US inflation expectations and heterogeneous loss functions, 1968–2010

Michael P. Clements

No 986

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF WARWICK

US inflation expectations and heterogeneous loss functions, 1968–2010

Michael P. Clements
Department of Economics,
University of Warwick,
UK. CV4 7AL.
m.p.clements@warwick.ac.uk.

March 28, 2012

Abstract

The recent literature has suggested that macroeconomic forecasters may have asymmetric loss functions, and that there may be heterogeneity across forecasters in the degree to which they weigh under and over-predictions. Using an individual-level analysis that exploits the SPF respondents' histogram forecasts, we find little evidence of asymmetric loss for the inflation forecasters.

Journal of Economic Literature classification: C53, E31, E37

Keywords: Disagreement, forecast uncertainty, asymmetric loss, Survey of Professional Forecasters.

1 Introduction

In this paper we examine US inflation expectations over the last forty years (1969 to 2010) making use of the Survey of Professional Forecasters data. Our particular interest is in whether the forecasters have asymmetric loss functions. In the extreme case, asymmetric loss permits forecasters to be 'full-information rational-expectations' (FI-RE) forecasters while still accounting for the empirical observation that 'forecasters disagree', provided that their loss functions exhibit different degrees of asymmetry.

There are a number of other explanations of why forecasters disagree, such as sticky information (see, e.g., Mankiw and Reis (2002), Mankiw, Reis and Wolfers (2003) and Coibion and Gorodnichenko (2011)), noisy information (Woodford (2001), Sims (2003)) as well as different views about the long-run values of variables like inflation and output growth (e.g., Patton and Timmermann (2010, 2011)) and different signals and interpretation of those signals (e.g., Kandel and Zilberfarb (1999), Lahiri and Sheng (2008) and Manzan (2011)). Our focus is the explanation proposed by Capistrán and Timmermann (2009) which stresses heterogeneity in the degree of asymmetry in loss functions, whereby individuals weigh costs to over and under-predictions differently.

We consider the evidence for asymmetric loss using a testing procedure originally applied by Clements (2009) to the output growth forecasts. This approach makes use of the inflation histograms reported by the SPF respondents. Clements (2010) applies the approach of Clements (2009) to the SPF inflation forecasts but in a panel data context. That analysis generally was not supportive of the hypothesis that the SPF respondents are FI-RE forecasters with asymmetric loss functions. We re-visit the issue of the inflation forecasts and asymmetric loss using an individual-level analysis (as in the application by Clements (2009) to the output growth forecasts) and allow the degree of asymmetry to differ across respondents. A second innovation is to fit the generalized beta distribution to the histograms to calculate forecast moments, as opposed to normal distributions as in earlier work. This should lessen possible distortions to the moment estimates from fitting a symmetric distribution such as the normal. As argued by Giordani and Söderlind (2003), a normal distribution may provide a reasonable approximation to the aggregate histograms, but may be less appropriate for the individual histograms. We follow Engelberg, Manski and Williams (2009) in fitting generalized beta distributions to the SPF histograms, which we regard as current 'best practice', in recognition that our results on the empirical relevance of loss function asymmetry may be sensitive to the way in which forecast moments are calculated from the histograms. We also consider only those histograms which provide enough information on the underlying subjective distribution to enable such a distribution to be fit. We would expect this to further reduce measurement error in our estimates of forecast moments. We also check the sensitivity of the results to the distribution we fit to the histograms. Thirdly, we consider the point predictions and histogram means as rival forecasts in a forecast encompassing framework, where the expectation is that the latter should forecast encompass the former in terms of predicting outcomes, if agents have asymmetric loss.

The SPF is chosen because it is provides information on individuals' forecast moments, which will be shown to greatly facilitate the testing of the asymmetry hypothesis, and because it has been running for long enough to support an individual-level analysis with a reasonable number of respondents.

The plan of the rest of the paper is as follows. Section 2 sets out the Capistrán and Timmermann (2009) explanation of the empirical correlation between disagreement and uncertainty (see, e.g., Zarnowitz and Lambros (1987)), and the framework for testing for asymmetric loss of Clements (2009). Section 3 describes the SPF forecast data, and how it is used to test the asymmetric loss hypothesis. Section 4 describes the results, and section 5 concludes.

2 Asymmetric Loss

The Capistrán and Timmermann (2009) explanation of the observed correlation between disagreement and uncertainty rests on the SPF point forecasts being generated by individuals seeking to minimize asymmetric loss functions.[1] Although Capistrán and Timmermann (2009) assume a specific asymmetric loss function, namely linex loss (see Varian (1975)), the results in Patton and Timmermann (2007) show that we can obtain essentially the same conclusions for a general asymmetric loss function, provided only weak restrictions are put on the form of that loss function and the data generating process. The requirement of the data generating process is that the variable of interest is conditionally location-scale distributed, and the requirement of the loss function is that it is homogeneous in the forecast error.[2] Letting $E_t(y_{t+h}) \equiv E(y_{t+h} \mid \Omega_t)$, and $V_t(y_{t+h}) \equiv Var(y_{t+h} \mid \Omega_t)$, and $e_{t+h,t}$ denote the h-step ahead forecast error, then formally we are assuming that:

$$y_{t+h} \mid \Omega_t \sim D\left(E_t\left[y_{t+h}\right], V_t\left[y_{t+h}\right]\right),$$

for some constant distribution function D, and:

$$L\left(a \cdot e_{t+h,t}\right) = g\left(a\right) L\left(e_{t+h,t}\right),$$

[1] Following the key insight of Granger (1969) and Zellner (1986) that optimal forecasts will be biased if the loss function is asymmetric, much of the recent literature has sought to test whether forecasts are rational once we allow forecasters to have asymmetric loss functions: see, for example, Elliott, Komunjer and Timmermann (2005), Elliott, Komunjer and Timmermann (2008) and Patton and Timmermann (2007).

[2] Homogeneity of the loss function rules out linex loss. A practical implication of adopting the Patton and Timmermann (2007) framework rather than linex loss (together with the assumption that the data generating process is conditionally normal, in order to obtain an expression for the optimal predictor) is that the optimal predictor should depend on the conditional standard deviation, rather than the conditional variance.

for some positive function g, and all $a \neq 0$. Patton and Timmermann (2007, Proposition 2) show that the optimal forecast is given by:

$$f_{t+h,t} = E_t\left(y_{t+h}\right) + \phi_h \cdot \sqrt{V_t\left(y_{t+h}\right)} \tag{1}$$

where ϕ_h is a constant that depends on the form of D and L. If $\phi_h < 0$ then over-predictions are more costly than under-predictions (and vice versa for $\phi_h^* > 0$). The more likely over-prediction (because the more uncertain the outlook, captured by the 'variance' term), then the more the forecaster will aim to under-predict on average. Hence the bias of a rational forecaster should depend on the forecast standard deviation[3] but not on other variables known at time t:

$$\begin{aligned} E\left(y_{t+h} - f_{t+h,t} \mid \Omega_t\right) &= E\left(y_{t+h} - \left(E_t\left(y_{t+h}\right) + \phi_h \cdot \sqrt{V_t\left(y_{t+h}\right)}\right) \mid \Omega_t\right) \\ &= -\phi_h \cdot \sqrt{V_t\left(y_{t+h}\right)}. \end{aligned}$$

Hence Pesaran and Weale (2006) suggest testing for rational expectations with asymmetric losses by running a regression such as:

$$e_{t+h,t} \equiv y_{t+h} - f_{t+h,t} = \zeta_1\sqrt{V_t\left(y_{t+h}\right)} + \boldsymbol{\zeta}_2'\mathbf{Z}_t + \epsilon_{t+h} \tag{2}$$

where under the null we would expect to find $\boldsymbol{\zeta}_2' = 0$, but $\zeta_1 \neq 0$ if loss is asymmetric.

