectations. In addition, even when such information exists, the length and frequency of the series do not allow to adequately investigate how the cross-sectional distribution of expectations varies over time, business cycles, and with economic policy. Consequently, there has

¹⁰Large parts of this section of the thesis (except Section 4.2.5) are based on a paper that I wrote jointly with two co-authors (Dovern et al., 2009).

¹¹See for instance Bernanke (2004), Morris and Shin (2005a), Woodford (2005) and the discussion in Section 1.2.

been little work on joint analyses of individual survey expectations across countries and variables with micro data.

We investigate determinants of disagreement (cross-sectional dispersion of forecasts) for six key economic indicators in G7 countries roughly over the past twenty years. Using a unique data set with individual expert forecasts from *Consensus Economics* (see Section 2.1), we provide a set of statistics that capture the key features of dynamics of disagreement and are consistently calculated across countries and variables. Although it is often challenging in large data sets like ours to find consistent results, to summarize them, and to interpret the findings, a number of findings emerge quite clearly from our analysis.

We find that disagreement about real variables (GDP, consumption, investment and unemployment) has a distinct dynamic from disagreement about nominal variables (inflation and interest rate). Disagreement about real variables intensifies strongly (by about 40 percent) during recessions. Disagreement about nominal variables rises with their level, has fallen after 1998 or so (by 30 percent), and is considerably lower under independent central banks (by 35 percent). For both groups cross-sectional dispersion increases with uncertainty about the underlying indicators, and disagreement is more strongly cross-correlated among variables within the groups than between them.

While we provide simple and transparent reduced-form estimates, we believe our statistics also suggest a causal relationship: central bank independence reduces disagreement about nominal variables. Country-by-country regressions for inflation and interest rates reveal that both the level of disagreement and its sensitivity to macroeconomic variables tend to be larger in Italy, Japan and the United Kingdom, where central banks became independent only around the mid-1990s. These findings suggest that more credible monetary policy can substantially contribute to the anchoring of expectations about nominal variables. In contrast, its effects on disagreement about real variables are moderate.

We believe our results could be of interest to both policy-makers and researchers. A key consensus result of the large literature on monetary theory and policy is that anchored inflation expectations are of utter importance for safeguarding of price stability. Much work – including Cogley and Sargent (2001) and Stock and Watson (2005) – has documented that the inflation and GDP processes in G7 countries became more stable in the late 1980s and their volatility has fallen further until recently.¹²

¹²More precisely, the work typically finds that the variance of the permanent component of inflation and GDP was declining before 2006.

Our data set confirms the existing finding that consensus (mean) expectations have become more stable for most countries and variables. However, for expectations to be perfectly anchored it is necessary that also their cross-sectional dispersion—disagreement—disappears. Our results document, across several countries and variables, the extent to which this has been the case and suggest how economic shocks and monetary policy settings contributed to the fall in disagreement we often find after 1998. Researchers can use the stylized facts we report to calibrate, test and improve models with heterogenous beliefs, learning or information processing constraints, which have recently become quite widespread.

Our work builds on two strands of literature on survey expectations. The first and larger area analyzes the central tendency in expectations about inflation, GDP and interest rates.¹³

A large literature exists on extracting inflation expectations from prices of indexed bonds. For example, Gürkaynak et al. (2006), Ehrmann et al. (2007) and Beechey et al. (2008) provide evidence similar to ours, on anchoring of long-run inflation expectations in the US, the euro area, the UK and Sweden using high-frequency financial data.

The second, more recent and more closely related body of work investigates heterogeneity in expectations, often using micro data. The key inspiration for our work is a recent seminal paper of Mankiw et al. (2003), which analyzes central tendency and dispersion of inflation expectations using several US survey data sets. In this work the authors test some theories of disagreement.

Separate work of Souleles (2004) uses the Michigan Survey Consumer Sentiment to examine the ability of various groups in the population to forecast consumption expenditure. Blanchflower and Kelly (2008) study determinants of inflation expectations in the Bank of England's Inflation Attitudes Survey and the European Commission's consumer survey. Carroll (2003) bridges

¹³Branch (2004) estimates a model of boundedly rational agents on inflation expectations from the Survey of Consumer Attitudes and Behavior of the University of Michigan. Ang et al. (2007) find that survey expectations provide better inflation forecasts than macro variables or asset markets. Using consensus forecasts from the same data set as ours, Levin et al. (2004) investigate the degree to which inflation expectations are anchored in industrial countries. Also in the *Consensus Economics* data set, Patton and Timmermann (2008a) study how uncertainty about macroeconomic variables is resolved using forecasts of the US inflation and GDP growth. Bernanke and Boivin (2003) and Faust and Wright (2007) compare the Greenbook inflation and GDP forecasts (produced by the US Federal Reserve) to predictions generated by reduced-form econometric models. Kim and Orphanides (2005), Piazzesi and Schneider (2008) and others use interest rate expectations from the US Survey of Professional Forecasters to improve on existing yield curve models.

the two strands of literature by proposing and testing a model of average inflation and unemployment expectations in which households interact with experts. But joint analyses of individual survey expectations across countries and variables are non-existent.

4.2.2 Data Issues

The Data Set

For the analysis in this section we use data from the *Consensus* survey data set on professional macroeconomic forecasts that has been described in more detail in Chapter 2.

In this section, we focus on forecasts for the following six principal macroeconomic indicators: consumer-price inflation, nominal three-month interest rate, GDP growth, consumption growth, investment growth and unemployment rate.¹⁴ Although the survey contains information on other variables (most importantly, industrial/manufacturing production, producer prices, wages, current account and budget balance), their coverage in terms of time period, countries and number of respondents is less complete. These additional indicators are also arguably less important and often less closely followed by forecasters than those we focus on.

Fixed Event and Fixed Horizon Forecasts

Except for interest rates, the respondents answer questions about their expectations over the current and the next calendar year; the survey data thus provide series of *fixed event* forecasts.¹⁵ However, we believe *fixed horizon* (e.g., one-year-ahead) forecasts are preferable for the analysis of disagreement because the forecasting horizon of fixed event forecasts varies from month to month and consequently uncertainty and cross-sectional dispersion are strongly seasonal.¹⁶

Following the method suggested in the previous section, we approximate fixed horizon forecasts as a weighted average of fixed event forecasts as fol-

¹⁴We have also investigated expectations about industrial production. These results are broadly consistent with those for GDP and not reported here.

¹⁵Once every quarter the survey includes additional questions for selected variables (CPI inflation, GDP, consumption) on the fixed horizon predictions for roughly the following two years (we used this data for the analysis presented in Section 3). However, these questions are not useful for the analysis of disagreement because only the *consensus (mean) forecasts* are published.

¹⁶In addition, we use fixed horizon forecasts, because we want to provide comparable results to much of the literature, including Mankiw et al., 2003.

lows. Denote $\mathbf{f}_{y_0,m,y_1}^{\text{fe}}(x)$ the fixed event forecast of variable x for year y_1 made in month m of previous year $y_0 = y_1 - 1$, and $\mathbf{f}_{y_0,m,12}^{\text{fh}}(x)$ the fixed horizon, twelve-month-ahead forecast made at the same time. For example, the November 2008 forecast for year 2009 is $\mathbf{f}_{2008,11,2009}^{\text{fe}}(x)$. Our approximation of the fixed horizon forecast is:

$$\mathbf{f}_{y_0,m,12}^{\text{fh}}(x) = \frac{12 - m + 1}{12} \mathbf{f}_{y_0,m,y_0}^{\text{fe}}(x) + \frac{m - 1}{12} \mathbf{f}_{y_0,m,y_0+1}^{\text{fe}}(x). \quad (4.8)$$

This means that we compute the fixed horizon forecast for the next twelve months as an average of the forecasts for the current and next calendar years weighted by their share in the forecasting horizon. For example, the November 2008 forecast of the inflation rate between November 2008 and November 2009 is approximated by the sum of $\mathbf{f}_{2008,11,2008}^{\text{fe}}(\pi)$ and $\mathbf{f}_{2008,11,2009}^{\text{fe}}(\pi)$ weighted by $\frac{2}{12}$ and $\frac{10}{12}$ respectively.

We use this procedure for all variables except the interest rate, which is reported as the fixed horizon forecast for the interest rate between now and three months from now.

Because the disagreement series is typically used only as the dependent variable, the remaining approximation/measurement error in series $\mathbf{f}_{y_0,m,12}^{\text{fh}}(x)$ from (4.8) does not affect the consistency of the regression estimates obtained below as long as the error is not correlated with the regressors. Such correlation should be relatively low also given the monthly frequency of the data.

Ultimately, it is an empirical question how well our approximation performs. Dovern and Fritsche (2008) investigate the ability of approach (4.8) to capture cross-sectional dispersion of predictions in the US Survey of Professional Forecasters, which contains both fixed event and fixed horizon forecasts. Dovern and Fritsche find that approximation (4.8) captures well the heterogeneity in fixed horizon forecasts.¹⁷

As the final issue, we need to decide about our preferred measure of cross-sectional dispersion of forecasts. For most parts of the remainder of this section, we use the interquartile range (IQR). We do so to be consistent with the previous work (Mankiw et al., 2003) and because the IQR is also less likely to be subject to outliers than the cross sectional standard deviation. The results for disagreement measured with standard deviation are consistent

¹⁷Correlation between cross-sectional dispersion in (4.8) and the true dispersion of fixed horizon forecasts is roughly 0.8–0.9 when measured with standard deviation and 0.6–0.9 for the interquartile range. The remaining nine methods that Dovern and Fritsche (2008) investigate, including several specifications with unobserved components and seasonal adjustment, typically correlate with the true dispersion at 0.5–0.9 for standard deviation and 0.2–0.8 for the interquartile range.

Table 4.5: Correlation between the Two Disagreement Measures

	Canada	France	Germany	Italy	Japan	UK	US
Inflation	0.75	0.61	0.58	0.82	0.68	0.92	0.70
Interest Rate	0.87	0.78	0.80	0.84	0.88	0.86	0.73
GDP	0.79	0.62	0.78	0.69	0.76	0.66	0.83
Consumption	0.78	0.51	0.69	0.68	0.76	0.72	0.82
Investment	0.72	0.67	0.77	0.80	0.78	0.71	0.75
Unemployment	0.79	0.66	0.93	0.84	0.75	0.85	0.71

Notes: Numbers denote the correlation between the cross sectional standard deviation and the cross sectional IQR computed over the full sample.

with those presented below, which is not surprising given the relatively high correlation between the two measures shown in Table 4.5.¹⁸

Descriptive Statistics

In Appendix C, Figures C-1–C-6 give a first visual impression of the behavior of expectations as obtained by the method explained in the previous section. They compare the reported forecasts and actual variables over time. The actual series are shifted backward by twelve months so that the vertical difference between them and the expectations is the expectation error (for example, for November 2003, the dots denote expectations of one-year ahead inflation rate and the actual series is inflation between November 2003 and November 2004.) The shaded areas denote recessions as identified by the Economic Cycle Research Institute (using the business cycle method, which mirrors the NBER procedure).

Three findings appear for all six expectation series. First, expectations are more stable than the actual series as the actual series contain substantial unpredictable and volatile components.

Second, expectations are sensitive to current conditions. This is perhaps not surprising in case of inflation, interest rate and unemployment, which are generally thought to be quite persistent (so that the last observation is a good predictor for the future one(s)). However, the sensitivity to current conditions is also apparent – although to a lesser extent – for variables like GDP growth, which are not highly serially correlated.

¹⁸For a normal distribution standard deviation (std) is proportional to the interquartile range, $\text{std} = 1.349 \times \text{IQR}$ because the 75th percentile of the standard normal distribution is 0.6745. This scaling on average roughly holds in our data, e.g. for inflation in Canada average $\text{IQR} = 0.34$ and $\text{std} \times 1.349 = 0.26 \times 1.349 = 0.35$.

Third, expectations are sluggish in that they typically overestimate the developments when the underlying variable is falling. This finding is apparent for example during the disinflations of the early 1990s when inflation expectations errors were on average positive. The result is clearer for more persistent variables – inflation, interest rates and unemployment – than for those subject to large transitory fluctuations (GDP, consumption and investment).

Table 4.6 summarizes the key descriptive statistics about the fixed horizon forecasts obtained by the method described in the previous section and the actual series.

The average number of forecasters, displayed in the first line of each panel, typically ranges between 15 and 35. It shows little systematic variation over time: while in Canada, Japan and the US it is approximately constant, it rises somewhat in France, Germany and Italy and falls in the UK. The number of respondents does not correlate with the phase of the business cycle, and varies little across variables (in a given country). Observations for each forecaster are available for about half of the time on average.

The second line in each panel shows the mean expectation error averaged across forecasters and time periods. The individual forecasts are not biased significantly, partly because the standard deviation of expectation errors is quite large (the bias of consensus, or mean, forecasts is significant for a few variables in some countries). Average expectation errors are typically positive, which may reflect forecasters' optimism or sluggishness (where the trend in the underlying variable is falling most of the time, such as in the case of inflation and interest rates).¹⁹ There are few systematic differences across countries (for example, while the US respondents do well in case of interest rate, they are doing worst for consumption).

The lines three, four and five give the average mean squared errors of forecasts, average levels of the underlying variable and its variance respectively (taken over the sample period October 1989–October 2006). Level and variance of economic variables are likely to be positively correlated.²⁰ Both are positively correlated with the mean squared error (MSE) and also disagreement across forecasters. However, the evidence in Table 4.6, based on variables averaged over the sample and summarized in the top panels of Figure C-7 implies if anything a negative correlation between MSE and level

¹⁹Bias of inflation forecasts tends to be quite high and positive before 1999 and negative afterwards.

²⁰For empirical and theoretical investigation for inflation, see Ball and Cecchetti (1990) and Ball (1992), who proposes a model in which the level of inflation and its uncertainty are positively correlated because when inflation is high, policy-makers face a dilemma: they would like to disinflate but fear the resulting recession.

Table 4.6: Summary Statistics on Individual Forecasts and Targeted Variables

Statistic	Canada	France	Germany	Italy	Japan	UK	US
INFL							
Average # Forecasters	16.40	18.07	28.02	14.53	18.27	32.84	28.16
Average Forecast Error [†]	0.21	0.15	-0.01	0.04	0.14	0.28	0.05
Average MSE	1.87	0.43	0.81	0.75	0.66	1.19	0.92
Average Level of INFL	2.31	1.89	2.19	3.49	0.59	3.20	2.94
Variance of INFL	2.05	0.59	1.73	2.62	1.76	3.96	1.08
R3M							
Average # Forecasters	16.36	17.98	25.61	11.99	19.09	31.05	26.77
Average Forecast Error [†]	0.55	0.03	0.25	0.20	0.48	0.57	0.35
Average MSE	2.71	1.39	0.93	2.53	1.00	1.60	2.14
Average Level of R3M	5.35	5.20	4.72	6.76	1.78	6.75	4.55
Variance of R3M	7.76	8.64	5.97	16.29	6.37	9.08	3.77
GDP							
Average # Forecasters	16.41	18.18	27.65	14.56	18.39	33.03	28.19
Average Forecast Error [†]	-0.02	0.45	0.17	0.59	0.38	-0.26	-0.25
Average MSE	3.52	2.06	3.10	2.82	3.63	1.90	2.59
Average Level of GDP	2.65	1.85	1.85	1.37	1.46	2.33	2.89
Variance of GDP	3.98	1.49	3.38	1.90	3.34	2.02	2.03
CONS							
Average # Forecasters	16.40	18.16	27.90	14.51	18.37	32.65	28.00
Average Forecast Error [†]	-0.23	0.16	-0.06	0.40	0.18	-0.26	-0.54
Average MSE	2.23	1.66	2.38	3.90	2.19	2.46	1.86
Average Level of CONS	2.68	2.03	1.71	1.43	1.71	2.59	3.23
Variance of CONS	2.39	1.54	2.77	2.91	2.55	2.71	1.46
INV							
Average # Forecasters	16.29	17.81	27.64	14.53	17.98	31.99	27.86
Average Forecast Error [†]	2.43	1.38	2.61	1.35	1.74	-0.57	1.31
Average MSE	46.67	20.90	32.41	22.20	42.06	16.89	23.66
Average Level of INV	3.66	2.26	1.67	1.81	-0.19	2.89	4.13
Variance of INV	35.83	15.07	19.73	22.32	13.84	23.31	20.11
UN							
Average # Forecasters	16.41	18.07	27.65	13.99	17.98	32.32	28.12
Average Forecast Error [†]	-0.02	0.43	0.04	0.79	-0.01	0.63	0.07
Average MSE	0.81	0.89	1.29	0.97	0.18	1.21	0.42
Average Level of UN	8.54	9.98	9.58	9.95	3.75	5.34	5.50
Variance of UN	2.50	1.30	2.99	2.06	1.28	5.88	0.93

Notes: Averages taken across forecasters and time periods. † : {*,**,***}=Statistical significance at {10,5,1} percent.

of the underlying variable.²¹ In contrast, imprecision of expectations tends to increase with the variance of the underlying variable. While Figure C-7 is subject to many criticisms – such as that the relationship is only bivariate

²¹Note that investment, being an outlier due to large MSEs, is excluded from the figure to be able to assess if the positive relationship between MSE and variance holds even without it.

and based on time-averaged statistics – we believe it is an interesting starting point for more careful regression analyses of disagreement below.²²

4.2.3 Drivers of Disagreement

The previous section summarizes some key properties of the proxies for individual fixed horizon forecasts. As opposed to this, this section focuses on the *disagreement* among forecasters its evolution over time and its relationship to the business cycle and monetary policy.

A First Look at Disagreement

Table 4.7 summarizes the average disagreement by country and variable.²³ Disagreement about inflation is relatively low for France, Germany and Italy. Cross-sectional dispersion of interest rates is quite high in Canada and the US, relatively low in France and Germany, and extremely low in Japan. The last finding is driven by effectively zero interest rates for much of the time since 2000. Forecasters in France, Germany and Italy agree to a large extent on GDP growth, compared to their counterparts in the UK, Canada and in particular in Japan (where the dynamics are again dominated by the recession part of the sample). While the series for consumption growth is smoother than that for GDP growth, disagreement about consumption tends to be somewhat higher, driven perhaps by the lower attention that some forecasters pay to the consumption series.²⁴

However, the two disagreement series correlate quite strongly, which is perhaps not surprising given the large share of consumption in GDP (see also Figures C-10–C-11 for the time perspective). Disagreement about investment is substantially larger than for other series because of its high volatility. Unemployment on the other hand is smooth (and predictable), which translates into little disagreement.²⁵

Figures C-8–C-13 illustrate the evolution of disagreement over time. Perhaps unsurprisingly given the monthly frequency of our sample, disagreement

²²Because the focus of this paper is on disagreement, we will not investigate the determinants of MSEs below.

²³Detailed results for recessions, booms, pre-1999 and post-1998 are given in Table B-13 Appendix B.

²⁴Similarly, probably for the same reason average disagreement about industrial production (not reported in the table) is much higher (more than twice as high) than about GDP.

²⁵The somewhat higher mean for Germany is driven by the uncertainty about labor market statistics during the re-unification.

Table 4.7: Average Disagreement across Countries and Variables

Variable	Canada	France	Germany	Italy	Japan	UK	US
Inflation	0.34	0.21	0.24	0.26	0.33	0.38	0.33
Interest Rate	0.76	0.50	0.46	0.58	0.27	0.71	0.56
GDP	0.46	0.30	0.34	0.26	0.69	0.43	0.38
Consumption	0.47	0.31	0.40	0.34	0.65	0.55	0.38
Investment	2.60	1.17	1.56	1.01	2.42	1.66	1.93
Unemployment	0.29	0.24	0.41	0.33	0.25	0.31	0.22

Notes: Averages are based on the cross sectional IQR and taken across all time periods.

is subject to much transitory variation. However, two findings arise in several countries and series. First, disagreement tends to rise during recessions. Second, there is a downward time trend in disagreement.

Disagreement about inflation in Figure C-8 is roughly constant in France and Germany but falls steadily after 1992 in Italy (as the country was expected to join the euro area), in Japan, and in the UK. The series is quite strongly anti-cyclical (in terms of the difference between its average in recessions and booms) in Canada, Italy, Japan, the UK and the US.

As shown in Figure C-9, disagreement about interest rates tends to trend downward in all countries except for the US and its dynamics are strongly anti-cyclical (except for Japan where there was little disagreement when the interest rates were close to zero).

Except for France, disagreement about GDP growth in Figure C-10 is again anti-cyclical and typically 30–50% higher in recessions than in booms. Disagreement about the remaining real variables (consumption, investment and unemployment) broadly tracks that of GDP.

One can think of at least two structural breaks in our sample: the introduction of the euro in January 1999 and the German re-unification in October 1990. The expectations of the first event seem to have affected disagreement about inflation in Italy, which started to fall following the breakdown of the European Exchange Rate Mechanism in September 1992. Disagreement in the remaining two euro area members, France and Germany, has been roughly constant, perhaps because the inflation rate in these two countries has been low and stable (see also Table 4.8 below). Figures C-8–C-13 show, the structural break due to the German re-unification in October 1990 temporarily elevated disagreement about real variables (GDP, consumption, investment and unemployment), but not about inflation and interest rates. To a large extent unrelated to these two events, there have been much dynamics in dis-

Table 4.8: Descriptive Statistics about Inflation, GDP Growth and Interest Rates

Variable	Statistic	Canada	France	Germany	Italy	Japan	UK	US
Time Range: 1970–2007								
CPI Inflation	Mean	4.76	5.14	3.04	7.93	3.19	6.87	4.67
	Variance	11.51	17.78	3.99	39.24	23.25	31.30	9.04
GDP Growth	Mean	3.38	2.05 [†]	2.14	2.23	2.87	2.46	3.09
	Variance	6.43	1.49 [†]	3.91	5.39	8.27	4.40	4.59
Short-term IR	Mean	7.63	7.55	6.79 [‡]	9.95	4.51	8.92	7.03
	Variance	14.26	13.23	3.67 [‡]	28.72	13.60	12.91	12.53
Time Range: 1970–1989								
CPI Inflation	Mean	7.14	8.27	3.92	12.29	5.72	10.24	6.35
	Variance	9.08	13.99	4.63	34.92	30.76	34.29	10.94
GDP Growth	Mean	3.63	2.24 [†]	2.34	2.85	4.26	2.45	3.33
	Variance	5.45	1.24 [†]	4.49	7.20	8.09	6.69	7.10
Short-term IR	Mean	9.81	9.79	7.98 [‡]	13.21	7.09	11.02	9.22
	Variance	10.73	7.63	1.90 [‡]	19.39	6.46	8.08	10.11

Notes: Calculated from quarterly data. †: GDP data for France start in 1978, ‡: interest rate data for Germany start in 1973:Q2. All numbers are given in percentages. The data has been obtained from the OECD Main Economic Indicators, the IMF International Financial Statistics, and the Deutsche Bundesbank.

agreement on various series, in particular the clear downward trend in the UK and cyclical dynamics in most countries. We investigate these developments in more detail below using simple regression analyses.

