

The ECB Survey of Professional Forecasters

Aidan Meyler and Ieva Rubene
European Central Bank*

February 2009

*The views and opinions expressed are those of the presenters and not necessarily those of the ECB

Outline of the presentations

- 1) An Overview of the ECB Survey of Professional Forecasters: 1999-2008 (Aidan)
- 2) Results from the ECB's SPF Special Questionnaire (Ieva)

An overview of the ECB Survey of Professional Forecasters (SPF): 1999-2008

Aidan Meyler

European Central Bank

(an update of Bowles et al 2007)

February 2009

Outline of the presentation

- (1) main features of the ECB SPF**
- (2) evaluation of short-term forecasts and performance to date**
- (3) forecast uncertainty as viewed by SPF participants**
- (4) longer-term expectations**

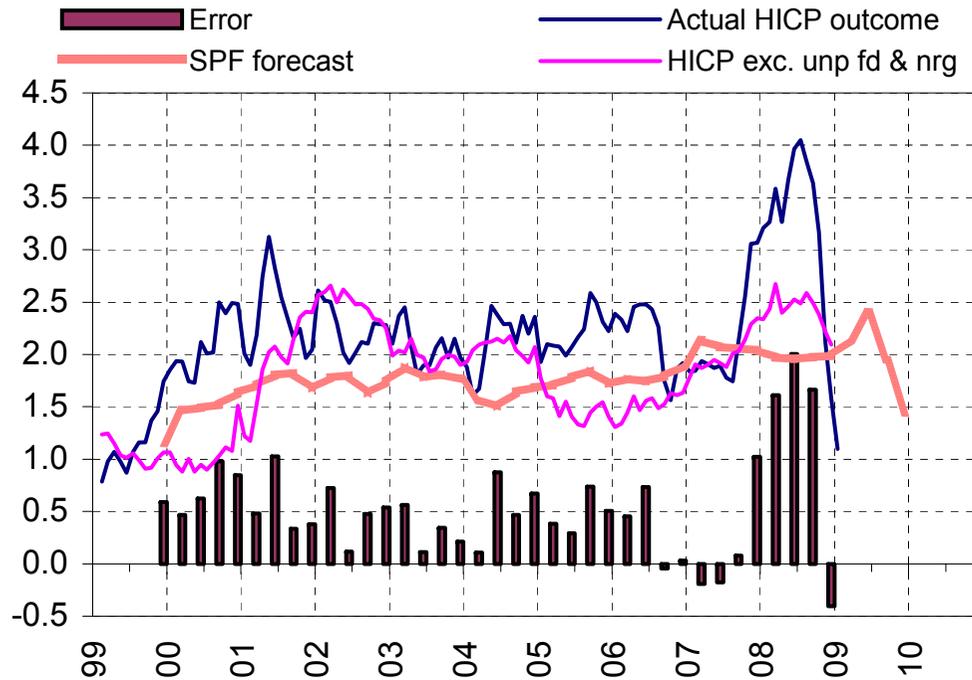
Main features of the SPF panel

- **quarterly survey**
- **conducted since 1999Q1**
- **euro area macroeconomic expectations for HICP inflation, real GDP growth and the unemployment rate**
- **short- and more medium- and longer-term horizons surveyed (including rolling and calendar year horizons)**
- **probability distributions**
- **qualitative answers also possible**
- **survey panel of financial and non-financial institutions**
- **75 active panellists (with average response of around 60) located throughout the EU**

(2) evaluation of short-term forecasts and performance to date

Summary: inflation forecasting performance

Chart: Inflation - 12 month ahead rolling horizons



Sample statistics (1999Q1-2008Q4)*

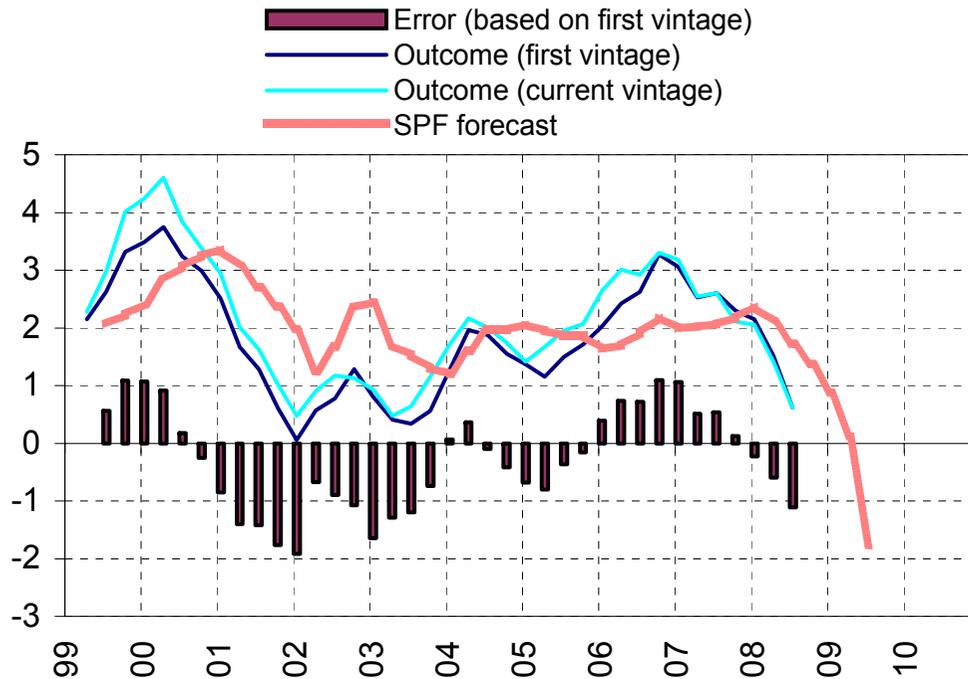
	Inflation	
Actual value		
Mean	2.2 (2.2)	
Std dev. (in pp)	0.6 (0.6)	
Forecast value	<i>1 year ahead</i>	<i>2 years ahead</i>
Mean	1.8	1.8
Std dev. (in pp)	0.2	0.1
Forecast error statistics	<i>1 year ahead</i>	<i>2 years ahead</i>
ME (in pp)	0.6 (0.5)	0.6 (0.5)
MAE (in pp)	0.6 (0.6)	0.6 (0.6)
RMSE (in pp)	0.8 (0.7)	0.8 (0.7)
Theil's U	1.0 (0.9)	1.0 (0.9)

* data in brackets are current vintage data; otherwise data refer to real-time

- inflation persistently under-forecast
- results broadly similar for one- and two-year ahead forecasts
- can largely be explained by specific shocks to food and energy

Summary: GDP growth forecasting performance

Chart: GDP - 12 month ahead rolling horizons



- persistence in growth forecast errors
- larger for two-year ahead errors (which are smoother)
- mean error sensitive to data vintage

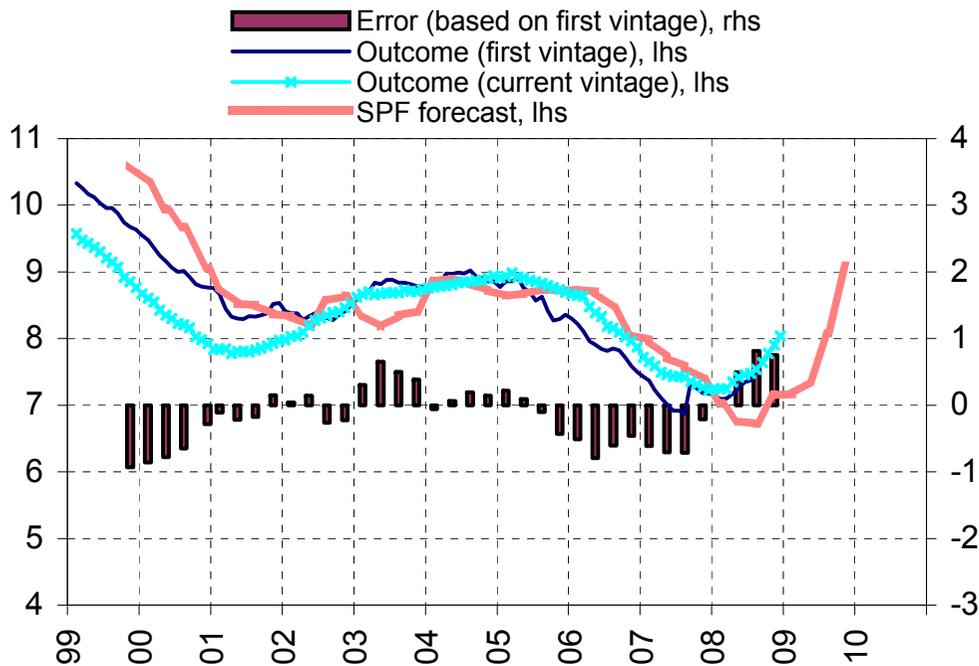
Sample statistics (1999Q1-2008Q4)*

	GDP	
Actual value		
Mean	1.8 (2.1)	
Std dev. (in pp)	1.0 (1.1)	
Forecast value	<i>1 year ahead</i>	<i>2 years ahead</i>
Mean	1.9	2.3
Std dev. (in pp)	0.9	0.4
Forecast error statistics	<i>1 year ahead</i>	<i>2 years ahead</i>
ME (in pp)	-0.3 (0.0)	-0.7 (-0.5)
MAE (in pp)	0.8 (0.8)	1.1 (1.0)
RMSE (in pp)	0.9 (1.0)	1.3 (1.2)
Theil's U	0.7 (0.7)	0.8 (0.7)