Forecasters with the same conditional mean will 'disagree' (have different point forecasts) when either i) they differ with regard to their conditional standard deviation of future inflation ('inflation uncertainty') or ii) they attach different relative costs to under and over-predictions ($\phi_i \neq \phi$). Moreover, for given values of ϕ_i, the extent of disagreement across forecasters will vary positively with 'uncertainty'. Allowing individuals to have different loss functions, we can test for rational expectations via regressions for each respondent such as (2):

$$e_{t+h,t,i} \equiv y_{t+h} - f_{t+h,t,i} = \zeta_{1,i}\sqrt{V_{t,i}\left(y_{t+h}\right)} + \boldsymbol{\zeta}_{2,i}'\mathbf{Z}_t + \epsilon_{t+h,i} \tag{3}$$

where $\boldsymbol{\zeta}_{2,i} = 0$ under rational expectations, and $\zeta_{1,i} = 0$ under symmetric loss, but we would expect to

[3] Under linex loss the optimal predictor is given by $f_{t+h,t}^* = E_t\left(y_{t+h}\right) + \frac{\phi}{2}V_t\left(y_{t+h}\right)$, where the parameter ϕ determines the degree of asymmetry in the linex loss function: $L\left(e_{t+h,t}, \phi\right) = \phi^{-2}\left[\exp\left(\phi e_{t+h,t}\right) - \phi e_{t+h,t} - 1\right]$.

find $\zeta_{1,i} \neq 0$ under asymmetric loss. In terms of forecasting macroaggregates such as inflation it seems reasonable to suppose that the relevant information is public, so accessible to all, so the test regressors $\mathbf{Z}_t$ typically will not have an i subscript. The main problem with (3) is obtaining a measure of $\sqrt{V_{t,i}\left(y_{t+h}\right)}$.[4] Capistrán and Timmermann (2009) assume that forecast uncertainty is the same across all individuals, and measure it as the one-quarter ahead conditional variance from a GARCH model fitted to the errors of an AR model for inflation. The finding of Rich and Tracy (2003) that there is no relationship between the conditional variance of the consensus errors and an *ex ante* measure of uncertainty casts doubt on this step. Model-based forecasts of the conditional variance are necessarily backward-looking and will adjust slowly when the economic environment changes relative to forecasters perceptions. It seems likely that assuming a constant-parameter GARCH model for a period that spans the Great Moderation might lead to the degree of persistence in inflation volatility being over-estimated.[5]

Clements (2009) proposes testing (1) directly for the SPF forecasts of output growth, for which histograms forecasts are reported which can be used to estimate the required forecast moments. The hypothesis of asymmetry is that $E_{t,i}\left(y_{t+h}\right) - f_{t+h,t,i}$ is systematically related to the forecast standard deviation (or the forecast variance, in the case of linex loss), but rationality requires it should not vary systematically with any variables in the individual's information set. This suggests the following regression:

$$E_{t,i}\left(y_{t+h}\right) - f_{t+h,t,i} = \delta_{1,i}\sqrt{V_{t,i}\left(y_{t+h}\right)} + \boldsymbol{\delta}_{2,i}'\mathbf{Z}_t + v_{t+h,i}. \tag{4}$$

The null of rationality and quadratic loss for individual i is that $\delta_{1,i} = 0$ and $\boldsymbol{\delta}_{2,i}' = \mathbf{0}$, against the alternative that any of these coefficients are non-zero. A rejection of the null due to $\delta_{1,i} \neq 0$, with $\boldsymbol{\delta}_{2,i}' = \mathbf{0}$, indicates asymmetry (and rationality), while $\boldsymbol{\delta}_{2,i}' \neq \mathbf{0}$ indicates irrationality (conditional on the assumed form of the loss function). In order to carry out these tests, as well as the point forecasts we require the individual's predictive distributions $P_{t+h,t,i}\left(y\right) = \Pr\left(Y_{t+h} < y \mid \Omega_{t,i}\right)$ so that $E_{t,i}\left(Y_{t+h}\right)$ and $V_{t,i}\left(Y_{t+h}\right)$ can be derived (or that these conditional moments are available directly). Estimates of these distributions are available in the SPF in the form of histograms, as described below.

The problem with testing based on (3) can be viewed as one of measurement error, as the dependent

[4] Or equivalently, of $V_{t,i}\left(y_{t+h}\right)$ if instead we assume a linex loss function.

[5] The Capistrán and Timmermann (2009) explanation of the dispersion of inflation expectations in terms of heterogeneity in asymmetric loss is grounded on their estimates of inflation volatility, which for most of the results they report are derived from an AR(4)-GARCH(1,1) model of US inflation. They find the coefficients on lagged volatility and the lagged squared error term sum to 0.98, indicating a highly persistent process.

variable in (4) is related to that in (3) by:

$$e_{t+h,i} = y_{t+h} - f_{t+h,t,i} = \left(E_{t,i}\left(y_{t+h}\right) - f_{t+h,t,i}\right) + \left(y_{t+h} - E_{t,i}\left(y_{t+h}\right)\right).$$

We let $\xi_{t+h,} = y_{t+h} - E_{t,i}\left(y_{t+h}\right)$. In this interpretation, $\xi_{t+h,i}$ is a measurement error, such that the dependent variable in (3) is a noisy proxy for the unobserved dependent variable $E_{i,t}\left(Y_{t+h}\right) - f_{t+h,t,i}$ in (4). Standard analysis suggests that if $E\left(\xi_{t+h,i} \mid \Omega_t\right) = 0$, so that $\xi_{t+h,i}$ is uncorrelated with any variables that might be included as explanatory variables, then inference based on (3) is less precise, but tests of both $\zeta_{1,i} = 0$ and $\boldsymbol{\zeta}'_{2,i} = 0$ remain valid. But when $E\left(\xi_{t+h,i}, \mathbf{X}_{ti}\right) \neq \mathbf{0}$, where $\mathbf{X}_{ti} = \left[\sqrt{V_{t,i}\left(y_{t+h}\right)} \quad \mathbf{Z}'_t\right]'$, then inference based on (3) is invalid.

As argued by Clements (2009), estimating (4) avoids the pitfalls of analyzing realized forecast errors directly. Suppose the forecast error (constructed from the outcomes and point forecasts, i.e., $e_{t+h,i}$) is regressed on variables known at the time the forecast was made, and the conditional standard deviation. Were there a series of negative shocks to inflation over the period, the individual's point forecasts may overstate the realized rates of inflation. This may show up as evidence of asymmetric loss - that the respondent has deliberately engineered a negative bias in their forecasts to avoid costly under-predictions. By instead considering the deviation between the point forecast and the conditional expectation, the conditional expectation would control for the negative shocks - it too would be higher than warranted based on an information set that includes the shocks, but unless loss truly were asymmetric there is no reason why the deviation should systematically differ from zero.