Disagreement over Time

To provide quantitative insights Tables B-1–B-6 use the fixed effects panel estimator (in which coefficients other than the constant are restricted to be the same in all countries²⁶) to assess general trends common in all countries. We discuss the results for the three most important indicators – inflation, interest rates and GDP – in more detail and summarize the results for the remaining variables (consumption, investment and unemployment) only briefly.

The top panel (Panel A) of each of Tables B-1–B-6 investigates how disagreement (“disagr”) varies over time and during recessions using four versions of regression:

$$\text{disagr}_t = \beta_0 + \beta_1 \times \text{rec}_t + \beta_2 \times \text{post-1998}_t + \varepsilon_t,$$

where “rec” denotes the recession dummy and “post-1998” is the dummy for the second part of the sample.

²⁶The constant term β_0 in the tables is normalized to the average of country-specific intercepts.

Disagreement about inflation is analyzed in Table B-1. Row 1 reports that the cross-sectional interquartile range averaged across countries and time is about 0.3, which suggests that half of the forecasters typically lie within 0.15 percentage points of the consensus. Row 2 shows that disagreement rises by about 20 percent during recessions, a fact that can be due to the increase in general macroeconomic uncertainty. Row 3 documents that disagreement is much lower – by 25% – in the second part of the sample, after 1998.

Qualitatively similar findings are obtained for disagreement about interest rates and GDP growth and reported in Tables B-2 and B-3 respectively. For both variables, disagreement rises during recessions and falls after 1998. While the effects for interest rates are quantitatively similar to those for inflation, the increase in disagreement about GDP during recessions is almost twice as large – 44 percent (and the fall after 1998 is less pronounced). This seems reasonable as macroeconomic uncertainty during a recession is skewed toward GDP (and less evident for inflation and interest rates).

The results for consumption growth, investment growth and unemployment rate in Tables B-4, B-5 and B-6 are again qualitatively in line with those findings.

Qualitatively, the estimates (together with those of Tables B-1–B-2) suggest that the recession differential in disagreement – the difference between average disagreement in a recession and in a boom – is generally larger for real variables (GDP, consumption, investment and unemployment) than for the two nominal variables. In contrast, the fall in disagreement after 1998 tends to be smaller for real variables than for nominal ones.

Two important broad factors behind the variation in disagreement can be the shocks to economic variables and economic policy. Larger shocks boost the volatility of the underlying variables and make them less predictable. As a result, forecasters are more likely to disagree about future outcomes (because of using different models, priors, subjective probabilities or data). More credible economic policies can make economic indicators easier to forecast. An obvious example is the introduction of an explicit numerical inflation target, which can contribute to a better anchoring of inflation expectations. Similarly, independent central banks are often perceived as better safeguards to price stability (and can indirectly also contribute to the stabilization of output). We investigate these two factors – economic shocks and policies – in a simple reduced-form setup below.

Disagreement and Macro Variables

The middle panel (Panel B) of Tables B-1–B-6 investigates how disagreement correlates with the underlying variables, its squared change – a proxy of

uncertainty about the series x_t , the output gap²⁷, and the squared change in the policy interest rate ($\Delta \text{policy rate}_t^2$) – a proxy of the variation in monetary policy:

$$\text{disagr}_t = \beta_0 + \beta_2 \times x_t + \beta_3 \times \Delta_{12}x_t^2 + \beta_4 \times \text{output gap}_t + \beta_5 \times \Delta \text{policy rate}_t^2 + \varepsilon_t,$$

where x_t denotes the level of the underlying variable and $\Delta_{12}x_t^2 \equiv (x_t - x_{t-12})^2$ denotes its uncertainty.

Disagreement about inflation increases with its level: one percentage point increase in inflation raises the cross-sectional interquartile range by 0.026, or by about 10 percent (with respect to the mean 0.299). The direct effect of inflation uncertainty (the term $\Delta_{12}INFL_t^2$), which is also highly significant but smaller, suggests large changes in inflation rate disproportionately increase disagreement. The coefficient on output gap is negative, which is in line with the previous evidence that disagreement increases during recessions. Finally, disagreement about inflation rises when monetary policy rates change, which again tends to coincide with recessions. (But the positive coefficient on interest rates is significant even when output gap is included.) In addition, including interest rates among the explanatory variables substantially increases the explanatory power of the regression.

Disagreement about interest rates shown in Table B-2 rises with the level and squared change of the interest rates. These findings are in line with the fact reported in panel A that disagreement about interest rates fell after 1998, as both level and variation in rates is much lower in the second part of the sample (see also Figure C-2). In addition, disagreement also tends to move inversely to the output gap. While the coefficients in these regressions are comparable to those of inflation and of GDP growth, their explanatory power is considerably higher.

Table B-3 analyzes drivers of disagreement about GDP growth. In contrast to inflation and interest rates but in line with the evidence of panel A, disagreement about GDP growth moves inversely with its level: disagreement rises in periods of weak economic growth. Arguably, the effects of GDP growth on disagreement are non-linear: disagreement can be expected to rise during periods of heightened uncertainty, which likely occur during recessions, but also when economic growth accelerates considerably (however, the latter periods are virtually absent in our sample as GDP growth only rarely exceeds 5 percent). To proxy economic uncertainty in the fifth model for GDP, the model includes the squared change in GDP growth over the last

²⁷The output gap used here is the ex post estimate taken from OECD's Economic Outlook. The series is quarterly, interpolated constant within each quarter, and starts in 1991:Q1.

year, which turns out to be positive but insignificant. As for disagreement about inflation and interest rates, variation in interest rates analyzed in model 7 also improves the performance of the regression (measured with adjusted R^2).

Given the large share of consumption expenditure in output, it is not surprising that the findings for consumption in Table B-4 mirror those for GDP quite closely. The results are qualitatively similar for investment and unemployment rate although the explanatory power of investment regressions is smaller (as disagreement about investment tends to move more, much of which is unrelated to macro variables).

The results in panel B are also broadly agree with the bivariate illustration of the relationship between time-averaged disagreement and level/variance of the underlying variable in the bottom panels of Figure C-7. While the first correlation is close to zero (for reasons outlined above), the correlation to the variance, which proxies better for underlying uncertainty, is positive and quite strong.

Our findings in this and the previous sections are in line with Mankiw et al. (2003) and Döpke and Fritsche (2006). Mankiw et al. (2003) report that in the US disagreement about inflation increases with its level and absolute value of its change. This is true, in particular when the change is sharp, and though it shows an anti-cyclical pattern after 1975 for consumers, its dependence on the phase of the business cycle is less clear for experts. Döpke and Fritsche (2006) find that dispersion of inflation and growth expectations in Germany is high before and during recessions and that it correlates positively with macroeconomic uncertainty.

Disagreement and Central Bank Independence

It might be a priori expected that better macroeconomic policy alleviates economic uncertainty and disagreement. Without going in much detail into measuring the quality of economic institutions, panel C of Tables B-1–B-6 provides a simple illustration of how a better and more credible monetary policy affects disagreement about various variables. We attempt to capture credibility of monetary policy using an indicator of central bank independence as defined in Table 4.9.²⁸

²⁸We intentionally use a simple indicator, which transparently tracks central bank independence throughout our sample. The indicator is broadly in line with a measure of political autonomy of central banks newly calculated by Arnone et al. (2007), who use the methodology proposed by Grilli et al. (1991) and Cukierman (1992). Their approach defines political autonomy as the ability of central banks to select the objectives of monetary policy. They measure independence using a combination of eight criteria related to how the governor and board of directors are appointed, the relations with government and the

Table 4.9: Central Bank Independence in G7 Countries, 1989–2006

Country	Monetary Policy Setting	Dummy Variable	Fraction of Sample with Indep. CB
Canada	De facto independent central bank;*	1 full sample	1
France [†]	Independent central bank since August 4, 1993	0 before August 1993, 1 otherwise	$\frac{159}{205} = 0.78$
Germany [†]	Independent central bank since August 1, 1957	1 full sample	1
Italy [†]	Independent central bank effectively since January 1, 1994 [‡]	0 before January 1994, 1 otherwise	$\frac{154}{205} = 0.75$
Japan	Independent central bank since June 18, 1997	0 before July 1997, 1 otherwise	$\frac{112}{205} = 0.55$
UK	Independent central bank since June 1, 1998; inflation targeting since October 1992	0 before June 1998, 1 otherwise	$\frac{101}{205} = 0.49$
US	Independent central bank since December 23, 1913	1 full sample	1

Notes: *: Bank of Canada was technically independent *until* 1967 and has been *de jure* dependent on the Minister of Finance since then. (Inflation targeting since February 1991.) †: Refers to the period before the country joined the euro area. ‡: Formally since February 7, 1992, however, granting of independence was effectively not completed until 1994. Information has been obtained from the web pages of the respective central banks.

We estimate two versions of the following regression

$$\begin{aligned} \text{disagr}_t &= \beta_0 + \beta_1 \times \text{CB Independence}_t + \beta_2 \times x_t + \beta_3 \times \Delta_{12}x_t^2 + \\ &+ \beta_4 \times \text{output gap}_t + \beta_5 \times \Delta \text{policy rate}_t^2 + \varepsilon_t. \end{aligned}$$

The dummy for central bank independence (CB Independence_t) is negative, large and highly statistically significant for all six variables.²⁹ Quantitatively, the reduction in disagreement related to central bank independence is largest for the two nominal variables, interest rates and inflation, 41% and 35% respectively; for real variables it ranges between 5% and 20%. In addition, the explanatory power of these regressions is substantially larger than of those with recession dummies of model 2 (and even marginally larger than those

nature of the laws relevant for central banks. Appendix Table 8 of Arnone et al. (2007) implies that political autonomy of central banks was relatively low in Japan, the UK, France, Canada and Italy in the late 1980s, has generally risen in G7 countries between the late 1980s and 2003, but still remained relatively low in 2003 in Japan, Canada and the UK.

²⁹The dummy is only marginally significant for investment.

with recession and post-1998 dummies of model 3) for nominal variables. The opposite is true, however, for real variables (where adjusted R^2 s of model 8 lie below those of model 2).

Model 9 attempts to separate the effects of central bank independence and other factors (by including macroeconomic control variables of Panel B jointly). The estimates imply that the monetary policy indicator remains overwhelmingly significant for nominal variables, but less so for real indicators. For most variables, the point estimate of β_1 changes only modestly (relative to model 8). At the same time some other regressors, in particular economic uncertainty turns out to be statistically significant (and broadly comparable in size to estimates of β_3 in Panel B).

These findings suggest that (i) higher central bank independence coincides with a substantial decline in disagreement and (ii) the effect is particularly pronounced for nominal variables. While the first result, the quantification of effects of central bank independence on disagreement, is to our knowledge new, it bears some relation to the large literature on economic effects of central bank independence (Rogoff, 1985; Alesina and Summers, 1993; Alesina and Gatti, 1995 and many others). Most empirical work in the field agrees that central bank independence promotes price stability although its effects on real economic performance are hard to pin down, which is broadly in line with our second finding.

The second result can also be explained with the introduction of inflation targeting (in several countries in our sample) and more generally with the adoption of more predictable monetary policy and increased and improved communication of central bankers with other economic agents. The effect of these developments is stronger for nominal variables, which are directly affected by explicit inflation targets or communication about possible future paths of policy rates. On the other hand, disagreement about real variables, whose future dynamics are typically communicated less extensively, is less sensitive to the institutional setting of monetary policy.

The explanatory power of our regressions is quite low; the adjusted R^2 often ranges between 0.1 and 0.2. This is perhaps not surprising because Figures C-1–C-6 show that disagreement is subject to much transitory variation, which cannot easily be captured with our explanatory variables and simple models. The disagreement series we construct is subject to much measurement and sampling uncertainty: First, questions that aim at capturing expectations about economic variables can be challenging to answer even for professional forecasters. Second, we use monthly data, which are generally known to be noisy. Finally, we attempt to extract cross-sectional variation from a sample of only about 20–30 experts. However, we believe the data

still do provide interesting information because many of the coefficients we estimate are overwhelmingly significant and reasonable in size.

4.2.4 Disagreement Across Countries

While panel analysis is useful in summarizing the general trends in determinants of disagreement, averaging wipes out information about cross-country heterogeneity. The analysis of this section attempts to capture and interpret such heterogeneity.

Country-by-Country Regression Analysis

Tables B-7–B-12 summarize the cross-country differences in the drivers of disagreement by estimating models 3 and 7 (of Tables B-1–B-6) separately country by country.

The results are broadly consistent with the panel analysis of the previous section. While most coefficients are signed in line with our priors, some of them are insignificant because of their smaller size and because of their larger standard errors caused in part by fewer observations (roughly 200 compared to roughly 7×200 for panel estimates).

The top panel of Table B-7 divides countries into two groups depending on how much the disagreement about inflation varies over time. In Canada, France, Germany and the United States this variable is rather insensitive to the phase of the business cycle and constant over the two time periods (pre-1999 and post-1998). Hence, while the coefficients β_1 and β_2 are sometimes significant (and almost always have the correct sign), the explanatory power of the regressions in these countries is rather low (about 0.1 or less in terms of adjusted R^2), because the coefficients tend to be smaller than in the remaining countries. In contrast, in Italy, Japan and the United Kingdom the two variables explain up to 40 percent of the variation in disagreement about inflation (in adjusted R^2 terms). Findings consistent with the top panel are shown in the bottom panel: Disagreement in Italy, Japan and the UK is much more sensitive to macro variables (inflation level and its variation, output gap and variation in policy interest rates) than in the rest of the sample, where coefficients are often insignificant or small.

Analogous regressions for short-run interest rates shown in Table B-8 are broadly consistent with those for inflation, even though the difference in sensitivity of disagreement for the two country groups is less pronounced. Adjusted R^2 s for Italy, Japan and the United Kingdom range around 0.4–0.5; those for other countries average to about 0.25. This finding seems reasonable, as the existence of independent central banks and explicit nu-

meric inflation targets can help stabilize inflation expectations and reduce disagreement about inflation. In contrast, such targets are not announced for interest rates (or other variables).

The results for the remaining, real variables – GDP, consumption, investment and unemployment rate, shown in Tables B-9, B-10, B-11 and B-12 respectively – exhibit little systematic variation across countries. The finding that the link between monetary policy institutions and sensitivity of disagreement about real variables is not particularly pronounced is explained by the fact that the key (and typically sole) goal for monetary policy is safeguarding of price stability. In contrast, central banks usually affect output stability only indirectly.

The level and sensitivity of disagreement about inflation and interest rates relate quite closely to the fraction of the sample with an independent central bank, displayed in the right-most column of Table 4.8. Canada, France, Germany and the United States have had an independent monetary authority for most of the sample period, Japan and the U.K. only for about half of the time.

Credibility of monetary policy affects disagreement through (at least) two channels: First, it stabilizes inflation rates, so that economic shocks do not affect prices. Second, given inflation volatility, it anchors inflation expectations, so that the shocks which are reflected in inflation do not affect inflation expectations (and disagreement). The first channel could be the reason why in model 9 of Tables B-7 and B-8 macroeconomic indicators remain significant.

However, the case of Italy suggests that the measure is not good enough to capture the whole story. Italy had, according to our measure, credible monetary policy for three quarters of the sample, but its level and sensitivity of disagreement is high. In our view, it is likely that disagreement could be affected by current and past economic uncertainty.

Although it is beyond the scope of this analysis to investigate this hypothesis more formally, we would like to provide some suggestive evidence. Table 4.8 shows averages and variances of some key economic variables (inflation, GDP growth and interest rates).³⁰ The two countries with high volatility of inflation and interest rates are Japan and, in particular, Italy where the two variances are almost twice as large as in any other country (except for inflation in the UK). In contrast, economic conditions have generally been

³⁰The top panel displays the baseline statistics for 1970–2007. The bottom panel shows that the statistics calculated for the period before the beginning of our estimation sample (October 1989) are broadly consistent.

stable in France, Germany and the US. Consequently, the history of economic shocks seems to affect the dynamics of disagreement.

Cross-Variable and Cross-Country Links between Disagreement

We have also investigated the cross-country and cross-variable links in disagreement. We only summarize them briefly as our results are not clear-cut enough to justify a more detailed exposition.³¹ First, for a given variable, disagreement is typically moderately positively correlated across countries with correlation of around 0.2 (at monthly frequency).³² Cross-country correlations tend to be higher for more persistent variables: interest rates and unemployment.

Cross-variable correlations (within a given country) are in the same ballpark as the cross-country ones and suggest that disagreement about inflation and GDP growth is strongly correlated with other variables. Disagreement is more synchronized in the UK and the US. Multivariate analyses of conditional correlations – regressions of disagreement about one variable on disagreement about the remaining variables – suggest that disagreement tends to co-move for real variables. This holds especially for GDP, consumption and unemployment, and for nominal variables, inflation and interest rate. Links between real and nominal variables are less important (conditional on correlations between variables from the same group). We found little systematic patterns between countries in cross-country conditional multivariate regressions (i.e., regressions of disagreement in one country on disagreement in others for a given variable).

4.2.5 Robustness Check Using Non-Transformed Raw Data³³

Although in Section 4.1 we have shown for data from the Survey of Professional Forecasters that a moving average of fixed event forecasts is a good proxy for corresponding (implied) fixed horizon forecasts (see also Dovern and Fritsche, 2008), the work presented in the previous section (see also Dovern et al., 2009) might be challenged by the critique that this result may not apply automatically to every other survey data set. To address this concern, we provide some additional evidence on the determinants of disagreement

³¹Cross-variable and cross-country correlations in disagreement are given in Tables B-14 and B-15 in the Appendix.

³²Correlations increase to around 0.3–0.4 once the data are aggregated to quarterly frequency by averaging.

³³This section of the thesis is based on an unpublished note (Dovern, 2009b).

among forecasters using the non-transformed raw data from the *Consensus Forecast* data set in this section. The results are based on an empirical study that analyzes the forecasts for GDP, consumption, industrial production, and the inflation rate.

The Horizon-Effect

It is a well known fact that the dispersion of a cross-section of fixed event forecasts decreases as the forecast horizon shrinks, since more and more information becomes available and the forecasts tend towards the realized outcome. Depending how much information about an outcome of a variable is revealed during the period for which the forecast is made this decline in disagreement might be more or less pronounced.

We capture this horizon effect by regressing the time series of cross-sectional standard deviations (derived from forecasts for one variable), σ_t^i , on a constant and a trend variable, $horizon_t$, that captures the decline of the forecast horizon:

$$\sigma_t^i = c + \alpha horizon_t + \epsilon_t, \quad (4.9)$$

where ϵ_t denotes the residual term and α is the effect of a one month change on the degree of disagreement. $horizon_t$ is equal to the remaining forecast horizon in each month, e.g. equal to 12 in January or equal to 2 in November.

Note that we run separate regressions for the forecasts with forecast horizon of one year or less and those with forecast horizon of more than one year. The reason why we do this, is that a visual inspection of the disagreement data shows that disagreement declines sharply during the year for which the forecast is made, while it is fairly constant if there is more than one year to go until the outcome is released.

In Appendix B, we show the estimates for α for all variables and countries in the sample. Table B-16 shows the results for the forecasts with a horizon of more than a year and Table B-17 shows the results for the forecasts with a horizon of one year or less. The corresponding t-statistics show that the effect is highly significant for the current year forecasts in all cases. While this is also true for inflation forecasts, for forecasts with a horizon of more than a year, the effect is not significant for the wide majority of next year forecasts in case of the three other variables.

To see how important the decline in disagreement is relative to the average degree of disagreement about the outlook for a specific variable, we compute how large the change of disagreement in a given month is as a percentage of the average level of disagreement. The numbers are presented in Table B-18 and Table B-19. While the numbers are quite low for the forecasts with a

horizon of more than a year, they are substantially high for lower horizon forecasts. Two observations are worth noting: First, the relative reduction of disagreement is highest for inflation forecasts. Second, the relative reduction of disagreement is highest in the US.

To see whether the decline in disagreement occurs in a linear fashion or whether there is a nonlinear effect, we run a similar regression, in which we include a quadratic trend in addition:

$$\sigma_t^i = c + \alpha_1 \text{horizon}_t + \alpha_2 \text{horizon}_t^2 + \epsilon_t \quad (4.10)$$

Note that $\alpha_2 > 0$ ($\alpha_2 < 0$) would imply that the reduction of disagreement follows a concave (convex) pattern, as the forecast horizon shrinks.

The results in Tables B-20 and B-21 show a significant result for forecasts for the next year only in one case (Inflation in Germany). The effects of the quadratic term are somewhat larger for the forecasts for the current year though they remain insignificant in most cases. A general feature is, however, that they tend to be negative with the exception of the US where all estimates of α_2 are positive.³⁴ That means that the reduction of disagreement evolves in a convex way, i.e. it starts slowly and accelerates the closer the forecast horizon shrinks to zero. These shapes are broadly in line with what Patton and Timmermann (2008a) find for inflation and growth forecasts. Figure C-14 in Appendix C visualizes the typical pattern on hands of an example.

To see whether there are secular trends in disagreement over the years (that we could show to exist in the previous section), we included a time trend in (4.9). The results are given in Tables B-22 and B-23. The results show that in some cases a significant decline in the level of disagreement can be observed for the forecasts for the current year as well as for the forecasts for the next year. This might be interpreted as a side-effect of the Great Moderation. The economy has become less volatile, there is a better understanding of implications of economic policy actions, and monetary policy has become more transparent. All this has led to more agreement among forecasters about the economic outlook, i.e. the relative importance of idiosyncratic forecasting errors has declined.

The same kind of analysis can be conducted with a measure of disagreement that simultaneously takes all four variables into account. The development of the vector variances for all countries are shown in Figure C-15.³⁵ A

³⁴In fact, with one exception (Inflation in the US) all significant effects are negative without exception.

