Predicting inflation: Professional experts versus no-change forecasts

Thordis L. Thorarinsdottir
Heidelberg University

Eltville am Rhein, June 1 2012

Joint work with Tilmann Gneiting

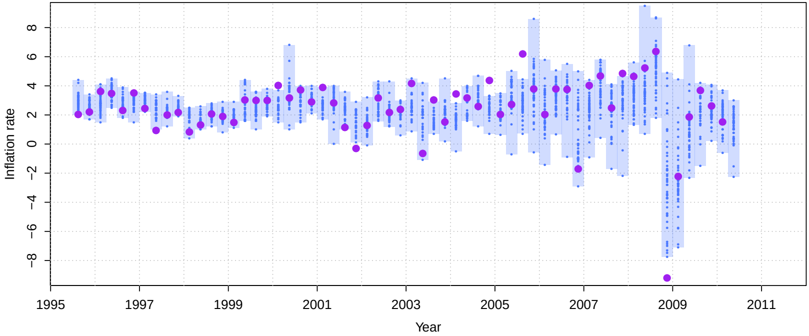
Survey of Professional Forecasters

The Survey of Professional Forecasters is the oldest quarterly survey of macroeconomic forecasts in the United States. The survey began in 1968 and was conducted by the American Statistical Association and the National Bureau of Economic Research. The Federal Reserve Bank of Philadelphia took over the survey in 1990.

The survey asks panelists to give their estimate for various macroeconomic variables such as GDP, inflation, unemployment, etc. for quarters and years ahead.

We focus on inflation forecasts, that is, quarter-over-quarter changes of the CPI expressed in annualized percentage points.

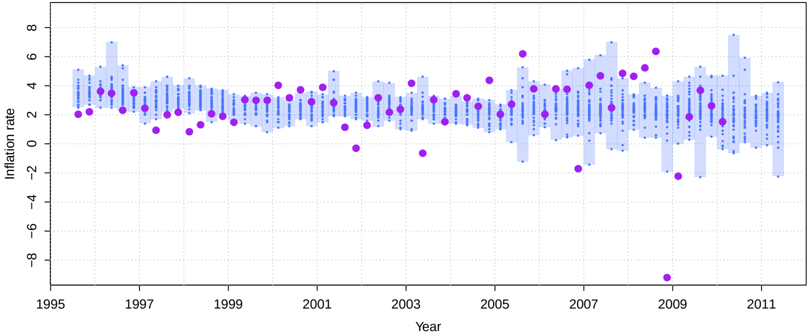
The forecasts are available for prediction horizons of 1 to 5 quarters with an average of about 30 individual forecasts each quarter.



One quarter ahead SPF forecasts of U.S. inflation rates and the realized values.

We focus on inflation forecasts, that is, quarter-over-quarter changes of the CPI expressed in annualized percentage points.

The forecasts are available for prediction horizons of 1 to 5 quarters with an average of about 30 individual forecasts each quarter.



Five quarters ahead SPF forecasts of U.S. inflation rates and the realized values.

To assess the predictive performance of the SPF forecasts we need

1. statistically sound evaluation methods that are based on decision theoretic principles
2. appropriate reference forecasts

Proper scoring rules

If $\mathcal{F}$ denotes a class of probabilistic forecasts on $\mathbb{R}$, a **proper scoring rule** is any function

$$\mathrm{R} : \mathcal{F} \times \mathbb{R} \to \mathbb{R}$$

such that

$$\mathbb{E}_F \, \mathrm{R}(F, Y) \leq \mathbb{E}_F \, \mathrm{R}(G, Y) \quad \text{for all} \quad F, G \in \mathcal{F}.$$

An example of a proper scoring rule is the **continuous ranked probability score**,

$$\begin{aligned} R_{CRP}(F, y) &= \int [F(x) - \mathbb{1}\{x \geq y\}]^2 dx \\ &= \mathbb{E}_F |X - y| - \frac{1}{2} \mathbb{E}_F \mathbb{E}_F |X - X'|. \end{aligned}$$

Proper scoring rules based on scoring functions

The **absolute error** of a forecast x and observation y is the scoring function

$$S(x, y) = |x - y|.$$

If the forecast is given by a distribution F, we use the Bayes predictor for F under the absolute error loss,