3 The Survey Data

3.1 The SPF survey

The SPF quarterly survey is our source of inflation expectations. The SPF began as the NBER-ASA survey in 1968:4 and runs to the present day.[6] It is a quarterly survey of macroeconomic forecasters of the US economy. The survey questions elicit information from the respondents on their point forecasts for a number of variables and their histograms for inflation and output growth. When the respondents make their forecasts they will know the advance national accounts release for the previous quarter. The

[6] Since June 1990 it has been run by the Philadelphia Fed, renamed as the Survey of Professional Forecasters (SPF): see Zarnowitz (1969) and Croushore (1993).

Federal Reserve Bank of Philadelphia maintains a quarterly Real Time Data Set for Macroeconomists (RTDSM): see Croushore and Stark (2001). This allows us to recreate the date on 'output price'[7] that the respondent would have had access to, including the data for the quarter immediately prior to the survey quarter.[8]

The histograms we use refer to annual inflation in the current year relative to the previous year, so that for first-quarter surveys the forecast horizon is just under a year, whereas for fourth-quarter surveys it is just under one quarter. In the statistical analyses we consider the surveys separately by quarter of the year. We construct matching series of point forecasts from the SPF data. So for Q1 surveys, for example, we sum the forecast of the current quarter and the forecasts of the next three quarters, and divide by the data for the previous year's four quarters.[9,10] For Q2 surveys the approach is the same except the value for the preceding quarter (Q1) is now data, and similarly for surveys made in the third and fourth quarters of the year. So we have forecasts of annual inflation made in Q1 through to Q4 of that year, that match the histograms in terms of forecast target (the annual change) and forecast horizon.

We have 169 quarterly surveys from 1968:4 to 2010:4. We exclude the surveys for which there is some doubt about the year that the histogram returns refer to.[11]

The methods of calculating forecast moments are described in the following sub-section.

3.2 Fitting continuous distributions to the SPF inflation histograms

We require measures of the forecast mean and variance of the annual rate of inflation. Calculating means and variances from histograms is not straightforward. Diebold, Tay and Wallis (1999) assume the probability mass is uniform within each interval (in the context of calculating probability integral

[7] Both the definition and base year of the output-price index has changed over time. The vintages of data in the RTDSM match the indices for which probability assessments and point forecasts were requested in the SPF. For 1969 to 1991, this was the implicit GNP deflator, for 1992 to 1995 the implicit GDP deflator, and for 1996 to 2002 the chain-weighted deflator.

[8] Later vintages of data contain revisions and definitional changes (see e.g., Landefeld, Seskin and Fraumeni (2008) for a discussion of the revisions to US national accounts data).

[9] As of 1981:3, forecasts of the level of the output price for the current year were provided. Summing the quarterly forecasts allows us to use data back to 1968:4. For the period from 1981:3 to 2005:1 we found that the differences in the two methods of calculating annual inflation were small.

[10] The point forecasts of the growth rate are calculated using the actual data for the previous year from the RTDSM available in the quarter of the survey. The one exception is that the RTDSM for 1996Q1 is missing the value for 1995Q4. In constructing the year-on-year point forecast growth rates for the respondents to the 1996Q1 survey we use the previous-quarter forecasts (of 1995Q4).

[11] See the online documentation provided by the Philadelphia Feb: 'Documentation for the Philadelphia Fed's Survey of Professional Forecasters', http://www.phil.frb.org/econ/spf/. The problematic survey quarters are 1985.1,1986.1, 1968.4, 1969.4, 1970.4, 1971.4, 1972.3, 1972.4, 1973.4, 1975.4, 1976.4, 1977.4, 1978.4, 1979.2, 1979.3, 1979.4.

Table 1: Example of a histogram return.

Interval	Probability	t	$F(t)$
'< -2'	0.0	-2	0
-2 to -1.1	0.0	-1	0
-1 to -0.1	0.0	0	0
0 to 0.9	0.0	1	0
1 to 1.9	0.0	2	0
2 to 2.9	0.20	3	0.2
3 to 3.9	0.50	4	0.7
4 to 4.9	0.30	5	1.0
5 to 5.9	0.0	6	1.0
'6+'	0.0	7	1.0

transforms), while Giordani and Söderlind (2003) fit normal distributions to the histograms. Clements (2009, 2010) uses both approaches. The uniformity assumption may tend to overstate the dispersion of the distribution especially when there is a large difference in the probability mass attached to adjacent intervals, where it might be thought desirable to attach higher probabilities to points in the lower interval near the boundary with the high probability interval. One might question the assumption of symmetry implied by fitting a normal distribution, and so we follow Engelberg *et al.* (2009) and fit (unimodal) generalized beta distribution. This distribution uses two parameters to describe the shape of beliefs, and two more to give their support.

The approach can best be understood by an example. Suppose a respondent's histogram is reported as the first two columns of table 1. When, as here, probability is assigned to three or more intervals, the histogram can be approximated by a generalized beta distribution, as in Engelberg *et al.* (2009). To illustrate, let $t_1, \ldots, t_{10}$ denote the right endpoints of the histogram intervals, so that $F(t_1),, \ldots, F(t_{10})$ are points on the individual's CDF (recorded in the fourth column of table 1). Thus we have $F(t_5) = 0$, $F(t_6) = 0.2$, $F(t_7) = 0.7$, and $F(t_8) = 1$.

The generalized beta CDF is given by:

$$\text{Beta}(t; a, b, l, r) = \left\{ \begin{array}{cc} 0 & \text{if } t \leq l \\ \frac{1}{B(a,b)} \int_l^t \frac{(x-l)^{a-1}(r-x)^{b-1}}{(r-l)^{a+b-1}} dx & \text{if } l < t \leq r \\ 1 & \text{if } t > r \end{array} \right\}$$

where the support of the distribution is determined by the parameters l and r, and where $B(a, b) =$

$(\Gamma(a)\Gamma(b))/\Gamma(a+b)$, and $\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}dx$ (see, e.g., Balakrishnan and Nevzorov (2003) for technical details). We can impose unimodality by restricting a, b such that $a > 1$ and $b > 1$. The support will depend on the distribution of probability across the histogram intervals. Suppose probability is only attached to interior intervals, as in our example. We then set l and r equal to the left and right endpoints of the intervals with positive probability. In our example, $l = t_5$, $r = t_8$. Then we minimize only over a and b:

$$\min_{a>1,b>1} \sum_{i=1}^{10} \left[\text{Beta}\left(t_i; a, b, t_5, t_8\right) - F\left(t_i\right)\right]^2$$

If there is mass in the lower tail interval, the we allow the support to extend below the left endpoint of the lower interval, and l is a free parameter (similarly r if probability is assigned to the upper tail interval. For example, if $F(t_1) = 0.2$, $F(t_2) = 0.5$, $F(t_3) = 0.7$, and $F(t_4) = 1$, so there is a 20% chance that inflation will be less than t_1 (the lower open-ended interval), we estimate:

$$\min_{a>1,b>1,l>l^*} \sum_{i=1}^{10} \left[\text{Beta}\left(t_i; a, b, l, t_4\right) - F\left(t_i\right)\right]^2$$

where l^* is the lowest historical value of inflation.