* data in brackets are current vintage data; otherwise data refer to real-time

Summary: unemployment forecasting performance

Chart: Unemp. - 12 month ahead rolling horizons



- little bias on average but persistent
- errors on real-time data slightly lower on average

Sample statistics (1999Q1-2008Q4)*

	Unemployment	
Actual value		
Mean	8.5 (8.3)	
Std dev. (in pp)	0.8 (0.6)	
Forecast value	<i>1 year ahead</i>	<i>2 years ahead</i>
Mean	8.4	8.1
Std dev. (in pp)	0.9	0.8
Forecast error statistics	<i>1 year ahead</i>	<i>2 years ahead</i>
ME (in pp)	-0.1 (-0.2)	-0.1 (-0.1)
MAE (in pp)	0.4 (0.5)	0.6 (0.7)
RMSE (in pp)	0.5 (0.7)	0.8 (0.9)
Theil's U	0.8 (1.2)	0.8 (0.9)

* data in brackets are current vintage data; otherwise data refer to real-time

Unbiasedness generally rejected using aggregate and panel regression

Regression tests for bias in the SPF aggregate forecasts

Variable	α	β	R^2	F-test (p-value)	μ (p-value)
GDP growth one year ahead	0.08 (0.68) [-0.11 (0.62)]	0.97 (0.34) [0.92 (0.30)]	0.22 [0.23]	0.992 [0.578]	0.946 [0.291]
GDP growth two years ahead	4.71 (1.37) * [5.28 (1.20) *]	-1.18 (0.57) * [-1.52 (0.50) *]	0.18 [0.30]	0.001 [0.000]	0.164 [0.045]
Unemployment one year ahead	4.73 (1.12) * [2.25 (0.94) *]	0.41 (0.13) [0.72 (0.10) *]	0.40 [0.70]	0.000 [0.001]	0.204 [0.231]
Unemployment two years ahead	8.73 (1.99) * [6.03 (1.78) *]	-0.06 (0.23) * [0.26 (0.19)] *	0.00 [0.06]	0.000 [0.000]	0.623 [0.594]

Note: Calculated using actual outcomes during the period 1999-2008. Standard errors are in round brackets (corrected using the Newey-West procedure). * indicates that the coefficient is significantly different from the null hypothesis (i.e. $\alpha = 0$ or $\beta = 1$) at the 10% level or lower. The p-value is from the F-test; bold indicates rejection of unbiasedness at the 10% level or lower. Figures in squared brackets refer to calculations made using the first estimate of data; otherwise the data refer to the current vintage of data.

Results from bias tests using pooled data of the SPF balanced panel

Variable	α	β	μ (p-value)
GDP one year ahead	1.64 (0.12) * [1.19 (0.11) *]	0.19 (0.03) * [0.28 (0.03) *]	0.770 [0.000]
GDP two years ahead	2.19 (0.16) * [2.07 (0.16) *]	-0.09 (0.06) * [-0.15 (0.06) *]	0.000 [0.000]
Unemp. one year ahead	6.64 (0.15) * [5.68 (0.18) *]	0.19 (0.02) * [0.32 (0.02) *]	0.000 [0.089]
Unemp. two years ahead	7.96 (0.25) * [8.25 (0.20) *]	-0.02 (0.01) * [-0.01 (0.02) *]	0.001 [0.022]

Note: Standard errors are in round brackets. * indicates that the coefficient is significantly different from the null hypothesis (i.e. $\alpha = 0$ or $\beta = 1$) at the 10% level or lower. Bold indicates rejection of the unbiasedness hypothesis at the 10% level or better.

* Note: Figures in square parentheses refer to calculations made using the first estimate of data; otherwise the data refer to the current vintage of data.

- similar results for inflation, although bias much smaller vis-à-vis HICP inflation excluding unprocessed food and energy

Considerable heterogeneity in expectations and forecast accuracy

Heterogeneity of forecast performance (rolling horizons)

(Root mean squared errors)

	HICP inflation		Real GDP growth		Unemployment rate	
	1-year ahead	2-years ahead	1-year ahead	2-years ahead	1-year ahead	2-years ahead
Minimum	0.5 (0.6)	0.4 (0.4)	0.7 (0.8)*	1.0 (1.1)	0.4 (0.4)	0.5 (0.5)
Average	0.7 (0.8)	0.7 (0.8)	1.0 (0.9)	1.2 (1.3)	0.7 (0.5)	0.9 (0.8)
Maximum	1.0 (1.0)	1.3 (1.3)	1.2 (1.2)	1.5 (1.7)	0.9 (0.6)	1.2 (1.0)

* Note: Figures in parentheses refer to calculations made using the first estimate of data; otherwise the data refer to the current vintage of data.

- but difficult to identify individual forecasters that are consistently good across forecast variables, horizons and sub-samples

Short-term forecasting and performance: Main conclusions

- Large and persistent errors with evidence of bias
- However, sample period (1999-2008) characterised by substantial shocks
- Considerable heterogeneity across forecasters, but difficult to identify consistently good or bad ones
- Little evidence of systematic differences across types of forecaster or nationality of forecaster

(3) Macroeconomic uncertainty according to the SPF

Different measures of uncertainty

The SPF provides several dimensions to measure uncertainty e.g.

- Using information about **point estimates** e.g. spread or standard deviation of point estimates
- Using information about **probability distribution** e.g. average std. dev. of individual or aggregation distributions, skew or kurtosis of distributions, event probability, etc.

A selection of alternative measures of uncertainty using SPF data

Measure	Calculation/description
1 disagreement	(quasi-) std. dev. of point est.
2 individual uncertainty	avg. std. dev. of individual dist.
3 aggregate uncertainty	std. dev. of aggregate dist.
4 combined uncertainty*	ind. uncertainty – disagreement
5 skew	mean of agg. dist – avg. point est
6 individual kurtosis	avg. kurtosis of individual dist
7 aggregate kurtosis	kurtosis of aggregate dist

* see *Giordani and Söderlind, 2003*

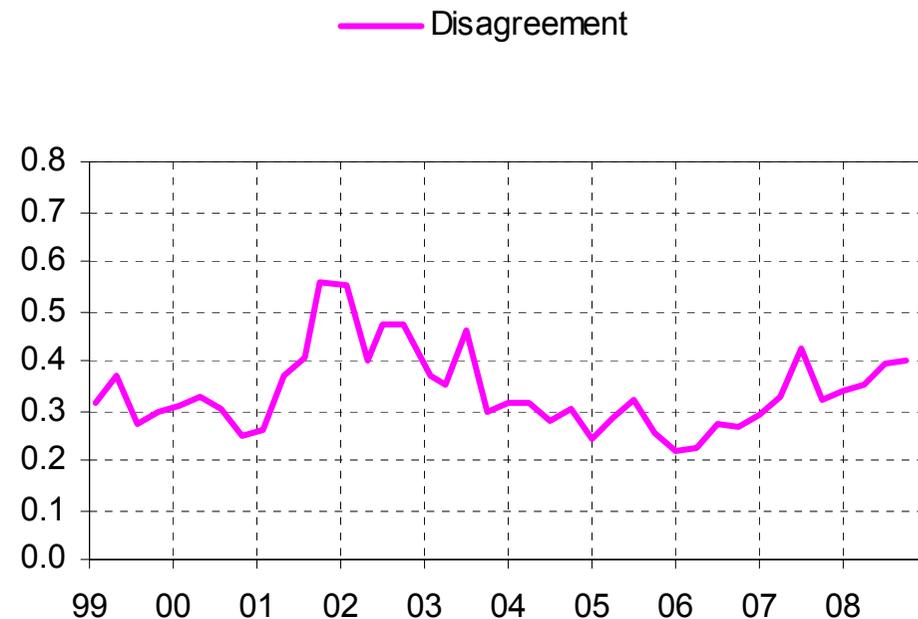
Our preferred measure of uncertainty

Standard deviation of aggregate distribution as it is a function of:

- (a) standard deviation of point estimates
(i.e. **disagreement** among forecasters)
- (b) average standard deviation of individual distributions
(i.e. **average individual uncertainty**)

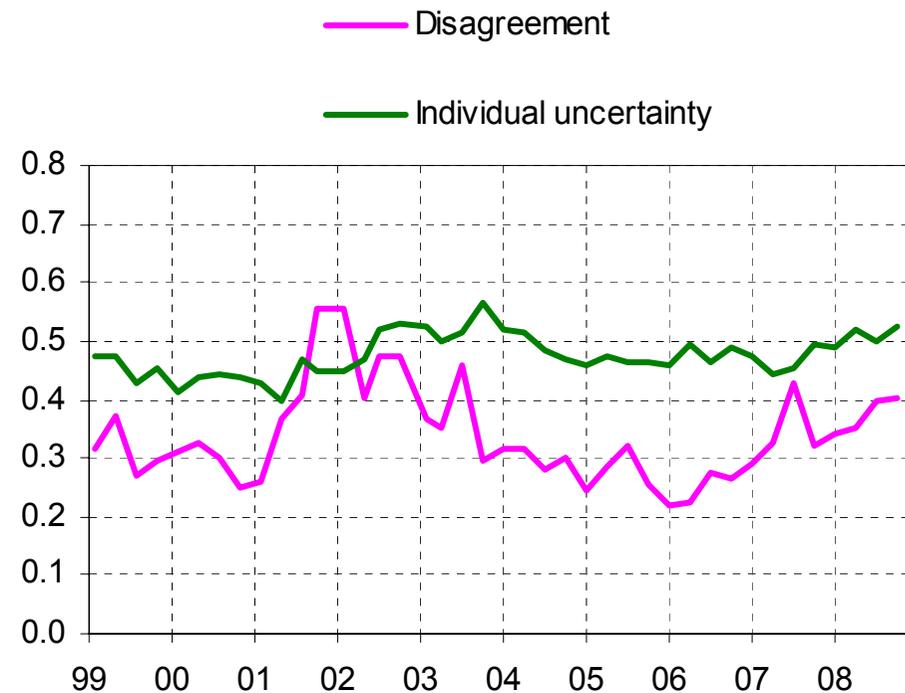
Short-term GDP growth uncertainty

- Level of disagreement 0.35 p.p. on average, but fluctuated in range 0.2 p.p. to 0.6 p.p.



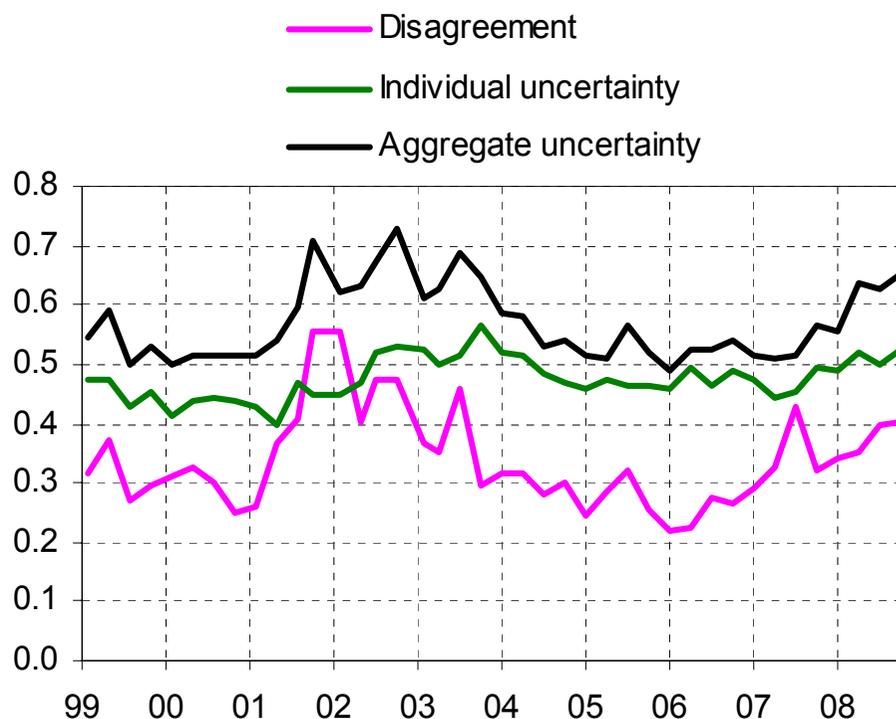
Short-term GDP growth uncertainty

- Level of **disagreement** 0.35 p.p. on average, but fluctuated in range 0.2 p.p. to 0.6 p.p.
- **Individual uncertainty** has been slightly higher, around 0.5 p.p. on average, and more stable

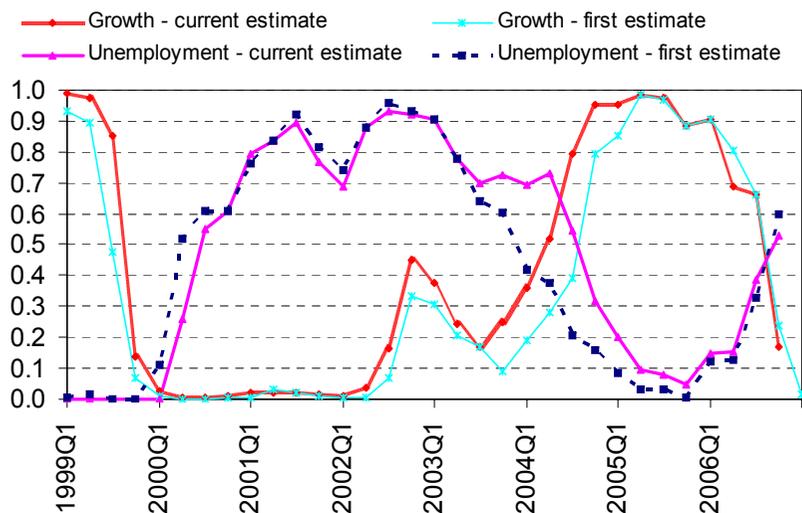
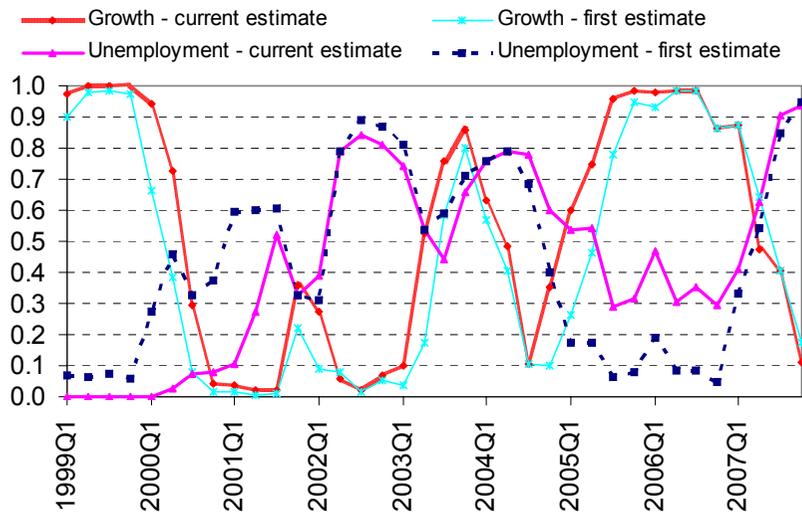


Short-term GDP growth uncertainty

- Level of **disagreement** 0.35p.p. on average, but fluctuated in range 0.2p.p. to 0.6p.p.
 - **Individual uncertainty** has been slightly higher, around 0.5p.p. on average, and more stable
 - **Aggregate uncertainty** has averaged around 0.6p.p., with cyclical movements mainly driven by disagreement
- Respondents appear not to have captured fully nature and extent of actual uncertainty, particularly when one takes into consideration upward bias (see *Stuart & Ord 1994*) in uncertainty measures above – this is a general result across variables and horizons²²