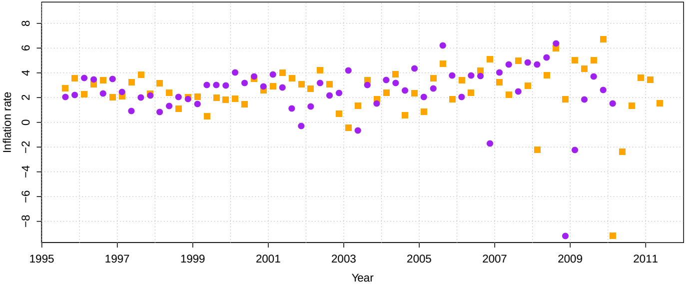
$$\arg\min_x \mathbb{E}_F |x - Y| = \mathrm{median}(F).$$

This results in the proper scoring rule

$$R_{AE}(F, y) = |\mathrm{median}(F) - y|.$$

Traditional no-change forecast (TNC)

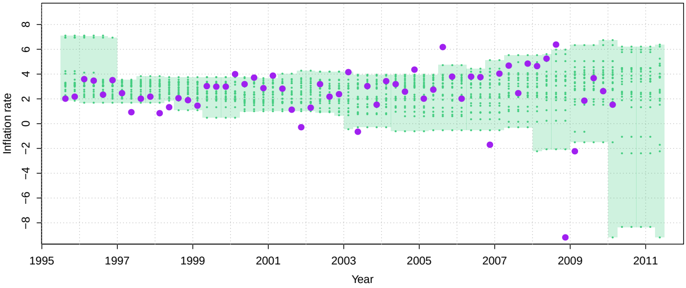
The most recent observation available at the issue time



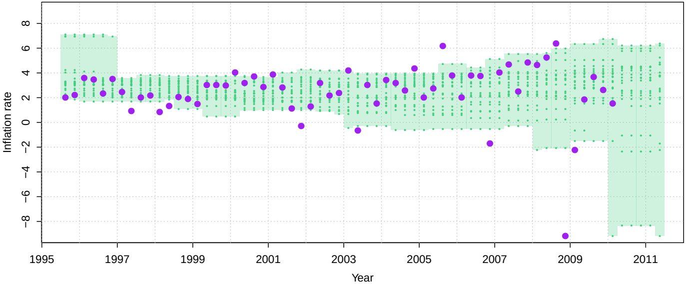
- Is the Bayes predictor under a symmetric **random walk model**
- Is constant across prediction horizons
- Can be made probabilistic with a Gaussian kernel which variance equals the mean squared error of recent forecasts

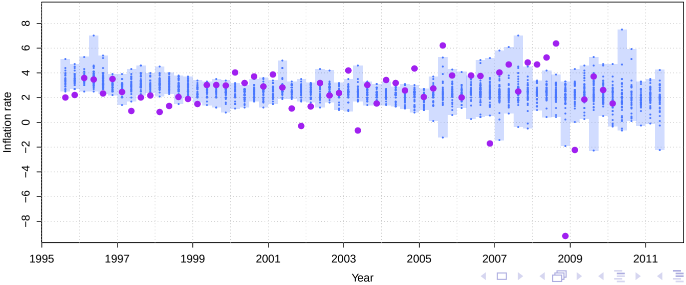
Probabilistic no-change forecast (PNC)

An ensemble of the *m* most recent observations available



- Corresponds to a **white noise model**
- Is identical for all prediction horizons

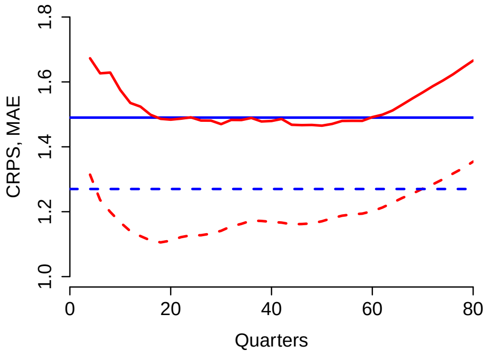

Inflation rate
8
6
4
2
0
−2
−4
−6
−8
1995
1997
1999
2001
2003
2005
2007
2009
2011
Year


Inflation rate
8
6
4
2
0
−2
−4
−6
−8
1995
1997
1999
2001
2003
2005
2007
2009
2011
Year

	Prediction horizon in quarters				
	1	2	3	4	5
Mean absolute error					
SPF	0.89	1.44	1.51	1.49	1.49
PNC (m=20)	1.45_{99}	1.46_{58}	1.45_{18}	1.48_{39}	1.48_{44}
TNC	1.81_{99}	2.06_{96}	2.00_{91}	2.06_{95}	2.03_{97}
Mean continuous ranked probability score					
SPF	0.69	1.16	1.25	1.26	1.27
PNC (m=20)	1.08_{99}	1.10_{21}	1.10_{00}	1.10_{00}	1.11_{00}
TNC	1.56_{99}	1.66_{97}	1.51_{82}	1.57_{90}	1.47_{87}