When $X \sim beta\left(a, b, l, r\right)$, the first two moments are given by:

$$EX = l + \frac{\left(r - l\right)a}{a + b}$$

and:

$$Var\ X = \frac{ab\left(r - l\right)^2}{\left(a + b\right)^2\left(a + b + 1\right)}.$$

When probability is assigned to fewer bins, the histogram less clearly reveals the individual's underlying subjective distribution. Formally, we are unable to fit the generalized beta distribution when there are fewer than 3 bins, and although there are other solutions (e.g., joining the points on the CDF joined by straight lines) we consider only those histograms with probability assigned to three or more intervals. This means that we only consider those histograms which provide sufficient points on the CDF to reasonably accurately reveal the underlying subjective distribution, when we wish to allow that the distribution need not be symmetric.

As a robustness check, we also fit normal distributions following Giordani and Söderlind (2003), by

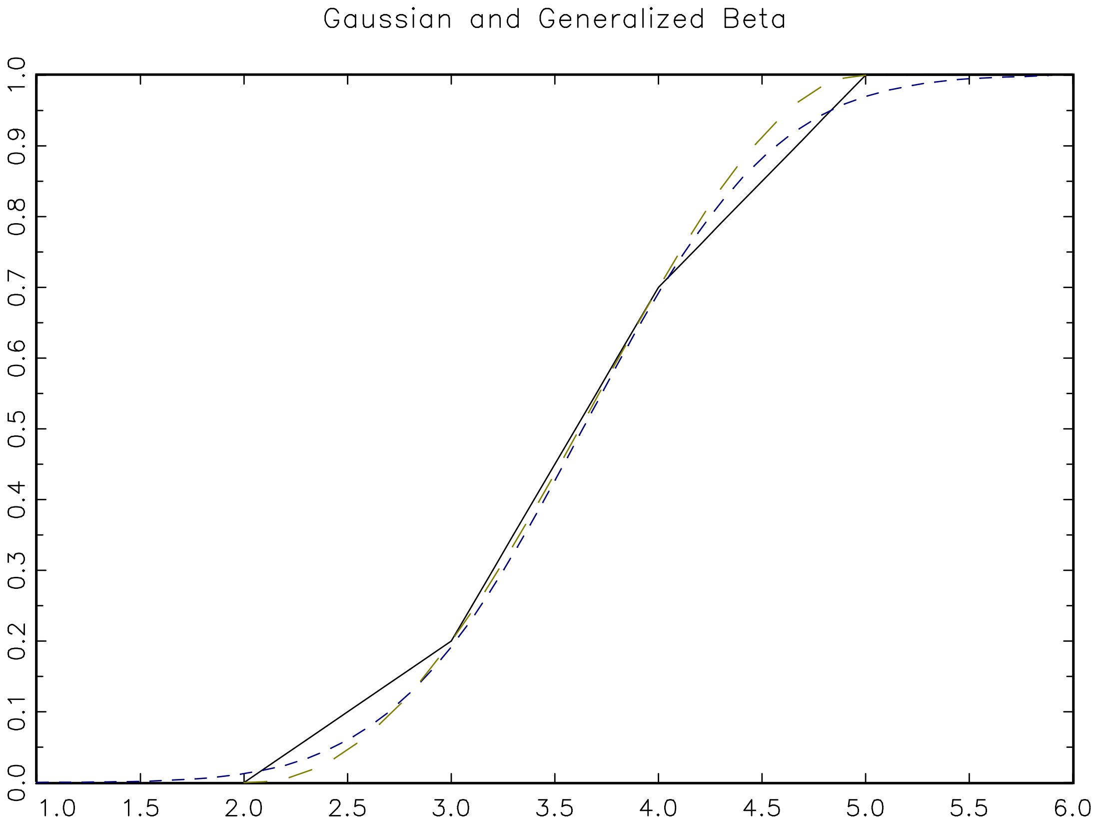

Figure 1: Estimated CDFs obtained by i) joining up CDP points with straight lines (solid line), ii) a nomal approximation (dashed line), and iii) fitting a generalized beta distribution (longer-dashed line).

minimizing:

$$\min_{\mu,\sigma^2} \sum_{i=1}^{10} \left[F_n\left(t_i;\mu,\sigma^2\right) - F\left(t_i\right)\right]^2$$

where $F_n\left(\mu,\sigma^2\right)$ is the normal CDF. The benefit of imposing symmetry is that we only require the histogram to have two non-zero probability intervals. Figure 1 reports the results of fitting both the beta and gaussian to the histogram in table 1. As seen the supports of the distribution differ - the normal extends below 2 and above 5 ($F_n\left(t_5;\hat{\mu},\widehat{\sigma}^2\right) > 0$, and $F_n\left(t_8;\hat{\mu},\widehat{\sigma}^2\right) < 1$). Of interest is whether the results are sensitive to using beta distributions (and only histograms with 3 or more non-zero probability bins) compared to imposing symmetry and using all histograms with 2 or more non-zero bins (by fitting normal distributions).

4 Results

We estimate (3) and (4) for each individual respondent and for each forecast horizon, conditional on there being at least ten observations. Given the relatively small number of observations for a respondent for a given horizon, $\mathbf{Z}_t$ was restricted to a constant. For both (3) and (4) we recorded whether the null that $\delta_{2,i} = 0$ was rejected at the 5% level, both with and without the forecast standard deviation; the sign of the constant when significant; whether the forecast standard deviation was significant ($\delta_{2,i} = 0$); the sign of this coefficient when significant; and whether the null that $\delta_{2,i} = 0$ was rejected but the null that $\delta_{1,i} = 0$ was not (i.e., the forecast standard deviation was significant but the constant was not). To estimate (3) requires an assumption about what it is the forecasters are trying to forecast - the first announcement of the data or an earlier release. Following much of the literature we assume the latter, specifically, the second release of output-price data. Table 2 reports the proportion of times the various hypotheses were rejected across all the regressions for a given horizon.

Firstly, the null of forecast bias ($\delta_{2,i} = 0$ in (3) with the forecast s.d. omitted) was rejected for 10% of the respondents across all horizons, with more rejections at the longer than the shorter horizons.[12] Over all horizons, roughly half the significant biases (in the two-sided tests) are of each sign. When we use the beta distribution to calculate moments, the proportion of rejections of $\delta_{2,i} = 0$ halves whether we include the forecast s.d. or the variance. The null that $\delta_{1,i} = 0$ is rejected only 10% for the variance (and less often for the s.d.), always with a negative coefficient. In summary, the 10% rejection rate of the test for bias (for a 5% level test) suggests little evidence of biased forecasts, and the proportion of rejections equals the size when the uncertainty term is included. Testing for rationality (allowing for asymmetric loss) as recommended by Pesaran and Weale (2006), for example, gives little cause for concern for rational expectations proponents.

Testing $\delta_{2,i} = 0$ in (4) tells a different story. Omitting the uncertainty term, we find the null is rejected in over half the cases, and it is always positive, suggesting $E_{t,i}(Y_{t+h}) > f_{t+h,t,i}$. This rejects rationality based on squared error loss, and suggests that this finding was largely masked by measurement error[13] when we considered (3). According to the asymmetric loss story, the significance of the constant term may be because of the omission of forecast uncertainty, which is expected to drive a wedge between

[12] When forecast moments are calculated using the beta distribution there are 73 regressions in total. The smaller number of Q4 survey regressions reflect the different timing conventions in the 1970s. When we use the normal distribution to estimate moments, there are 144 regressions. The use of histograms with two non-zero intervals gives rise to twice as many sets of 10 or more observations by individual and horizon.