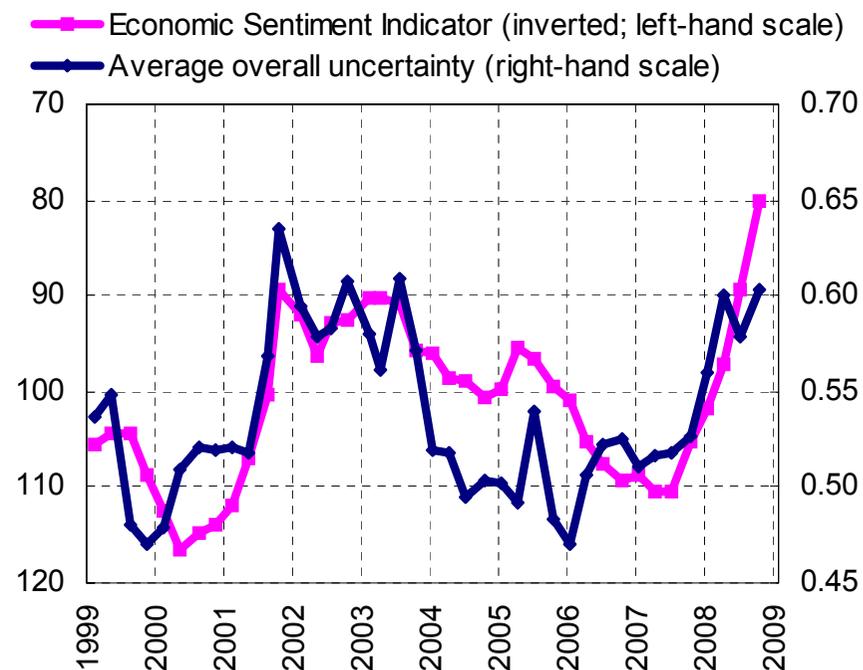
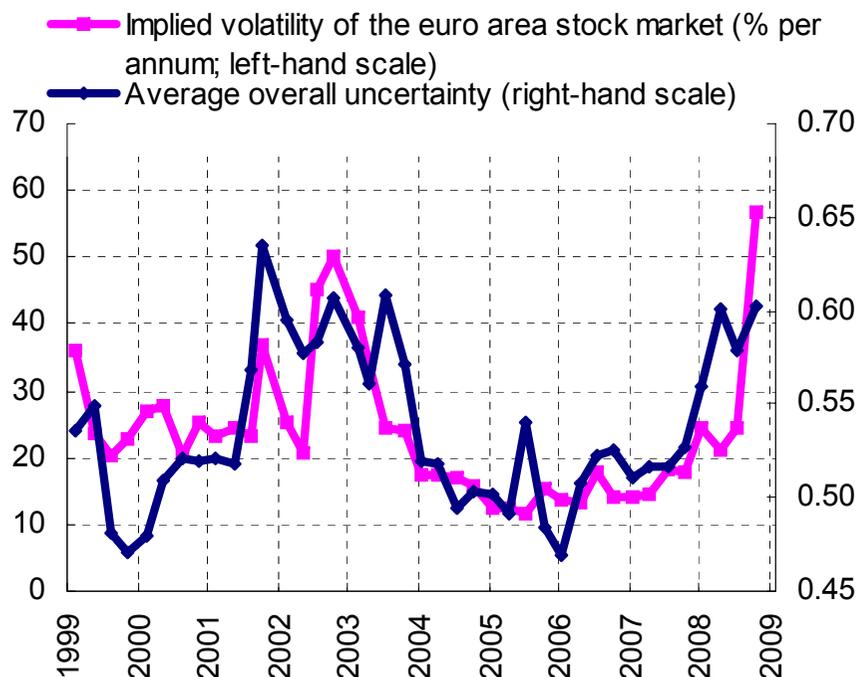
Forecast uncertainty and the Probability Integral Transform (PIT)...



- PIT analysis confirms that respondents representation of uncertainty surrounding their forecasts did not capture fully the nature and extent of the uncertainty as revealed by eventual outcomes

- Even when allowing for fact of overlapping errors, PIT shows long-lasting autocorrelation and a preponderance of outcomes in upper or lower

Uncertainty correlates strongly with other business cycle indicators



- GDP growth and unemployment uncertainty indicators from SPF correlate strongly with other proxies such as:
 - implied stock market volatility
 - business/consumer tendency surveys

Quantitative measures of skew and kurtosis may also be informative

Skew and Kurtosis indicators, Sample averages: 1999-2008

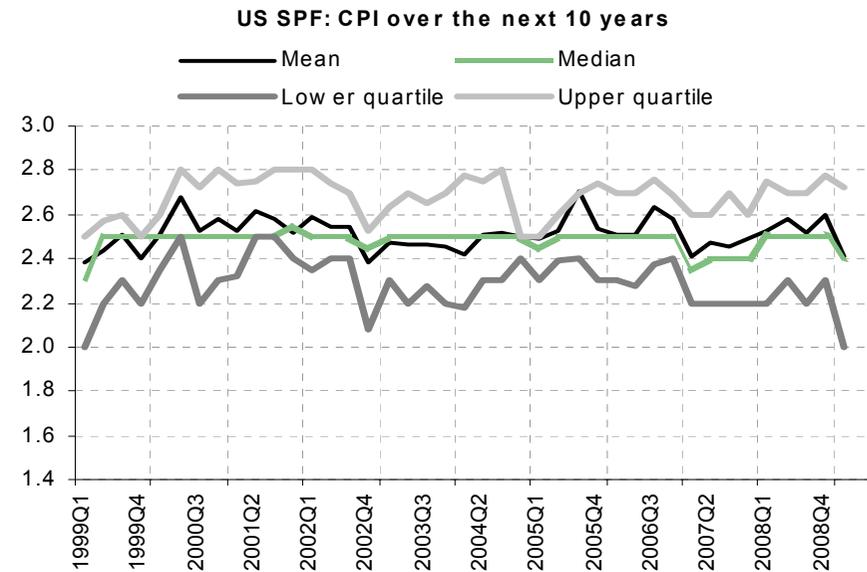
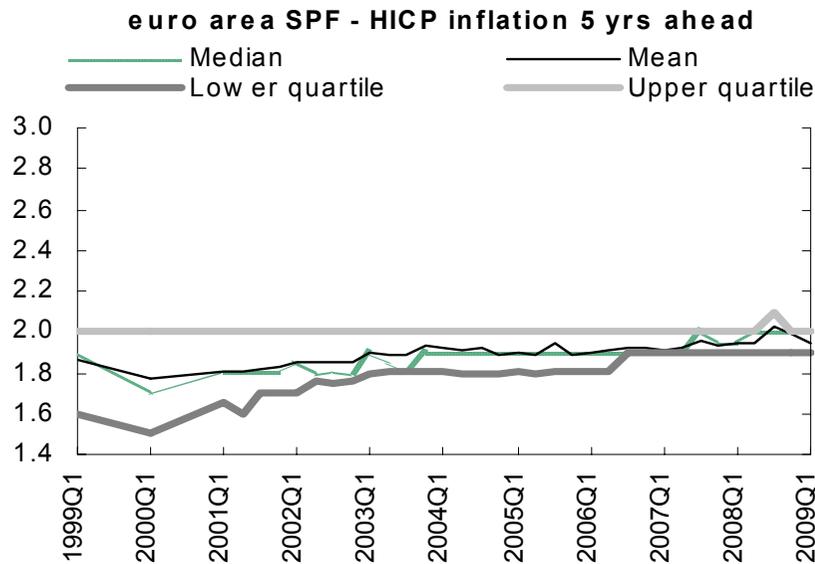
	HICP inflation	GDP growth*	Unemployment*
1. “Proxy skew”	0.07	-0.04	0.03
2. “Individual kurtosis”	-0.17	-0.35	-0.16
3. “Aggregate kurtosis”	0.91	1.17	1.86

Average value since the first quarter of 1999 SPF round for rolling one-year ahead horizon.

- A general result across variables is that negative individual kurtosis highlights a relatively low weight assigned by individuals to more extreme or tail events that is captured by aggregate distribution
- Measures of skew are more variable and horizon dependent. For the short-term inflation distribution has been skewed to the upside. There is less skew evident for longer-term inflation expectations. GDP growth distribution systematically skewed to the downside both for short- and longer-term horizons. Results for unemployment are opposite of GDP growth

