Federal Reserve Bank of New York
Staff Reports

The Relationship between Expected Inflation, Disagreement, and Uncertainty: Evidence from Matched Point and Density Forecasts

Robert Rich
Joseph Tracy

Staff Report no. 253
July 2006

The Relationship between Expected Inflation, Disagreement, and Uncertainty: Evidence from Matched Point and Density Forecasts

Robert Rich and Joseph Tracy

Federal Reserve Bank of New York Staff Reports, no. 253
July 2006
JEL classification: C12, C22, E37

Abstract

This paper examines matched point and density forecasts of inflation from the Survey of Professional Forecasters to analyze the relationship between expected inflation, disagreement, and uncertainty. We extend previous studies through our data construction and estimation methodology. Specifically, we derive measures of disagreement and uncertainty by using a decomposition proposed in earlier research by Wallis and by applying the concept of entropy from information theory. We also undertake the empirical analysis within a seemingly unrelated regression framework. Our results offer mixed support for the propositions that disagreement is a useful proxy for uncertainty and that increases in expected inflation are accompanied by heightened inflation uncertainty. However, we document a robust, quantitatively and statistically significant positive association between disagreement and expected inflation.

Key words: Survey of Professional Forecasters, density forecasts, point forecasts, inflation predictions, seemingly related regression

---

Rich: Federal Reserve Bank of New York (e-mail: robert.rich@ny.frb.org). Tracy: Federal Reserve Bank of New York (e-mail: joseph.tracy@ny.frb.org). The authors thank J. S. Butler, Timothy Cogley, John Ham, Bart Hobijn, John Leahy, Jose Lopez, Simon Potter, Til Schuermann, Tom Stark, and Giorgio Topa for helpful comments and suggestions. Bess Rabin and Ariel Zetlin-Jones provided excellent research assistance. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

1. Introduction

There is widespread agreement that inflation expectations are important for understanding the behavior of individuals and observed macroeconomic outcomes. While a great deal of research continues to focus on how people form expectations, there is also interest in examining other aspects of predictive behavior and characterizing their relationships. For example, Zarnowitz and Lambros (1987) and Giordani and Söderlind (2003) investigate the linkage between the dispersion of individual mean forecasts of inflation (a measure of disagreement over inflation forecasts) and the average dispersion of corresponding density forecast distributions (a measure of uncertainty over inflation forecasts). This issue bears upon the validity of using disagreement as a proxy for inflation uncertainty in empirical investigations. Other studies seek to determine if changes in anticipated inflation are associated with parallel changes in uncertainty about inflation. If this relationship holds, then an additional cost of rising inflation is the adverse real effects associated with increased uncertainty. More recently, Mankiw, Reis and Wolfers (2003) explore the relationship between the dispersion of individual mean forecasts and expected inflation to test predictions of the 'sticky-information' model of Mankiw and Reis (2002).

This paper examines matched point and density forecasts of inflation from the Survey of Professional Forecasters (SPF) to analyze the relationship between (aggregate) expected inflation, disagreement and uncertainty. Our study improves upon previous studies in terms of data construction and estimation methodology. With regard to data construction, we derive empirical measures of disagreement and uncertainty using two alternative approaches. One approach draws upon the work of Wallis (2004, 2005) and uses a decomposition of the variance of the aggregate density forecast distribution. The second approach applies the concept of entropy from information theory. While we argue that each approach has its own merits, the use of both approaches has the added benefit of allowing us to assess the sensitivity of the results to different data constructs.

With regard to estimation methodology, the matched point and density inflation forecasts from the SPF involve four forecast horizons. Previous studies have either selected a single horizon for analysis or examined the horizons separately. We adopt a seemingly unrelated regression (SUR) approach in which we group the equations for each horizon. This choice of estimation strategy not only stems from theoretical considerations suggesting the regression residuals should be correlated across horizons, but also from formal statistical tests that confirm this feature of the data. The SUR framework provides efficiency gains relative to conventional estimation methods and also allows us to assess the robustness of the results across different forecast horizons.

Our findings offer mixed evidence concerning the nature of the relationships between disagreement (across inflation forecasts) and inflation uncertainty as well as between expected inflation and inflation uncertainty. Specifically, when we employ the Wallis-based measures of disagreement and uncertainty, the relationships between disagreement and uncertainty as well as between expected inflation and uncertainty display little economic importance. On the other hand, the entropy-based measures of disagreement and uncertainty reveal a positive association between the variables in these two relationships that is economically and statistically significant. While we are unable to offer a compelling argument that would favor one set of findings over the other, we can nevertheless draw some conclusions concerning the use of disagreement as a proxy for inflation uncertainty. The analysis not only raises questions about the validity of this practice, but also suggests that the measures of disagreement commonly adopted within this practice (i.e., Wallis-type measures) may be particularly problematic.[1]

In contrast, the nature of the relationship between disagreement and expected inflation is robust across both data constructs. Specifically, we find strong evidence that more diversity among respondents' point predictions of inflation coincides with increases in expected inflation, with the linkage between the variables displaying both economic and statistical significance. While we are cautious about the interpretation and implications of these findings at the aggregate level for specific models of expectations formation, we acknowledge that the positive co-movement between disagreement and expected inflation appears to be an important feature of predictive behavior and an issue warranting greater attention on the part of researchers.[2]

In the next section of the paper, we provide an overview of the SPF inflation data. Section 3 describes our econometric methodology. We present the empirical results in Section 4. We then conclude with a short summary of our findings.

2. Data

This section begins with a description of the statistical frameworks that underlie our measures of expected inflation, disagreement and uncertainty for the SPF inflation

[1] There is an extensive literature that has used forecast dispersion measures from surveys of inflation expectations as a proxy for inflation uncertainty. Zarnowitz and Lambros (1987) and Giordani and Söderlind (2003) contain references to various studies that have sought to determine the effect of inflation uncertainty on macroeconomic and financial variables such as output growth, unemployment, nominal interest rates, and labor contract durations.

[2] This finding initially might be viewed as corroborating evidence in support of the 'sticky-information' model of Mankiw and Reis (2002). In a related paper, however, Rich and Tracy (2004) argue that another implication of the 'sticky-information' model is that there should be no persistent differences across SPF respondents in their forecast behavior. When we examine the SPF inflation data at the individual level, we strongly reject the model's prediction that there are no significant fixed effects associated with either the respondents' ex ante forecast uncertainty or their ex post forecast accuracy.

data. We then provide details on the construction of the variables for the empirical analysis and also discuss particular features of the SPF inflation data that bear upon estimation of the relationships of interest. We conclude by comparing our approach to that in Zarnowitz and Lambros (1987) and Giordani and Söderlind (2003).

A. Variable Definitions

The SPF has undergone significant changes throughout its history. The survey was jointly initiated in late 1968 by the National Bureau of Economic Research (NBER) and the American Statistical Association (ASA), and was first known as the NBER-ASA Economic Outlook Survey. The survey is mailed four times a year, on the day after the first release of the National Income and Product Accounts data for the preceding quarter. Over time, the number of respondents declined, and in early 1990 the NBER-ASA Economic Outlook Survey was discontinued. However, later that year the Federal Reserve Bank of Philadelphia revived the survey and renamed it the SPF.

The survey originally asked respondents to provide point forecasts for 10 variables over a range of forecast horizons. Unlike other surveys, the questionnaire also solicits density forecasts for aggregate output and inflation in the form of histograms. That is, respondents are asked to attach a probability to each of a number of pre-assigned intervals, or bins, in which output growth and inflation might fall. Because these forecasts relate to the spread of a probability distribution of possible outcomes, they provide a unique basis from which to derive empirical measures of uncertainty.

We will restrict our attention to data on the inflation forecasts due to the lack of a homogeneous sample for the output forecasts.[3] With regard to the density forecasts of inflation, in the fourth quarter the survey asks respondents about the annual average percentage change in prices between the current year and the following year. In the first, second and third quarters, however, the survey asks respondents about the annual average percentage change in prices between the current year and the previous year. Consequently, the target variable for the density forecasts remains fixed for four consecutive surveys (from the fourth quarter of year *t* through the third quarter of year *t*+1), with a corresponding forecast horizon (*h*) that declines from approximately $4\frac{1}{2}$ quarters to $1\frac{1}{2}$ quarters.[4] For convenience, we refer to these horizons as $h = 4,\ldots,1$.

Defining notation, let ${}_j\phi_{h,t}(\pi)$ denote respondent *j*'s *h*-quarter-ahead density forecast of inflation (π) in year *t*. Therefore, ${}_j\phi_{4,t}(\pi)$ will denote respondent *j*'s density forecast in the fourth quarter (*h*=4) of year *t*, while ${}_j\phi_{3,t+1}(\pi)$ will denote the subsequent density forecast in the first quarter (*h*=3) of year *t*+1. We will then let ${}_j\phi^e_{h,t}(\pi)$ and ${}_j\sigma^2_{h,t}(\pi)$ denote, respectively, the mean and variance of the corresponding density forecasts.

With regard to the point forecasts, the SPF asks respondents for predictions of the price level for the current quarter and the next four quarters. Because data is available on the price index in preceding quarters, a point forecast, ${}_j f^e_{h,t}$, can be constructed that matches each density forecast. Therefore, we will let ${}_j f^e_{4,t}$ denote respondent *j*'s point forecast of the annual average percentage change in prices in the fourth quarter (*h*=4) of year *t*. The subsequent point forecast of the annual average percentage change in prices in the first quarter (*h*=3) of year *t*+1 will be denoted by ${}_j f^e_{3,t+1}$.

[3] Specifically, respondents switched from forecasting nominal output to real output in the early 1980s.