[13] In section 2 we argued that inference based on (3) would be less precise than on (4).

the conditional mean and the point prediction. Including the uncertainty term results in $\delta_{2,i} = 0$ being rejected much less often (15% of the time for the forecast standard deviation), as would be expected under asymmetric loss. However, the insignificance of the forecast uncertainty term is rejected only one fifth of the time (for the variance, less often for the s.d.) which is not consistent with asymmetry loss, but might reflect the relatively small samples and low power.

The results using the normal distribution to estimate forecast means and variances indicate fewer rejections of $\delta_{2,i} = 0$ when forecast uncertainty is omitted, but a similar fraction when this term is included.

So far, our results are reasonably supportive of the asymmetric loss story, especially if the failure to find $\delta_{1,i}$ significant is attributed to low power or little variation in the uncertainty measure in the sample. However, the asymmetric loss explanation assumes that the forecast means are unbiased, $E\left(y_{t+h} - E_{t,i}\left(y_{t+h}\right)\right) = 0$, and the expected squared errors of the forecast means should be less than the expected squared errors of the point predictions (by construction, the mean minimizes squared-error loss and the point prediction is optimal for an assumed non-symmetric loss function).

We investigate these issues with a further set of regressions, where we regard $E_{t,i}\left(Y_{t+h}\right)$ and $f_{t+h,t,i}$ as rival forecasts in a forecast encompassing framework (see, e.g., Clements and Harvey (2009) for a recent review and exposition). Forecast encompassing relates to whether or not one forecast encapsulates all the useful predictive information contained in a second forecast. If we use a squared error loss function to measure forecast accuracy, then under the hypothesis of rationality and asymmetric loss we would expect $E_{t,i}\left(Y_{t+h}\right)$ to forecast encompass $f_{t+h,t,i}$ in the sense that in a pooled or combined forecast of the two $f_{t+h,t,i}$ should receive zero weight. In terms of the general formulation of Fair and Shiller (1989), $E_{t,i}\left(Y_{t+h}\right)$ encompasses $f_{t+h,t,i}$ if $\beta_2 = 0$ in the regression:

$$y_{t+h} = \alpha + \beta_1 E_{t,i}\left(Y_{t+h}\right) + \beta_2 f_{t+h,t,i} + \varepsilon_{t+h}.$$

We report the results of tests based on a version due to Chong and Hendry (1986):

$$y_{t+h} - E_{t,i}\left(Y_{t+h}\right) = \alpha + \beta\left(E_{t,i}\left(Y_{t+h}\right) - f_{t+h,t,i}\right) + \varepsilon_{t+h} \tag{5}$$

and:

$$y_{t+h} - f_{t+h,t,i} = \alpha + \beta\left(f_{t+h,t,i} - E_{t,i}\left(Y_{t+h}\right)\right) + \varepsilon_{t+h} \tag{6}$$

where we would expect i) not to reject the null that $\beta = 0$ in (5) but ii) to reject $\beta = 0$ in (6). We also report the number of times we reject $\alpha = 0$ in (5), when the slope regressor is excluded, which is a test for forecast bias of the histogram mean (the results for the point predictions are recorded in table 2). When $\beta = 0$ in (5) it is not possible to improve the (squared-error loss) accuracy of the mean forecast by combining it with the point prediction. Under rationality and asymmetric loss, the point prediction differs from the mean by a term which is independent of y_{t+h}, so we would expect the optimal weight on the point prediction to be zero. In (6) we would expect $\beta = 1$, so that the null that $\beta = 0$ should be rejected. The results in table 4 are exactly the opposite of what would be expected assuming RE forecasters with asymmetric loss functions. We reject the null that $E_{t,i}(Y_{t+h})$ encompasses $f_{t+h,t,i}$ for two thirds of the individuals and horizons, but the null that $f_{t+h,t,i}$ encompasses $E_{t,i}(Y_{t+h})$ is rejected for only one tenth of the individual/horizon pairs. Moreover, the unbiasedness of the mean forecasts is rejected in around one third of cases. (The comparable figure for the point predictions was one tenth of cases).

4.1 Robustness checks

A possible explanation of our findings is that the way we estimate individuals' mean forecasts results in poor estimates relative to the means of the (unobserved) underlying subjective distributions. This is possible, but we obtain similar results (table 2) whether we fit gaussian or generalized beta distributions. The results for the latter require (by construction) that positive probabilities are assigned to three or more intervals. This reduces the scope for mis-measuring means compared to analyses that estimate means for one and two-interval histograms. That the results using the two different distributions match quite closely would suggest that mis-measurement of means is unlikely to be the explanation.

A second explanation is that we have used data from the earliest days of the survey and have analyzed the data by individual. The location and widths of the histogram intervals has changed over time. For most of the period, the inflation histogram intervals were one percentage point, but for the period 1981:Q3 to 1991:Q4 this was increased to two percentage points, so that these histograms are less informative about the underlying continuous distributions, other things being equal. Further, the SPF documentation warns of the possibility that the individual identifiers might have been assigned to more than one person during the earlier period.[14]

[14] See http://www.phil.frb.org/research-and-data/real-time-center/survey-of-professional-forecasters/spf-

For these two reasons we truncate the sample of forecasts, and only consider the surveys from 1992:Q1 onwards. This covers a 19 year period and 76 surveys. The results are recorded in table 3 for testing for rationality allowing asymmetric. This table is directly comparable to the full-sample results table 2. Results are reported only for the forecast standard deviation measure of uncertainty (results for the whole sample varied little with the measure of uncertainty). There are fewer instances of respondents providing more than 10 returns to surveys in a given quarter of the year (and therefore of a given horizon), as expected. Nevertheless, the overall picture is unchanged. The regression of the histogram mean minus the point prediction yields a significant difference between the two for around a half of the regressions, but this is reduced to around a tenth when the forecast standard deviation is included, as would be expected under the asymmetric loss story. Table 5 indicates that the forecast encompassing tests are also qualitatively similar. On the shortened sample, we reject the null that the mean encompasses the point prediction for around two thirds of the regressions, but the reverse hypothesis is rejected in only one tenth of cases. We conclude that the message from the SPF surveys from 1992:1 onwards is the same as for the whole period.

We have used real-time actual values to calculate forecast errors, rather than the latest-available values. Specifically, we use the so-called 'final values' which are the second-quarterly release in the RTDSMs. The last panel of table 5 indicates that the results of the forecasting encompassing tests are qualitatively unchanged if we use the latest-available (2010:Q4 vintage) instead.