(4) Longer-term inflation expectations of SPF participants

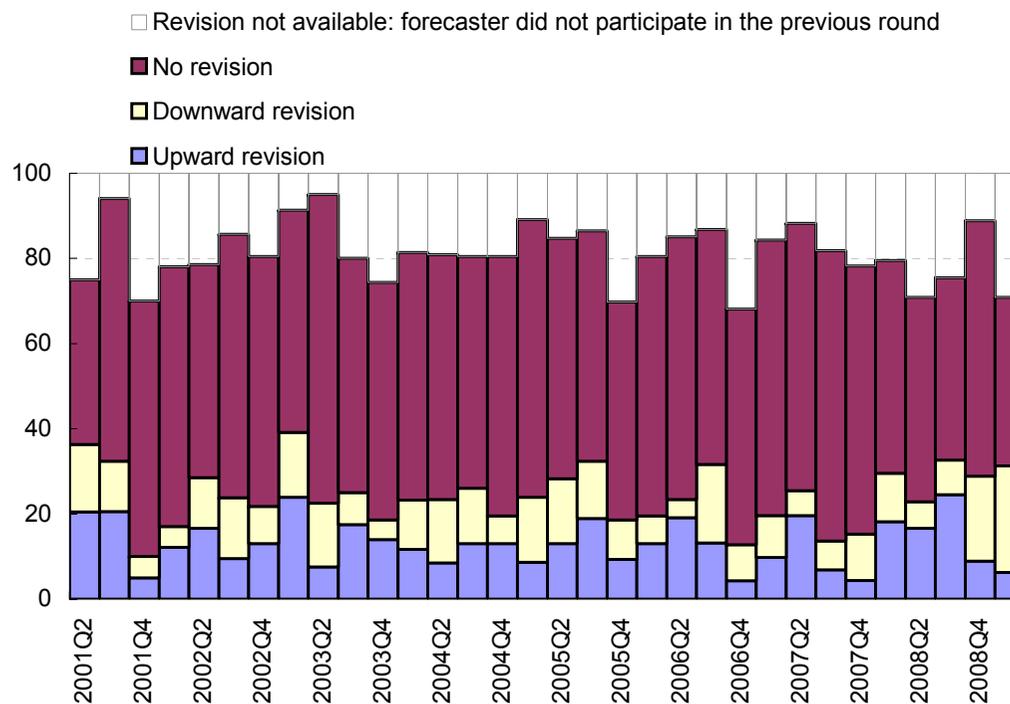
Longer-term inflation expectations



- Average longer-term inflation expectations in line with ECB definition of price stability; also less dispersed over time

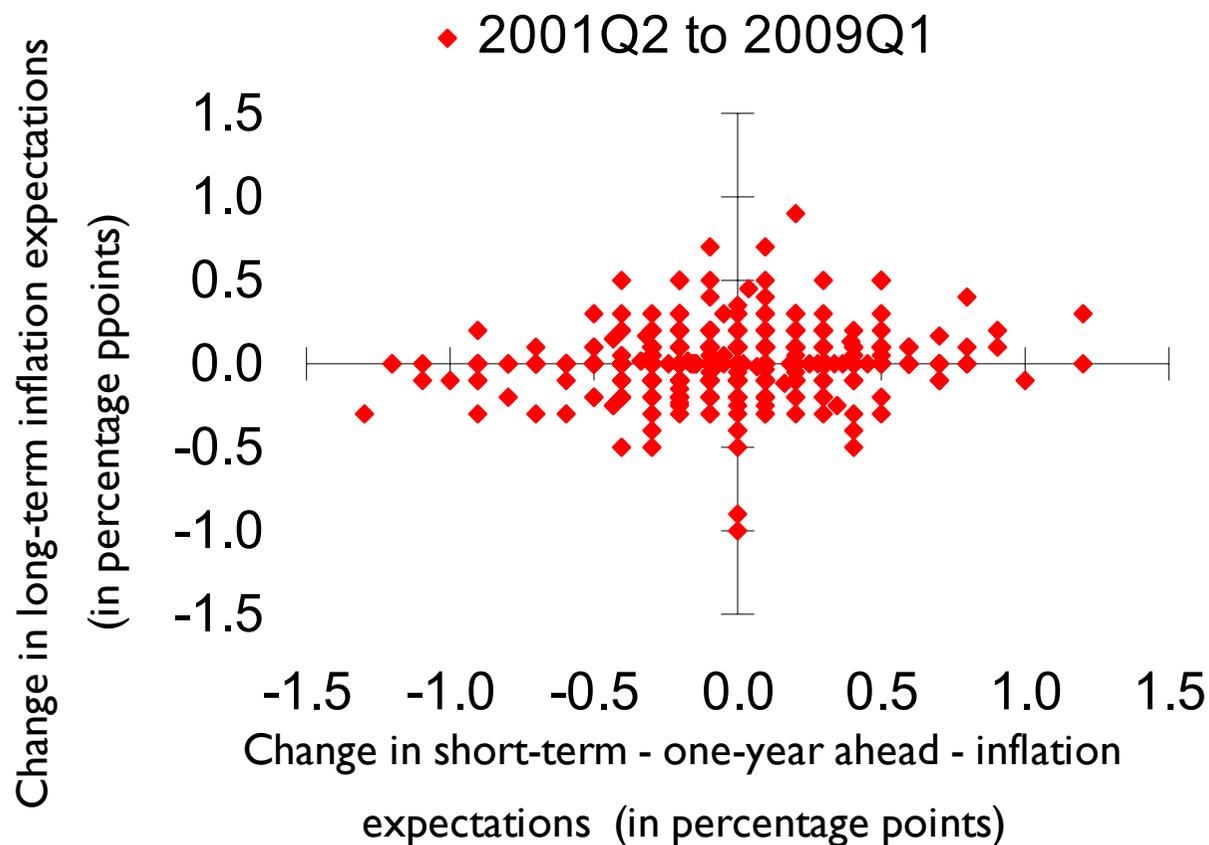
But stability does not mean inertia

Revisions to longer-term inflation expectations from one SPF round to the next
(as a percentage of respondents who have provided a point estimate)



- On average around 30% of respondents change their longer-term inflation forecast each round

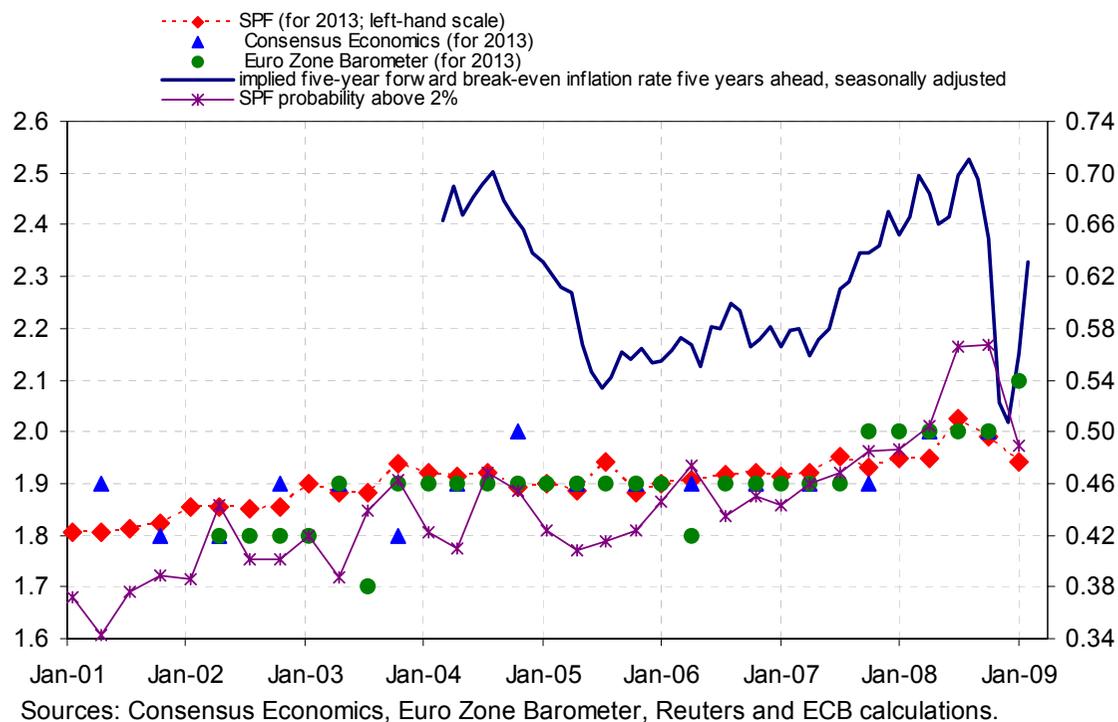
Well-anchored LT inflation expectations



- No significant relationship between changes in short-term inflation forecasts and those in longer-term expectations (see *Castlenuovo et al., 2003*)

Cross-checking survey and financial market based measures of longer-term inflation expectations