³⁵The vector variance at time t is the norm of the vectorized covariance matrix of all forecasts made at that point in time. Formally, if Σ_t denotes the covariance matrix in month t , the vector variance, VV_t is given by $VV_t = \|\text{vec}(\Sigma_t)\|$, where $\text{vec}(\cdot)$ denotes the vec-operator (see e.g. Djauhari, 2007).

visual inspection reveals that also for the multivariate measure the horizon-induced time-variability is much more apparent for the forecasts with a forecast horizon of one year or less. This is also confirmed by a formal regression; Table B-24 shows all estimates for the horizon effect in (4.9). Again, as one would expect there is not much of such an effect for the forecasts with horizon of more than a year, while the effect is highly significant for the forecasts with horizon of one year or less. The relative change per month is again found to be highest for the US forecasts.

Also here, we checked for a quadratic effect using (4.10). The results show, however, that no such nonlinear effect is significantly estimated in the multivariate setup.

Dispersion and the Business Cycle

To measure the sensitivity of disagreement to the stance of the economy, we can augment (4.9) by a measure of the output gap:

$$\sigma_t^i = c + \alpha \text{horizon}_t + \beta (y_t - y_t^*) + \epsilon_t, \quad (4.11)$$

where y_t is the log of industrial production at time t and y_t^* is the log of a measure of potential production that we estimate by a simple HP-filter. The results are shown in Tables B-25 and B-26. An increase (decrease) in the output gap tends to decrease (increase) dispersion about the outlook of the economy for forecasts with a horizon of one year or more while the opposite (though not significantly in most cases) is true for forecasts with a lower forecast horizon. In addition, there are also some significant differences across variables. While the disagreement across forecasts for the real variables – especially GDP and industrial production – seems to have a strong counter-cyclical element, we find less of such evidence for disagreement across inflation forecasts. This finding is in line with our results from Section 4.2.3 and the evidence that Patton and Timmermann (2008b) provide for disagreement across GDP and inflation forecasts in the US.

To cross-check these results – which might be influenced by the way we construct the output gap – we include a variable that is equal to 1 if there is at most one month of recession in the year for which a forecast is made, and 0 otherwise (instead of the output gap measure).³⁶ Using this variable, we can check whether the occurrence of a recession during the time, for which the forecast is targeted, influences the degree of disagreement. The results are given in Tables B-27 and B-28. They show that disagreement about the outcome of variables for the ongoing year tends to be higher in almost all

³⁶The recession dating was taken from the Economic Cycle Research Institute (ECRI).

cases if there is a recession during that year. On the contrary, there is no such clear-cut result for longer horizon forecasts. Only some of the results are significant; in two cases (inflation Canada and Italy) the effect is negative, and the magnitude of the effect is much lower than in Table B-28 in most cases.

We analyze the correlation between business cycle and dispersion also using the multivariate measure of dispersion. If we replace σ_t^i in (4.11) by the multivariate measure VV_t , we obtain again the results that while a negative output gap increases disagreement about the outlook for the next year, it tends to lower disagreement for the current year (though again not significantly in about half of the cases). Detailed estimates are shown in Table B-29.

All in all, the results support the widely expressed perception that it is notoriously hard to forecast turning points which are usually only foreseen by a minority of forecasters. Instead, once facing a recession it is relatively easier to anticipate the trajectory of the recovery such that forecasters mostly agree on the figures for the recession year.

4.3 Concluding Remarks

In Section 4.1, we have shown how well different methods perform that can be used to extract a measure of disagreement from fixed event forecasts, which is not distorted by horizon-effects. We concluded that the approach via a simple moving average of fixed event forecasts performs best in general and is easily applicable and transparent.

Using this method in Section 4.2 to construct and analyze forecast dispersion in the G7 countries, we document a dichotomy between disagreement about real variables (GDP, consumption, investment and unemployment), which is more strongly affected by real factors, and disagreement about nominal variables (inflation and interest rate), which react to the institutional setting of monetary policy (in particular central bank independence). Disagreement about real variables intensifies strongly during recessions. Disagreement about nominal variables is considerably lower under independent central banks. Cross-sectional dispersion for both groups increases with uncertainty about the underlying indicators. Country-by-country regressions for inflation and interest rates reveal that both the level of disagreement and its sensitivity to macroeconomic variables tend to be larger in Italy, Japan and the United Kingdom, where central banks became independent only around the mid-1990s.

Our findings also suggest that more credible monetary policy can substantially contribute to anchoring of expectations about nominal variables; its effects on disagreement about real variables are moderate. While our analysis uses data on expectations of professional forecasters, qualitatively similar results may be obtained also for other economists (in industry, government and academia) and households. This could be the case if our data are viewed as a proxy for expectations of the rest of population, or if news spread epidemiologically from experts to other agents (as proposed by Carroll, 2003).

Some of these findings were tested for robustness in Section 4.2.5 using a different approach that does not rely on transformed data; the results usually supported the findings qualitatively. We interpret this as evidence in favor of the proposed way of working with the transformed measures of disagreement to circumvent the horizon-effect. In addition, using the non-transformed data, we find some features of the horizon-effect that are in line with findings of other papers on single countries.

To our knowledge, the results of the studies presented in this part of the thesis are one of the first joint analyses of dispersion of individual survey expectations across countries and variables using micro data. The strength of some signals we use to investigate disagreement has been relatively weak: Following the Great Moderation, economic shocks in our sample (1989–2006) have been quite modest. Further insights about expectations and disagreement will be gained once data covering the recent global turbulences become available.

Chapter 5

Summary

In this thesis we use survey data to analyze expectations about macroeconomic variables and the way these expectations are formed. Thereby, we concentrate on inflation (expectation) dynamics and the analysis of forecast dispersion. Most of the analysis is based on a newly constructed data set that covers individual forecasts for a bunch of macroeconomic variables for all G7 countries.

Using a new approach for modeling forecast errors in a structural way, we show that most of the individual forecasts in the *Consensus* survey data set have fairly good properties with respect to unbiasedness and efficiency.

We provide empirical evidence in favor of partial equilibrium models of the sticky information type that explain the development of inflation dynamics and the formation process of inflation expectations of households respectively. The model proposed by Carroll (2003) as well as the SIPC proposed by Mankiw and Reis (2002) fit European data reasonably well.

In the last part of the thesis, we present an innovative analysis of the determinants of forecast dispersion that is unique in the literature. We show how forecast dispersion can be estimated from fixed event data and use the method that was found to outperform the other methods to analyze the behavior of forecast dispersion in the G7 countries. We show that there is a dichotomy between disagreement about real variables, which is more strongly affected by real factors, and disagreement about nominal variables, which is related to the institutional setting of monetary policy. Disagreement about real variables intensifies strongly during recessions, while disagreement about nominal variables is considerably lower under independent central banks.

In all those analyses, we present evidence for several countries – thus increasing the reliability of our findings. By analyzing data for multiple countries, we were able to present more general evidence than would have been the case if we had concentrated only on a specific country.

There are several directions for extensions to some of the issues covered in this thesis. First, the structural form by which we modeled the forecast errors in section 2.3 is still based on some strong assumptions (e.g. linear accumulation of individual shocks) and could certainly be generalized once data over a longer time span becomes available.¹ Second, using our findings that the average forecasts – while usually outperforming individual forecasts in terms of forecast accuracy – show a high degree of inefficiency, one could use the autocorrelation of revisions to generate “modified” consensus forecasts that have better forecasting precision. Third, survey data might also help to assess the validity of different concepts for modeling inflation dynamics in a general equilibrium framework which is done so far only under the assumption of model consistent rational expectations (e.g. Trabandt, 2007; Korenok, 2008). Fourth, while we analyze the determinants of dispersion in a two-steps procedure (first we generated a measure of dispersion and then we used it as the regressand in the regression analysis), one could imagine that analyzing the effects in one model might bring efficiency gains. The latter could maybe be implemented as some form of heteroscedasticity model with exogenous drivers of the conditional dispersion of forecasts. Finally, once some stylized facts about the development of dispersion of expectations and its relation to uncertainty have been established, it may be worthwhile to incorporate this heterogeneity into theoretical general equilibrium models as recently demonstrated by Bloom et al. (2008). This way, the models could potentially reflect the heterogeneity of agents and the impact of uncertainty on investment decisions better than current macroeconomic models.

¹See also the discussion in Dovern (2009a).

Bibliography

- Ager, P., Kappler, M., Osterloh, S., 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.
- Alesina, A., Gatti, R., 1995. Independent central banks: Low inflation at no cost? *American Economic Review, Papers and Proceedings* 85, 196–200.
- Alesina, A., Summers, L. H., 1993. Central bank independence and macroeconomic performance: Some comparative evidence. *Journal of Money, Credit, and Banking* 25, 151–162.
- Andolfatto, D., Hendry, S., Moran, K., 2008. Are inflation expectations rational? *Journal of Monetary Economics* 55 (2), 406–422.
- Ang, A., Bekaert, G., Wei, M., 2007. Do macro variables, asset markets or surveys forecast inflation better? *Journal of Monetary Economics* 54, 1163–1212.
- Arnone, M., Laurens, B. J., cois Segalotto, J.-F., Sommer, M., 2007. Central bank autonomy: Lessons from global trends. Working Paper 88, IMF.
- Ashiya, M., 2003. Testing the rationality of Japanese GDP forecasts: the sign of forecast revision matters. *Journal of Economic Behavior & Organization* 50, 263–269.
- Ball, L., 1992. Why does higher inflation raise inflation uncertainty? *Journal of Monetary Economics* 29, 371–378.
- Ball, L., Cecchetti, S., 1990. Inflation uncertainty at short and long horizons. *Brookings Papers on Economic Activity* 1, 251–254.
- Ball, L., Romer, D., 1990. Real rigidities and the nonneutrality of money. *Review of Economic Studies* 57, 539–552.

- Ball, R. J., 1962. The prediction of wage-rate changes in the United Kingdom economy 1957-60. *Economic Journal* 72, 27–44.
- Batchelor, R., 2001. How useful are the forecasts of intergovernmental agencies? The IMF and OECD versus the consensus. *Applied Economics* 33, 225–235.
- Batchelor, R., 2007. Bias in macroeconomic forecasts. *International Journal of Forecasting* 23, 189–203.
- Batchelor, R., Dua, P., 1990. Forecaster ideology, forecasting technique, and the accuracy of economic forecasts. *International Journal of Forecasting* 6, 3–10.
- Batchelor, R., Dua, P., 1992. Conservatism and consensus-seeking among economic forecasters. *Journal of Forecasting* 11, 169–181.
- Batchelor, R. A., Orr, A. B., 1988. Inflation expectations revisited. *Economica* 55 (219), 317–331.
- Bates, J. M., Granger, C. W. J., 1969. The combination of forecasts. *Operations Research Quarterly* 20, 451–468.
- Beechey, M. J., Johannsen, B. K., Levin, A. T., 2008. Are long-run inflation expectations anchored more firmly in the euro area than in the united states? FEDS discussion paper 23, Federal Reserve Board.
- Bernanke, B. S., December 2 2004. The logic of monetary policy. Remarks before the National Economists Club, Federal Reserve Board.
- Bernanke, B. S., Boivin, J., 2003. Monetary policy in a data-rich environment. *Journal of Monetary Economics* 50 (3), 525–546.
- Blanchflower, D. G., Kelly, R., 2008. Macroeconomic literacy, numeracy and the implications for monetary policy. Mimeo, Bank of England.
- Bloom, N., Floetotto, M., Jaimovich, N., 2008. Really uncertain business cycles. Mimeo, Stanford University.
- Blundell, R., Bond, S., Devereux, M., Schiantarelli, F., 1992. Investment and Tobin's Q: Evidence from company panel data. *Journal of Econometrics* 51, 233–257.
- Bomberger, W. A., 1996. Disagreement as a measure of uncertainty. *Journal of Money, Credit, and Banking* 28, 318–392.

PhD-Thesis written by: Jonas Dovern

- Bonham, C. S., Cohen, R. H., 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., Friz, R., Genre, V., Kenny, G., Meyler, A., Rautanen, T., 2007. The ECB survey of professional forecasters (SPF). Review after eight years' experience. Occasional Paper Series 59, European Central Bank.
- Branch, W. A., 2004. The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations. *The Economic Journal* 114, 592–621.
- Branch, W. A., Evans, G. W., 2006. Intrinsic heterogeneity in expectation formation. *Journal of Economic Theory* 127, 264–295.
- Brock, W. A., Hommes, C. H., 1997. A rational route to randomness. *Econometrica* 5, 1059–1095.
- Calvo, G., 1983. Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics* 12 (3), 383–398.
- Capistran, C., Timmermann, A., October 2006. Disagreement and biases in inflation expectations, mimeo.
- Carlson, J. A., Parkin, M., 1975. Inflation expectations. *Economica* 42 (166), 123–137.
- Carroll, C. D., 1997. Buffer-stock saving and the life cycle/permanent income hypothesis. *Quarterly Journal of Economics* 112 (1), 1–55.
- Carroll, C. D., 2003. Macroeconomic expectations of households and professional forecasters. *Quarterly Journal of Economics* 118 (1), 269–298.
- Carroll, C. D., Slacalek, J., 2008. International evidence on sticky consumption growth. NBER Working Paper Series 13876, National Bureau of Economic Research.
- Christiano, L. J., Fitzgerald, T. J., 2003. The band pass filter. *International Economic Review* 44 (2), 435–466.
- Christoffel, K., Linzert, T., 2005. The role of real wage rigidity and labor market frictions for unemployment and inflation dynamics. Working Paper Series 556, European Central Bank.

- Clemens, M. P., 1995. Evaluating the rationality of fixed-event forecasts. *Journal of Forecasting* 16, 225–239.
- Clemens, M. P., Joutz, F., Stekler, H. O., 2007. An evaluation of the forecasts of the Federal Reserve: A pooled approach. *Journal of Applied Econometrics* 22 (1), 121–136.
- Cogley, T., Sargent, T. J., 2001. Evolving post-world war ii u.s. inflation dynamics. In: Bernanke, B. S., Rogoff, K. (Eds.), *NBER Macroeconomics Annual*. Vol. 16. pp. 331–388.
- Croushore, D., 1993. Introducing: The Survey of Professional Forecasters. *Business review* Nov/Dec, Federal Reserve Bank of Philadelphia.
- Croushore, D., Stark, T., 2001. A real-time data set for macroeconomists. *Journal of Econometrics* 105, 111–130.
- Cukierman, A., 1992. *Central Bank Strategy, Credibility, and Independence*. MIT Press.
- Curtin, R. T., 1996. Procedure to estimate price expectations. Mimeo, University of Michigan Survey Research Center.
- Davies, A., Lahiri, K., 1995. A new framework for analysing survey forecasts using three-dimensional panel data. *Journal of Econometrics* 68, 205–227.
- Djauhari, M. A., 2007. A measure of multivariate data concentration. *Journal of Applied Probability and Statistics* 2, 139–157.
- Döpke, J., Dovern, J., Fritsche, U., Slacalek, J., 2008a. The dynamics of European inflation expectations. *The B.E. Journal of Macroeconomics* 8 (1).
- Döpke, J., Dovern, J., Fritsche, U., Slacalek, J., 2008b. Sticky information phillips curves: European evidence. *Journal of Money, Credit, and Banking* 40 (7), 1513–1519.
- Döpke, J., Fritsche, U., 2006. When do forecasters disagree? An assessment of German growth and inflation forecast dispersion. *International Journal of Forecasting* 22, 125–135.
- Dovern, J., 2009a. Are common models for idiosyncratic errors of fixed event forecasts backed by the data? Mimeo, Kiel Economics.
- Dovern, J., 2009b. Disagreement in the Consensus data set. A characterization. Mimeo, Kiel Economics.

PhD-Thesis written by: Jonas Dovern

- Dovern, J., Fritsche, U., 2008. Estimating fundamental cross-section dispersion from fixed event forecasts. Discussion Papers 787, DIW.
- Dovern, J., Fritsche, U., Slacalek, J., 2009. Disagreement among forecasters in G7 countries. Mimeo, European Central Bank.
- Dovern, J., Weisser, J., 2008. Are they really rational? Assessing professional macro-economic forecasts from the G7-countries. Kiel Working Papers 1447, The Kiel Institute for the World Economy.
- Durbin, J., Koopman, S. J., 2001. Time Series Analysis by State Space Methods. Oxford University Press, Oxford.
- ECB, 2005. Inflation persistence and price setting behaviour in the Euro area. report, European Central Bank, available at <http://www.ecb.int/home/pdf/research/inflationpersistencepricesettingreport.pdf>.
- Edwards, W., 1968. Conservatism in human information processing. In: Kleinmuntz, B. (Ed.), Formal Representations of Human Judgements. Wiley, New York.
- Ehrbeck, T., Waldmann, R. J., 1996. Why are professional forecasters biased? Agency versus behavioral explanations. Quarterly Journal of Economics 111 (1), 21–40.
- Ehrmann, M., Fratzscher, M., 2005. Communication and decision-making by central bank committees: Different strategies, same effectiveness? Working Paper Series 488, European Central Bank.
- Ehrmann, M., Fratzscher, M., Gürkaynak, R. S., Swanson, E. T., 2007. Convergence and anchoring of yield curves in the euro area. Working Paper 817, European Central Bank.
- Faust, J., Wright, J. H., 2007. Comparing greenbook forecasts and reduced form forecasts using a large realtime dataset. NBER Working Paper Series 13397, National Bureau of Economic Research.
- Figlewski, S., Wachtel, P., 1981. The formation of inflationary expectations. Review of Economics and Statistics 63, 1–10.
- Fisher, R. A., 1925. Statistical Methods for Research Workers. Oliver and Boyd, Edinburgh.
- Friedman, M., 1977. Nobel lecture: Inflation and unemployment. The Journal of Political Economy 85, 451–472.

- Fuller, D., Geide-Stevenson, D., 2003. Consensus among economists: Revisited. *The Journal of Economic Education* 34, 369–387.
- Gali, J., 2008. The New Keynesian approach to monetary policy analysis: Lessons and new directions. Mimeo, Universitat Pompeu Fabra.
- Gali, J., Gertler, M., 1999. Inflation dynamics: A structural econometric approach. *Journal of Monetary Economics* 4 (2), 195–222.
- Gallo, G. M., Granger, C. W., Jeon, Y., 2002. Copycats and common swings: The impact of the use of forecasts in information sets. *IMF Staff Papers* 49, 4–21.
- Gerberding, C., 2001. The information content of survey data on expected price developments for monetary policy. Discussion Paper 9/2001, Deutsche Bundesbank.
- Giordani, P., Söderlind, P., 2003. Inflation forecast uncertainty. *European Economic Review* 47, 1037–1059.
- Granger, C. W. J., Ramanathan, R., 1984. Improved methods of combining forecasts. *Journal of Forecasting* 3, 197–204.
- Grilli, V., Masciandaro, D., Tabellini, G., 1991. Political and monetary institutions and public financial policies in the industrial countries. *Economic Policy* 13, 341–392.
- Gürkaynak, R. S., Levin, A. T., Swanson, E. T., 2006. Does inflation targeting anchor long-run inflation expectations? evidence from long-term bond yields in the us, uk and sweden. Working Paper 9, Federal Reserve Bank of San Francisco.
- Harvey, A. C., 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge.
- Harvey, D. I., Leybourne, S. J., Newbold, P., 2001. Analysis of a panel of UK macroeconomic forecasts. *Econometrics Journal* 4, 37–55.
- Hendry, D. F., Clements, M. P., 2004. Pooling of forecasts. *The Econometrics Journal* 7 (1), 1–31.
- Heppke-Falk, K., Hüfner, F., 2004. Expected budget deficits and interest rate swap spreads - evidence for France, Germany and Italy. Discussion Paper Series 1 40/2004, Deutsche Bundesbank.

PhD-Thesis written by: Jonas Dovern

- Iacoviello, M., 2005. House prices, borrowing constraints and monetary policy in the business cycle. *American Economic Review* 95 (3), 739–764.
- Isiklar, G., 2005. On aggregation bias in fixed-event forecast efficiency tests. *Economic Letters* 89 (3), 312–316.
- Isiklar, G., Lahiri, K., Loungani, P., 2006. How quickly do forecasters incorporate news? Evidence from cross-country surveys. *Journal of Applied Econometrics* 21, 703–725.
- Kahneman, D., 2003. Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review* 93 (5), 1449–1475.
- Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47 (2), 263–292.
- Keane, M. P., Runkle, D. E., 1990. Testing the rationality of price forecasts: New evidence from panel data. *American Economic Review* 80, 714–735.
- Keynes, J. M., 1936. *The General Theory of Employment, Interest and Money*. Macmillan, London.
- Khan, H., Zhu, Z., 2002. Estimates of the sticky information Phillips curve for the United States, Canada, and the United Kingdom. Working Paper 19, Bank of Canada.
- Khan, H., Zhu, Z., 2006. Estimates of the sticky-information Phillips curve for the United States. *Journal of Money, Credit, and Banking* 38 (1), 195–207.
- Kiley, M. T., 2007. A quantitative comparison of sticky-price and sticky-information models of price setting. *Journal of Money, Credit, and Banking* 39 (1), 101–125.
- Kim, D. H., Orphanides, A., 2005. Term structure estimation with survey data on interest rate forecasts. FEDS Discussion Paper 48, Federal Reserve Board.
- Kirchgässner, G., 1999. On the political economy of economic policy advice. In: Mohr, E. (Ed.), *The Transfer of Economic Knowledge*. Edward Elgar, Cheltenham.
- Korenok, O., 2008. Empirical comparison of sticky price and sticky information models. *Journal of Macroeconomics* 30 (3), 906–927.

- Krause, M. U., Lubik, T. A., 2007. The (ir)relevance of real wage rigidity in the New Keynesian model with search frictions. *Journal of Monetary Economics* 54 (3), 706–727.
- Krusell, P., Smith, A. A., 1998. Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy* 106 (5), 867–896.
- Krusell, P., Smith, A. A., 1999. On the welfare effects of eliminating business cycles. *Review of Economic Dynamics* 2, 245–272.
- Lamont, O. A., 2002. Macroeconomic forecasts and microeconomic forecasters. *Journal of Economic Behavior & Organization* 48, 265–280.
- Laster, D., Bennett, P., Geom, I. S., 1999. Rational bias in macroeconomic forecasts. *The Quarterly Journal of Economics*, 293–318.
- Lechthaler, W., Merkl, C., Snower, D., 2008. Monetary persistence and the labor market: A new perspective. Kiel Working Paper 1409, Kiel Institute for the World Economy.
- Levin, A. T., Natalucci, F. M., Piger, J. M., 2004. Explicit inflation objectives and macroeconomic outcomes. Working Paper 383, European Central Bank.
- Loungani, P., 2001. How accurate are private sector forecasts? Cross-country evidence from consensus forecasts of output growth. *International Journal of Forecasting* 17, 419–432.
- Lucas, R. J. E., 1973. Some international evidence on inflation-output trade-offs. *American Economic Review* 63, 326–334.
- Mankiw, N. G., Reis, R., 2002. Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve. *Quarterly Journal of Economics* 117, 1295–1328.
- Mankiw, N. G., Reis, R., 2003. Sticky information: A model of monetary non-neutrality and structural slumps. In: Aghion, P. (Ed.), *Knowledge, Information, and Expectation in Modern Macroeconomics*. MIT Press, Cambridge, MA, pp. 64–86.
- Mankiw, N. G., Reis, R., 2006. Pervasive stickiness. *American Economic Review, Papers and Proceedings* 96 (2), 164–169.
- Mankiw, N. G., Reis, R., Wolfers, J., 2003. Disagreement on inflation expectations. *NBER Macroeconomics Annual*, 209–248.