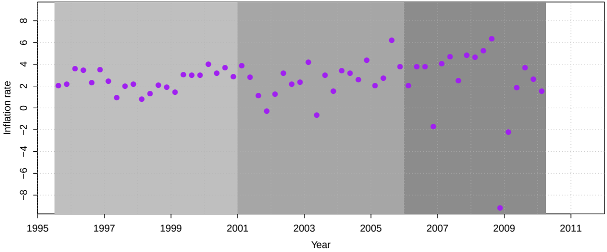
Average results over the third quarter of 1995 to the first quarter of 2010.

Robustness of the PNC forecast



Average performance of the PNC forecast for five quarter ahead forecasts as a function of m, compared to the corresponding scores for the SPF forecast.

However...



- Scores are generally low in the 1st period; TNC performs well
- PNC is a better reference forecast for the 2nd and 3rd period
- PNC outperforms SPF for longer lead times in the 2nd period but not in the 3rd period

Can we do better?

If there is a structure in the forecasts errors over time, statistical postprocessing methods can be successfully applied to improve the forecasts.

Can we do better?

If there is a structure in the forecasts errors over time, statistical postprocessing methods can be successfully applied to improve the forecasts.

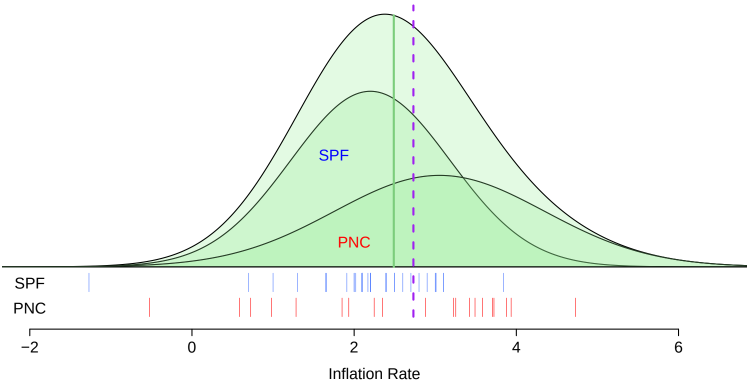
One option is to combine the SPF and the PNC forecasts using a **Gaussian mixture model**,

$$F_{GM}(y) = \alpha\Phi\left(\frac{y-\mu_1}{\sigma_1}\right) + (1-\alpha)\Phi\left(\frac{y-\mu_2}{\sigma_2}\right),$$

where $\alpha \in [0,1]$, $\sigma_1, \sigma_2 > 0$ and

$$\mu_1 = \mu_{SPF}, \quad \mu_2 = \mu_{PNC}.$$

In each quarter, we estimate $\alpha, \sigma_1, \sigma_2$ using data from the 40 previous quarters.



The Gaussian mixture forecast for the second quarter of 2005 at a prediction horizon of two quarters with the corresponding observation.

	Prediction horizon in quarters				
	1	2	3	4	5
Mean absolute error					
SPF	0.89	1.44	1.51	1.49	1.49
PNC (m=20)	1.45_{99}	1.46_{58}	1.45_{18}	1.48_{39}	1.48_{44}
GM	0.89_{91}	1.47_{69}	1.50_{34}	1.48_{39}	1.45_{16}
Mean continuous ranked probability score					
SPF	0.69	1.16	1.25	1.26	1.27
PNC (m=20)	1.08_{99}	1.10_{21}	1.10_{00}	1.10_{00}	1.11_{00}
GM	0.68_{36}	1.10_{20}	1.11_{00}	1.12_{00}	1.10_{00}
$N(\mu_{SPF}, \sigma_{SPF})$	0.70_{53}	1.10_{39}	1.15_{36}	1.14_{34}	1.13_{31}

Average results over the third quarter of 1995 to the first quarter of 2010.

- Forecast evaluation methods should be based on decision theoretic principles
  $\longrightarrow$ *This can be attained by applying proper scoring rules*
- For a meaningful result concerning predictive performance, we must compare our prediction model to an appropriate reference forecast
  $\longrightarrow$ *In situations like ours, both TNC and PNC should be considered*
- Statistical postprocessing methods can significantly improve the predictive performance of a forecasting procedure if we can identify a structure in the forecast errors