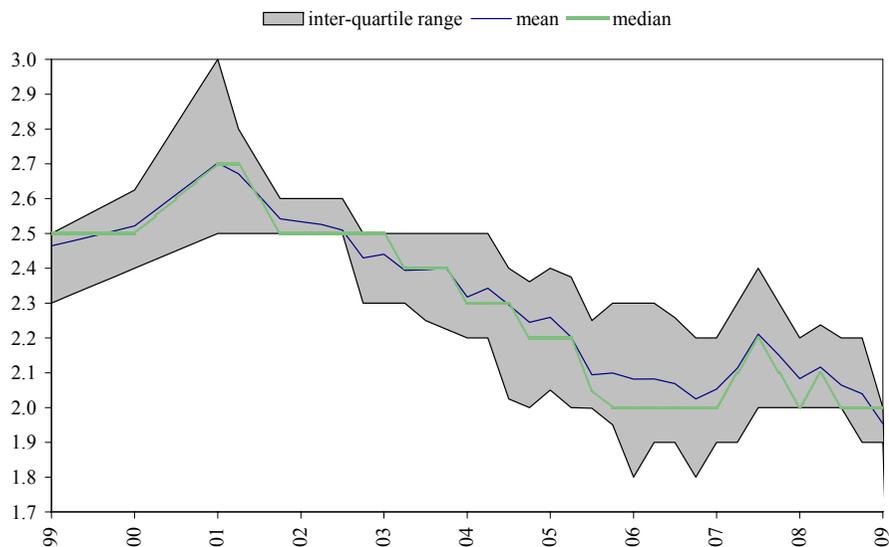
Longer-term inflation expectations from surveys and break-even inflation rates



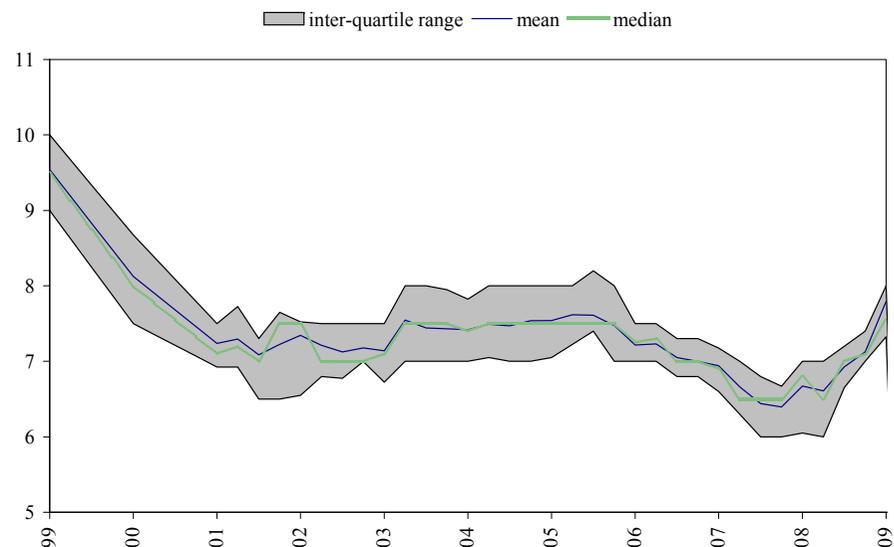
- Although average longer-term inflation expectations from surveys have remained stable, there has been some movement in risk assessment; consistent with inflation risk premium in measures of inflation expectations extracted from index-linked bonds (see, e.g., *Kwan 2005*)

Longer-term growth and unemployment expectations

GDP growth



Unemployment rate



- Balance of risks to longer-term growth expectations has generally been judged to the downside, while those for the unemployment rate have been to the upside
- For longer-term unemployment expectations there has been a considerable correlation between revisions to short-term expectations and longer-term expectations, indicating SPF respondents perceive significant hysteresis

Concluding comments

- Evidence of persistent errors and possible bias in inflation, growth and unemployment rate expectations from professional forecasters (small sample caveat). Errors are comparable with those in other surveys (e.g. Consensus Economics).
- Notwithstanding significant heterogeneity in forecast accuracy (MSE) across survey participants, errors are largely common across forecasters. Dispersion of point forecasts provides a poor indication of the true level of uncertainty.
- Beyond point forecasts, SPF also provides insight into other aspects such as uncertainty and longer-term expectations. Some evidence that SPF respondents appear not to have captured fully nature and extent of actual macroeconomic risks
- Overall empirical expectations in line with the recent emphasis on learning and sticky information rather than fully rational expectations.

Future work

- Much depends on a longer time span of data
 - Revisit conclusions when more data are available
 - Use in macroeconomic models with forward looking behaviour
- Also given overview nature of paper, many more focused questions such as forecast heterogeneity and combination, properties of probability distributions remain to be investigated in more detail
- Investigate and improve understanding of expectations formation process

An overview of the ECB Survey of Professional Forecasters (SPF): 1999-2008

Aidan Meyler
European Central Bank

February 2009

Results from the ECB's SPF Special Questionnaire

Ieva Rubene

Philadelphia, February 2009

Questionnaire prepared and conducted jointly with Aidan Meyler

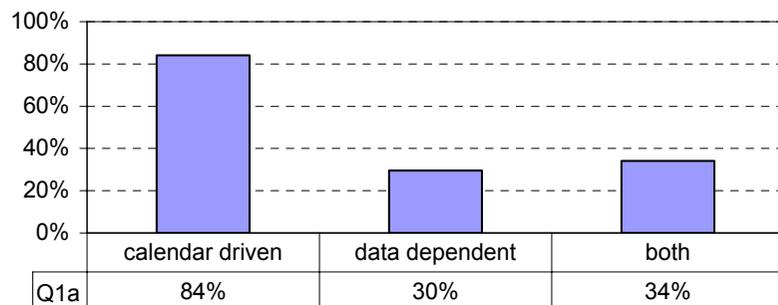
We thank Tom Stark (Federal Reserve Bank of Philadelphia), Prof. Kenneth Wallis (University of Warwick), and Aurelio Maccario (UniCredit Group)

Motivation

- **10 years of the survey of professional forecasters**
- **Insights into forecasting methods and models**
- **Open issues on:**
 - **timeliness of the survey**
 - **probability distributions:**
 - **which central tendency reported (mean, median, mode)**
 - **estimation method (model versus judgment)**
 - **symmetric versus asymmetric loss function**
 - **importance of point estimate versus probability distribution**
- **Feedback to forecasters and forum for discussions**

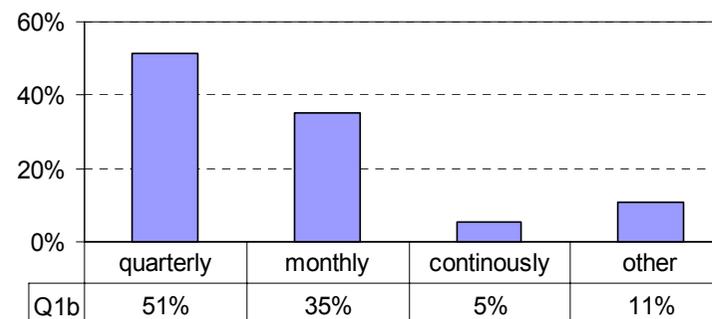
I Frequency of the updates and data

Chart 1a When do you update your forecasts?



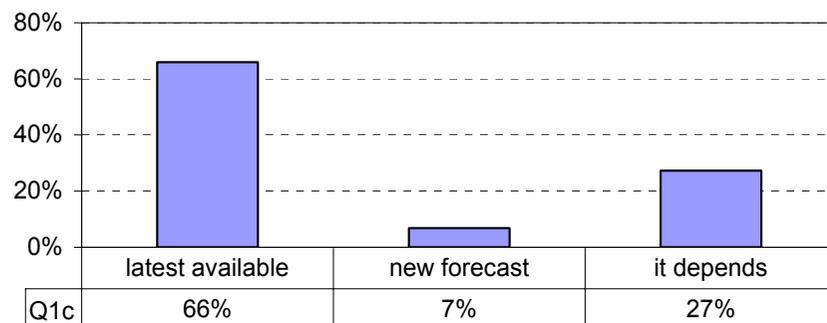
Note: may sum to more than 100% as some respondents report both categories. In addition, some of those who responded calendar-driven also stated they might occasionally update based on new data/shocks.

Chart 1b If it is calendar driven, how often do you update your forecasts?



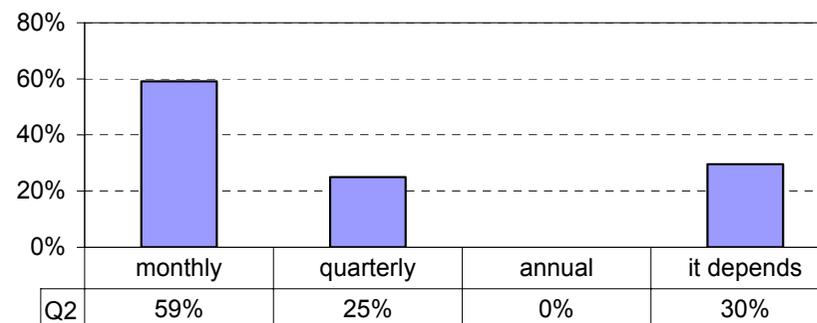
Note: may sum to more than 100% as some respondents report both categories.