- Mincer, J., Zarnowitz, V., 1969. The evaluation of economic forecasts. In: Mincer, J. (Ed.), *Economic Forecasts and Expectations*. National Bureau of Economic Research, New York.
- Morris, S., Shin, H. S., 2005a. Central bank transparency and the signal value of prices. *Brookings Papers on Economic Activity* 2, 1–66.
- Morris, S., Shin, H. S., 2005b. The social value of public information. *American Economic Review* 92, 1521–1534.
- Muth, J. F., 1961. Rational expectations and the theory of price movements. *Econometrica* 29, 315–335.
- Nordhaus, W. D., 1987. Forecasting efficiency concepts and applications. *Review of Economics and Statistics* 69, 667–674.
- O’Reilly, G., Whelan, K., 2005. Has Euro-area inflation persistence changed over time? *Review of Economics and Statistics* 87 (4), 709–720.
- Paloviita, M., 2005. Comparing alternative Phillips curve specifications: European results with survey-based expectations. Discussion Paper 22, Bank of Finland.
- Patton, A., Timmermann, A., 2007. How quickly is macroeconomic uncertainty resolved? Theory and empirical evidence from the term structure of forecast errors. Mimeo, LSE and UCSD.
- Patton, A., Timmermann, A., 2008a. The resolution of macroeconomic uncertainty: Evidence from survey forecasts. Mimeo.
- Patton, A., Timmermann, A., 2008b. Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion. Mimeo.
- Pesaran, M. H., Weale, M., 2006. Survey expectations. In: Elliott, G., Granger, C. W., Timmermann, A. (Eds.), *Handbook of Economic Forecasting Volume 1*. North-Holland, Amsterdam.
- Phillips, L. D., Edwards, W., 1966. Conservatism in a simple probability inference task. *Journal of Experimental Psychology* 72, 346–357.
- Piazzesi, M., Schneider, M., 2008. Bond positions, expectations, and the yield curve. working paper 2, Federal Reserve Bank of Atlanta.
- Pigou, A. C., 1927. *Industrial Fluctuations*. Macmillan, London.

- Reis, R., 2006a. Inattentive consumers. *Journal of Monetary Economics* 53 (8), 1761–1800.
- Reis, R., 2006b. Inattentive producers. *The Review of Economic Studies* 73 (3), 793–821.
- Rich, R., Tracy, J., 2003. Modelling uncertainty: Predictive accuracy as a proxy for predictive confidence. Federal Reserve Bank of New York Staff Reports 161, Federal Reserve Bank of New York, New York.
- Rogoff, K., 1985. The optimal degree of commitment to an intermediate monetary target. *Quarterly Journal of Economics* 100, 1169–1190.
- Rotemberg, J. J., 1982. Sticky prices in the United States. *Journal of Political Economy* 90 (6), 1187–1211.
- Scheinkman, J., Xiong, W., 2003. Heterogeneous beliefs, speculation and trading in financial markets. In: *Paris–Princeton Lectures on Mathematical Finance*. Vol. 1847. Springer.
- Sims, C. A., 1980. Macroeconomics and reality. *Econometrica* 1, 1–48.
- Sims, C. A., 2003. Implications of rational inattention. *Journal of Monetary Economics* 50 (3), 665–690.
- Smant, D. J. C., 2002. Has the European Central Bank followed a Bundesbank policy? Evidence from the early years. *Kredit und Kapital* 35, 327–343.
- Souleles, N. S., 2004. Expectations, heterogeneous forecast errors, and consumption: Micro evidence from the michigan consumer sentiment surveys. *Journal of Money, Credit, and Banking* 36 (1), 39–72.
- Stege, C., 1989. Zur Beurteilung der Prognosefähigkeit der "führenden Wirtschaftswissenschaftlichen Forschungsinstitute" in der Bundesrepublik Deutschland. Verlag Dr. Kovac.
- Stock, J. H., Watson, M. W., 2005. Understanding changes in international business cycle dynamics. *Journal of the European Economic Association* 3, 968–1006.
- Timmermann, A., 2007. An evaluation of the World Economic Outlook Forecasts. *IMF Staff Papers* 54 (1), 1–33.

- Townsend, R. M., 1983. Forecasting the forecasts of others. *The Journal of Political Economy* 91, 546–588.
- Trabandt, M., 2007. Sticky information vs. sticky prices: A horse race in a DSGE framework. Working Paper 209, Sveriges Riksbank.
- US-Census-Bureau, 2007. The X-12-ARIMA seasonal adjustment program. Retrieved from the World Wide Web at July 26, 2007.
URL <http://www.census.gov/srd/www/x12a/index.html>
- Walsh, C. E., October 2005. Labor market search, sticky prices, and interest rate policies. *Review of Economic Dynamics* 8 (4), 829–849.
- Woodford, M., 2003. *Interest and prices: Foundations of a theory of monetary policy*. Princeton University Press, Princeton.
- Woodford, M., 2005. Central bank communication and policy effectiveness. NBER Working Paper Series 11898, National Bureau of Economic Research.
- Zarnowitz, V., Braun, P. A., 1993. Twenty-two years of the NBER-ASA quarterly economic outlook surveys: Aspects and comparisons of forecasting performance. In: Stock, J. H., Watson, M. M. (Eds.), *Business Cycles, Indicators and Forecasting*. University of Chicago Press, Chicago.
- Zarnowitz, V., Lambros, L. A., 1987. Consensus and uncertainty in economic prediction. *Journal of Political Economy* 95, 591–621.

Appendix

A Carlson-Parking-Method for Quantifying Survey Data

In section 3.1, we have to re-scale the balance statistics from the qualitative survey on households' inflation expectations to obtain a measure of inflation expectations. The standard method follows Carlson and Parkin (1975) and its extensions. In our case, the observed data are from a pentachotomous survey, i.e., they classify the responses into five subgroups:

Consumer prices will:

- Increase more rapidly,
- Increase at the same rate,
- Increase at a slower rate,
- Stay about the same,
- Fall.

Batchelor and Orr (1988) derive how responses from a pentachotomous survey can be transformed into a measure of inflation expectations: ${}_t\mu_{t+1} = \tilde{\mu}_t \times f({}_tA_{t+1}, {}_tB_{t+1}, \dots, {}_tE_{t+1})$, where ${}_tA_{t+1}, \dots, {}_tE_{t+1}$ are the fractions of respondents answering each option and f is a known function (see Batchelor and Orr, 1988, p. 322, formula (11)) and $\tilde{\mu}_t$ is the current perceived inflation rate that has to be specified.

B Detailed Results about Determinants of Disagreement

Table B-1: Disagreement and Business Cycle – Panel Results for Inflation

Model	β_0	β_1	β_2	β_3	β_4	β_5	\bar{R}^2
Panel A: Disagreement over Time $\text{disagr}_t = \beta_0 + \beta_1 \times \text{rec}_t + \beta_2 \times \text{post-1998}_t + \varepsilon_t$							
1.	0.299*** (0.003)						-0.004
2.	0.287*** (0.002)	0.064*** (0.011)					0.034
3.	0.327*** (0.013)	0.050*** (0.014)	-0.082*** (0.023)				0.133
Panel B: Disagreement and Macro Variables $\text{disagr}_t = \beta_0 + \beta_2 \times \text{INFL}_t + \beta_3 \times \Delta_{12}\text{INFL}_t^2 + \beta_4 \times \text{output gap}_t + \beta_5 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$							
4.	0.237*** (0.007)		0.026*** (0.003)				0.081
5.	0.237*** (0.008)		0.024*** (0.003)	0.004*** (0.000)			0.082
6.	0.232*** (0.006)		0.023*** (0.004)	0.003*** (0.001)	-0.015 (0.011)		0.077
7.	0.232*** (0.007)		0.021*** (0.003)	0.003*** (0.001)	-0.014*** (0.002)	0.056*** (0.010)	0.095
Panel C: Disagreement and Central Bank Independence $\text{disagr}_t = \beta_0 + \beta_1 \times \text{CB Independence}_t + \beta_2 \times \text{INFL}_t + \beta_3 \times \Delta_{12}\text{INFL}_t^2 + \beta_4 \times \text{output gap}_t + \beta_5 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$							
8.	0.420*** (0.007)	-0.152*** (0.009)					0.173
9.	0.388*** (0.025)	-0.147*** (0.011)	0.007 (0.008)	0.002*** (0.001)	-0.017*** (0.004)	0.032 (0.026)	0.229

Notes: Fixed effects estimators, clustered standard errors (groups of countries: (i) France, Germany, Italy, UK, (ii) Canada, US, (iii) Japan). β_0 denotes the average of country-specific intercepts. “post-1998_t” denotes a dummy variable which equals 0 before 1999 and 1 after 1998. “recession_t” denotes a dummy variable which equals 1 during recession set by the Economic Cycle Research Institute (ECRI) and 0 otherwise. “output gap_t” denotes the ex-post output gap estimated in the OECD Economic Outlook quarterly output gap revisions database (in August 2008). $\Delta_{12}\text{INFL}_t^2 \equiv (\text{INFL}_t - \text{INFL}_{t-12})^2$. “MP Independence_t” denotes a 0–1 indicator of independent monetary policy defined in table 4.9.

Table B-2: Disagreement and Business Cycle – Panel Results for the Interest Rate

Model	β_0	β_1	β_2	β_3	β_4	β_5	\bar{R}^2
Panel A: Disagreement over Time							
$\text{disagr}_t = \beta_0 + \beta_1 \times \text{rec}_t + \beta_2 \times \text{post-1998}_t + \varepsilon_t$							
1.	0.549*** (0.007)						-0.004
2.	0.519*** (0.016)	0.161* (0.083)					0.057
3.	0.624*** (0.022)	0.124* (0.070)	-0.214*** (0.024)				0.236
Panel B: Disagreement and Macro Variables							
$\text{disagr}_t = \beta_0 + \beta_2 \times R3M_t + \beta_3 \times \Delta_{12}R3M_t^2 + \beta_4 \times \text{output gap}_t + \beta_5 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$							
4.	0.319*** (0.013)		0.046*** (0.003)				0.279
5.	0.321*** (0.006)		0.042*** (0.002)	0.006*** (0.001)			0.274
6.	0.330*** (0.032)		0.038*** (0.007)	0.006** (0.003)	-0.018 (0.024)		0.225
7.	0.333*** (0.012)		0.036*** (0.002)	0.006*** (0.001)	-0.018*** (0.004)	0.054*** (0.017)	0.227
Panel C: Disagreement and Central Bank Independence							
$\text{disagr}_t = \beta_0 + \beta_1 \times \text{CB Independence}_t + \beta_2 \times R3M_t + \beta_3 \times \Delta_{12}R3M_t^2 + \beta_4 \times \text{output gap}_t + \beta_5 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$							
8.	0.783*** (0.013)	-0.294*** (0.007)					0.169
9.	0.540*** (0.088)	-0.180*** (0.059)	0.024*** (0.009)	0.005*** (0.001)	-0.022 (0.014)	0.042 (0.029)	0.278

Notes: Fixed effects estimators, clustered standard errors (groups of countries: (i) France, Germany, Italy, UK, (ii) Canada, US, (iii) Japan). β_0 denotes the average of country-specific intercepts. “post-1998_t” denotes a dummy variable which equals 0 before 1999 and 1 after 1998. “recession_t” denotes a dummy variable which equals 1 during recession set by the Economic Cycle Research Institute (ECRI) and 0 otherwise. “output gap_t” denotes the ex-post output gap estimated in the OECD Economic Outlook quarterly output gap revisions database (in August 2008). $\Delta_{12}R3M_t^2 \equiv (R3M_t - R3M_{t-12})^2$. “MP Independence_t” denotes a 0–1 indicator of independent monetary policy defined in table 4.9.

Table B-3: Disagreement and Business Cycle – Panel Results for GDP

Model	β_0	β_1	β_2	β_3	β_4	β_5	\bar{R}^2
Panel A: Disagreement over Time							
$\text{disagr}_t = \beta_0 + \beta_1 \times \text{rec}_t + \beta_2 \times \text{post-1998}_t + \varepsilon_t$							
1.	0.410*** (0.004)						-0.004
2.	0.378*** (0.007)	0.165*** (0.038)					0.130
3.	0.394*** (0.013)	0.160*** (0.041)	-0.032** (0.015)				0.138
Panel B: Disagreement and Macro Variables							
$\text{disagr}_t = \beta_0 + \beta_2 \times GDP_t + \beta_3 \times \Delta_{12}GDP_t^2 + \beta_4 \times \text{output gap}_t + \beta_5 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$							
4.	0.482*** (0.010)		-0.035*** (0.005)				0.111
5.	0.463*** (0.027)		-0.033*** (0.006)	0.004 (0.005)			0.117
6.	0.453*** (0.024)		-0.030*** (0.004)	0.004 (0.004)	-0.002 (0.003)		0.099
7.	0.447*** (0.010)		-0.029*** (0.003)	0.004*** (0.001)	-0.002 (0.003)	0.047*** (0.013)	0.103
Panel C: Disagreement and Central Bank Independence							
$\text{disagr}_t = \beta_0 + \beta_1 \times \text{CB Independence}_t + \beta_2 \times GDP_t + \beta_3 \times \Delta_{12}GDP_t^2 + \beta_4 \times \text{output gap}_t + \beta_5 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$							
8.	0.458*** (0.010)	-0.061 (0.054)					0.015
9.	0.494*** (0.004)	-0.063*** (0.022)	-0.026*** (0.005)	0.004 (0.004)	-0.005 (0.003)	0.036*** (0.013)	0.120

Notes: Fixed effects estimators, clustered standard errors (groups of countries: (i) France, Germany, Italy, UK, (ii) Canada, US, (iii) Japan). β_0 denotes the average of country-specific intercepts. “post-1998_t” denotes a dummy variable which equals 0 before 1999 and 1 after 1998. “recession_t” denotes a dummy variable which equals 1 during recession set by the Economic Cycle Research Institute (ECRI) and 0 otherwise. “output gap_t” denotes the ex-post output gap estimated in the OECD Economic Outlook quarterly output gap revisions database (in August 2008). $\Delta_{12}GDP_t^2 \equiv (GDP_t - GDP_{t-12})^2$. “MP Independence_t” denotes a 0–1 indicator of independent monetary policy defined in table 4.9.

Table B-4: Disagreement and Business Cycle – Panel Results for Consumption

Model	β_0	β_1	β_2	β_3	β_4	β_5	\bar{R}^2
Panel A: Disagreement over Time							
$\text{disagr}_t = \beta_0 + \beta_1 \times \text{rec}_t + \beta_2 \times \text{post-1998}_t + \varepsilon_t$							
1.	0.445*** (0.004)						-0.004
2.	0.414*** (0.008)	0.164*** (0.040)					0.131
3.	0.419*** (0.008)	0.161*** (0.040)	-0.012 (0.011)				0.132
Panel B: Disagreement and Macro Variables							
$\text{disagr}_t = \beta_0 + \beta_2 \times \text{CONS}_t + \beta_3 \times \Delta_{12}\text{CONS}_t^2 + \beta_4 \times \text{output gap}_t + \beta_5 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$							
4.	0.512*** (0.018)		-0.031*** (0.008)				0.077
5.	0.491*** (0.018)		-0.027*** (0.007)	0.004 (0.003)			0.092
6.	0.477*** (0.005)		-0.021*** (0.003)	0.004 (0.003)	-0.001 (0.003)		0.076
7.	0.472*** (0.010)		-0.020*** (0.003)	0.004*** (0.001)	-0.001 (0.003)	0.052*** (0.013)	0.082
Panel C: Disagreement and Central Bank Independence							
$\text{disagr}_t = \beta_0 + \beta_1 \times \text{CB Independence}_t + \beta_2 \times \text{CONS}_t + \beta_3 \times \Delta_{12}\text{CONS}_t^2 + \beta_4 \times \text{output gap}_t + \beta_5 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$							
8.	0.469*** (0.018)	-0.031 (0.032)					0.004
9.	0.492*** (0.010)	-0.027*** (0.009)	-0.019*** (0.004)	0.004 (0.003)	-0.002 (0.005)	0.047*** (0.005)	0.088

Notes: Fixed effects estimators, clustered standard errors (groups of countries: (i) France, Germany, Italy, UK, (ii) Canada, US, (iii) Japan). β_0 denotes the average of country-specific intercepts. “post-1998_t” denotes a dummy variable which equals 0 before 1999 and 1 after 1998. “recession_t” denotes a dummy variable which equals 1 during recession set by the Economic Cycle Research Institute (ECRI) and 0 otherwise. “output gap_t” denotes the ex-post output gap estimated in the OECD Economic Outlook quarterly output gap revisions database (in August 2008). $\Delta_{12}\text{CONS}_t^2 \equiv (\text{CONS}_t - \text{CONS}_{t-12})^2$. “MP Independence_t” denotes a 0–1 indicator of independent monetary policy defined in table 4.9.

Table B-5: Disagreement and Business Cycle – Panel Results for Investment

Model	β_0	β_1	β_2	β_3	β_4	β_5	\bar{R}^2
Panel A: Disagreement over Time $\text{disagr}_t = \beta_0 + \beta_1 \times \text{rec}_t + \beta_2 \times \text{post-1998}_t + \varepsilon_t$							
1.	1.764*** (0.017)						-0.004
2.	1.701*** (0.004)	0.329*** (0.024)					0.035
3.	1.702*** (0.045)	0.329*** (0.010)	-0.002 (0.101)				0.034
Panel B: Disagreement and Macro Variables $\text{disagr}_t = \beta_0 + \beta_2 \times \text{INV}_t + \beta_3 \times \Delta_{12}\text{INV}_t^2 + \beta_4 \times \text{output gap}_t + \beta_5 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$							
4.	1.784*** (0.023)		-0.025*** (0.009)				0.032
5.	1.700*** (0.038)		-0.021** (0.010)	0.002*** (0.001)			0.059
6.	1.704*** (0.055)		-0.019 (0.014)	0.002** (0.001)	-0.003 (0.048)		0.048
7.	1.697*** (0.025)		-0.018*** (0.004)	0.002*** (0.000)	-0.002 (0.013)	0.092 (0.075)	0.044
Panel C: Disagreement and Central Bank Independence $\text{disagr}_t = \beta_0 + \beta_1 \times \text{CB Independence}_t + \beta_2 \times \text{INV}_t + \beta_3 \times \Delta_{12}\text{INV}_t^2 + \beta_4 \times \text{output gap}_t + \beta_5 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$							
8.	1.833*** (0.023)	-0.086 (0.086)					0.001
9.	1.720*** (0.111)	-0.028 (0.093)	-0.018 (0.013)	0.002** (0.001)	-0.003 (0.046)	0.088** (0.035)	0.048

Notes: Fixed effects estimators, clustered standard errors (groups of countries: (i) France, Germany, Italy, UK, (ii) Canada, US, (iii) Japan). β_0 denotes the average of country-specific intercepts. “post-1998_t” denotes a dummy variable which equals 0 before 1999 and 1 after 1998. “recession_t” denotes a dummy variable which equals 1 during recession set by the Economic Cycle Research Institute (ECRI) and 0 otherwise. “output gap_t” denotes the ex-post output gap estimated in the OECD Economic Outlook quarterly output gap revisions database (in August 2008). $\Delta_{12}\text{INV}_t^2 \equiv (\text{INV}_t - \text{INV}_{t-12})^2$. “MP Independence_t” denotes a 0–1 indicator of independent monetary policy defined in table 4.9.

Table B-6: Disagreement and Business Cycle – Panel Results for Unemployment

Model	β_0	β_1	β_2	β_3	β_4	β_5	\bar{R}^2
Panel A: Disagreement over Time $\text{disagr}_t = \beta_0 + \beta_1 \times \text{rec}_t + \beta_2 \times \text{post-1998}_t + \varepsilon_t$							
1.	0.293*** (0.005)						-0.004
2.	0.261*** (0.011)	0.169*** (0.059)					0.129
3.	0.297*** (0.004)	0.156*** (0.049)	-0.074** (0.030)				0.173
Panel B: Disagreement and Macro Variables $\text{disagr}_t = \beta_0 + \beta_2 \times UN_t + \beta_3 \times \Delta_{12}UN_t^2 + \beta_4 \times \text{output gap}_t + \beta_5 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$							
4.	0.299*** (0.098)		-0.001 (0.013)				-0.001
5.	0.324*** (0.090)		-0.009 (0.013)	0.054*** (0.012)			0.085
6.	0.413*** (0.031)		-0.022*** (0.005)	0.055*** (0.013)	-0.020*** (0.004)		0.109
7.	0.414*** (0.028)		-0.022*** (0.004)	0.055*** (0.005)	-0.020*** (0.004)	0.020 (0.014)	0.106
Panel C: Disagreement and Central Bank Independence $\text{disagr}_t = \beta_0 + \beta_1 \times \text{CB Independence}_t + \beta_2 \times UN_t + \beta_3 \times \Delta_{12}UN_t^2 + \beta_4 \times \text{output gap}_t + \beta_5 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$							
8.	0.344*** (0.098)	-0.064 (0.064)					0.016
9.	0.535*** (0.106)	-0.081* (0.049)	-0.029*** (0.010)	0.049*** (0.009)	-0.028*** (0.002)	0.006 (0.023)	0.128

Notes: Fixed effects estimators, clustered standard errors (groups of countries: (i) France, Germany, Italy, UK, (ii) Canada, US, (iii) Japan). β_0 denotes the average of country-specific intercepts. “post-1998_t” denotes a dummy variable which equals 0 before 1999 and 1 after 1998. “recession_t” denotes a dummy variable which equals 1 during recession set by the Economic Cycle Research Institute (ECRI) and 0 otherwise. “output gap_t” denotes the ex-post output gap estimated in the OECD Economic Outlook quarterly output gap revisions database (in August 2008). $\Delta_{12}UN_t^2 \equiv (UN_t - UN_{t-12})^2$. “MP Independence_t” denotes a 0–1 indicator of independent monetary policy defined in table 4.9.