Our study considers two alternative approaches to derive measures of disagreement and uncertainty. The first is based on the statistical framework of Wallis (2004, 2005) that yields a formal relationship among measures of disagreement and uncertainty. Specifically, let $\bar{\phi}_{h,t}(\pi)$ denote the *h*-quarter-ahead aggregate density forecast of inflation in year *t* defined as:

$$\bar{\phi}_{h,t}(\pi) = (1/N_{h,t})\sum_{j=1}^{N_{h,t}} {}_j\phi_{h,t}(\pi)\,, \tag{1}$$

which averages the density forecasts across all $N_{h,t}$ respondents. As Wallis notes, the combined density forecast in equation (1) is an example of a finite mixture distribution.

If we assume that the individual point forecasts $({}_j f^e_{h,t})$ are the means of the individual forecast densities $({}_j\phi^e_{h,t}(\pi))$, then the first two moments of the aggregate density forecast about the origin are given, respectively, by:[5]

$$\mu_1' = (1/N)\sum_{j=1}^{N} {}_j f^e = \bar{f}^e \tag{2}$$

and

$$\mu_2' = (1/N)\sum_{j=1}^{N}\left[({}_j f^e)^2 + {}_j\sigma^2(\pi)\right], \tag{3}$$

where for convenience we temporarily suppress the subscripts denoting the specific forecast horizon and year. Consequently, the variance of $\bar{\phi}(\pi)$ is given by:

[4] Zarnowitz and Lambros select these values for the distances between the dates of the surveys and the end of the target year. As we demonstrate shortly, the horizons also reflect publication lags in the price index.
[5] Engelberg, Manski and Williams (2006) provide evidence that most SPF forecasters give point predictions that are consistent with the means/medians/modes of their density forecast distributions.

$$Var[\bar{\phi}(\pi)] = \mu_2' - (\mu_1')^2 = (1/N)\sum_{j=1}^{N}({}_j f^e - \bar{f}^e)^2 + (1/N)\sum_{j=1}^{N}{}_j\sigma^2(\pi). \\ = s_{f^e}^2 + \bar{\sigma}_{\phi(\pi)}^2 \tag{4}$$

The resulting decomposition of the variance of the aggregate distribution underlies our choice of this strategy to obtain measures of disagreement and uncertainty. The first term on the right-hand side of (4) is the cross-sectional variance of the point forecasts $(s_{f^e}^2)$ and provides the corresponding measure of disagreement. The second-term is the average individual variance $(\bar{\sigma}_{\phi(\pi)}^2)$ and provides a natural measure of aggregate uncertainty.

Our second approach to derive measures of disagreement and uncertainty draws upon information theory and the concept of entropy.[6] To better understand the motivation for the entropy-based measures, one can think about trying to assess the information in a message confirming the occurrence of a particular event. If the event was expected to occur with almost complete certainty, then the message causes little surprise and contains little information. On the other hand, if there was very little reason to believe the event would occur, then the message causes considerable surprise and contains a great deal of information. Thus, the informational content of the message is inversely related to the likelihood of the event.

The concept of entropy extends the previous illustration by computing the expected informational content of the message based on all possible events and their associated probabilities. As such, there is a direct connection between the expected information of the message and the notion of uncertainty. If there is little uncertainty prior to the message, due to the number of events being small or the existence of one

[6] Interested readers can consult The New Palgrave Dictionary of Economics for a useful summary of the history and development of information theory.

highly anticipated event, then its arrival is expected to convey little information. However, if there is greater uncertainty arising from an increase in the number of events and/or a greater uniformity of probabilities across events, then more information is expected from the message. While we have originally introduced the entropy as the expected information of the message, it is clear that it can also be regarded as a measure of the uncertainty associated with an empirical distribution, and hence with the SPF histograms.

Following convention in the information literature and continuing to suppress subscripts denoting forecast horizon and year, we calculate the entropy of an individual SPF histogram as:

$$ {}_j\sigma_H^2 = \sum_{k=1}^{n} {}_j p(k) \left[\log\left(\frac{1}{{}_j p(k)} \right) \right], \tag{5} $$

where ${}_j p(k)$ denotes the probability that individual *j* attaches to interval *k*. The entropy is nonnegative, and can attain a value of zero when $p(k) = 1$ for one of the *n* bins. If we hold the number of bins fixed at *n*, then the entropy is maximized when $p(k) = (1/n)$. However, this maximum increases when the number of possible outcomes (*n*) increases. Our entropy-based measure of aggregate uncertainty is then obtained by averaging the individual values of (5) across the *N* respondents:

$$ \bar{\sigma}_H^2 = (1/N) \sum_{j=1}^{N} {}_j\sigma_H^2 \tag{6} $$

While our previous discussion of entropy has been cast in terms of uncertainty, its close association with the notion of divergence suggests that it can also be used to measure disagreement. Consequently, we can derive an entropy-based measure of disagreement that parallels that for uncertainty. Using the same pre-assigned bins as those for the SPF density forecasts, an aggregate histogram for the individual point forecasts can be constructed. We can then recast the formula in (5) in terms of aggregate probabilities to obtain an entropy measure of disagreement which we denote by s_H^2.[7]

B. Variable Construction

While the expressions for expected inflation, disagreement and uncertainty in the previous section serve as useful definitions, they need to be made operational for our empirical analysis. We now provide details on the construction of these measures.

For our purposes, it is relatively straightforward to construct the individual point forecasts of inflation and the measures of disagreement. Recalling the structure of the target variable for the density forecasts, the matching point forecast for respondent *j* of the annual average percentage change in prices in the fourth quarter of year *t* is given by:

$$ {}_j f_{4,t}^{e} = 100 * \left[\frac{ {}_jP_{t+1,1}^{e} + {}_jP_{t+1,2}^{e} + {}_jP_{t+1,3}^{e} + {}_jP_{t+1,4}^{e} }{ P_{t,1} + P_{t,2} + P_{t,3} + {}_jP_{t,4}^{e} } - 1 \right], \tag{7} $$

where ${}_jP_{t,q}^{e}$ is respondent *j*'s predicted value of the price level in quarter *q* of year *t* and $P_{t,q}$ is the "actual" value of the price level in quarter *q* of year *t*.[8] The subsequent point forecast of the annual average percentage change in prices in the first quarter of year *t*+1 is then given by:

[7] We recognize that it would be useful if the entropy-based measures of disagreement and uncertainty could be constructed along the same lines as in Wallis (2004, 2005). However, while the entropy for the aggregate density forecast of inflation can be calculated and decomposed into two terms, their interpretation would not be identical to those in (4). One of the terms would correspond to average uncertainty, but the other term would correspond to the dispersion in respondents' forecast uncertainty and not in their inflation forecasts. This consideration accounts for the disconnect between the construct of the entropy-based measures of uncertainty and disagreement relative to the Wallis-based measures.

[8] The term "actual" value includes recently reported figures that the SPF provides to assist respondents with their forecasts.

$$ {}_j f^e_{3,t+1} = 100 * \left[\frac{{}_j P^e_{t+1,1} + {}_j P^e_{t+1,2} + {}_j P^e_{t+1,3} + {}_j P^e_{t+1,4}}{P_{t,1} + P_{t,2} + P_{t,3} + P_{t,4}} - 1 \right], \qquad (8) $$

where the P^e's and P's reflect the new quarterly price level predictions and realizations, respectively. A similar updating would occur for ${}_j f^e_{2,t+1}$ and ${}_j f^e_{1,t+1}$. The availability of the individual point forecasts then allows us to calculate the mean point forecast $(\overline{f}^e)$, the cross-sectional variance of the point forecasts $(s^2_{f^e})$, and the entropy-based measure of disagreement (s^2_H) with little effort.

Turning to the density forecast data, the construction of the entropy-based measure of average uncertainty $(\overline{\sigma}^2_H)$ is also relatively straightforward. However, the nature of the data does not immediately lend itself to deriving the remaining variables of interest. Therefore, we proceed by making additional assumptions and calculating moments of the aggregate distribution of inflation. The estimate of the mean will provide a measure of expected inflation $(\overline{\phi}^e(\pi))$ from the density forecast data. Given an estimate of the corresponding variance, we can then use the decomposition in (4) and the calculated values of the series $s^2_{f^e}$ to back out the Wallis-based measure of average uncertainty $(\overline{\sigma}^2_{\phi(\pi)})$:

$$ \overline{\sigma}^2_{\phi(\pi)} = Var[\overline{\phi}(\pi)] - s^2_{f^e} \qquad (9) $$

Continuing the previous discussion, there are two common approaches that have been used to estimate the mean and variance of the SPF aggregate histograms. The first approach assumes all the probability mass is located at the interval midpoints. The alternative approach assumes the probability mass is distributed uniformly across each interval. For the analysis, we adopt the first approach and apply the following formulas to compute the mean and variance of the aggregate density forecast, respectively:

$$\begin{aligned} \bar{\phi}^{e}(\pi) &= \sum_{k=1}^{n} p(k)\pi^{Mid}(k) \\ Var[\bar{\phi}(\pi)] &= \sum_{k=1}^{n} p(k)\left[\pi^{Mid}(k) - \bar{\phi}^{e}(\pi)\right]^{2} \end{aligned} \tag{10}$$

where $p(k)$ denotes the aggregate probability of interval *k*, and $\pi^{Mid}(k)$ denotes the midpoint of the corresponding interval. We omit the results for the uniform assumption because they are similar, although slightly weaker.[9]

C. Other Features of the SPF Inflation Data

The discussion up to this point has abstracted from a number of other important features of the SPF inflation data. For example, there have been occasional errors in the conduct of the survey where the probability variables have been subject to a mismatch between the intended and requested forecast horizon. As noted earlier, the matching of the point forecast and density forecast series is based on definitions in which the probability variables in the fourth quarter refer to the following year, whereas the probability variables in the first through third quarters refer to the current year. However, the surveys conducted in 1974:Q4 and 1980:Q4 mistakenly asked respondents for density forecasts of inflation between 1973-74 and 1979-80, respectively. Conversely, the surveys conducted in 1972:Q3, 1979:Q2-Q3, 1985:Q1 and 1986:Q1 mistakenly asked survey respondents for density forecasts of inflation between 1972-73, 1979-80, 1985-86, and 1986-87, respectively. Thus, these data are excluded from the analysis due to their forecast horizons not being comparable to those in related quarters.