Chart 1c When responding to the SPF do you provide...



Note: many of those providing qualitative information indicated that they may do a partial update when responding to the SPF, if changes are significant to do it.

Chart 2 What is the highest frequency of data at which you model / forecast?



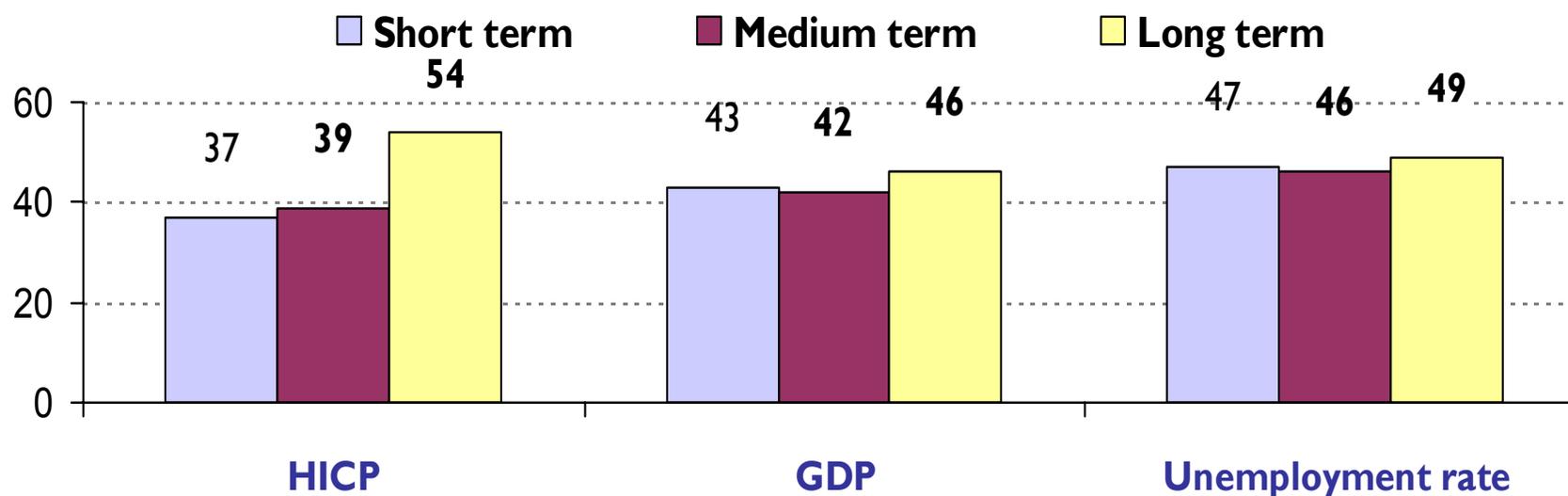
Note: Many respondents reported that they follow the frequency of the underlying variable being forecast (i.e. monthly for inflation, quarterly for GDP)

2a Role of judgment

Number of respondents reporting: judgment = 100%

	HICP inflation	GDP growth	Unemployment rate
Short term (one year or less)	5	5	8
Medium term (up to two years)	6	6	7
Long term (five years ahead)	10	8	9

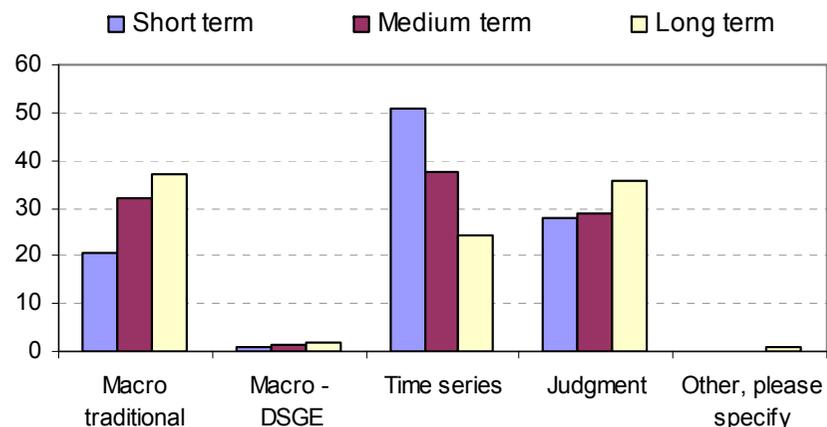
Chart 3a Judgment applied to the forecast, %



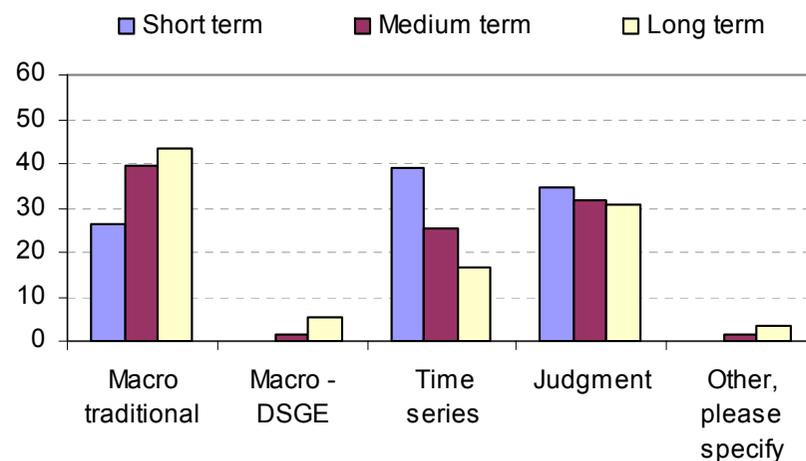
Note: calculations include responses with 100% judgement.