Table B-7: Determinants of Disagreement – Country-by-Country Results for Inflation

Country	β_0	β_1	β_2	β_3	β_4	\bar{R}^2
Panel A: Disagreement over Time disagreement _t = $\beta_0 + \beta_1 \times \text{recession}_t + \beta_2 \times \text{post-1998}_t + \varepsilon_t$						
CN	0.353*** (0.012)	0.073*** (0.026)	-0.054*** (0.017)			0.112
FR	0.212*** (0.008)	0.007 (0.015)	-0.006 (0.011)			-0.007
GE	0.224*** (0.009)	0.005 (0.011)	0.041*** (0.011)			0.056
IT	0.310*** (0.012)	0.096*** (0.027)	-0.137*** (0.016)			0.349
JP	0.376*** (0.012)	0.055*** (0.016)	-0.145*** (0.016)			0.321
UK	0.503*** (0.017)	-0.005 (0.037)	-0.268*** (0.023)			0.414
US	0.324*** (0.009)	0.105*** (0.023)	-0.005 (0.013)			0.088
Panel B: Disagreement and Macro Variables disagreement _t = $\beta_0 + \beta_1 \times INFL_t + \beta_2 \times \Delta_{12}INFL_t^2 + \beta_3 \times \text{output gap}_t + \beta_4 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$						
CN	0.191*** (0.043)	0.024*** (0.006)	0.002 (0.002)	-0.009** (0.004)	0.066*** (0.020)	0.124
FR	0.118*** (0.027)	-0.008 (0.008)	-0.005 (0.009)	0.004 (0.005)	0.005 (0.049)	0.089
GE	0.225*** (0.027)	-0.007 (0.006)	0.001 (0.004)	-0.013*** (0.005)	0.027 (0.084)	0.036
IT	0.038 (0.040)	0.048*** (0.006)	0.002 (0.005)	-0.019*** (0.006)	0.077*** (0.022)	0.463
JP	0.136*** (0.043)	0.017 (0.011)	0.012** (0.006)	0.010 (0.006)	0.014 (0.014)	0.337
UK	-0.046 (0.057)	-0.002 (0.010)	0.002 (0.002)	-0.138*** (0.011)	0.249*** (0.062)	0.512
US	0.199*** (0.034)	0.010 (0.008)	0.002 (0.003)	-0.014*** (0.005)	0.108* (0.058)	0.078

Notes: Fixed effects estimators. β_0 denotes the average of country-specific intercepts. “post-1998_t” denotes a dummy variable which equals 0 before 1999 and 1 after 1998. “recession_t” denotes a dummy variable which equals 1 during recession set by the Economic Cycle Research Institute (ECRI) and 0 otherwise. “output gap_t” denotes the ex-post output gap estimated in the OECD Economic Outlook quarterly output gap revisions database (in August 2008). $\Delta_{12}INFL_t^2 \equiv (INFL_t - INFL_{t-12})^2$.

Table B-8: Determinants of Disagreement – Country-by-Country Results for the Interest Rate

Country	β_0	β_1	β_2	β_3	β_4	R^2
Panel A: Disagreement over Time						
$\text{disagreement}_t = \beta_0 + \beta_1 \times \text{recession}_t + \beta_2 \times \text{post-1998}_t + \varepsilon_t$						
CN	0.885*** (0.032)	0.190*** (0.067)	-0.328*** (0.044)			0.295
FR	0.514*** (0.018)	0.231*** (0.037)	-0.101*** (0.026)			0.221
GE	0.477*** (0.018)	0.110*** (0.024)	-0.126*** (0.023)			0.191
IT	0.720*** (0.024)	0.149*** (0.055)	-0.341*** (0.034)			0.405
JP	0.417*** (0.015)	-0.052*** (0.019)	-0.267*** (0.019)			0.513
UK	0.827*** (0.024)	0.208*** (0.052)	-0.299*** (0.033)			0.399
US	0.549*** (0.017)	0.124*** (0.043)	0.008 (0.024)			0.031
Panel B: Disagreement and Macro Variables						
$\text{disagreement}_t = \beta_0 + \beta_1 \times R3M_t + \beta_2 \times \Delta_{12}R3M_t^2 + \beta_3 \times \text{output gap}_t + \beta_4 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$						
CN	-0.030 (0.087)	0.024** (0.012)	-0.003 (0.004)	-0.048*** (0.010)	0.055 (0.046)	0.356
FR	0.221*** (0.051)	0.039*** (0.005)	-0.003 (0.003)	-0.054*** (0.012)	0.365*** (0.114)	0.278
GE	0.045 (0.047)	0.044*** (0.006)	0.010 (0.006)	-0.031*** (0.010)	-0.052 (0.160)	0.323
IT	0.089 (0.070)	0.048*** (0.005)	-0.001 (0.002)	-0.047*** (0.013)	0.020 (0.049)	0.392
JP	0.044 (0.044)	-0.005 (0.007)	0.003 (0.005)	0.071*** (0.007)	0.021 (0.015)	0.414
UK	0.138** (0.061)	0.022*** (0.007)	0.018*** (0.003)	-0.096*** (0.015)	0.090 (0.079)	0.495
US	0.452*** (0.039)	-0.020*** (0.007)	0.012*** (0.003)	-0.004 (0.009)	0.037 (0.106)	0.135

Notes: Fixed effects estimators. β_0 denotes the average of country-specific intercepts. “post-1998_t” denotes a dummy variable which equals 0 before 1999 and 1 after 1998. “recession_t” denotes a dummy variable which equals 1 during recession set by the Economic Cycle Research Institute (ECRI) and 0 otherwise. “output gap_t” denotes the ex-post output gap estimated in the OECD Economic Outlook quarterly output gap revisions database (in August 2008). $\Delta_{12}R3M_t^2 \equiv (R3M_t - R3M_{t-12})^2$.

Table B-9: Determinants of Disagreement – Country-by-Country Results for GDP

Country	β_0	β_1	β_2	β_3	β_4	R^2
Panel A: Disagreement over Time						
$\text{disagreement}_t = \beta_0 + \beta_1 \times \text{recession}_t + \beta_2 \times \text{post-1998}_t + \varepsilon_t$						
CN	0.492*** (0.017)	0.170*** (0.037)	-0.114*** (0.024)			0.240
FR	0.283*** (0.009)	0.027 (0.018)	0.036*** (0.013)			0.037
GE	0.340*** (0.014)	0.114*** (0.019)	-0.086*** (0.019)			0.214
IT	0.241*** (0.012)	0.120*** (0.028)	0.025 (0.017)			0.076
JP	0.589*** (0.027)	0.207*** (0.035)	0.029 (0.035)			0.141
UK	0.477*** (0.014)	0.189*** (0.031)	-0.138*** (0.019)			0.387
US	0.348*** (0.013)	0.255*** (0.034)	0.015 (0.019)			0.211
Panel B: Disagreement and Macro Variables						
$\text{disagreement}_t = \beta_0 + \beta_1 \times GDP_t + \beta_2 \times \Delta_{12}GDP_t^2 + \beta_3 \times \text{output gap}_t + \beta_4 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$						
CN	-0.002 (0.052)	-0.012* (0.007)	0.008*** (0.002)	-0.022*** (0.007)	0.020 (0.030)	0.376
FR	0.204*** (0.027)	-0.015** (0.006)	0.009*** (0.002)	0.025*** (0.006)	0.061 (0.060)	0.126
GE	0.145*** (0.048)	-0.016** (0.008)	0.008*** (0.002)	-0.016 (0.010)	0.196 (0.145)	0.164
IT	0.041 (0.039)	-0.035*** (0.007)	0.009*** (0.002)	0.014** (0.007)	0.002 (0.021)	0.163
JP	0.435*** (0.086)	-0.037*** (0.014)	-0.010*** (0.004)	-0.004 (0.011)	0.024 (0.031)	0.087
UK	0.236*** (0.047)	-0.030*** (0.008)	0.014*** (0.004)	-0.042*** (0.011)	0.173*** (0.058)	0.318
US	0.115** (0.047)	-0.051*** (0.006)	0.000 (0.002)	0.023*** (0.007)	0.474*** (0.077)	0.259

Notes: Fixed effects estimators. β_0 denotes the average of country-specific intercepts. “post-1998_t” denotes a dummy variable which equals 0 before 1999 and 1 after 1998. “recession_t” denotes a dummy variable which equals 1 during recession set by the Economic Cycle Research Institute (ECRI) and 0 otherwise. “output gap_t” denotes the ex-post output gap estimated in the OECD Economic Outlook quarterly output gap revisions database (in August 2008). $\Delta_{12}GDP_t^2 \equiv (GDP_t - GDP_{t-12})^2$.

Table B-10: Determinants of Disagreement – Country-by-Country Results for Consumption

Country	β_0	β_1	β_2	β_3	β_4	\bar{R}^2
Panel A: Disagreement over Time						
$\text{disagreement}_t = \beta_0 + \beta_1 \times \text{recession}_t + \beta_2 \times \text{post-1998}_t + \varepsilon_t$						
CN	0.478*** (0.019)	0.231*** (0.041)	-0.071*** (0.027)			0.208
FR	0.285*** (0.010)	0.046** (0.019)	0.031** (0.013)			0.038
GE	0.377*** (0.015)	0.088*** (0.020)	-0.013 (0.019)			0.079
IT	0.321*** (0.014)	0.198*** (0.031)	0.006 (0.019)			0.166
JP	0.561*** (0.025)	0.206*** (0.031)	0.016 (0.031)			0.170
UK	0.548*** (0.017)	0.176*** (0.036)	-0.043* (0.023)			0.146
US	0.364*** (0.012)	0.213*** (0.031)	0.003 (0.018)			0.182
Panel B: Disagreement and Macro Variables						
$\text{disagreement}_t = \beta_0 + \beta_1 \times \text{CONS}_t + \beta_2 \times \Delta_{12}\text{CONS}_t^2 +$ $+ \beta_3 \times \text{output gap}_t + \beta_4 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$						
CN	0.095 (0.069)	-0.026*** (0.010)	0.012*** (0.002)	-0.008 (0.007)	0.008 (0.030)	0.214
FR	0.189*** (0.034)	-0.003 (0.006)	-0.001 (0.002)	-0.008 (0.007)	0.029 (0.068)	0.053
GE	0.210*** (0.051)	-0.001 (0.009)	0.004 (0.003)	-0.007 (0.011)	0.023 (0.151)	0.072
IT	0.118** (0.049)	-0.021*** (0.006)	0.004*** (0.001)	0.024*** (0.008)	0.039 (0.026)	0.168
JP	0.288*** (0.084)	-0.022 (0.018)	-0.004 (0.003)	0.001 (0.011)	0.041 (0.029)	0.145
UK	0.294*** (0.057)	-0.028*** (0.007)	0.006* (0.003)	0.005 (0.013)	0.202*** (0.070)	0.127
US	0.140*** (0.047)	-0.033*** (0.009)	0.013*** (0.003)	0.021*** (0.008)	0.246*** (0.080)	0.183

Notes: Fixed effects estimators. β_0 denotes the average of country-specific intercepts. “post-1998_t” denotes a dummy variable which equals 0 before 1999 and 1 after 1998. “recession_t” denotes a dummy variable which equals 1 during recession set by the Economic Cycle Research Institute (ECRI) and 0 otherwise. “output gap_t” denotes the ex-post output gap estimated in the OECD Economic Outlook quarterly output gap revisions database (in August 2008). $\Delta_{12}\text{CONS}_t^2 \equiv (\text{CONS}_t - \text{CONS}_{t-12})^2$.

Table B-11: Determinants of Disagreement – Country-by-Country Results for Investment

Country	β_0	β_1	β_2	β_3	β_4	R^2
Panel A: Disagreement over Time						
$\text{disagreement}_t = \beta_0 + \beta_1 \times \text{recession}_t + \beta_2 \times \text{post-1998}_t + \varepsilon_t$						
CN	2.929*** (0.099)	-0.369* (0.209)	-0.619*** (0.137)			0.083
FR	1.170*** (0.041)	0.001 (0.081)	-0.004 (0.057)			-0.010
GE	1.284*** (0.049)	0.293*** (0.066)	0.374*** (0.063)			0.201
IT	1.027*** (0.051)	0.224* (0.118)	-0.085 (0.072)			0.024
JP	2.141*** (0.093)	0.294** (0.119)	0.346*** (0.117)			0.061
UK	1.641*** (0.051)	0.727*** (0.113)	-0.147** (0.072)			0.224
US	1.845*** (0.053)	1.021*** (0.134)	-0.003 (0.076)			0.217
Panel B: Disagreement and Macro Variables						
$\text{disagreement}_t = \beta_0 + \beta_1 \times \text{INV}_t + \beta_2 \times \Delta_{12}\text{INV}_t^2 + \beta_3 \times \text{output gap}_t + \beta_4 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$						
CN	2.027*** (0.349)	0.017 (0.014)	0.001 (0.001)	-0.069* (0.041)	0.081 (0.179)	0.093
FR	0.820*** (0.147)	-0.027*** (0.008)	0.003*** (0.001)	0.043* (0.025)	0.039 (0.260)	0.084
GE	0.938*** (0.182)	-0.029*** (0.009)	0.001 (0.001)	0.048* (0.027)	-0.083 (0.496)	0.135
IT	-0.242 (0.175)	-0.035*** (0.008)	0.003*** (0.001)	0.126*** (0.030)	0.007 (0.088)	0.246
JP	1.087*** (0.328)	0.036 (0.025)	-0.005* (0.003)	-0.179*** (0.048)	2.103 (6.029)	0.104
UK	1.128*** (0.197)	-0.049*** (0.007)	0.001 (0.001)	0.053 (0.035)	0.291 (0.195)	0.117
US	0.565** (0.220)	-0.028*** (0.009)	0.007*** (0.001)	0.008 (0.029)	0.925*** (0.341)	0.214

Notes: Fixed effects estimators. β_0 denotes the average of country-specific intercepts. “post-1998_t” denotes a dummy variable which equals 0 before 1999 and 1 after 1998. “recession_t” denotes a dummy variable which equals 1 during recession set by the Economic Cycle Research Institute (ECRI) and 0 otherwise. “output gap_t” denotes the ex-post output gap estimated in the OECD Economic Outlook quarterly output gap revisions database (in August 2008). $\Delta_{12}\text{INV}_t^2 \equiv (\text{INV}_t - \text{INV}_{t-12})^2$.

Table B-12: Determinants of Disagreement – Country-by-Country Results for Unemployment

Country	β_0	β_1	β_2	β_3	β_4	\bar{R}^2
Panel A: Disagreement over Time						
$\text{disagreement}_t = \beta_0 + \beta_1 \times \text{recession}_t + \beta_2 \times \text{post-1998}_t + \varepsilon_t$						
CN	0.330*** (0.011)	0.121*** (0.023)	-0.112*** (0.015)			0.375
FR	0.238*** (0.009)	0.035** (0.017)	-0.004 (0.012)			0.011
GE	0.410*** (0.028)	0.341*** (0.038)	-0.261*** (0.036)			0.395
IT	0.344*** (0.021)	0.166*** (0.048)	-0.066** (0.029)			0.100
JP	0.208*** (0.011)	0.042*** (0.015)	0.051*** (0.014)			0.086
UK	0.337*** (0.010)	0.187*** (0.022)	-0.113*** (0.014)			0.498
US	0.217*** (0.007)	0.098*** (0.018)	-0.013 (0.010)			0.121
Panel B: Disagreement and Macro Variables						
$\text{disagreement}_t = \beta_0 + \beta_1 \times UN_t + \beta_2 \times \Delta_{12}UN_t^2 + \beta_3 \times \text{output gap}_t + \beta_4 \times \Delta \text{policy rate}_t^2 + \varepsilon_t$						
CN	0.032 (0.037)	0.022* (0.012)	0.026*** (0.006)	-0.009 (0.009)	0.003 (0.019)	0.347
FR	0.172*** (0.026)	-0.028*** (0.008)	0.046*** (0.011)	-0.041*** (0.008)	0.117** (0.055)	0.134
GE	0.085 (0.094)	-0.172*** (0.014)	0.047*** (0.016)	-0.054*** (0.016)	0.442* (0.245)	0.356
IT	-0.025 (0.064)	0.023** (0.011)	0.077*** (0.018)	-0.030*** (0.011)	-0.022 (0.040)	0.236
JP	0.241*** (0.031)	-0.011 (0.014)	0.064* (0.038)	-0.020*** (0.008)	-0.011 (0.012)	0.127
UK	0.152*** (0.033)	0.038*** (0.005)	0.021*** (0.005)	0.064*** (0.011)	0.103*** (0.038)	0.446
US	0.140*** (0.024)	-0.018** (0.008)	0.038*** (0.009)	-0.002 (0.006)	0.014 (0.045)	0.137

Notes: Fixed effects estimators. β_0 denotes the average of country-specific intercepts. “post-1998_t” denotes a dummy variable which equals 0 before 1999 and 1 after 1998. “recession_t” denotes a dummy variable which equals 1 during recession set by the Economic Cycle Research Institute (ECRI) and 0 otherwise. “output gap_t” denotes the ex-post output gap estimated in the OECD Economic Outlook quarterly output gap revisions database (in August 2008). $\Delta_{12}UN_t^2 \equiv (UN_t - UN_{t-12})^2$.

Table B-13: Average Disagreement Across Countries and Variables

Variable	Canada	France	Germany	Italy	Japan	UK	US
Full Sample							
Inflation	0.34	0.21	0.24	0.26	0.33	0.38	0.33
Interest Rate	0.76	0.50	0.46	0.58	0.27	0.71	0.56
GDP	0.46	0.30	0.34	0.26	0.69	0.43	0.38
Consumption	0.47	0.31	0.40	0.34	0.65	0.55	0.38
Investment	2.60	1.17	1.56	1.01	2.42	1.66	1.93
Unemployment	0.29	0.24	0.41	0.33	0.25	0.31	0.22
IP	1.24	0.75	0.81	0.75	1.71	0.89	0.84
Pre-1999							
Inflation	0.37	0.21	0.23	0.33	0.40	0.50	0.33
Interest Rate	0.93	0.55	0.52	0.75	0.40	0.87	0.56
GDP	0.53	0.29	0.38	0.26	0.67	0.52	0.37
Consumption	0.53	0.29	0.41	0.36	0.64	0.58	0.38
Investment	2.85	1.17	1.39	1.07	2.26	1.79	1.93
Unemployment	0.36	0.24	0.53	0.38	0.22	0.38	0.22
IP	1.40	0.72	0.76	0.70	1.60	0.98	0.79
1999+							
Inflation	0.30	0.21	0.27	0.17	0.25	0.24	0.33
Interest Rate	0.56	0.44	0.39	0.38	0.13	0.53	0.57
GDP	0.38	0.32	0.29	0.27	0.71	0.34	0.39
Investment	2.31	1.17	1.76	0.94	2.61	1.49	1.94
Consumption	0.41	0.32	0.39	0.33	0.67	0.50	0.39
Unemployment	0.22	0.24	0.27	0.28	0.28	0.22	0.21
IP	1.05	0.79	0.86	0.81	1.84	0.79	0.91
Booms							
Inflation	0.32	0.21	0.24	0.24	0.31	0.37	0.32
Interest Rate	0.71	0.47	0.42	0.55	0.30	0.67	0.55
GDP	0.43	0.30	0.30	0.25	0.60	0.41	0.36
Consumption	0.44	0.30	0.37	0.32	0.57	0.53	0.37
Investment	2.61	1.17	1.46	0.98	2.30	1.57	1.84
Unemployment	0.27	0.24	0.29	0.31	0.23	0.28	0.21
IP	1.20	0.74	0.74	0.75	1.53	0.85	0.82
Recessions							
Inflation	0.43	0.22	0.25	0.41	0.36	0.50	0.43
Interest Rate	1.07	0.71	0.53	0.87	0.24	1.04	0.68
GDP	0.66	0.32	0.42	0.36	0.81	0.67	0.61
Consumption	0.71	0.34	0.46	0.52	0.78	0.72	0.58
Investment	2.56	1.17	1.74	1.25	2.60	2.37	2.86
Unemployment	0.45	0.27	0.63	0.51	0.27	0.52	0.31
IP	1.52	0.81	0.93	0.74	1.96	1.25	1.08

Notes: Numbers are based on the cross sectional IQR as the disagreement measure. Averages taken across time periods.

Table B-14: Correlation of Disagreement Across Countries

Variable	Country	Canada	France	Germany	Italy	Japan	UK
INFL	France	0.16					
	Germany	0.01	-0.05				
	Italy	0.28	0.25	-0.18			
	Japan	0.16	0.18	0.03	0.48		
	UK	0.35	0.23	-0.04	0.64	0.56	
	US	0.16	0.24	0.07	0.19	0.09	0.07
GDP	France	0.12					
	Germany	0.25	0.08				
	Italy	0.19	0.21	0.32			
	Japan	0.13	0.15	0.26	0.15		
	UK	0.55	-0.04	0.12	0.15	0.02	
	US	0.37	0.29	0.01	0.26	-0.07	0.31
R3M	France	0.27					
	Germany	0.38	0.38				
	Italy	0.34	0.42	0.41			
	Japan	0.41	0.22	0.28	0.52		
	UK	0.52	0.45	0.47	0.46	0.54	
	US	0.28	0.01	0.25	0.01	-0.05	0.09
CONS	France	0.17					
	Germany	-0.02	0.15				
	Italy	0.14	0.04	0.22			
	Japan	0.05	0.17	0.18	0.37		
	UK	0.32	0.13	0.13	0.06	0.06	
	US	0.40	0.10	0.10	0.16	0.06	0.30
INV	France	0.01					
	Germany	0.00	0.20				
	Italy	0.31	0.30	0.27			
	Japan	0.11	0.10	0.30	0.19		
	UK	0.06	0.03	-0.02	0.03	0.03	
	US	0.14	0.11	0.16	0.29	0.22	0.36
UN	France	0.09					
	Germany	0.39	0.02				
	Italy	0.13	0.33	0.34			
	Japan	-0.13	0.19	-0.23	0.08		
	UK	0.53	-0.01	0.54	0.26	-0.29	
	US	0.36	-0.03	0.07	0.00	-0.01	0.27

Notes: Numbers are based on the cross sectional IQR as the disagreement measure.