[9] These results are available from the authors upon request.

There have also been changes in the price index used to define inflation in the survey as well as periodic changes in the base year of the relevant price indexes. There is also a question of whether to use real-time or final revised data. Another issue concerns the exclusion of respondents due to either their failure to provide matching point and density forecasts or due to discrepancies between their point and density forecasts that are judged to be excessive.[10] We refer the reader to Appendix A for further details.

D. Comparison To Other Studies

For our purposes, the statistical framework of Wallis (2004, 2005) is extremely attractive for analyzing the SPF inflation data. The decomposition of the variance of the aggregate density forecast and the resulting measures of disagreement and uncertainty correspond closely to the notions underlying previous studies. Moreover, and in contrast to other studies, there is a formal derivation underlying the measures of uncertainty and disagreement. For example, Zarnowitz and Lambros (1987) generate measures of uncertainty by calculating the average standard deviation from the individual density forecasts.[11] With regard to measures of disagreement, they calculate the cross-sectional standard deviation of the point forecasts. While these measures are analogous to the two terms on the right-hand side of (4), the use of standard deviations rather than variances breaks the link to the decomposition.

For their analysis, Zarnowitz and Lambros examine the same three relationships that are of interest to us. They find that disagreement and uncertainty display a weak positive relationship, while expected inflation contributes almost nothing to movements

[10] This is similar to Engelberg, Manski and Williams (2006) who also find there are some SPF forecasters whose point predictions appear to be inconsistent with the means/medians/modes of their density forecasts.
[11] Zarnowitz and Lambros make their calculations assuming the probability mass is distributed uniformly within bins.

in disagreement. They do, however, document an economically and statistically significant association between expected inflation and uncertainty. It should be noted that Zarnowitz and Lambros base their findings on a sample that runs from 1968:Q4–1981:Q2, resulting in estimated regressions for individual forecast horizons that only use 10-13 observations.

Giordani and Söderlind (2003) extend the work of Zarnowitz and Lambros by developing a statistical framework that features individual forecasters with private information. Their analysis yields the following expression that is similar to (4):

$$Var\left[\overline{\phi}(\pi)\right] = Var\left[({}_j\phi^e(\pi))\right] + E({}_j\sigma^2) \tag{11}$$

where E denotes the expectations operator. However, Giordani and Söderlind make no subsequent use of the variance of the aggregate distribution or the equality in (11). Rather, they elect to follow the approach of Zarnowitz and Lambros and calculate the measures of disagreement and uncertainty as standard deviations and not variances. Unlike Zarnowitz and Lambros, however, the standard deviation calculations are based on normal approximations to the individual forecast histograms. Consequently, Giordani and Söderlind exclude the individual point forecast data from their analysis.[12]

In contrast to Zarnowitz and Lambros, Giordani and Söderlind restrict their attention to the question of whether disagreement is a valid proxy for uncertainty. They principally focus on first quarter (h=3) data and find a correlation of 0.60 between their

[12] The normal approximation provides the estimates of the mean and standard deviation of each individual forecast histogram. As previously noted, it is straightforward to construct a measure of disagreement from the point forecast data. On the other hand, it is much more problematic to derive a measure of disagreement from the density forecast data. As we will discuss, the nature of the data may limit the ability to estimate a mean for each individual forecast histogram. The use of an estimate may also introduce a source of measurement error into the analysis. Abstracting from the previous two considerations, there is a more general question of relevance in that Giordani and Söderlind's approach is not consistent with the conventional practice of using disagreement across point forecast data as a proxy for uncertainty. We return to this latter issue in Section 3 where we discuss the specification of the regression equations.

measures of disagreement and uncertainty, although they report that correlations for the other quarters are similar and range from 0.46 to 0.68. They interpret their findings as showing that disagreement is a better proxy of inflation uncertainty than previously thought.

With regard to the methodology of Giordani and Söderlind, there are two issues that merit special discussion. The first is theoretical in nature and relates to the statistical foundation of their model. As noted by Wallis (2004, 2005), the pooling of disparate information sets actually presents conceptual difficulties and greatly complicates the issue of aggregation. This consideration may explain why Giordani and Söderlind are unable to provide an interpretation for the aggregate density forecast and may also underlie their acknowledgement that the expression in (11) may be problematic. In contrast, the finite mixture distribution proposed by Wallis provides an appropriate representation for combining the individual densities of the SPF respondents. Moreover, as Wallis notes, the sample average notation on the right-hand side of (4) is statistically more accurate than the use of E and Var on the right-hand side of (11).

The second issue is empirical in nature and relates to fitting distributions to the individual density forecasts. Specifically, an examination of the histograms reveals that respondents typically assign probabilities to only a few bins. As we discuss in greater detail in Appendix B, this concentration of probabilities raises concerns about the feasibility and reliability of estimating means and standard deviations based on fitted normal distributions.[13] While our approach also involves estimating moments of a

[13] The ability to fit a unique normal distribution to a histogram is only possible when a respondent uses three or more bins. The relevance of this condition is not trivial for the SPF inflation data, especially as the forecast horizon declines. Engleberg, Manski and Williams (2006) and D’Amico and Orphanides (2006)

distribution, we are much more comfortable working with aggregate histograms due to the greater diffusion of predictive probabilities. Moreover, we feel that the assumption concerning the location of probability mass in (10) is less tenuous than the maintenance of a particular distributional assumption.[14]

With regard to the entropy-based measures of disagreement and uncertainty, they are less formal than the Wallis-based measures. However, the entropy approach has the advantage of not requiring any assumption for the location of probability mass for the density forecasts and thereby circumvents concerns related to the accuracy of approximations to the underlying distributions. Moreover, our entropy-based measure of average uncertainty are derived using data on the individual density forecasts and therefore afford some comparability to the constructs in Zarnowitz and Lambros (1987) as well as Giordani and Söderlind (2003).

3. Empirical Framework

The previous discussion focused on the construction of measures of expected inflation, disagreement and uncertainty. We now turn our attention to evaluating the economic and statistical significance of the various relationships of interest. Specifically, we will consider the following model to gauge whether disagreement is a symptom of uncertainty:

$$\overline{\sigma}^2 = \alpha + \beta s^2 + \varepsilon \tag{12}$$

also question the appropriateness of using a normal distribution to approximate each repondent's probabilistic beliefs. Consequently, they consider alternative distribution fitting methods.

[14] Giordani and Söderlind cite the 'visual' normality of the aggregate density forecasts to motivate their approach. We will also discuss the issue of fitting normal distributions to the SPF histograms at the aggregate level in Appendix B. It is worth noting here, however, that we will report formal statistical tests that overwhelmingly reject the normality assumption for the aggregate density forecasts.

where $\bar{\sigma}^2$ is average uncertainty, s^2 is the degree of disagreement among forecasts, and ε is a mean-zero, random disturbance term. We will consider both data approaches in the course of estimating (12), although we will maintain a consistency across the measures of average uncertainty and disagreement. That is, we will examine the relationship between $\bar{\sigma}^2_{\phi(\pi)}$ and $s^2_{f^e}$ as well as the relationship between $\bar{\sigma}^2_H$ and s^2_H.[15]

With regard to analyzing the contribution of expected inflation to movements in uncertainty and disagreement, we differentiate between the use of the density forecast data and the point forecast data.[16] However, we allow for differences in the construction of the measures within each of the relationships we examine. Specifically, we adopt the following model to investigate the linkage between expected inflation and uncertainty:

$$\bar{\sigma}^2 = \alpha + \beta\bar{\phi}^e(\pi) + \varepsilon \qquad (13)$$

where we consider both $\bar{\sigma}^2_{\phi(\pi)}$ and $\bar{\sigma}^2_H$ as measures of inflation uncertainty, and where $\bar{\phi}^e(\pi)$ again denotes the mean of the aggregate density forecast.

In the case of the linkage between expected inflation and disagreement, we adopt the following model:

$$s^2 = \alpha + \beta\bar{f}^e + \varepsilon \qquad (14)$$

where we consider both $s^2_{f^e}$ and s^2_H as measures of disagreement, and where $\bar{f}^e$ again denotes the mean of the point forecasts.

[15] With the exception of using variances rather than standard deviations, equation (12) is identical to the model used in Zarnowitz and Lambros. This similarity also includes the use of the density forecast data to construct the uncertainty measure and the use of the point forecast data to construct the disagreement measure. These selected measures are the appropriate choice to assess the validity of using measures of disagreement across point forecasts as a proxy for uncertainty.

[16] Zarnowitz and Lambros adopted this same approach.

While we have touched on differences between the analyses of Zarnowitz and Lambros (1987) and Giordani and Söderlind (2003), it is worth noting these studies share one important feature. Specifically, they almost exclusively base their analysis on data for a single horizon or for individual horizons. We will argue, however, that the nature of the data lends itself to applying the method of seemingly unrelated regression (SUR).

As previously discussed, the forecasting horizon for the SPF inflation data is not constant and instead declines from the fourth quarter of year *t* through the third quarter of year *t*+1. Because of the variation in forecast horizons, it is more reasonable to treat the data as annual observations on four different series than as quarterly observations on a homogenous series. By itself, this consideration would suggest estimation of the following regression equations across the individual horizons:

$$Y_{h,t} = \alpha_h + \beta_h X_{h,t} + \varepsilon_{h,t}, \qquad h = 4,3,2,1 \tag{15}$$

where $Y_{h,t}$ and $X_{h,t}$ denote, respectively, the relevant independent and dependent variables specified in (12), (13) and (14), and where we allow the intercept and slope coefficients to vary across forecast horizons.