4.2 Related results

Elliott *et al.* (2008) consider the SPF 1-step ahead point forecasts of inflation and output growth, and find results which are more favourable to the asymmetric loss hypothesis. They identify 51 forecasters who made 20 or more one-step ahead inflation forecasts, and find that rationality is rejected at the 5% level for 19 of the forecasters (around 40%) assuming symmetric loss (see their Table 1). This is higher than the 10% we find. A possible explanation is that they have larger average samples by individual. Because we use the histograms to calculate the forecast moments to test for rationality, and because the histograms refer to the survey year relative to the previous year, each survey will provide forecasts of one of four horizons (depending on the quarter of the year the survey belongs to), so that on average we

documentation.pdf

will have only a quarter as many observations per individual as Elliott *et al.* (2008).[15] However, their rejections are based on testing 'efficiency', $\alpha = 0$ and $\beta = 1$ in:

$$y_{t+1} = \alpha + \beta f_{t+1,t} + v_t$$

following Mincer and Zarnowitz (1969), rather than testing for forecast bias (as in column (3) of our table 2). Testing our data for efficiency gave rejection rates of 25 to 30% (at the 5% significance level), bringing our findings more in line with their results. There was more evidence against rationality at the longer-horizons (i.e., for forecast returns to Q1 surveys), and fewer rejections at shorter horizons. However, testing the same null ($\alpha = 0$ and $\beta = 1$) in a model that includes a measure of forecast uncertainty (namely the forecast standard deviations calculated from the histograms) had virtually no effect on the rate at which the null was rejected, contrary to what would be expected under asymmetric loss.

Elliott *et al.* (2008) test for rationality allowing for asymmetric loss using the approach of Elliott *et al.* (2005). They allow for a general loss function:

$$L_p(e_{t+1};\alpha) = \left[\alpha + (1-2\alpha)\, 1_{(e_{t+1}<0)}\right] |e_{t+1}|^p \tag{7}$$

where $p = 2$ gives a piecewise quadratic loss function, with under and over-predictions being penalized differently when $\alpha \neq \frac{1}{2}$. Given (7), the forecast $f_{t+1,t}$ minimizes $E\left[L_p(e_{t+1};\alpha) \mid \mathcal{F}_t\right]$ and so solves $E\left[L'_p(e_{t+1};\alpha) \mid \mathcal{F}_t\right] = 0$, where (when $p = 2$), $L'_2(e_{t+1};\alpha) = e_{t+1} - (1-2\alpha)|e_{t+1}| = 2\left(\alpha - 1_{(e_{t+1}<0)}\right)|e_{t+1}|$. Letting v_t be a vector of instruments, $v_t \in \mathcal{F}_t$, the test of rationality allowing for loss L_2 is given by the GMM over-identification test where the population moments are $E\left\{v_t.L'_2(e_{t+1};\alpha)\right\} = 0$. They find that rationality allowing for asymmetric loss is rejected for single-figure numbers of respondents, the precise number depending on the instrument set. However, given the small samples involved there are concerns over the power of GMM tests of overidentifying restrictions: the failure to reject the null of rationality might simply reflect low power.

[15] On the plus side, we consider 1 to 4-quarter ahead forecasts of annual year-on-year inflation, as opposed to simply 1-step ahead forecasts of next quarters inflation rate. Patton and Timmermann (2010) have shown that forecasters tend to put relatively more weight on new information at shorter horizons than at longer horizons when expectations about the long-run value receive greater weight. One might expect that the degree to which forecasters might aim to under or over-predict would also depend on the forecast horizon.

5 Conclusions

In the recent literature it has been suggested that symmetric loss ought to be regarded as just a special case of more general loss functions, and that the presumption of symmetric loss might be misplaced for macro-forecasters. If forecasters have asymmetric loss functions, then we would expect their point forecasts to be biased, and the bias should depend on the conditional forecast uncertainty. However, the small samples of forecasts we have by individual for a given horizon might explain why we often fail to reject the null of unbiasedness (irrespective of whether loss is really asymmetric). Our preferred test analyses the difference between the point predictions and (estimates of) the conditional means and appears to have greater power, and is consistent with the asymmetry story in that the difference is zero mean once we allow for forecast uncertainty. However, the outcomes of the forecast encompassing tests provide telling evidence against asymmetric loss: we tend to reject the null that the mean forecasts encompass the point predictions, under squared error loss, but not vice versa.

We have assumed throughout that the histograms accurately reflect the individuals' true beliefs, in the sense that the histograms are not intentionally 'biased' representations of the individuals' probability assessments. Note that this is distinct from the measurement issues of estimating forecast moments from the histograms, which we have tackled by reporting a sensitivity analysis. If forecasters report probability distributions in the form of histograms with 'in-built' bias, then the point predictions could be optimal for an asymmetric loss function, and could simultaneously be more accurate than the 'conditional mean' forecasts, when both are judged by squared-error loss. It is recognized that forecasters may have incentives not to report their true beliefs (see, e.g., Ehrbeck and Waldmann (1996), Laster, Bennett and Geoum (1999) and Ottaviani and Sorensen (2006)): they may balance the minimization of forecast errors against conflicting aims, such as convincing the market that they are well-informed, or of attracting media attention. However, one might expect these incentives to relate to the relatively higher-profile point predictions, rather than histogram forecasts (or probability assessments more generally) which tend to receive less attention.

Asymmetric loss requires forecasters bias their forecasts away from what they expect the outcome to be to an increasing extent as the degree of uncertainty about the future increases, that is, at just those times when they are liable to make large forecast errors anyway. One might suppose that credibility issues arise when one's forecasts are sufficiently 'poor' (when judged by conventional criteria such as bias and squared-error loss) such that the assumption of symmetric loss might not be unreasonable.

Table 2: Testing for rationality and asymmetric loss by individual and horizon, 1968:4 to 2010:4

		$y_{t+h} - f_{t+h,t,i}$ as the dep. var. (eqn. 3)						$E_{t,i}(Y_{t+h}) - f_{t+h,t,i}$ as the dep. var. (eqn. 4)					
Survey	# regns.	ζ_2	$\zeta_2 < 0$	ζ_2 $\zeta_1 \neq 0$	ζ_1	$\zeta_1 < 0$	$\zeta_2 = 0$ $\zeta_1 \neq 0$	ζ_2	$\zeta_2 < 0$	ζ_2 $\zeta_1 \neq 0$	ζ_1	$\zeta_1 < 0$	$\zeta_2 = 0$ $\zeta_1 \neq 0$
1	2	3	4	5	6	7	8	9	10	11	12	13	14
Generalized beta - forecast standard deviation													
All	73	0.11	0.38	0.05	0.07	1.00	0.04	0.52	0.00	0.12	0.15	0.18	0.10
Q1	29	0.17	0.00	0.00	0.03	1.00	0.03	0.55	0.00	0.07	0.07	0.00	0.07
Q2	24	0.13	1.00	0.13	0.08	1.00	0.04	0.54	0.00	0.13	0.13	0.00	0.13
Q3	13	0.00	0.00	0.08	0.15	1.00	0.08	0.54	0.00	0.23	0.31	0.25	0.08
Q4	7	0.00	0.00	0.00	0.00	0.00	0.00	0.29	0.00	0.14	0.29	0.50	0.14
Generalized beta - forecast variance													
All	73	0.11	0.38	0.05	0.11	1.00	0.10	0.52	0.00	0.15	0.21	0.33	0.14
Q1	29	0.17	0.00	0.07	0.07	1.00	0.07	0.55	0.00	0.14	0.14	0.25	0.10
Q2	24	0.13	1.00	0.08	0.08	1.00	0.04	0.54	0.00	0.17	0.21	0.40	0.13
Q3	13	0.00	0.00	0.00	0.23	1.00	0.23	0.54	0.00	0.23	0.31	0.25	0.15
Q4	7	0.00	0.00	0.00	0.14	1.00	0.14	0.29	0.00	0.00	0.29	0.50	0.29
Gaussian - forecast standard deviation													
All	144	0.10	0.60	0.13	0.15	0.77	0.08	0.29	0.00	0.09	0.13	0.47	0.06
Q1	46	0.13	0.17	0.15	0.15	0.86	0.09	0.35	0.00	0.04	0.07	0.33	0.07
Q2	44	0.16	0.86	0.14	0.16	0.57	0.07	0.30	0.00	0.14	0.14	0.50	0.02
Q3	32	0.06	1.00	0.13	0.19	0.83	0.09	0.25	0.00	0.13	0.16	0.60	0.03
Q4	22	0.00	0.00	0.05	0.09	1.00	0.09	0.23	0.00	0.05	0.23	0.40	0.18
Gaussian - forecast variance													
All	144	0.10	0.60	0.11	0.19	0.74	0.13	0.29	0.00	0.17	0.18	0.42	0.11
Q1	46	0.13	0.17	0.15	0.22	0.80	0.11	0.35	0.00	0.20	0.09	0.50	0.04
Q2	44	0.16	0.86	0.14	0.18	0.63	0.11	0.30	0.00	0.16	0.23	0.30	0.18
Q3	32	0.06	1.00	0.06	0.16	0.80	0.16	0.25	0.00	0.13	0.22	0.57	0.13
Q4	22	0.00	0.00	0.05	0.18	0.75	0.18	0.23	0.00	0.18	0.23	0.40	0.09