Table B-15: Correlation of Disagreement Within Countries

Country	Variable	Inflation	GDP	Int. Rate	Cons	Inv
CN	GDP	0.19				
	Interest Rate	0.22	0.27			
	Consumption	0.22	0.53	0.27		
	Investment	0.11	0.25	0.05	0.15	
	Unemployment	0.31	0.57	0.34	0.48	0.28
FR	GDP	0.16				
	Interest Rate	0.26	0.14			
	Consumption	0.04	0.41	0.12		
	Investment	0.20	0.39	0.19	0.28	
	Unemployment	0.05	0.23	0.27	0.16	0.11
GE	GDP	-0.03				
	Interest Rate	0.09	0.18			
	Consumption	0.14	0.44	0.28		
	Investment	0.07	0.17	0.11	0.32	
	Unemployment	-0.04	0.34	0.22	0.23	-0.08
IT	GDP	0.17				
	Interest Rate	0.56	0.12			
	Consumption	0.25	0.39	0.16		
	Investment	0.29	0.36	0.23	0.32	
	Unemployment	0.31	0.32	0.12	0.32	0.26
JP	GDP	0.26				
	Interest Rate	0.47	0.06			
	Consumption	0.24	0.57	0.08		
	Investment	0.13	0.53	-0.03	0.38	
	Unemployment	0.04	0.50	-0.14	0.39	0.26
UK	GDP	0.49				
	Interest Rate	0.66	0.53			
	Consumption	0.31	0.59	0.33		
	Investment	0.37	0.56	0.41	0.52	
	Unemployment	0.47	0.63	0.61	0.45	0.56
US	GDP	0.41				
	Interest Rate	0.31	0.22			
	Consumption	0.36	0.74	0.05		
	Investment	0.32	0.55	0.11	0.50	
	Unemployment	0.18	0.48	0.12	0.50	0.37

Notes: Numbers are based on the cross sectional IQR as the disagreement measure.

Table B-16: Horizon Effect for Disagreement across Forecasts for the Next Year

	GDP	Consumption	Inflation	Ind.Prod.
Germany	0.007 (2.8***)	0.011 (3.0***)	0.010 (6.4***)	-0.010 (-1.9*)
Canada	0.005 (1.8*)	0.005 (1.4)	0.008 (4.1***)	-0.006 (-0.6)
France	0.001 (1.0)	0.002 (1.0)	0.004 (3.2***)	-0.009 (-1.7*)
Italy	0.003 (1.9*)	0.000 (0.2)	0.011 (3.4***)	-0.013 (-2.4**)
Japan	-0.007 (-1.3)	0.011 (2.4**)	0.014 (4.1***)	-0.025 (-1.8*)
United Kingdom	0.003 (1.1)	0.004 (0.9)	0.017 (2.9***)	0.001 (0.2)
United States	0.001 (0.5)	-0.001 (-0.5)	0.011 (6.5***)	0.025 (4.2***)

Notes: Numbers below the estimates are t-statistics.

Table B-17: Horizon Effect for Disagreement across Forecasts for the Current Year

	GDP	Consumption	Inflation	Ind.Prod.
Germany	0.017 (9.0***)	0.017 (5.8***)	0.015 (21.6***)	0.027 (5.1***)
Canada	0.032 (17.1***)	0.030 (12.9***)	0.021 (17.8***)	0.066 (9.0***)
France	0.017 (13.3***)	0.012 (9.0***)	0.013 (13.4***)	0.025 (6.0***)
Italy	0.012 (6.6***)	0.009 (4.4***)	0.017 (10.7***)	0.016 (3.5***)
Japan	0.034 (10.3***)	0.029 (10.2***)	0.022 (11.5***)	0.110 (14.1***)
United Kingdom	0.028 (12.5***)	0.031 (10.3***)	0.025 (10.7***)	0.057 (13.3***)
United States	0.030 (16.7***)	0.027 (18.2***)	0.024 (19.4***)	0.065 (16.8***)

Notes: Numbers below the estimates are t-statistics.

Table B-18: Rel. Change of Disagreement per Month for Forecasts for the Next Year

	GDP	Consumption	Inflation	Ind.Prod.
Germany	1.9	2.3	3.4	-1.3
Canada	1.1	1.0	2.2	-0.6
France	0.5	0.6	1.8	-1.2
Italy	1.3	0.1	3.5	-1.9
Japan	-1.0	1.7	3.6	-1.6
United Kingdom	0.6	0.6	3.3	0.1
United States	0.3	-0.3	2.8	2.6

Notes: Numbers indicate the change in disagreement per month relative to the average level of disagreement calculated across all forecasts with a horizon of more than a year.

Table B-19: Rel. Change of Disagreement per Month for Forecasts for the Current Year

	GDP	Consumption	Inflation	Ind.Prod.
Germany	7.1	5.2	11.5	4.0
Canada	11.8	9.4	11.0	7.8
France	8.2	5.4	8.4	4.1
Italy	5.8	3.3	10.7	2.5
Japan	7.1	6.6	10.0	10.0
United Kingdom	9.8	7.8	12.5	9.4
United States	13.2	11.7	13.1	12.7

Notes: Numbers indicate the change in disagreement per month relative to the average level of disagreement calculated across all forecasts with a horizon of less than a year.

Table B-20: Quadratic Trend Effect per Month for Disagreement across Forecasts for the Next Year

	GDP	Consumption	Inflation	Ind.Prod.
Germany	0.0007 (0.9)	-0.0007 (-0.6)	-0.0013 (-2.5**)	-0.0002 (-0.1)
Canada	-0.0004 (-0.4)	-0.0009 (-0.8)	0.0008 (1.3)	0.0012 (0.4)
France	0.0002 (0.4)	-0.0002 (-0.3)	0.0000 (0.0)	0.0032 (1.9*)
Italy	0.0007 (1.1)	0.0008 (1.1)	-0.0006 (-0.6)	-0.0006 (-0.4)
Japan	-0.0011 (-0.7)	0.0000 (0.0)	-0.0004 (-0.4)	0.0009 (0.2)
United Kingdom	-0.0002 (-0.2)	-0.0007 (-0.5)	0.0001 (0.0)	0.0007 (0.4)
United States	0.0005 (0.6)	0.0004 (0.5)	-0.0008 (-1.3)	0.0014 (0.7)

Notes: Numbers below the estimates are t-statistics.

Table B-21: Quadratic Trend Effect per Month for Disagreement across Forecasts for the Current Year

	GDP	Consumption	Inflation	Ind.Prod.
Germany	-0.0011 (-1.8*)	-0.0022 (-2.5**)	0.0001 (0.3)	-0.0034 (-2.0**)
Canada	-0.0005 (-0.8)	-0.0012 (-1.5)	0.0001 (0.4)	-0.0070 (-2.3**)
France	-0.0001 (-0.3)	-0.0008 (-1.8*)	-0.0002 (-0.6)	-0.0015 (-1.1)
Italy	-0.0007 (-1.2)	0.0001 (0.1)	-0.0001 (-0.1)	-0.0021 (-1.4)
Japan	-0.0024 (-2.3**)	-0.0019 (-2.0**)	-0.0008 (-1.5)	-0.0025 (-0.8)
United Kingdom	-0.0013 (-1.8*)	-0.0005 (-0.6)	0.0009 (1.2)	-0.0033 (-2.5**)
United States	0.0006 (1.0)	0.0005 (1.0)	0.0007 (1.9*)	0.0011 (0.9)

Notes: Numbers below the estimates are t-statistics.

Table B-22: Time Trends in Disagreement across Forecasts for the Next Year

	GDP	Consumption	Inflation	Ind.Prod.
Germany	-0.0008 (-6.0***)	-0.0007 (-3.3***)	-0.0001 (-0.6)	0.0013 (4.0***)
Canada	-0.0010 (-6.4***)	-0.0011 (-5.8***)	-0.0011 (-7.6***)	-0.0022 (-2.5**)
France	0.0003 (2.7***)	0.0000 (0.2)	-0.0003 (-3.8***)	0.0013 (4.0***)
Italy	0.0002 (2.2**)	0.0003 (2.1**)	-0.0018 (-14.0***)	0.0031 (10.4***)
Japan	-0.0002 (-0.6)	0.0002 (0.8)	-0.0018 (-12.2***)	0.0047 (7.2***)
United Kingdom	-0.0003 (-1.3)	-0.0006 (-1.7*)	-0.0049 (-16.8***)	-0.0001 (-0.5)
United States	-0.0005 (-3.5***)	-0.0001 (-0.6)	-0.0003 (-2.3**)	0.0006 (1.7*)

Notes: Numbers below the estimates are t-statistics.

Table B-23: Time Trends in Disagreement across Forecasts for the Current Year

	GDP	Consumption	Inflation	Ind.Prod.
Germany	-0.0011 (-13.1***)	-0.0001 (-0.6)	0.0001 (2.7***)	-0.0001 (-0.4)
Canada	-0.0008 (-5.8***)	-0.0009 (-4.6***)	-0.0001 (-1.5)	-0.0036 (-5.6***)
France	-0.0003 (-4.2***)	-0.0004 (-5.6***)	0.0000 (0.3)	-0.0010 (-3.5***)
Italy	-0.0001 (-1.5)	0.0001 (0.6)	-0.0006 (-5.7***)	-0.0001 (-0.4)
Japan	-0.0014 (-6.5***)	-0.0008 (-4.1***)	-0.0007 (-6.5***)	-0.0009 (-1.4)
United Kingdom	-0.0004 (-2.1**)	-0.0007 (-3.5***)	-0.0009 (-7.2***)	-0.0004 (-1.1)
United States	-0.0006 (-5.8***)	-0.0005 (-5.0***)	-0.0002 (-2.2**)	-0.0005 (-1.9*)

Notes: Numbers below the estimates are t-statistics.

Table B-24: Horizon Effects for Multivariate Measure of Disagreement

	Germany	Canada	France	Italy	Japan	UK	US
Next Year							
$\hat{\alpha}$	-0.0061 (-0.5)	-0.0140 (-0.6)	-0.0088 (-1.0)	-0.0091 (-1.1)	-0.0726 (-1.2)	0.0075 (0.6)	0.0477 (3.0***)
<i>rel.change</i>	-0.0070	-0.0088	-0.0133	-0.0137	-0.0199	0.0061	0.0419
Current Year							
$\hat{\alpha}$	0.0468 (5.3***)	0.1342 (7.9***)	0.0409 (5.5***)	0.0220 (3.3***)	0.2836 (10.4***)	0.0905 (9.7***)	0.0853 (12.1***)
<i>rel.change</i>	0.0772	0.1281	0.0874	0.0455	0.1641	0.1592	0.2154

Notes: Numbers below the estimates are t-statistics. The numbers in labeled *rel.change* indicate the change in disagreement per month relative to the average level of disagreement calculated across all forecasts with a horizon of more than a year or less than a year respectively. All results are based on the concept of vector variance (*VV*).

Table B-25: Business Cycle Effect on Disagreement across Forecasts for the Next Year

	GDP	Consumption	Inflation	Ind.Prod.
Germany	-1.8868 (-2.6***)	-1.1901 (-0.9)	-0.3193 (-0.7)	-4.5207 (-3.7***)
Canada	-2.8134 (-3.5***)	-2.7834 (-3.4***)	1.3343 (2.6**)	-6.4181 (-2.8***)
France	-1.1913 (-2.2**)	-0.9505 (-1.7*)	-0.4457 (-1.1)	-7.1767 (-4.3***)
Italy	-1.3665 (-2.8***)	-2.4018 (-3.5***)	-0.8878 (-1.1)	-3.4127 (-2.3**)
Japan	-3.4107 (-3.3***)	-1.7145 (-1.8*)	-0.6963 (-1.3)	-10.3581 (-3.7***)
United Kingdom	1.1201 (0.9)	-0.4951 (-0.3)	0.4811 (0.2)	-5.6249 (-2.9***)
United States	-3.3163 (-2.4**)	-1.4689 (-1.0)	0.1139 (0.1)	-6.0481 (-2.7***)

Notes: Numbers below the estimates are t-statistics.

Table B-26: Business Cycle Effect on Disagreement across Forecasts for the Current Year

	GDP	Consumption	Inflation	Ind.Prod.
Germany	0.4240 (0.7)	0.0755 (0.1)	0.5488 (2.8***)	7.1651 (4.2***)
Canada	0.9088 (2.0*)	0.9110 (1.5)	-0.1969 (-0.6)	4.0694 (1.6)
France	1.3360 (2.1**)	1.6560 (3.0***)	1.0148 (3.2***)	3.3790 (2.1**)
Italy	0.8345 (1.6)	0.6930 (1.1)	1.1788 (2.4**)	1.1645 (0.9)
Japan	0.0717 (0.1)	0.4988 (1.0)	-0.0923 (-0.2)	2.5940 (1.4)
United Kingdom	0.1317 (0.1)	-0.5301 (-0.4)	-0.1910 (-0.2)	2.0097 (1.0)
United States	1.9230 (2.7***)	0.5513 (0.7)	-0.3288 (-0.7)	4.5154 (2.8***)

Notes: Numbers below the estimates are t-statistics.

Table B-27: Effect of Recession in Next Year on Disagreement of Forecasts for the Next Year

	GDP	Consumption	Inflation	Ind.Prod.
Germany	0.0550 (4.1***)	0.0879 (3.1***)	0.0740 (7.7***)	0.1667 (4.8***)
Canada	0.0461 (1.8*)	0.0371 (1.1)	-0.0664 (-3.1***)	0.0730 (0.6)
France	0.0330 (2.5**)	0.0448 (3.3***)	-0.0024 (-0.3)	0.0212 (0.6)
Italy	0.0229 (1.1)	0.0669 (1.5)	-0.0177 (-0.5)	-0.0480 (-0.9)
Japan	0.0035 (0.1)	0.0519 (1.6)	0.0082 (0.5)	-0.2348 (-2.6***)
United Kingdom	0.0251 (0.8)	0.1678 (4.3***)	-0.2480 (-6.2***)	0.2388 (5.2***)
United States	0.0240 (1.0)	0.0575 (2.6***)	0.0370 (1.7*)	0.0123 (0.2)

Notes: Numbers below the estimates are t-statistics.

Table B-28: Effect of Recession in Current Year on Disagreement of Forecasts for the Current Year

	GDP	Consumption	Inflation	Ind.Prod.
Germany	0.0288 (3.4***)	0.0197 (1.0)	0.0104 (2.3**)	0.1449 (3.9***)
Canada	0.0607 (3.5***)	0.1001 (4.1***)	0.0389 (2.7***)	0.1427 (1.4)
France	0.0462 (3.2***)	0.0263 (2.3**)	-0.0021 (-0.4)	0.0810 (2.1**)
Italy	0.1120 (5.2***)	0.1352 (4.4***)	0.0425 (2.1**)	0.0298 (0.6)
Japan	0.0882 (3.8***)	0.0391 (2.0**)	0.0190 (1.7*)	0.1649 (2.7***)
United Kingdom	0.1234 (5.5***)	0.1889 (8.0***)	0.0624 (2.6**)	0.2331 (4.8***)
United States	0.0868 (6.1***)	0.0502 (3.2***)	0.0406 (3.9***)	0.1216 (3.7***)

Notes: Numbers below the estimates are t-statistics.

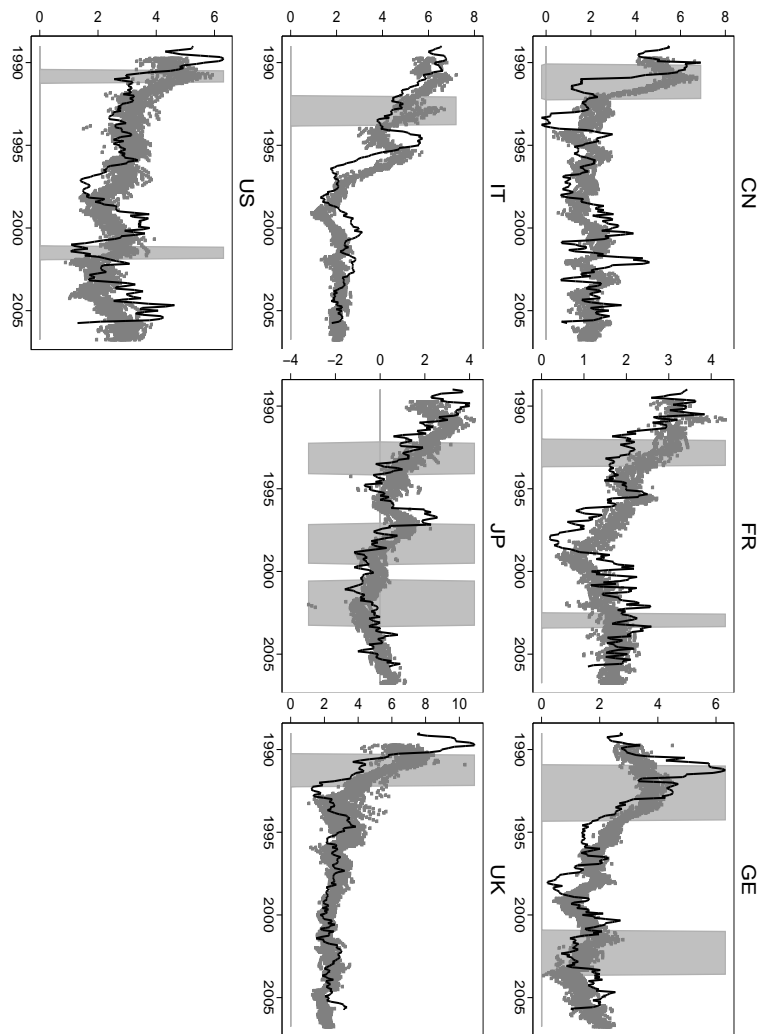
Table B-29: Business Cycle Effect on Multivariate Measure of Disagreement

For Next Year							
	Germany	Canada	France	Italy	Japan	UK	US
$\hat{\beta}$	-7.86	-16.27	-10.78	-6.04	-44.57	-8.68	-15.90
	(-2.7***)	(-3.0***)	(-4.3***)	(-2.7***)	(-3.4***)	(-1.9*)	(-2.6***)
For Current Year							
$\hat{\beta}$	12.12	10.05	7.57	3.34	5.30	3.01	5.72
	(3.5***)	(1.5)	(2.4**)	(1.8*)	(0.8)	(0.7)	(2.2**)

Notes: Numbers below the estimates are t-statistics. All results are based on the concept of vector variance (VV).

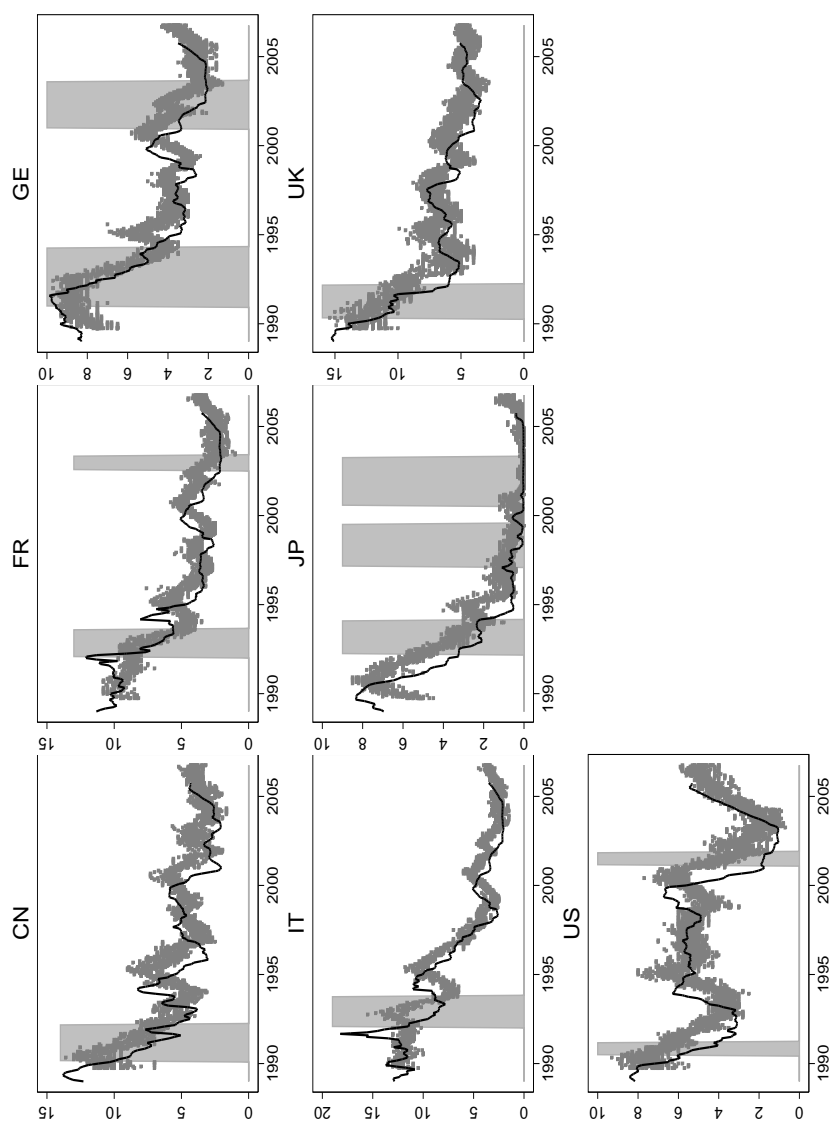
C Graphical Appendix

Figure C-1: Expected and Actual Inflation Rate

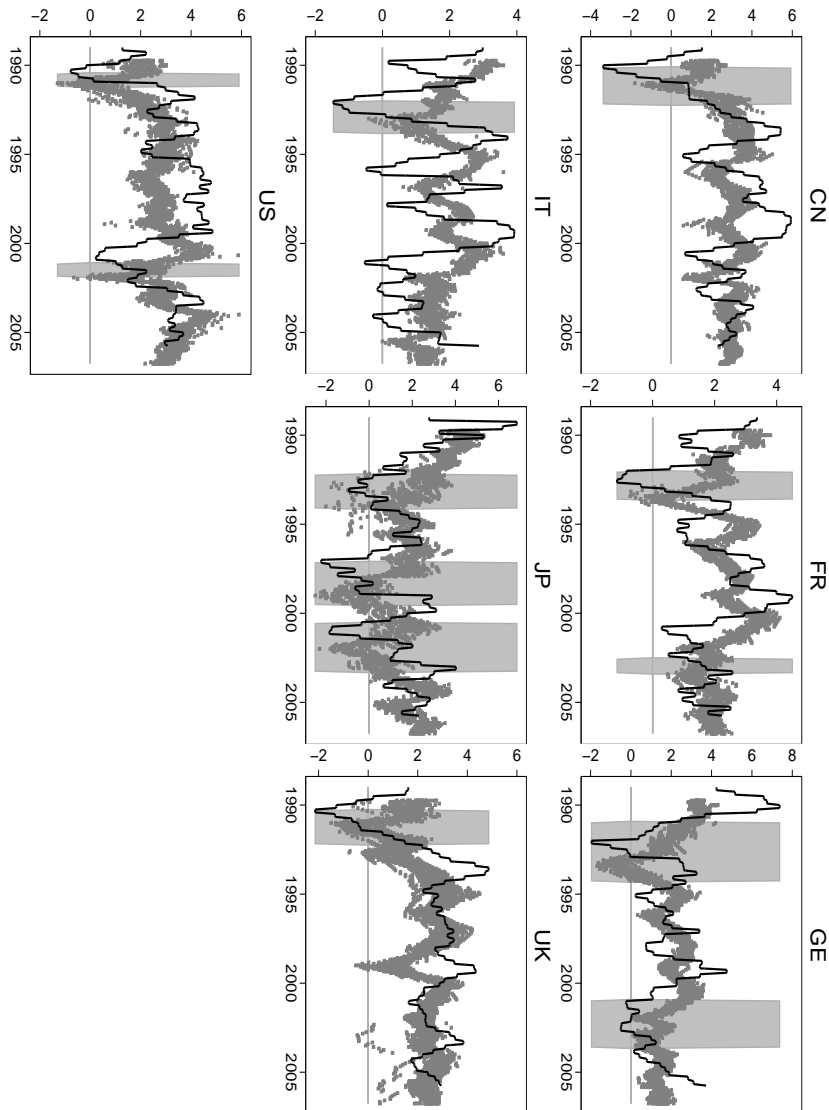


Notes: All series are measured in percent. The solid line represents the actual history of the series (shifted by 12 months to ease the comparison with the forecasts). Each individual forecast is represented by one grey dot.