While the different forecast horizons argue for separate equations for the data, it does not seem reasonable to view the equations as completely unrelated due to their sharing a common inflation target over four contiguous quarters. This feature of the survey suggests that the corresponding error terms $\left[\left(\varepsilon_{4,t}, \varepsilon_{3,t+1}, \varepsilon_{2,t+1}, \varepsilon_{1,t+1}\right)\right]$ are likely correlated with each other. If this is the case, then it is possible to exploit the correlation structure of the error terms and apply the generalized-least squares estimators proposed by Zellner (1962) to generate more efficient parameter estimates than those obtained by the application of ordinary least squares (OLS) to each equation individually.[17]

Our seemingly unrelated regression (SUR) estimation strategy is standard except for one minor modification. Specifically, we group the equations based on their affiliation with the forecast horizon and target rate of inflation. In particular, we stack the four time series regressions as follows:

$$\begin{bmatrix} Y_{1,2} \\ \vdots \\ Y_{1,T} \\ \vdots \\ Y_{4,1} \\ \vdots \\ Y_{4,T-1} \end{bmatrix} = \begin{bmatrix} \alpha_1 + \beta_1 X_{1,2} + \varepsilon_{1,2} \\ \vdots \\ \alpha_1 + \beta_1 X_{1,T} + \varepsilon_{1,T} \\ \vdots \\ \alpha_4 + \beta_4 X_{4,1} + \varepsilon_{4,1} \\ \vdots \\ \alpha_4 + \beta_4 X_{4,T-1} + \varepsilon_{4,T-1} \end{bmatrix} \tag{16}$$

where we order the equations from horizon *h*=1 to horizon *h*=4.[18] We will follow convention with regard to the structure of the variance-covariance matrix Ω. Specifically, we assume the disturbance term in any single equation is conditionally homoscedastic and non-autocorrelated, although allowance is made for the data to be conditionally heteroskedastic across equations. These assumptions imply the following correlation pattern for the errors:

$$\begin{gathered} E[\varepsilon_{i,t}\ \varepsilon_{j,t}] = \sigma_j \ , i = j \\ E[\varepsilon_{i,t}\ \varepsilon_{j,t+\tau}] = \delta_{ij} \ , i \neq j \ , \tau = 0 \\ E[\varepsilon_{i,t}\ \varepsilon_{j,t+\tau}] = 0 \ , \textit{otherwise} \end{gathered} \tag{17}$$

Consequently, our estimate of Ω will have the following form:

[17] It should also be noted that the explanatory variables will not be identical across the different forecast horizons. If this condition did not hold, then no gains in efficiency could be realized from the SUR estimator over the OLS estimator.

$$\Omega = \begin{bmatrix} Q_1 & R_{12} & \cdots & R_{14} \\ R_{21} & Q_2 & \cdots & R_{24} \\ \vdots & \vdots & \ddots & \vdots \\ R_{41} & R_{42} & \cdots & Q_4 \end{bmatrix} \tag{18}$$

where

$$Q_j = \begin{bmatrix} \sigma_j & 0 & \cdots & 0 \\ 0 & \sigma_j & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \sigma_j \end{bmatrix} \tag{19}$$

and

$$R_{ij} = R_{ji}^{'} = \begin{bmatrix} \delta_{ij} & 0 & \cdots & 0 \\ 0 & \delta_{ij} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \delta_{ij} \end{bmatrix} \tag{20}$$

Following Breusch and Pagan (1980), we can construct the following Lagrange multiplier test to formally test for non-zero correlations between the disturbance terms in the four equations:

$$\lambda = T \sum_{m=1}^{i} \sum_{n=1}^{i-1} \rho_{mn}^2 \tag{22}$$

where ρ_{mn} is the estimated correlation between the OLS residuals of the i=4 equations and T is the number of observations in each equation. The tests statistic is distributed asymptotically as a chi-square random variable with $i(i-1)/2$ degrees of freedom under the null hypothesis of zero correlation between the disturbance terms.[19]

[18] We assume the number of observations on each equation is the same, which accounts for the slight difference in the time subscripts for the data associated with horizons h=1, 2, 3 (t=2, . . . , T) and horizon h=4 (t=1, . . . , T-1).

[19] The assumptions underlying the specification of Ω are broadly consistent with the data. Our decision not to incorporate additional own- and cross-covariance processes was based on further inspection of the OLS residuals as well as degrees of freedom considerations.

4. Empirical Results

A. Measures of Expected Inflation, Disagreement and Uncertainty

There is one additional feature of the SPF inflation data that merits special attention. Specifically, there have been periodic changes in the number of intervals and their widths in the SPF's survey instrument. As shown in Table 1, the survey initially provided 15 intervals. From 1981:Q3-1991:Q4, however, the number of intervals was reduced to 6. Since 1992:Q1, there have been 10 intervals. The interval widths also varied from 1 percentage point before 1981:Q3 and after 1991:Q4 to 2 percentage points in the intervening period.

The presence of varying interval widths poses a particular concern because it will impact on some of our summary measures and their movements across sub-periods. Therefore, we redefine the intervals to impose a common 2 percentage point width throughout the whole sample period.[20] To understand the importance of this consideration, the upper and lower panels of Figure 1 depict, respectively, the entropy of the aggregate density forecast distribution and the entropy of the aggregate point forecast distribution using both the raw and adjusted data.[21] The profiles for the entropy differ markedly before 1981:Q3 and after 1991:Q4, especially in the case of the aggregate density forecast distribution. Thus, the use of the raw data would result in an artificial increase in the entropy during the sub-periods associated with the narrower interval widths.

[20] Due to the odd number of intervals used over the sub-period 1968:Q4-1981:Q2, we use a unit interval length for the middle interval. As an alternative to imposing a common 2% width, one might think about redefining the intervals from 1981:Q3-1991:Q4 to have a unit interval width. We found this adjustment procedure to be much less satisfactory due to the difficulty of determining how to allocate the probabilities across the subdivided intervals.

As shown in Figure 2, the changes in interval widths also affect the profile for the estimated variance of the aggregate density forecast distribution used for the Wallis decomposition.[22] Specifically, while the estimates of the variance are generally higher during the 1980s, the use of the adjusted data partly reduces the differential during the pre-1981:Q3 and post-1991:Q4 sub-periods. Consequently, the use of the raw data would lead to lower estimates of inflation uncertainty during the pre-1981:Q3 and post-1991:Q4 sub-periods.

Figures 3-5 present the time profiles for the measures of disagreement, uncertainty and expected inflation used in the empirical analysis. As shown in Figure 3, the behavior of the two disagreement measures is qualitatively similar and indicates a greater diversity of opinion about expected inflation during the earlier part of the sample period. The entropy-based measure (s_H^2) displays slightly more variability, although the cross-sectional variance of the point forecast $(s_{f^e}^2)$ is characterized by occasional spikes in disagreement. The sawtooth pattern evident in both measures speaks to the greater unanimity across point forecasts as the forecast horizon shrinks.

In contrast to the measures of disagreement, there is a marked difference in the features of the average uncertainty measures across the two data approaches. In particular, the Wallis-based measures are generally higher and more variable than the entropy-based measure.[23] Nevertheless, both measures depict a decline in inflation

[21] The missing observations in Figure 1 (as well as in subsequent figures) reflect the excluded survey dates discussed in Section 2.C. We only display one series during the middle period due to the coincidence of the raw and adjusted data.

[22] The impact of the changes in interval widths on the estimated mean of the aggregate density forecast distribution turns out to be negligible. The measure of disagreement using the point forecast data is not affected by changes in interval width.

[23] The Wallis-based measure of uncertainty is about twice as high on average with a standard deviation that is more that twice that for the entropy-based measure.

uncertainty starting around 1990. As expected, they also tend to reflect a greater dispersion of intrapersonal probabilistic beliefs as the forecast horizon increases, although there is a surprising slight decline at the *h*=4 quarter horizon. Comparing the average levels of disagreement and uncertainty across the same data approach, disagreement understates uncertainty to a considerable extent. Specifically, the uncertainty measure is larger by a factor of five using the Wallis approach and is nearly twice as large using the entropy approach. While uncertainty displays greater variability than disagreement for the Wallis-based measures, the opposite is true for the entropy-based measures.

When we examine the measures of expected inflation in Figure 5, however, we observe that the series display a high degree of conformity and are practically indistinguishable from each other. The two inflation expectations series display the same pronounced rise and subsequent decline as actual inflation during the course of the sample period.

B. Estimated Relationships

The sample covers the surveys conducted from 1968:Q4 through 2003:Q3, so that the values on the realized annual rate of inflation cover the periods 1968-69 through 2002-2003. We begin by examining correlations and goodness-of-fit measures from OLS estimation of equations (12)–(14) reported, respectively, in Tables 2-4.[24] Because we will subsequently address the issue of estimation efficiency, we defer for the moment from any discussion of statistical significance and instead focus our initial attention on the economic significance of the relationships.

As shown, the variables display a positive association across all of the relationships. With regard to any systematic pattern to the correlations, they do not behave in a monotonic fashion as the forecast horizon increases. Rather, the correlations tend to be highest at the h = 2 and 3 quarter horizons. There are, however, several other notable findings that that emerge from the analysis.

One is immediately struck by the extremely low explanatory content of disagreement for movements in average uncertainty when disagreement is measured by the variance of the point forecasts. With the exception of the regression associated with the h=3 quarter horizon in the upper panel of Table 2, disagreement accounts for less than 10 percent of the variation in uncertainty.[25] The results are qualitatively similar when we turn to the linkage between expected inflation and the Wallis-based measure of uncertainty in the upper panel of Table 3. The lack of any meaningful co-movement between the variables suggests the issue of the statistical significance of these relationships is largely irrelevant for the remaining analysis.