Notes. Column 3 headed ζ_2 reports the proportion of individual regressions for which we reject $\zeta_{2,i} = 0$ at the 5% level in regression (3) with the standard deviation omitted. Column 4 (headed $\zeta_2 < 0$) reports the proportion of these rejections for which the constant term is negative. Column 5 is the proportion of regressions for which we reject $\zeta_{2,i} = 0$ at the 5% level with the standard deviation included. Column 6 (headed ζ_1) reports the proortion of rejections of the null that the uncertainty measure is insignificant, and column 7 the proportion of these for which $\zeta_1 < 0$. Column 8 is the proportion of regressions for which the constant is not significant and the standard deviation is significant ($\zeta_{2,i} = 0$ and $\zeta_{1,i} \neq 0$). Columns 9 to 14 then repeat columns 3 to 8 but for the regressions based on equation (4).

Table 3: Testing for rationality and asymmetric loss by individual and horizon, Restricted sample, 1992:1 to 2010:4

		$y_{t+h} - f_{t+h,t,i}$ as the dep. var. (eqn. 3)						$E_{t,i}(Y_{t+h}) - f_{t+h,t,i}$ as the dep. var. (eqn. 4)					
Survey	# regns.	ζ_2	$\zeta_2 < 0$	ζ_2 $\zeta_1 \neq 0$	ζ_1	$\zeta_1 < 0$	$\zeta_2 = 0$ $\zeta_1 \neq 0$	ζ_2	$\zeta_2 < 0$	ζ_2 $\zeta_1 \neq 0$	ζ_1	$\zeta_1 < 0$	$\zeta_2 = 0$ $\zeta_1 \neq 0$
1	2	3	4	5	6	7	8	9	10	11	12	13	14
Generalized beta - forecast standard deviation													
All	42	0.17	0.57	0.07	0.10	0.50	0.02	0.57	0.00	0.10	0.10	0.00	0.07
Q1	15	0.27	0.25	0.00	0.07	0.00	0.07	0.53	0.00	0.00	0.13	0.00	0.13
Q2	13	0.23	1.00	0.15	0.15	0.50	0.00	0.62	0.00	0.15	0.00	0.00	0.00
Q3	9	0.00	0.00	0.11	0.11	1.00	0.00	0.67	0.00	0.22	0.22	0.00	0.11
Q4	5	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.00	0.00	0.00	0.00	0.00
Gaussian - forecast standard deviation													
All	77	0.09	0.57	0.09	0.09	0.14	0.03	0.39	0.00	0.12	0.13	0.40	0.05
Q1	22	0.14	0.00	0.00	0.05	0.00	0.05	0.41	0.00	0.05	0.09	0.00	0.09
Q2	22	0.18	1.00	0.23	0.23	0.20	0.05	0.41	0.00	0.23	0.18	0.50	0.05
Q3	16	0.00	0.00	0.06	0.06	0.00	0.00	0.44	0.00	0.19	0.19	0.67	0.00
Q4	17	0.00	0.00	0.06	0.00	0.00	0.00	0.29	0.00	0.00	0.06	0.00	0.06

Notes. Layout is identical to table 2, repeated here for conveneience. Column 3 headed ζ_2 reports the proportion of individual regressions for which we reject $\zeta_{2,i} = 0$ at the 5% level in regression (3) with the standard deviation omitted. Column 4 (headed $\zeta_2 < 0$) reports the proportion of these rejections for which the constant term is negative. Column 5 is the proportion of regressions for which we reject $\zeta_{2,i} = 0$ at the 5% level with the standard deviation included. Column 6 (headed ζ_1) reports the proortion of rejections of the null that the uncertainty measure is insignificant, and column 7 the proportion of these for which $\zeta_1 < 0$. Column 8 is the proportion of regressions for which the constant is not significant and the standard deviation is significant ($\zeta_{2,i} = 0$ and $\zeta_{1,i} \neq 0$). Columns 9 to 14 then repeat columns 3 to 8 but for the regressions based on equation (4).

Table 4: Forecast encompassing tests, 1968:4 to 2010:4

Survey	# regns.	α in eqn. (5)	$\alpha < 0$ in eqn. (5)	β in eqn. (5)	β in eqn. (6)
1	2	3	4	5	6
			Generalized beta		
All	73	0.32	1.00	0.66	0.08
Q1	29	0.10	1.00	0.41	0.07
Q2	24	0.42	1.00	0.67	0.13
Q3	13	0.62	1.00	1.00	0.00
Q4	7	0.29	1.00	1.00	0.14
			Gaussian		
All	144	0.22	1.00	0.66	0.11
Q1	46	0.09	1.00	0.39	0.11
Q2	44	0.32	1.00	0.68	0.09
Q3	32	0.31	1.00	0.88	0.16
Q4	22	0.14	1.00	0.86	0.09

Notes. Column 3 reports the proportion of individual regressions for which we reject $\alpha_i = 0$ at the 5% level in regression (5). Column 4 reports the proportion of these rejections for which the constant term is negative. Column 5 is the proportion of regressions for which we reject $\beta_i = 0$ at the 5% level in eqn. (5) - the null that the mean forecast encompasses the point prediction, and column 6 is the same but for eqn. (6), where the null is that the point prediction encompasses the mean. The actual values of the inlfation rates are 'real-time' - the second-quarterly release vintage.