PhD-Thesis written by: Jonas Dovern

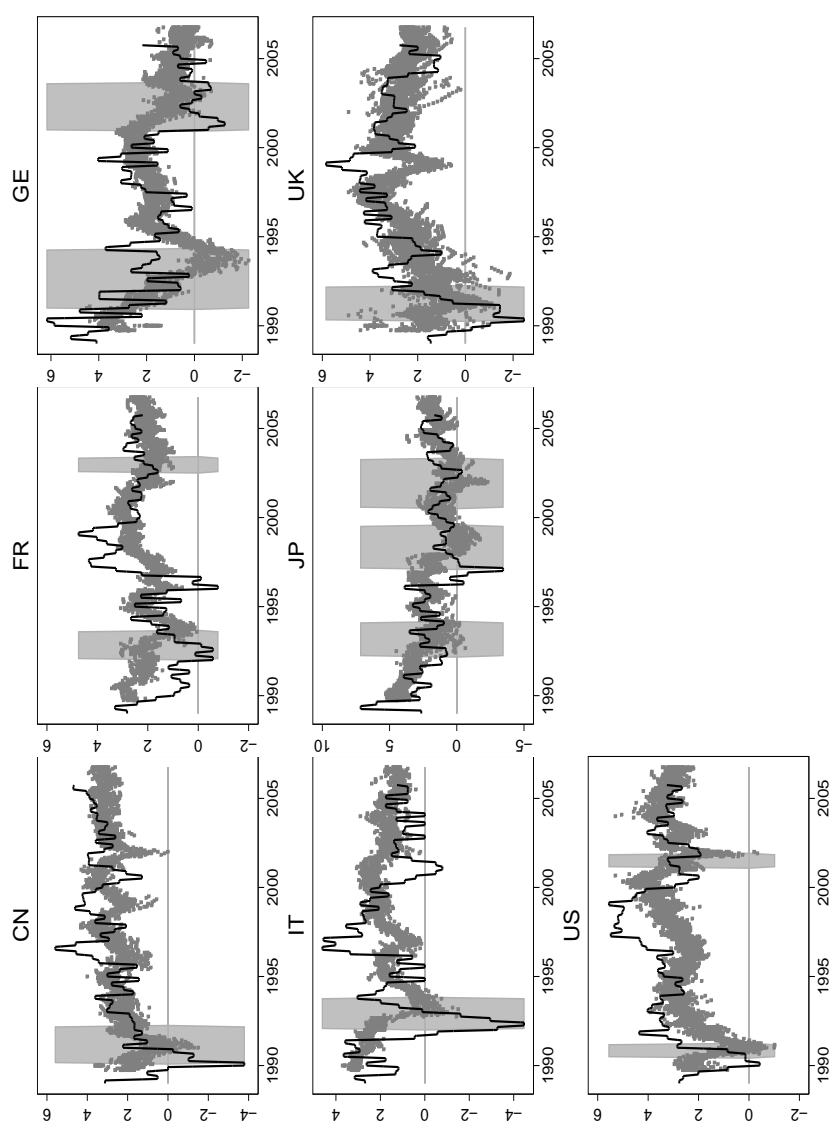
Figure C-2: Expected and Actual Short-term Interest Rates

Notes: All series are measured in percent. The solid line represents the actual history of the series (shifted by 12 months to ease the comparison with the forecasts). Each individual forecast is represented by one grey dot.

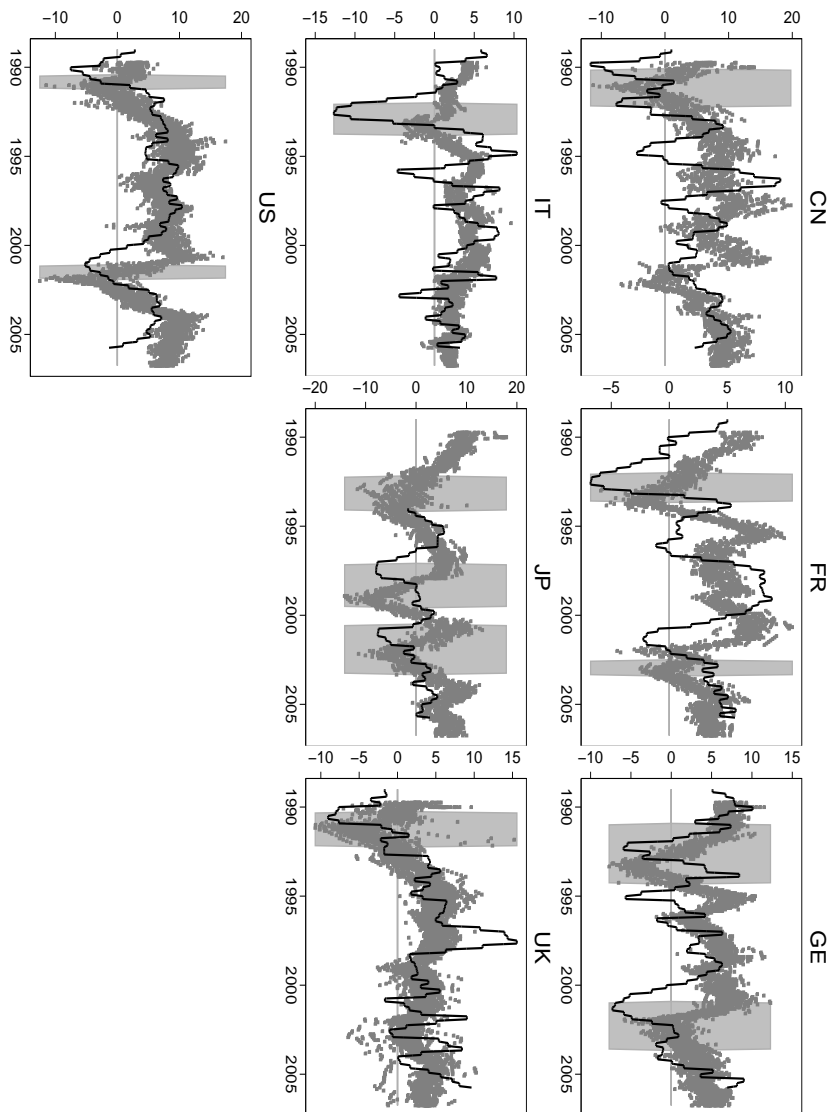
Figure C-3: Expected and Actual GDP Growth

Notes: All series are measured in percent. The solid line represents the actual history of the series (shifted by 12 months to ease the comparison with the forecasts). Each individual forecast is represented by one grey dot.

PhD-Thesis written by: Jonas Dovern

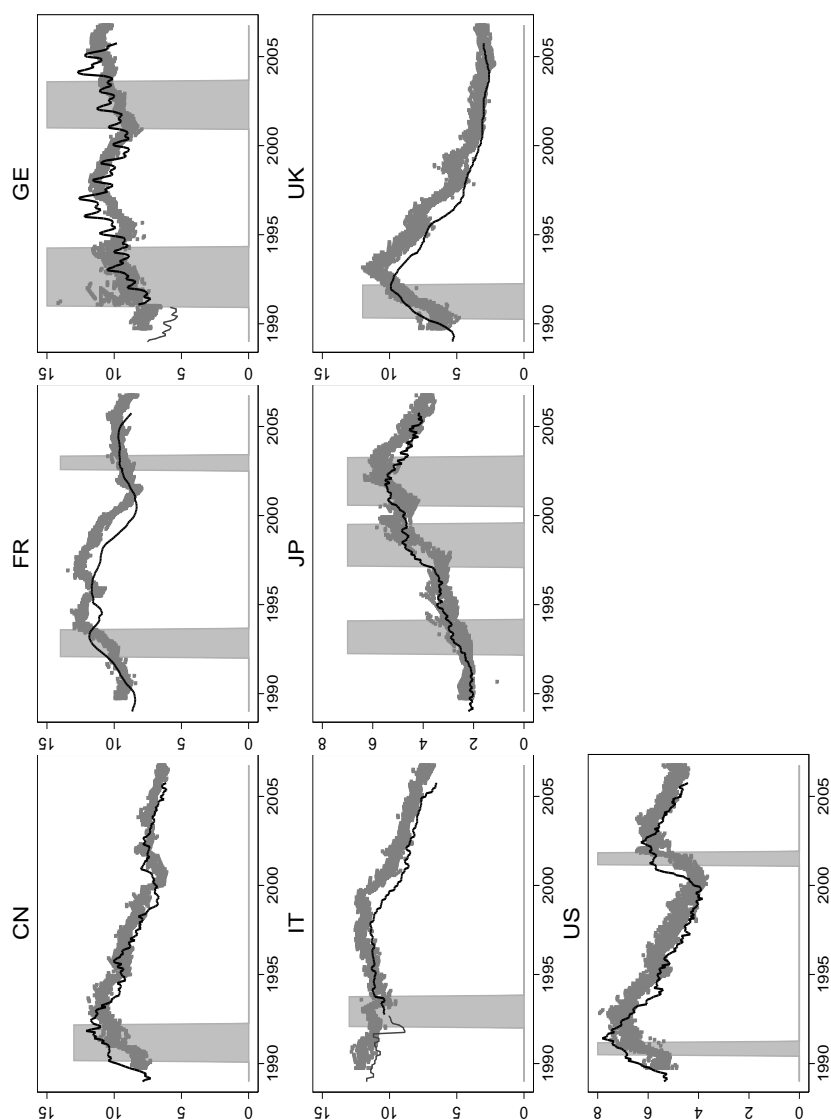
Figure C-4: Expected and Actual Consumption Growth

Notes: All series are measured in percent. The solid line represents the actual history of the series (shifted by 12 months to ease the comparison with the forecasts). Each individual forecast is represented by one grey dot.

Figure C-5: Expected and Actual Investment Growth

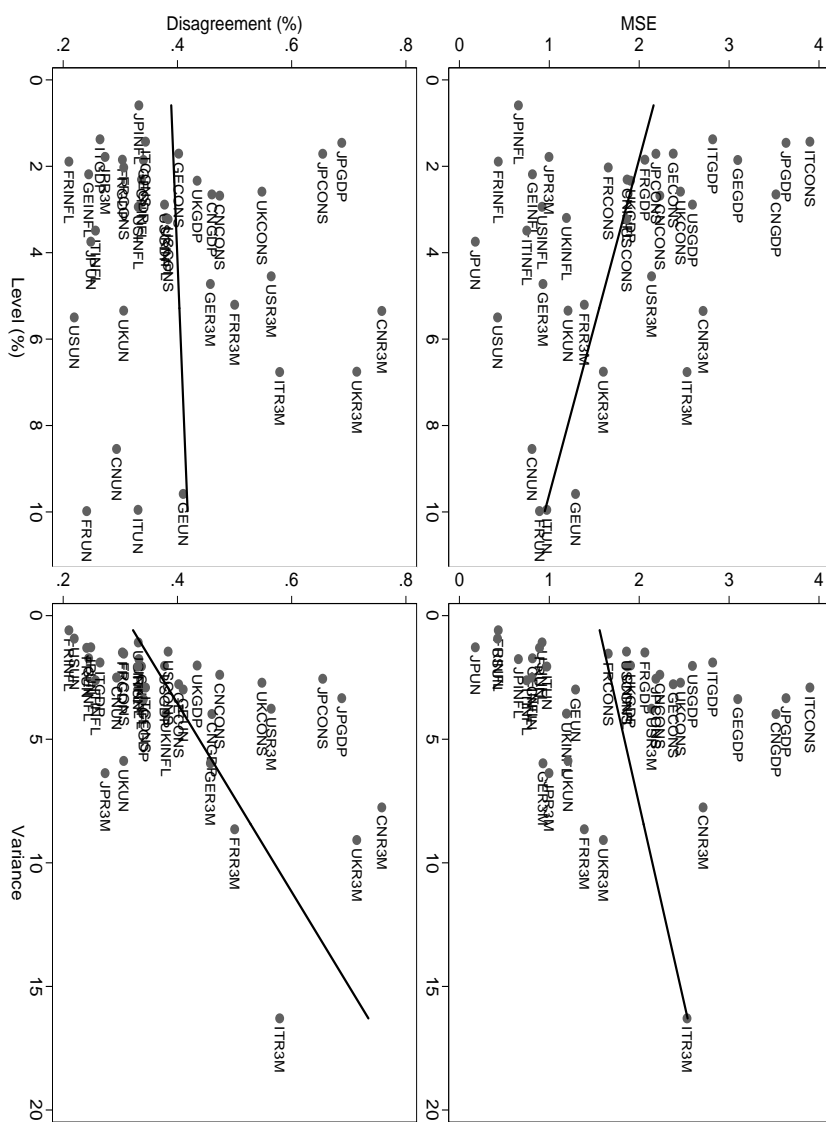
Notes: All series are measured in percent. The solid line represents the actual history of the series (shifted by 12 months to ease the comparison with the forecasts). Each individual forecast is represented by one grey dot.

PhD-Thesis written by: Jonas Dovern

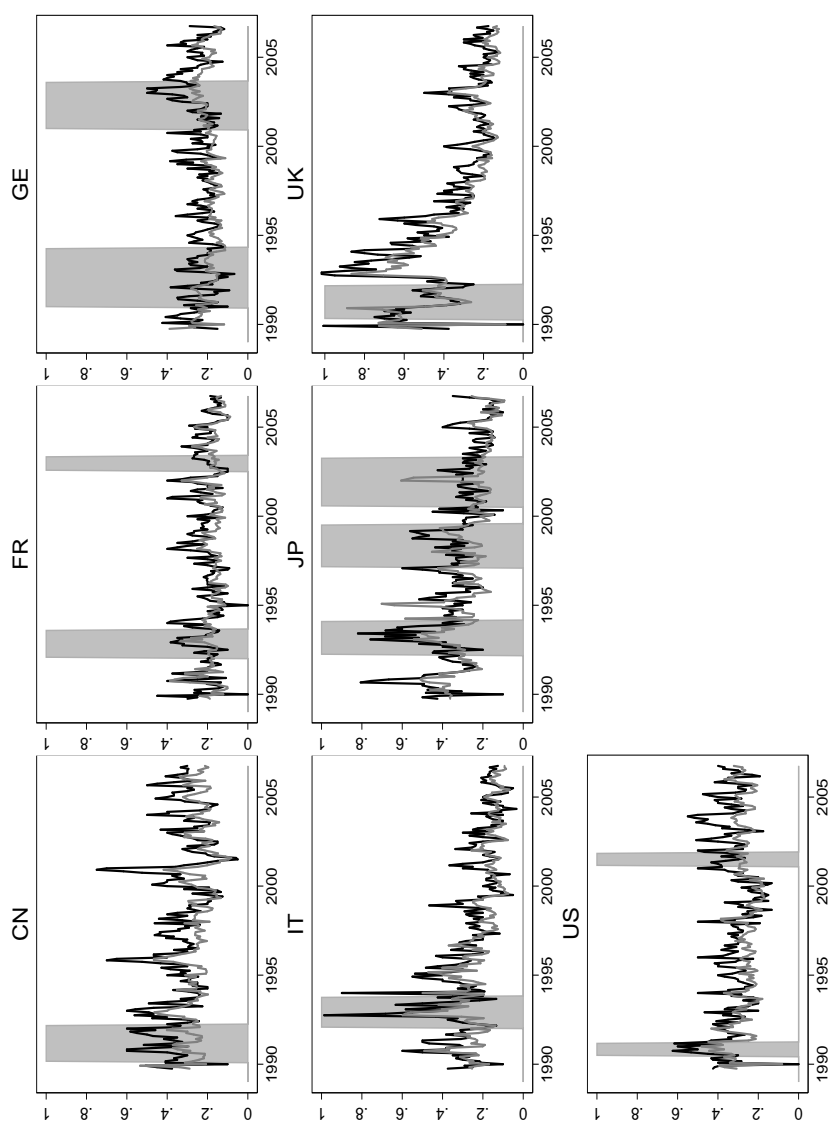
Figure C-6: Expected and Actual Unemployment Rate

Notes: All series are measured in percent. The solid line represents the actual history of the series (shifted by 12 months to ease the comparison with the forecasts). Each individual forecast is represented by one grey dot.

Figure C-7: Disagreement and MSE of Expectations vs. Level and Variance of the Underlying Variable

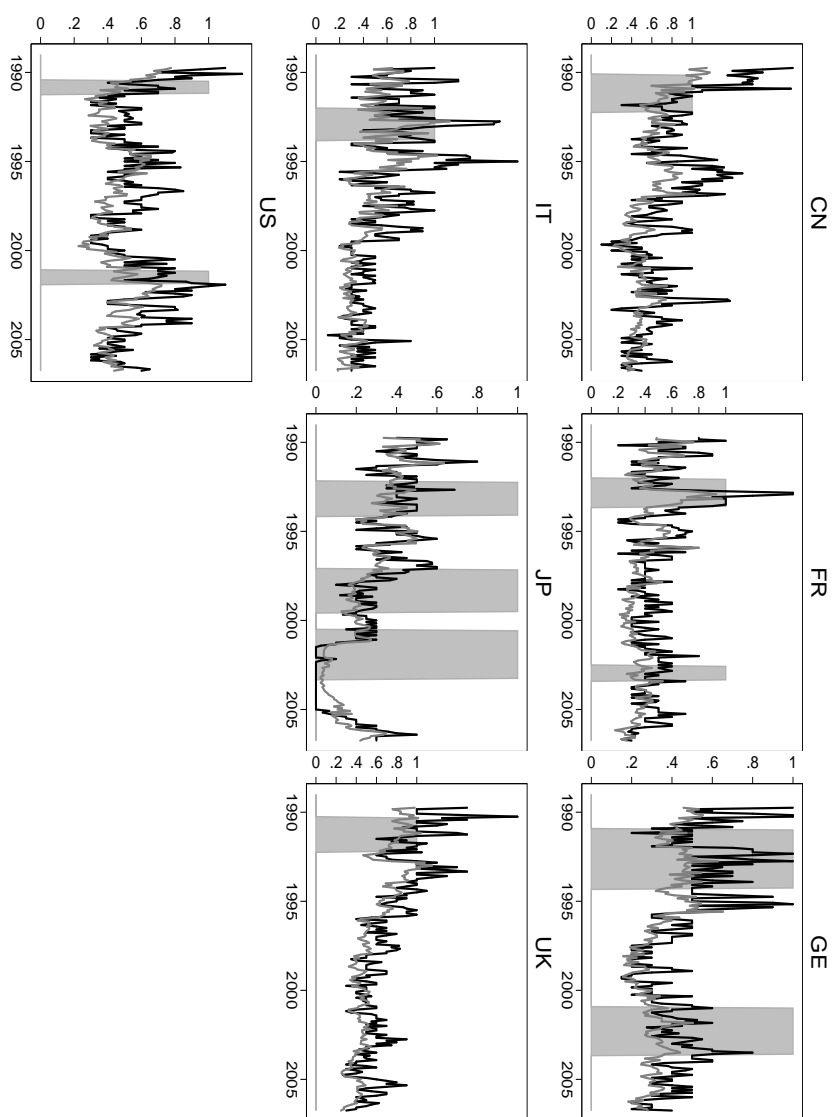


Notes: Disagreement is measured with the cross-sectional interquartile range.