Equally striking, however, is the marked increase in the strength of these same relationships when the entropy-based measures of disagreement and uncertainty are used in the regressions. For example, the correlations exceed 0.6 at the h=2 and 3 quarter horizons in the lower panel in Table 2 and Table 3, with the other correlations of moderate size. Last, an examination of Table 4 indicates the relationship between disagreement and expected inflation is much more robust to the construct of the measure of forecast dispersion. While the correlations indicate a somewhat weak relationship at

[24] We recognize there is little difference in the information conveyed by the reported correlations and the $\bar{R}^2$'s, as the latter simply involves squaring the former and adjusting for degrees of freedom. Nevertheless, we report both statistics to allow for a basis of comparison to the results of other studies.
[25] Recall that Söderlind and Giordani focus their analysis on the data for the h=3 quarter horizon.

the h=1 quarter horizon, the other horizons display reasonably strong correlations that are comparable to those associated with the entropy–based measures in Tables 2 and 3.

To address the issue of statistical significance in the relationships, we initially applied the Breusch-Pagan (1980) testing procedure to the OLS residuals within each system of four equations. As shown by the values of the test statistic reported in the last column of Tables 2-4, there is significant correlation between the equations' disturbance terms associated with the same inflation target. The one exception is the relationship between disagreement and expected inflation using the variance of the point forecasts to measure dispersion. Consequently, we retain the method of OLS for estimation in this case. In all other cases, we will estimate the relationships using the method of SUR.

Tables 5-7 report the estimates of the parameters and the corresponding standard errors.[26] Because the definition of an $\bar{R}^2$ statistic is not obvious in the case of SUR estimation, we do not attempt to report any type of goodness-of-fit statistic. Moreover, because most researchers typically posit a positive relationship between the variables, we conduct a one-tailed test for statistical significance. The conclusions, however, generally will not depend on the choice of a one- versus two-tailed test for statistical significance.

As shown, the findings typically document a statistically significant positive association between the variables in the relationships. Not surprisingly, the qualitative features of the results parallel those from the previous analysis in terms of forecast horizon and data construct. That is, the statistical significance of the relationships between disagreement and uncertainty as well as between expected inflation and uncertainty is less robust using the Wallis-based measures than the entropy-based

[26] There are an unequal number of observations across the equations. However, we ignored this difference and calculated the own- and cross-covariances using all available observations.

measures. Moreover, the relationship between disagreement and expected inflation remains highly statistically significant across both data constructs.

Taken together, there are several conclusions that can be drawn from the reported results in Tables 2-7. There is mixed support for the propositions that greater disagreement is indicative of heightened uncertainty and that higher expected inflation is accompanied by increased uncertainty. The lack of robustness of these results likely stems from the greater variability of the Wallis-based measure of uncertainty relative to the entropy-based measure of uncertainty (as well as the corresponding measures of disagreement). In light of the conflicting evidence, it is natural to ask if one set of results might be viewed as more persuasive. The answer to this question would be guided by selecting the approach that provides the better approximation to the measures of interest. Because we see advantages and disadvantages to each approach that are roughly equal on balance, we are unable to offer any resolution to this matter at present.

While we cannot resolve the disparity in the results, we can still comment on the results associated with a particular data approach. In this regard, the estimated relationship between disagreement and inflation uncertainty using the Wallis-based measures has particular relevance. This is because almost all empirical studies using disagreement as a proxy for uncertainty have measured disagreement by the variance (or standard deviation) of point forecasts. We find, however, that among all of the estimated relationships, the association between this measure of disagreement and inflation uncertainty is the weakest in terms of economic and statistical significance.[27]

[27] As discussed in Section 2.B, the results for the disagreement-inflation uncertainty and disagreement-expected inflation relationships using the Wallis-based measures are slightly weaker if we adopt the uniform assumption for the location of the probability mass within intervals. Thus, these results are being presented in a more favorable light.

Thus, there appears to be little justification for this conventional measure of disagreement to serve as a proxy for uncertainty.

The estimated relationship between disagreement and uncertainty using the Wallis-based measures also allows for a reasonable basis of comparison to the results of Zarnowitz and Lambros (1987) and Giordani and Söderlind (2003). Our findings are closer to those reported by Zarnowitz and Lambros and contrast sharply with the conclusions of Giordani and Söderlind. With regard to the latter study, their results likely differ because of a smoother measure of uncertainty due to the use of normal approximation methods as well as a measure of disagreement that pertains to the density forecast data. As noted in the text and Appendices, we have discussed various concerns about the logic and statistical basis of their methodology.

On the other hand, the evidence is much less ambiguous and quite favorable about a positive co-movement between disagreement and expected inflation. When we restrict our attention to the Wallis-based measures to allow for a basis of comparison to Zarnowitz and Lambros, our findings are much stronger in terms of economic and statistical significance. This may be a consequence of the longer sample period used in our analysis. It is also interesting to note that, of the three relationships examined in the paper, the linkage between disagreement and expected inflation has received the least attention on the part of researchers.

V. Conclusion

Our study uses matched point and density forecasts of inflation from the Survey of Professional Forecasters to revisit questions concerning the co-movement between aggregate expected inflation, the degree of disagreement among individual inflation forecasts, and the level of average inflation uncertainty. We attempt to improve upon previous studies in terms of the construction of the measures used for the empirical analysis as well as the statistical methods used to assess the nature of the relationships. As such, we derive measures of disagreement and uncertainty by using a statistical framework recently proposed by Wallis (2004, 2005) as well as by drawing upon the concept of entropy from the more established literature on information theory. We also adopt a seemingly unrelated regression framework to exploit efficiency gains that are afforded by the recurrent declining forecast horizon of the SPF inflation data.

The variables generally display a statistically significant association, although this feature varies somewhat across the particular relationships. Specifically, the incidence and level of statistical significance is highest for the linkage between disagreement and expected inflation, and somewhat lower for the other relationships. In terms of economic significance, however, we obtain markedly different results across the relationships and data constructs. We document that movements in disagreement and expected inflation display reasonably strong positive correlations, and would contend that an adequate model of expectations formation must be able to account for this co-movement.

On the other hand, the evidence of a meaningful relationship between disagreement and uncertainty as well as between expected inflation and uncertainty essentially disappears when we switch from the entropy-based measures of uncertainty and disagreement to the Wallis-based measures. The lack of robustness of these results leads us to conclude that the relevance of one of the posited channels of effect of expected inflation on real activity remains an open question. The same holds true concerning the validity of using disagreement as a proxy for uncertainty. With respect to the last point, we are especially cautious about the conclusions of empirical studies that have used conventional forecast dispersion measures to proxy inflation uncertainty.

References

Breusch, Trevor S., and Adrian R. Pagan. "The Lagrange Multiplier Test and Its Application to Model specification in Econometrics." *Review of Economic Studies* 47 (January 1980):239-253.

Engelberg, Joseph, Charles F. Manski, and Jared Williams. "Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters." Working paper, Northwestern University: 2006.

Giordani, Paul, and Paul Söderlind. "Inflation Forecast Uncertainty." *European Economic Review* 47 (December 2003), 1037-1059.

Lahiri, Kajal, and Christie Teigland. "On the Normality of Probability Distributions of Inflation and GNP Forecasts." *International Journal of Forecasting* 3 (1987): 269-279.

Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers. "Disagreement about Inflation Expectations." in *NBER Macroeconomics Annual 2003*, vol. 18: 209-248.

Mankiw, N. Gregory, and Ricardo Reis. "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics* 117 (November 2002), 1295-1328.

Rich, Robert, and Joseph Tracy. " 'Sticky Information' and Heterogeneity in Forecast Behavior: Further Implications and Evidence from the Survey of Professional Forecasters." Working paper, Federal Reserve Bank of New York: 2004.

Wallis, Kenneth F. "Forecast Uncertainty, Its Representation and Evaluation." Tutorial Lectures, IMS Singapore, May 3-6 2004.

Wallis, Kenneth F. "Combining Density and Interval Forecasts: A Modest Proposal." *Oxford Bulletin of Economics and Statistics* 67 (2005 Supplement): 983-994.

Zarnowitz, Victor and Louis A. Lambros. Consensus and Uncertainty in Economic Prediction." *Journal of Political Economy* 95 (June 1987): 591-621.

Appendix

A. Additional Data Considerations

The analysis takes into account changes in the price index used to define inflation in the survey. Specifically, the survey originally asked about inflation based on the GNP deflator (1968:Q4-1991:Q4), and then asked about inflation based on the GDP deflator (1992:Q1-1995:Q4). Presently, the survey asks about inflation as measured by the chain-weighted GDP index. We also account for periodic changes in the base year of the relevant price indexes.

To construct the point forecasts of inflation, we followed the formulas in equations (7)-(8) and combined the respondent's predictions with values of the 'actual' price index from the real-time macroeconomic data set collected by the Federal Reserve Bank of Philadelphia. The availability and use of the vintage data sets allows us the constructed inflation forecast to correspond to the same value that would have been computed at the time of the survey.

Last, we found it necessary to exclude some individual responses either due to our inability to generate matching point and density forecasts or due to discrepancies in the point and density forecasts that were judged to be excessive. Our sample initially covered 5547 respondents. However, 278 responses were excluded because they corresponded to 'bad' survey dates. We then excluded 436 responses because the individuals did not provide density forecasts, while an additional 79 responses were excluded because the probabilities assigned to the bins did not sum to unity. An additional 301 responses were excluded because individuals did not provide point forecasts. Finally, we wanted to try and safeguard against situations in which a respondent's point forecast and density forecast were at odds with each other. To do so, we applied the midpoint formula to the individual density forecasts to construct a forecasted mean of inflation. We then compared the mean of the density forecast to the corresponding point forecast and excluded those responses for which the differential (in absolute value) exceeded 1.5 percent. This resulted in an additional 226 responses being dropped from the survey. This left a total of 4,227 responses that were used for the analysis.