2b Model versus judgement

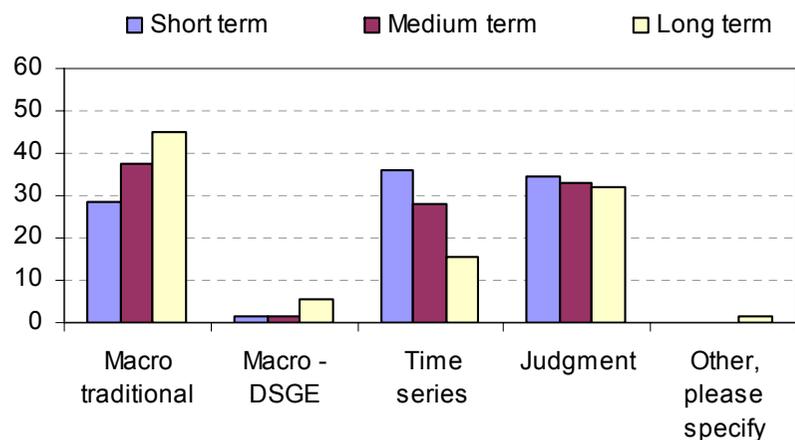
HICP Inflation



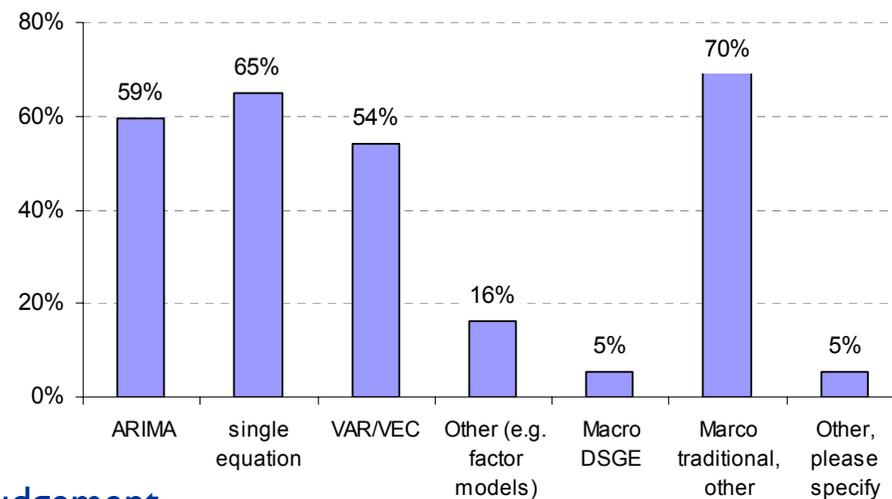
GDP growth



Unemployment rate



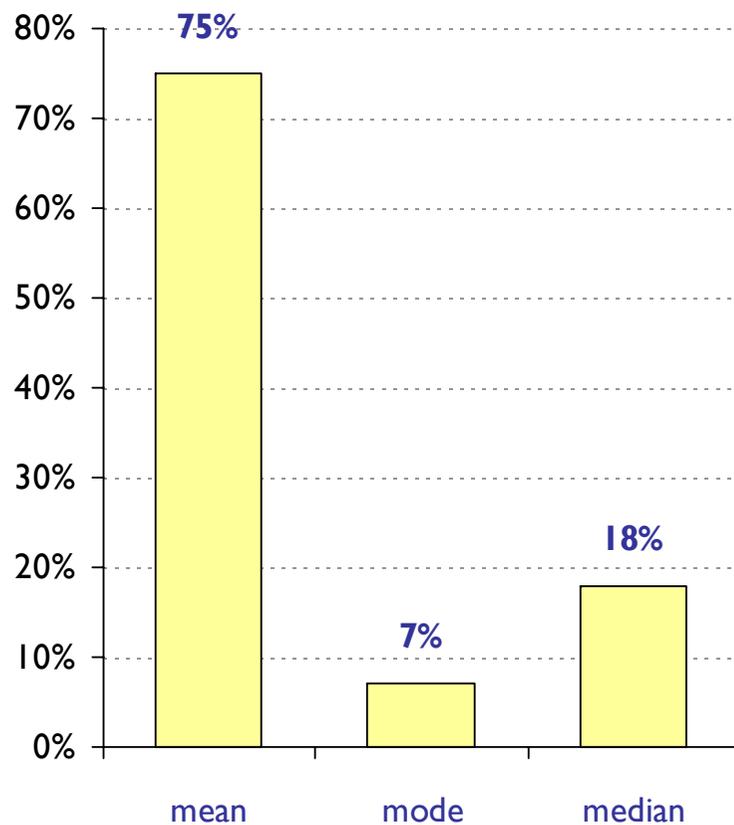
What model do you use?



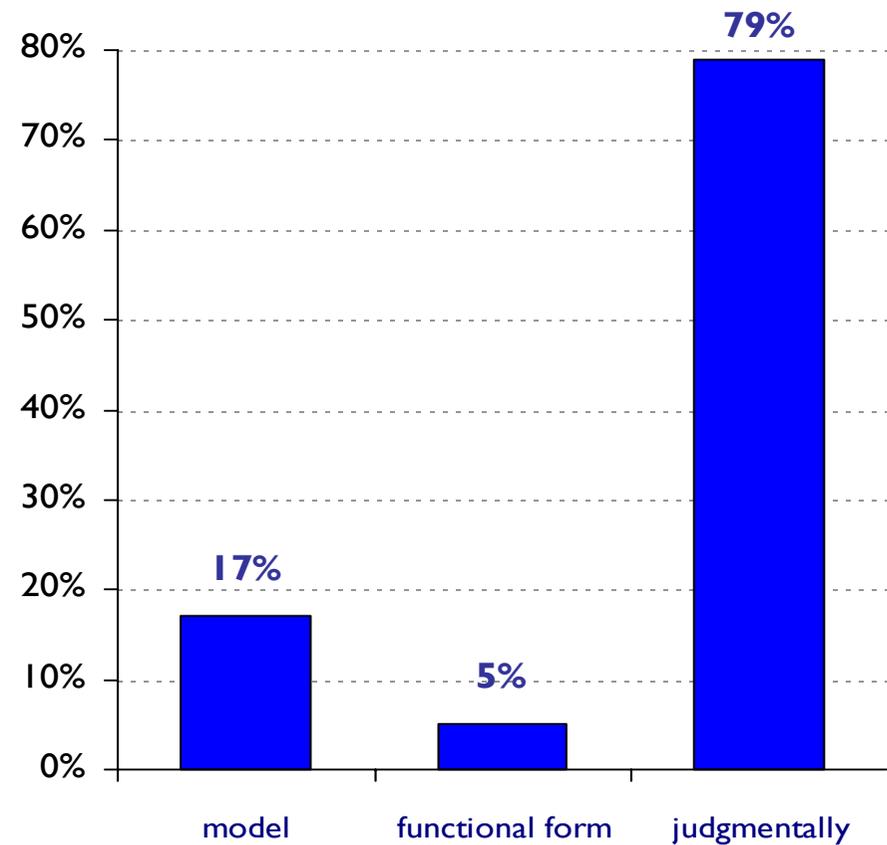
Note: calculations exclude responses with 100% judgement.

3 Probability distributions

Do you report the mean, mode or median of the probability distribution?

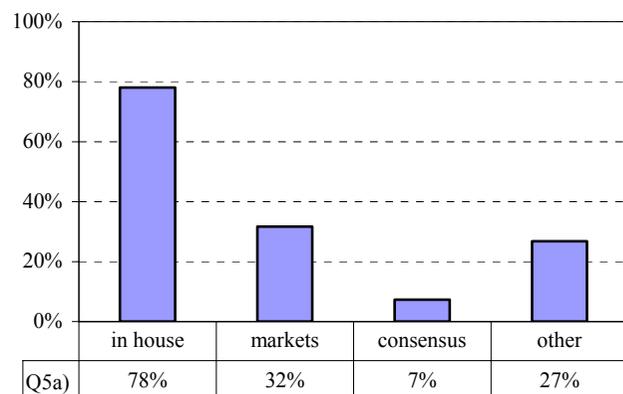


How do you generate your reported probability distribution?



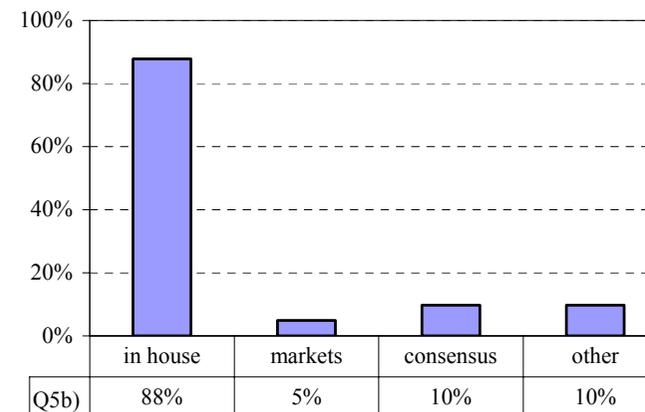
4 External assumptions – how derived?

Oil prices



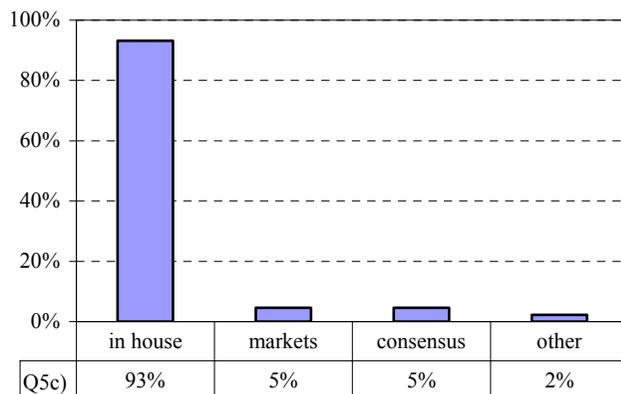
Note: may sum to more than 100% as some respondents report both categories.

Exchange rate



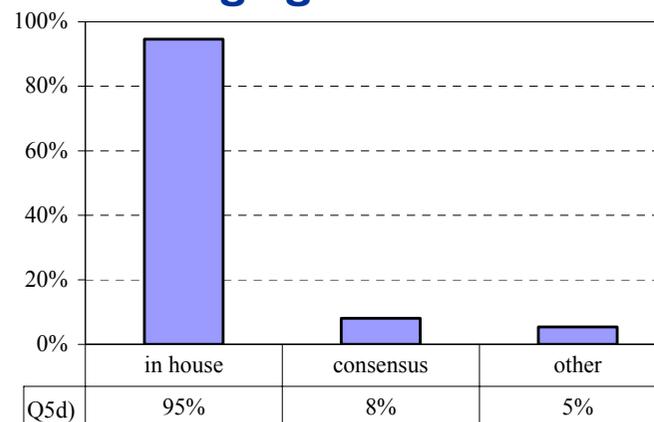
Note: may sum to more than 100% as some respondents report both categories.

ECB's interest rate



Note: may sum to more than 100% as some respondents report both categories.

Wage growth



Note: may sum to more than 100% as some respondents report both categories.

Conclusions

- **SPF responses likely to be timely despite many being updated on a calendar basis**
- **Key role for judgement both in point forecasts but more particularly for probability distributions**
 - **Don't over-interpret individual probability distributions**
- **Time series models dominate for shorter horizons; for longer-term horizons it is (traditional) macro models**
- **Forecasts based on heterogeneous assumptions, generated in-house**
 - *Lesson learned: one cannot be too careful when phrasing the questions*

The ECB Survey of Professional Forecasters

Aidan Meyler and Ieva Rubene
European Central Bank

February 2009