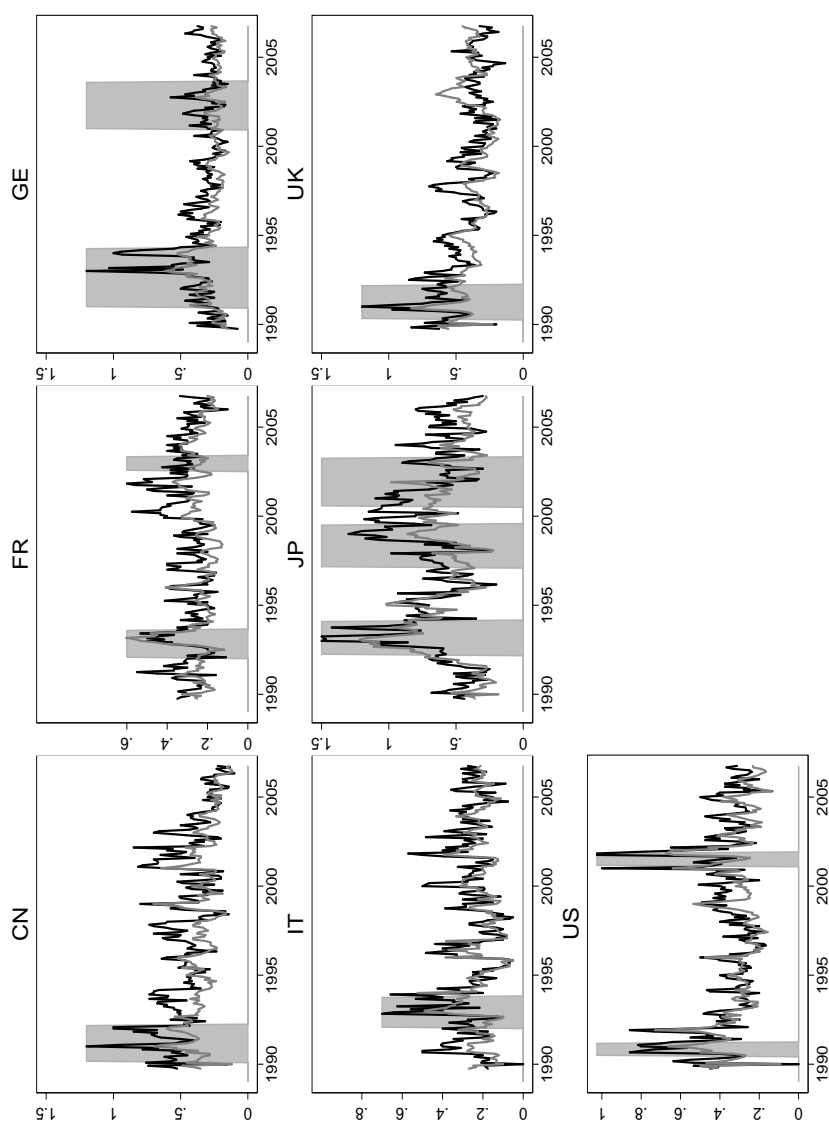
Figure C-8: Disagreement about Inflation Rate

Notes: All series are measured in percent. Disagreement is measured with the cross-sectional interquartile range.

Figure C-9: Disagreement about Short-term Interest Rate

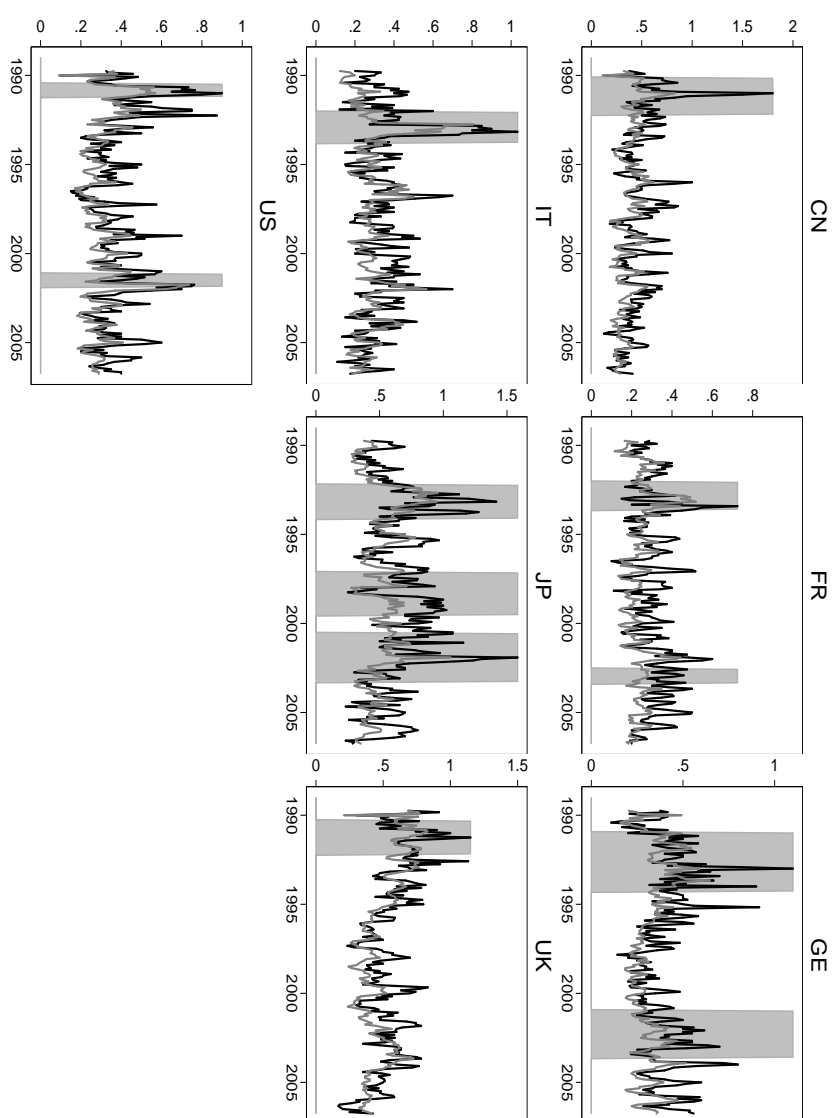


Notes: All series are measured in percent. Disagreement is measured with the cross-sectional interquartile range.

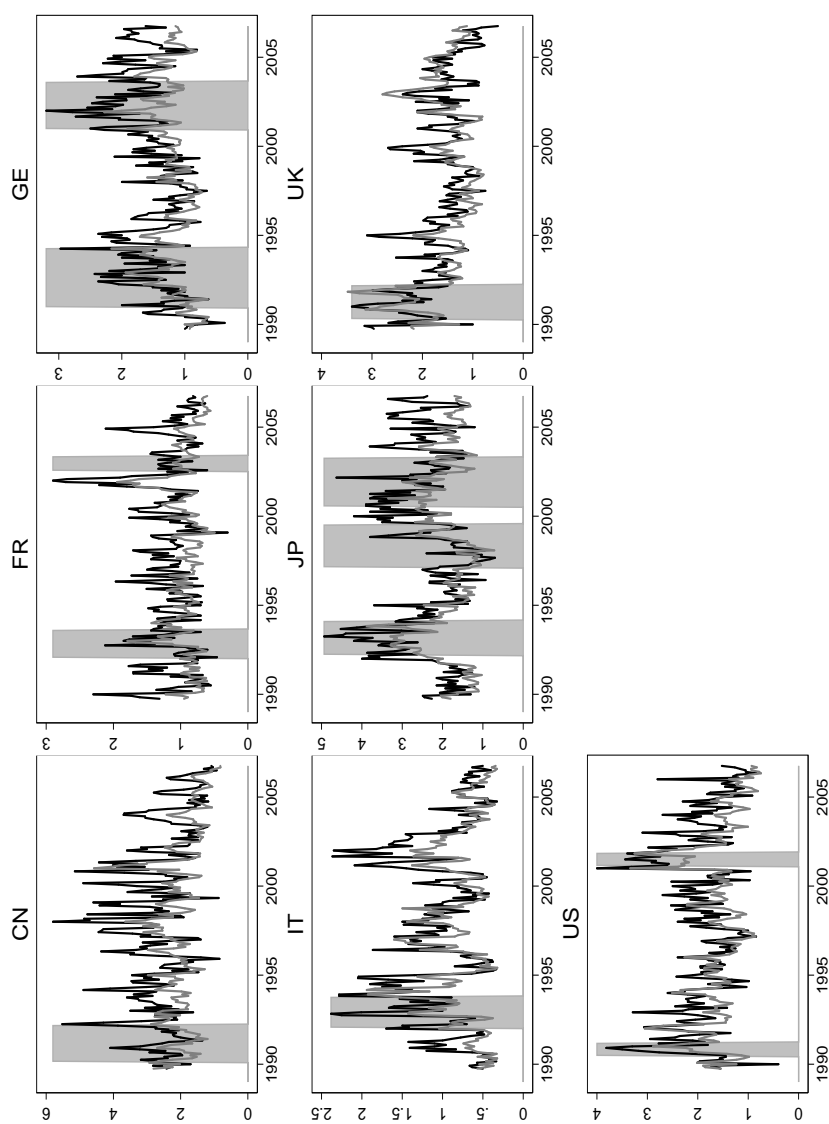
Figure C-10: Disagreement about GDP Growth

Notes: All series are measured in percent. Disagreement is measured with the cross-sectional interquartile range.

Figure C-11: Disagreement about Consumption Growth

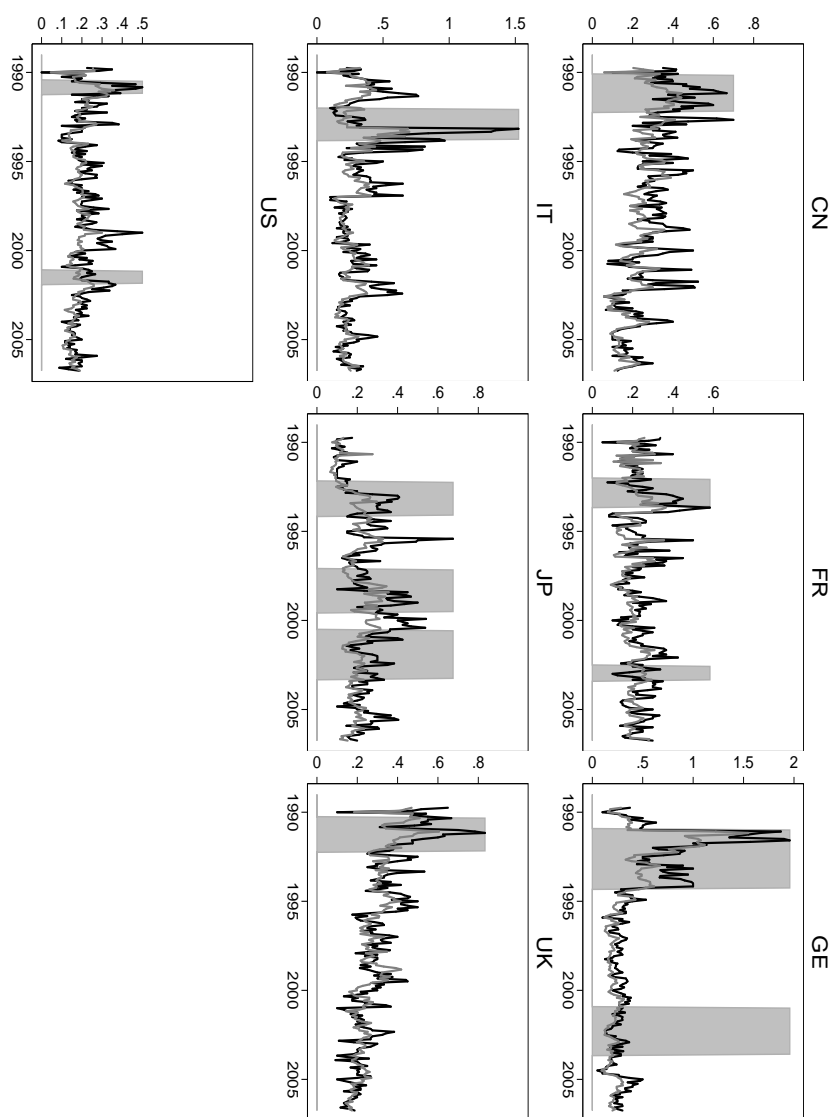


Notes: All series are measured in percent. Disagreement is measured with the cross-sectional interquartile range.

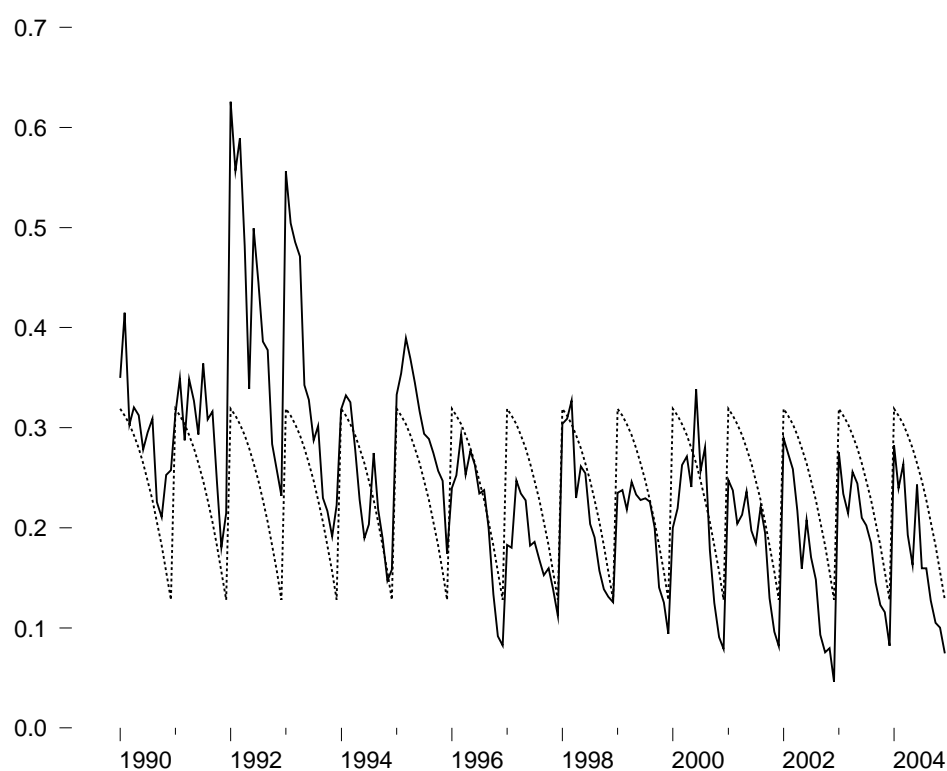
Figure C-12: Disagreement about Investment Growth

Notes: All series are measured in percent. Disagreement is measured with the cross-sectional interquartile range.

Figure C-13: Disagreement about Unemployment Rate



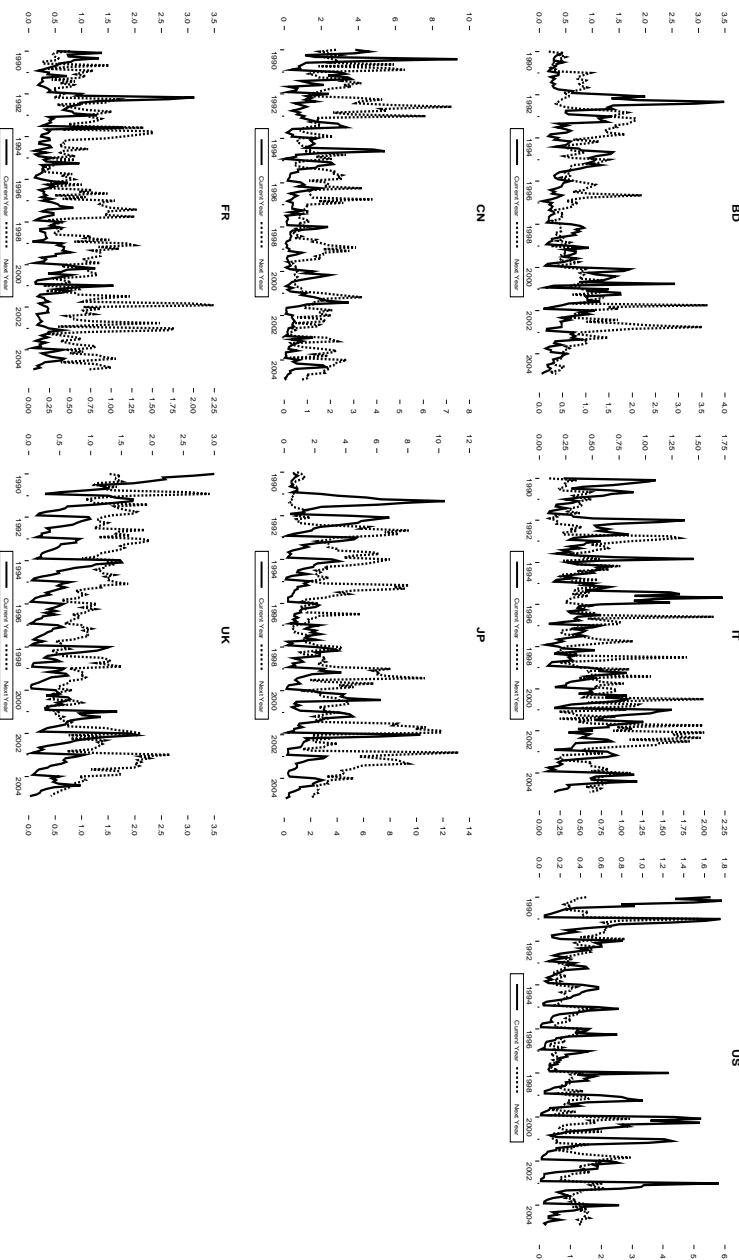
Notes: All series are measured in percent. Disagreement is measured with the cross-sectional interquartile range.

Figure C-14: Typical Shape of Horizon Effect

Notes: The Example refers to disagreement across GDP forecasts in Germany. The solid line refers to the disagreement across GDP forecasts for the current year in Germany (measured as the cross sectional standard deviation), while the dashed line are the fitted values that correspond to equation 4.10.

Figure C-15: Multivariate Disagreement Measure

Multivariate Disagreement Measures (Vector Variance)



Notes: Plots show the vector variance calculated at each point in time across all forecasts for GDP, consumption, industrial production, and inflation. The solid line refers to the disagreement across forecasts for the current year, while the dashed line refers to the disagreement across forecasts for the next year.

Curriculum Vitae

PERSONAL DATA

Name: Dovern
Surname: Jonas
Date of Birth: November 25, 1980
Place of Birth: Aachen, Germany
Nationality: German
Phone: +49-431-530349-7
Email: jonas.dovern[at]kiel-economics.de

EDUCATION

10/2006 – 07/2009 *Doctoral Programme „Quantitative Economics“ at Kiel University.*

08/2005 – 05/2006 *Advanced Studies Program (ASP) at the Kiel Institute for the World Economy (IfW), Kiel.*

09/2004 – 01/2005 *Study of European Studies at Maastricht University.*
Degree: Minor.

09/2001 – 06/2005 *Study of Econometrics at Maastricht University.*
Degree: Doctorandus, equivalent to the Master of Science (Msc).

09/1991 – 06/2000 *Viktoriaschule, Aachen*
Degree: Abitur, German high-school diploma.

PROFESSIONAL EXPERIENCE

06/2009 – Self-employed; founding Kiel Economics GmbH & Co. KG.

11/2005 – 12/2008 *Employment as a researcher at the Kiel Institute for the World Economy (IfW), Kiel.*

04/2005 – 06/2005 *Internship at the Department for Macro Analysis and Forecasting at the German Institute for Economic Research (DIW), Berlin.*

09/2000 – 06/2001 *Community service at the Kleebachschule, Aachen.*

PAPERS & PUBLICATIONS

- 01/2009 *Disagreement among Forecasters in the G7 Countries* (joint with U. Fritsche and J. Slacalek), mimeo, ECB, Frankfurt.
- 10/2008 *Predicting Growth Rates and Recessions. Assessing U.S. Leading Indicators Under Real-Time Conditions* (joint with Ch. Ziegler), *Applied Economics Quarterly*, 54(4).
- 10/2008 *The Sticky Information Phillips Curve. European Evidence* (joint with J. Döpke, U. Fritsche, and J. Slacalek), *Journal of Money, Credit, and Banking*, 40(7), 1513-1519.
- 05/2008 *How Resilient is the German Banking System to Macroeconomic Shocks?* (joint with C.-P. Meier and J. Vilsmeier), Kiel Working Paper 1419, IfW, Kiel.
- 04/2008 *Estimating Fundamental Cross-Section Dispersion from Fixed Event Forecasts* (joint with U. Fritsche), Discussion Papers, 787, DIW, Berlin.
- 01/2008 *The Dynamics of European Inflation Expectations* (joint with J. Döpke, U. Fritsche, and J. Slacalek), *The B.E. Journal of Macroeconomics*, 8(1).
- 08/2007 *Muss das hohe US-Leistungsbilanzdefizit zu einer massiven Aufwertung der europäischen Währungen führen? Eine andere Sicht* (joint with C.-P. Meier and J. Scheide), *Aussenwirtschaft*, 62, 233-254.
- 07/2007 *Aid and Growth Accelerations: An Alternative Approach to Assess the Effectiveness of Aid* (joint with P. Nunnenkamp). *Kyklos*, 60(3), 359-383.
- 09/2006 *Macroeconomic Aspects of Structural Labor Market Reforms in Germany* (joint with C.-P. Meier), Kiel Working Paper 1295, IfW, Kiel.
- 03/2006 – *Quarterly business cycle reports of the Kiel Institute*. Various issues (together with the colleagues of the forecasting center).
- 06/2005 Final Thesis of the econometrics study.
Title: *Ever Closer Union? An Analysis of the Level of Synchronization of the Business Cycles in the Economic and Monetary Union*.
-

CONTRIBUTIONS TO CONFERENCES AND SYMPOSIA

- 04/2009 Spring Meeting of Young Economists, Istanbul.
Rationality of Macroeconomic Expectations in G7-Countries.
- 05/2007 Spring Meeting of Young Economists, Hamburg.
Disagreement on Inflation Expectations in G7-Countries.

- 05/2006 Spring Meeting of Young Economists, Seville.
European inflation expectation dynamics.
- 09/2006 Annual meeting of the European Economic Association, Wien.
European inflation expectation dynamics.
- 10/2006 Meeting of the Association of European Conjunction Institutes
(AIECE), Brüssel.
*Macroeconomic Aspects of Structural Labor Market Reforms in
Germany.*
-

SCHOLARSHIPS

- 08/2005 – 05/2006 Covering of the tuition fees for the Advanced Studies Program by a
Deka-Bank grant.
- 11/2002 – 06/2005 Supported by a scholarship of the *Studienstiftung des Deutschen
Volkes* (German national academic foundation).
-

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides Statt, dass ich meine Doktorarbeit "Formation and Dispersion of Expectations - An Empirical Analysis with Special Focus on Inflation" selbstständig und ohne fremde Hilfe angefertigt habe und dass ich alle von anderen Autoren wörtlich übernommenen Stellen, wie auch die sich an die Gedanken anderer Autoren eng anlehenden Ausführungen meiner Arbeit, besonders gekennzeichnet und Quellen nach den mir angegebenen Richtlinien zitiert habe. Im Fall der Abschnitte, die auf einzelnen Aufsätzen basieren, welche ich in Zusammenarbeit mit anderen Autoren verfasst habe, erkläre ich, dass ich eine angemessene anteilige Leistung beim Verfassen der jeweiligen Aufsätze erbracht habe.

Datum

Unterschrift