B. Fitting Normal Distributions to the SPF Histograms

This Appendix summarizes our findings regarding the appropriateness of fitting normal distributions to the SPF histograms at both the individual level as well as at the aggregate level.[28] While we recognize that Giordani and Söderlind fit normal distributions to individual histograms and also recognize that the use of a normal approximation to the individual histograms does not carry any implications for the distribution of a combination of the density forecasts, we are interested in exploring this issue at the aggregate level for two reasons. First, Giordani and Söderlind appeal to the 'normal appearance' of the aggregate density forecasts to motivate their approach. Second, it would be relatively straightforward to incorporate this approach within the statistical framework of Wallis. Specifically, one could fit a normal distribution to the aggregate density forecasts and use the resulting estimate of the variance in equation (9) (along with the cross-sectional variance of the point forecasts) to derive an alternative average uncertainty series.

If we initially consider normal approximations to the aggregate density forecasts, then two interesting results emerge. First, statistical evidence overwhelmingly rejects the

[28] The mean and variance are estimated by minimizing the sum of the squared differences between the survey probabilities and the probabilities for the same intervals implied by the normal distribution.

assumption of normality for the aggregate probability distributions for inflation. While the distributions are characterized by occasional episodes of skewness (19 out of 133 distributions), the deviation from normality is principally due to the distributions being leptokurtic (68 out of 133 distributions).[29] That is, the distributions have higher peaks and fatter tails than those of a normal. These findings are consistent with those previously reported in Lahiri and Teigland (1987).

Second, the estimated moments of a distribution using the normal approximation can differ markedly from those based on other approaches. The upper panel of Figure 6 compares estimates of the variance of the aggregate density forecasts using a normal approximation to those derived under the assumption that the probability mass is at the midpoint of an interval. As shown, the variance estimates from the fitted normal distribution are consistently lower, much less variable, and occasionally move in an opposite direction. Not surprisingly, these disparate features carry over when we apply the Wallis decomposition and subtract the corresponding measure of disagreement $(s^2_{f^e})$ from the estimated variance series. There is, however, another concern that now emerges from this undertaking. As shown in the lower panel of Figure 6, the Wallis-based measures of uncertainty using the normal approximation are actually <u>negative</u> in 1980:Q1 and 1985:Q4.

When we examine the individual density forecasts, the idea of fitting normal distributions to the data becomes even more problematic. One concern is that the choice of a normal distribution is hard to justify given that few respondents place positive probabilities on the two tail intervals, suggesting that some sort of truncated distribution

[29] These calculations are based on a 5% significance level for the tests of skewness and kurtosis under the

would be a more appropriate choice. Another concern is that most respondents do not assign probabilities to more than a couple of bins. For example, 21% of the respondents assign non-zero probabilities to 2 bins or less which precludes us from fitting a unique normal distribution.

The previous findings relate to the individual histograms using the raw data. When we impose a common 2% interval width, this consideration only exacerbates the problems encountered by this method. In particular, if we were to try to fit a normal distribution to the 4,227 responses described in Appendix A, then there would be 2,083 responses that assign non-zero probabilities to 2 bins or less. Compared to our Wallis-based and entropy-based measures of disagreement and uncertainty, the adoption of Giordani and Söderlind's methodology would require us to exclude almost half of the respondents from the analysis.

Taken together, the evidence suggests to us that fitting appropriate distributions to histograms, at either the individual or aggregate levels, is not as straightforward as may be assumed on the part of researchers.

---

assumption that the aggregate density forecast distributions are normally distributed.

Table 1

Intervals for Density Forecasts of Inflation

Period	1968:Q4-1973:Q1	1973:Q2-1974:Q3	1974:Q4-1981:Q2	1981:Q3-1985:Q1	1985:Q2-1991:Q4	1992:Q1-Present
Intervals	≥ 10%	≥ 12%	≥ 16%	≥ 12%	≥ 10%	≥ 8%
	+9% to +9.9%	+11% to +11.9%	+15% to +15.9%	+10% to +11.9%	+8% to +9.9%	+7% to +7.9%
	+8% to +8.9%	+10% to +10.9%	+14% to +14.9%	+8% to +9.9%	+6% to +7.9%	+6% to +6.9%
	+7% to +7.9%	+9% to +9.9%	+13% to +13.9%	+6% to +7.9%	+4% to +5.9%	+5% to +5.9%
	+6% to +6.9%	+8% to +8.9%	+12% to +12.9%	+4% to +5.9%	+2% to +3.9%	+4 to +4.9%
	+5% to +5.9%	+7% to +7.9%	+11% to +11.9%	< +4%	< +2%	+3% to +3.9%
	+4% to +4.9%	+6% to +6.9%	+10% to +10.9%			+2% to +2.9%
	+3% to +3.9%	+5% to +5.9%	+9% to +9.9%			+1% to +1.9%
	+2% to +2.9%	+4% to +4.9%	+8% to +8.9%			0 to +0.9%
	+1% to +1.9%	+3% to +3.9%	+7% to +7.9%			< 0
	0% to +0.9%	+2% to +2.9%	+6% to +6.9%			
	-1% to -0.1%	+1% to +1.9%	+5% to +5.9%			
	-2% to -1.1%	0 to +0.9%	+4% to +4.9%			
	-3% to -2.1%	-1% to -0.1%	+3% to +3.9%			
	< -3%	< -1%	< +3%			

Table 2

$\bar{\sigma}^2_{\phi(\pi)} = \alpha + \beta s^2_{f^e} + \varepsilon$			Correlations		Breusch-Pagan
Horizon/Quarter	Observations	Estimator	r	$\bar{R}^2$	λ=55.750**
h=1/Q3	33	OLS	0.318	0.073	
h=2/Q2	34	OLS	0.107	-0.018	
h=3/Q1	33	OLS	0.398	0.134	
h=4/Q4	33	OLS	0.216	0.023	

$\bar{\sigma}^2_H = \alpha + \beta s^2_H + \varepsilon$			Correlations		Breusch-Pagan
Horizon/Quarter	Observations	Estimator	r	$\bar{R}^2$	λ=23.740**
h=1/Q3	33	OLS	0.504	0.287	
h=2/Q2	34	OLS	0.824	0.679	
h=3/Q1	33	OLS	0.731	0.554	
h=4/Q4	33	OLS	0.341	0.109	

Note: Breusch-Pagan test is distributed asymptotically as a $\chi^2(6)$ random variable.
** Significant at 1% level
* Significant at 5% level

Table 3

$\bar{\sigma}^2_{\phi(\pi)} = \gamma + \delta\bar{\phi}^e(\pi) + \varepsilon$			Correlations		Breusch-Pagan
Horizon/Quarter	Observations	Estimator	r	$\bar{R}^2$	λ=47.433**
h=1/Q3	33	OLS	0.216	0.021	
h=2/Q2	34	OLS	0.273	0.053	
h=3/Q1	33	OLS	0.497	0.227	
h=4/Q4	33	OLS	0.260	0.075	

$\bar{\sigma}^2_{H} = \gamma + \delta\bar{\phi}^e(\pi) + \varepsilon$			Correlations		Breusch-Pagan
Horizon/Quarter	Observations	Estimator	r	$\bar{R}^2$	λ=54.828**
h=1/Q3	33	OLS	0.486	0.239	
h=2/Q2	34	OLS	0.621	0.406	
h=3/Q1	33	OLS	0.656	0.421	
h=4/Q4	33	OLS	0.426	0.255	

Note: Breusch-Pagan test is distributed asymptotically as a $\chi^2(6)$ random variable.
** Significant at 1% level
* Significant at 5% level

Table 4

$s_{f^e}^2 = \alpha + \beta \bar{f}^e + \varepsilon$			Correlations		Breusch-Pagan
Horizon/Quarter	Observations	Estimator	r	$\bar{R}^2$	λ=4.211
h=1/Q3	33	OLS	0.374	0.130	
h=2/Q2	34	OLS	0.690	0.514	
h=3/Q1	33	OLS	0.591	0.331	
h=4/Q4	33	OLS	0.601	0.526	

$s_H^2 = \alpha + \beta \bar{f}^e + \varepsilon$			Correlations		Breusch-Pagan
Horizon/Quarter	Observations	Estimator	r	$\bar{R}^2$	λ=14.678*
h=1/Q3	33	OLS	0.394	0.149	
h=2/Q2	34	OLS	0.633	0.428	
h=3/Q1	33	OLS	0.589	0.329	
h=4/Q4	33	OLS	0.427	0.250	

Note: Breusch-Pagan test is distributed asymptotically as a $\chi^2(6)$ random variable.
** Significant at 1% level
* Significant at 5% level

Table 5

$\bar{\sigma}^2_{\phi(\pi)} = \alpha + \beta s^2_{f^e} + \varepsilon$			Regression Estimates		
Horizon/Quarter	Observations	Estimator	α	β	
h=1/Q3	33	SUR	0.851** (0.057)	1.97** (0.554)	
h=2/Q2	34	SUR	1.05** (0.089)	0.235 (0.514)	
h=3/Q1	33	SUR	1.120** (0.080)	0.390* (0.171)	
h=4/Q4	33	SUR	1.154** (0.093)	0.121 (0.217)	

$\bar{\sigma}^2_{H} = \alpha + \beta s^2_{H} + \varepsilon$			Regression Estimates		
Horizon/Quarter	Observations	Estimator	α	β	
h=1/Q3	33	SUR	0.493** (0.025)	0.365** (0.098)	
h=2/Q2	34	SUR	0.532** (0.018)	0.377** (0.050)	
h=3/Q1	33	SUR	0.596** (0.025)	0.272** (0.047)	
h=4/Q4	33	SUR	0.654** (0.037)	0.129* (0.070)	