Table 5: Forecast encompassing tests, Restricted sample, 1992:1 to 2010:4

Survey	# regns.	α in eqn. (5)	$\alpha < 0$ in eqn. (5)	β in eqn. (5)	β in eqn. (6)
1	2	3	4	5	6
Generalized beta, real-time actuals					
All	42	0.43	1.00	0.69	0.12
1	15	0.20	1.00	0.40	0.13
2	13	0.54	1.00	0.77	0.15
3	9	0.67	1.00	1.00	0.00
4	5	0.40	1.00	0.80	0.20
Gaussian, real-time actuals					
All	77	0.30	1.00	0.69	0.10
1	22	0.14	1.00	0.45	0.14
2	22	0.45	1.00	0.77	0.09
3	16	0.44	1.00	0.81	0.13
4	17	0.18	1.00	0.76	0.06
Generalized beta, 2010:4 vintage actuals					
All	42	0.12	1.00	0.60	0.12
1	15	0.07	1.00	0.27	0.13
2	13	0.23	1.00	0.69	0.08
3	9	0.11	1.00	0.89	0.00
4	5	0.00	0.00	0.80	0.40

Notes. Layout is identical to table 4, repeated here for convenience. Column 3 reports the proportion of individual regressions for which we reject $\alpha_i = 0$ at the 5% level in regression (5). Column 4 reports the proportion of these rejections for which the constant term is negative. Column 5 is the proportion of regressions for which we reject $\beta_i = 0$ at the 5% level in eqn. (5) - the null that the mean forecast encompasses the point prediction, and column 6 is the same but for eqn. (6), where the null is that the point prediction encompasses the mean.

References

Balakrishnan, N., and Nevzorov, V. B. (eds.)(2003). *A Primer on Statistical Distributions*: John Wiley & Sons, New Jersey.

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

Chong, Y. Y., and Hendry, D. F. (1986). Econometric evaluation of linear macro-economic models. *Review of Economic Studies*, **53**, 671–690. Reprinted in Granger, C. W. J. (ed.) (1990), *Modelling Economic Series*. Oxford: Clarendon Press.

Clements, M. P. (2009). Internal consistency of survey respondents' forecasts: Evidence based on the Survey of Professional Forecasters. In Castle, J. L., and Shephard, N. (eds.), *The Methodology and*

Practice of Econometrics. A Festschrift in Honour of David F. Hendry. Chapter 8, pp. 206–226. Oxford: Oxford University Press.

Clements, M. P. (2010). Explanations of the Inconsistencies in Survey Respondents Forecasts. *European Economic Review*, **54**, 536–549.

Clements, M. P., and Harvey, D. I. (2009). Forecasting combination and encompassing. In Mills, T. C., and Patterson, K. (eds.), *Palgrave Handbook of Econometrics, Volume 2: Applied Econometrics*, pp. 169–198: Palgrave MacMillan.

Coibion, O., and Gorodnichenko, Y. (2011). What can survey forecasts tell us about information rigidities?. College of William and Mary.

Croushore, D. (1993). Introducing: The Survey of Professional Forecasters. *Federal Reserve Bank of Philadelphia Business Review*, **November/December**, 3–13.

Croushore, D., and Stark, T. (2001). A real-time data set for macroeconomists. *Journal of Econometrics*, **105**, 111–130.

Diebold, F. X., Tay, A. S., and Wallis, K. F. (1999). Evaluating density forecasts of inflation: The Survey of Professional Forecasters. In Engle, R. F., and White, H. (eds.), *Cointegration, Causality and Forecasting: A Festschrift in Honour of Clive Granger*, pp. 76–90. Oxford: Oxford University Press.

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

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

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

Engelberg, J., Manski, C. F., and Williams, J. (2009). Comparing the point predictions and subjective probability distributions of professional forecasters. *Journal of Business and Economic Statistics*, **27**, 30–41.

Fair, R. C., and Shiller, R. J. (1989). The informational content of *ex ante* forecasts. *Review of Economics and Statistics*, **71**, 325–331.

Giordani, P., and Söderlind, P. (2003). Inflation forecast uncertainty. *European Economic Review*, **74**, 1037–1060.

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

Kandel, E., and Zilberfarb, B. Z. (1999). Differential interpretation of information in inflation forecasts. *The Review of Economics and Statistics*, **81**, 217–226.

Lahiri, K., and Sheng, X. (2008). Evolution of forecast disagreement in a Bayesian learning model. *Journal of Econometrics*, **144**, 325–340.

Landefeld, J. S., Seskin, E. P., and Fraumeni, B. M. (2008). Taking the pulse of the economy. *Journal of Economic Perspectives*, **22**, 193–216.

Laster, D., Bennett, P., and Geoum, I. S. (1999). Rational bias in macroeconomic forecasts. *The Quarterly Journal of Economics*, **114(1)**, 293–318.

Mankiw, N. G., and 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., and Wolfers, J. (2003). Disagreement about inflation expectations. mimeo, National Bureau of Economic Research, Cambridge MA.

Manzan, S. (2011). Differential interpretation in the Survey of Professional Forecasters. *Journal of Money, Credit and Banking*, **43**, 993–1017.

Mincer, J., and Zarnowitz, V. (1969). The evaluation of economic forecasts. In Mincer, J. (ed.), *Economic Forecasts and Expectations*. New York: National Bureau of Economic Research.

Ottaviani, M., and Sorensen, P. N. (2006). The strategy of professional forecasting. *Journal of Financial Economics*, **81**, 441–466.

Patton, A. J., and Timmermann, A. (2007). Testing forecast optimality under unknown loss. *Journal of the American Statistical Association*, **102**, 1172–1184.

Patton, A. J., and Timmermann, A. (2010). Why do forecasters disagree? lessons from the term structure of cross-sectional dispersion. *Journal of Monetary Economics*, **57**, 803–820.

Patton, A. J., and Timmermann, A. (2011). Predictability of output growth and inflation: A multi-horizon survey approach. *Journal of Business and Economic Statistics*, **29**, 397–410.

Pesaran, M. H., and Weale, M. (2006). Survey expectations. In Elliott, G., Granger, C., and Timmermann, A. (eds.), *Handbook of Economic Forecasting, Volume 1. Handbook of Economics 24*, pp.

715–776: Elsevier, Horth-Holland.

Rich, R., and Tracy, J. (2003). Modeling uncertainty: Predictive accuracy as a proxy for predictive confidence. *Federal Reserve Bank of New York Staff Reports*, **no. 161, February**.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, **50**, 665–690.

Varian, H. R. (1975). A Bayesian approach to real estate assessment. In Fienberg, S. E., and Zellner, A. (eds.), *Studies in Bayesian econometrics and statistics in honor of Leonard J. Savage*, pp. 195–208. Amsterdam: North Holland.

Woodford, M. (2001). Imperfect common knowledge and the effects of monetary policy. In Aghion, P., Frydman, R., Stiglitz, J., and Woodford, M. (eds.), *Knowledge, Information, and Expectations in Modern Macroeconomics: In honor of Edmund Phelps*: Princeton University Press.

Zarnowitz, V. (1969). The new ASA-NBER Survey of Forecasts by Economic Statisticians. *The American Statistician*, **23, No. 1**, 12–16.

Zarnowitz, V., and Lambros, L. A. (1987). Consensus and uncertainty in economic prediction. *Journal of Political Economy*, **95**, 591–621.

Zellner, A. (1986). Biased predictors, rationality and the evaluation of forecasts. *Economics Letters*, **21**, 45–48.