Note: One-tailed test for statistical significance of β

$H_0 : \beta_h = 0, H_1 : \beta_h > 0$

** Significant at 1% level

* Significant at 5% level

Table 6

$\bar{\sigma}^2_{\phi(\pi)} = \gamma + \delta\bar{\phi}^e(\pi) + \varepsilon$			Regression Estimates		
Horizon/Quarter	Observations	Estimator	α	β	
h=1/Q3	33	SUR	0.776** (0.111)	0.037 (0.022)	
h=2/Q2	34	SUR	0.837** (0.138)	0.055* (0.028)	
h=3/Q1	33	SUR	0.870** (0.122)	0.081** (0.023)	
h=4/Q4	33	SUR	0.945** (0.134)	0.058* (0.029)	

$\bar{\sigma}^2_H = \gamma + \delta\bar{\phi}^e(\pi) + \varepsilon$			Regression Estimates		
Horizon/Quarter	Observations	Estimator	α	β	
h=1/Q3	33	SUR	0.417** (0.044)	0.032** (0.009)	
h=2/Q2	34	SUR	0.440** (0.042)	0.043** (0.008)	
h=3/Q1	33	SUR	0.553** (0.036)	0.038** (0.006)	
h=4/Q4	33	SUR	0.548** (0.044)	0.041** (0.009)	

Note: One-tailed test for statistical significance of β

$H_0 : \beta_h = 0, H_1 : \beta_h > 0$

* Significant at 5% level

** Significant at 1% level

Table 7

$s^2_{f^e} = \alpha + \beta \bar{f}^e + \varepsilon$			Regression Estimates		
Horizon/Quarter	Observations	Estimator	α	β	
h=1/Q3	33	OLS	-0.002 (0.022)	0.011** (0.004)	
h=2/Q2	34	OLS	0.000 (0.021)	0.027** (0.004)	
h=3/Q1	33	OLS	-0.022 (0.089)	0.075** (0.018)	
h=4/Q4	33	OLS	-0.042 (0.068)	0.095** (0.015)	

$s^2_H = \alpha + \beta \bar{f}^e + \varepsilon$			Regression Estimates		
Horizon/Quarter	Observations	Estimator	α	β	
h=1/Q3	33	SUR	-0.015 (0.065)	0.037** (0.013)	
h=2/Q2	34	SUR	-0.073 (0.068)	0.075** (0.014)	
h=3/Q1	33	SUR	0.111 (0.085)	0.073** (0.017)	
h=4/Q4	33	SUR	0.103 (0.098)	0.084** (0.022)	

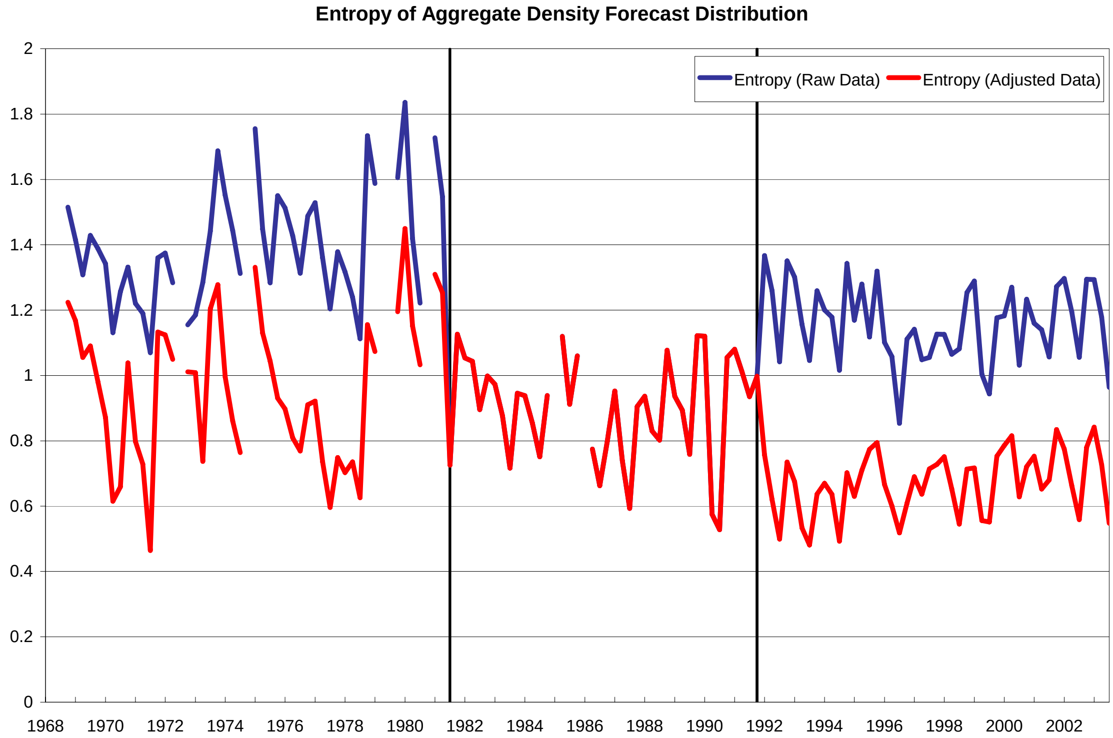
Note: One-tailed test for statistical significance of β

$H_0 : \beta_h = 0, H_1 : \beta_h > 0$

* Significant at 5% level

** Significant at 1% level

Figure 1:
Effects of Changing Interval Widths



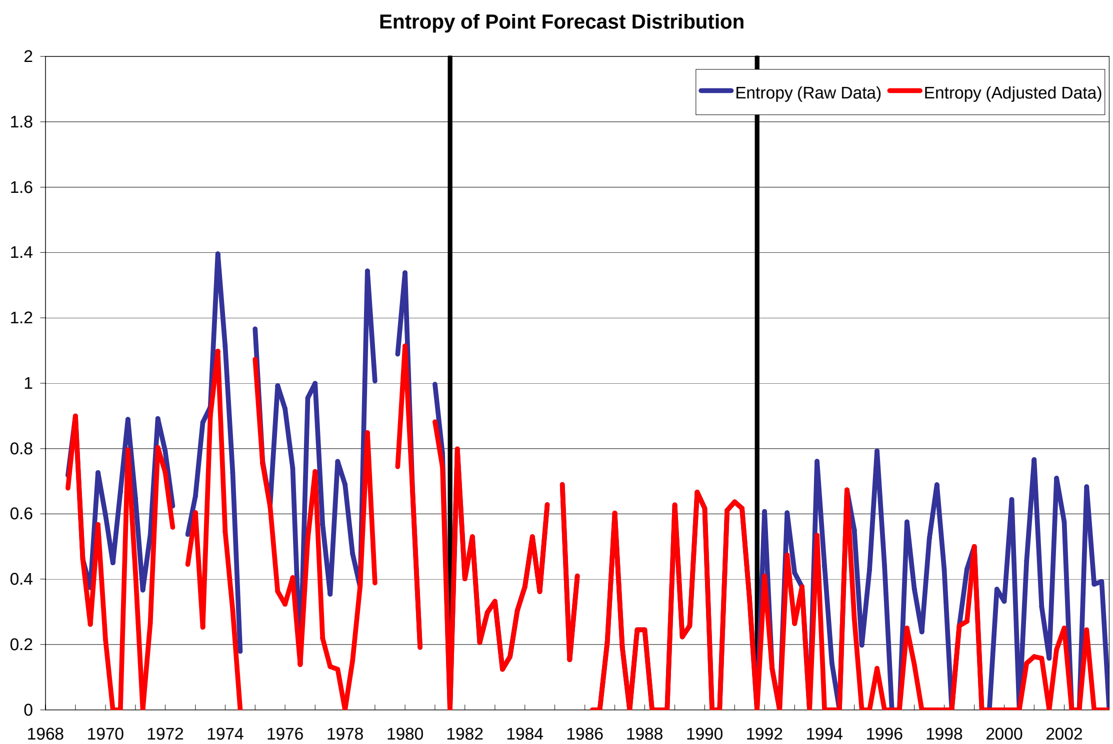


Figure 2:
Effects of Changing Interval Widths

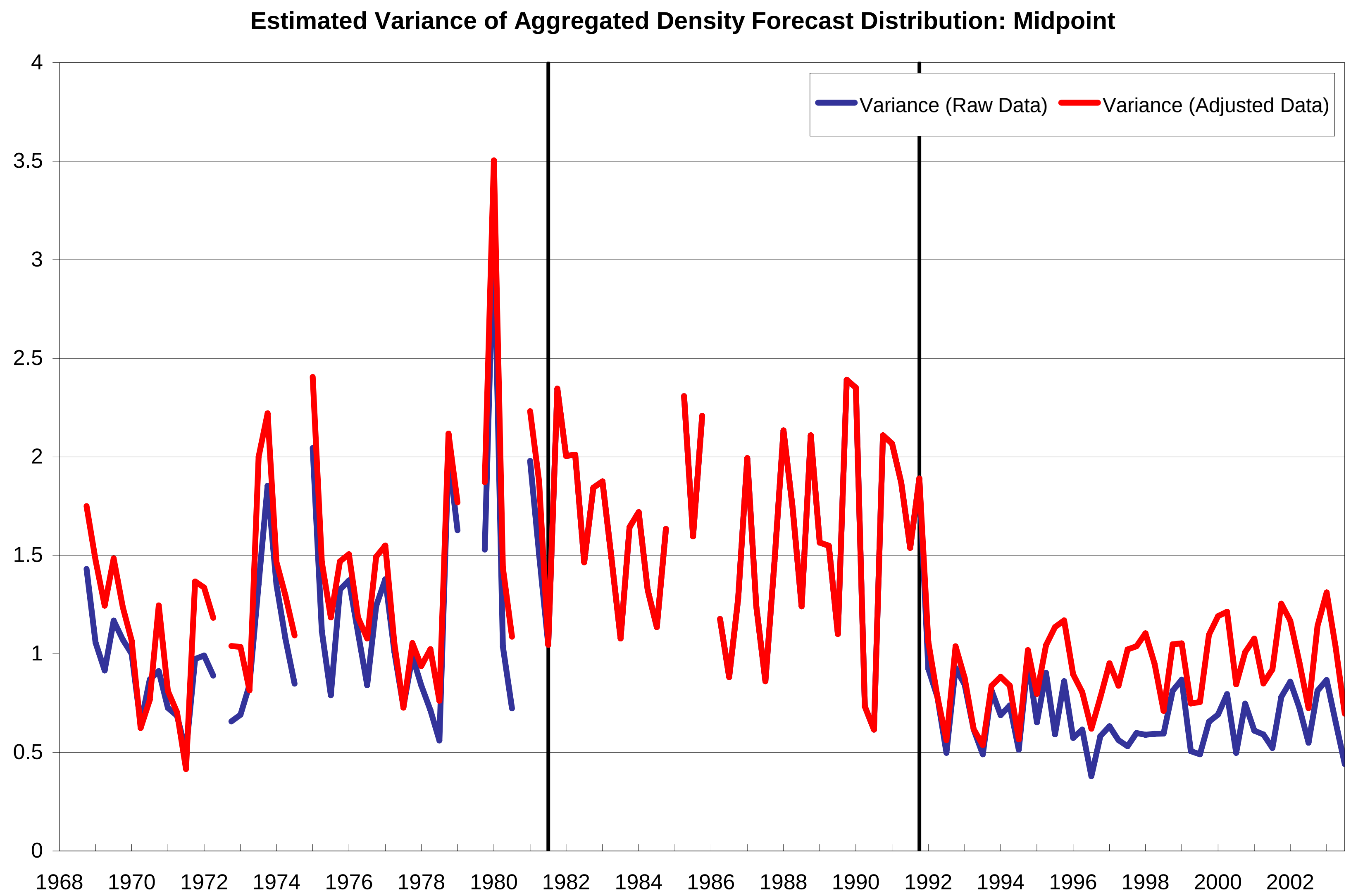
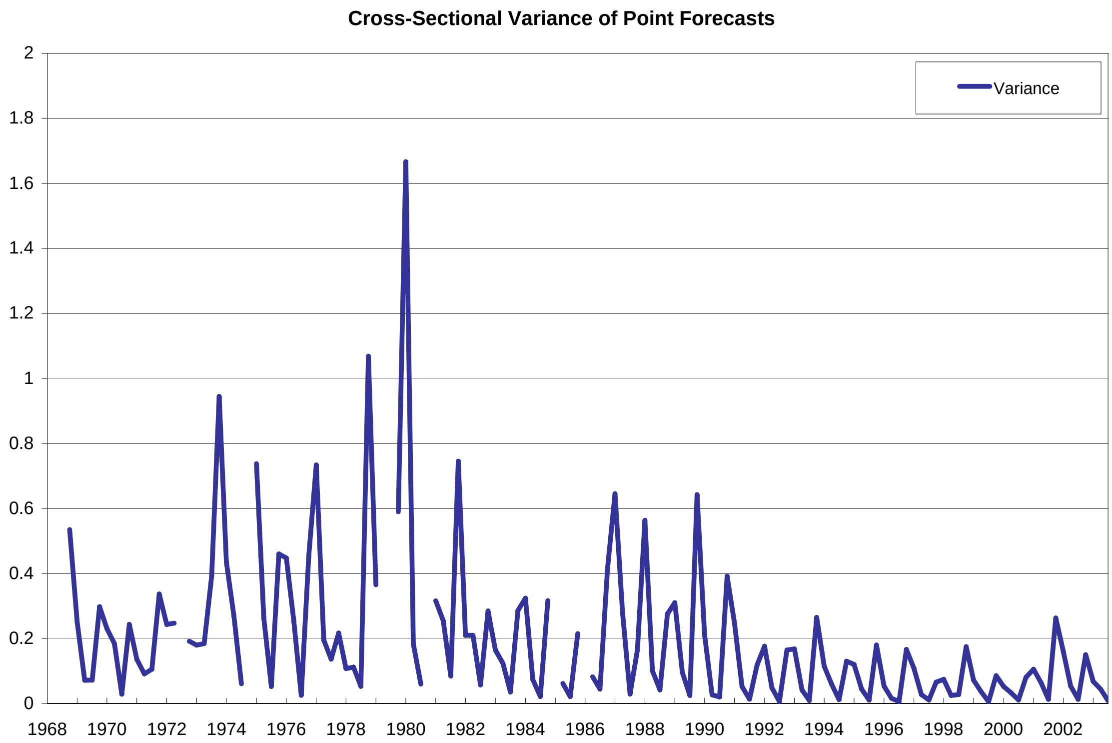


Figure 3:
Measures of Disagreement



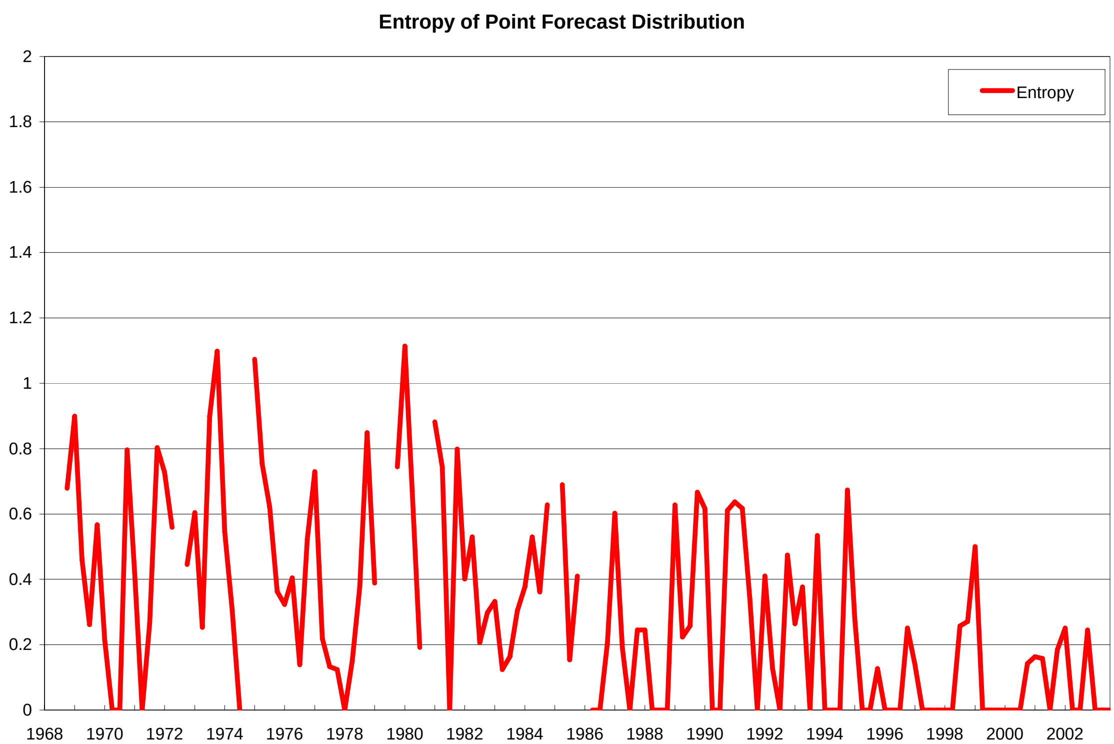
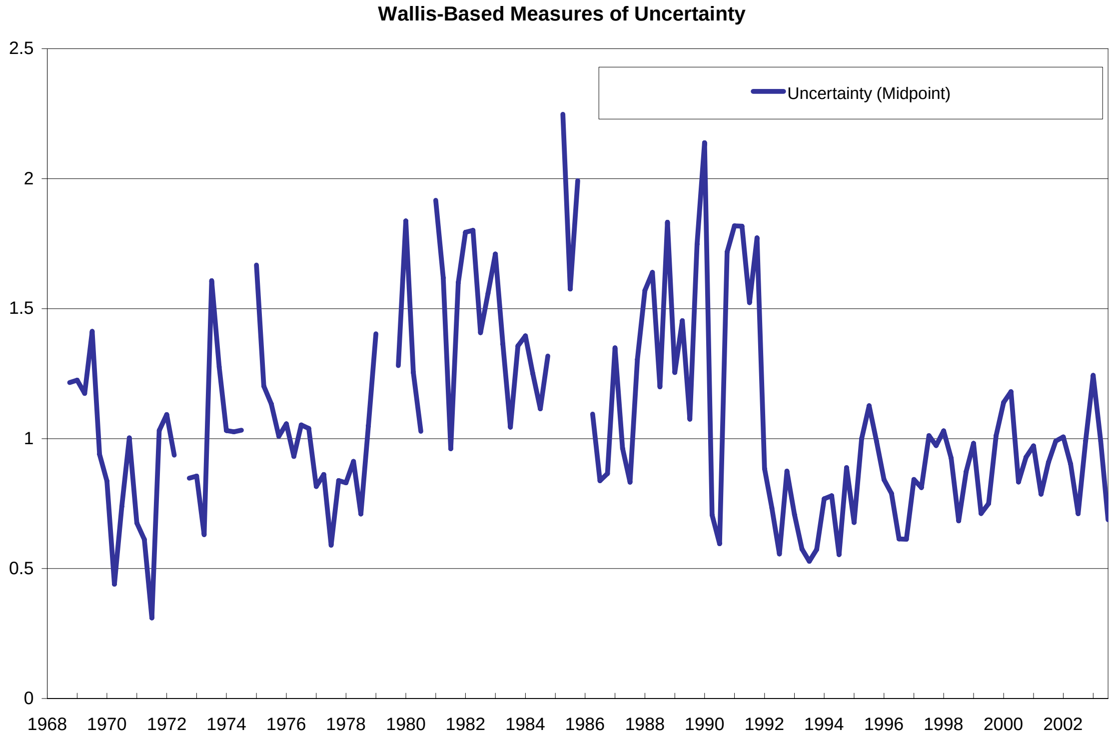


Figure 4:
Measures of Average Uncertainty



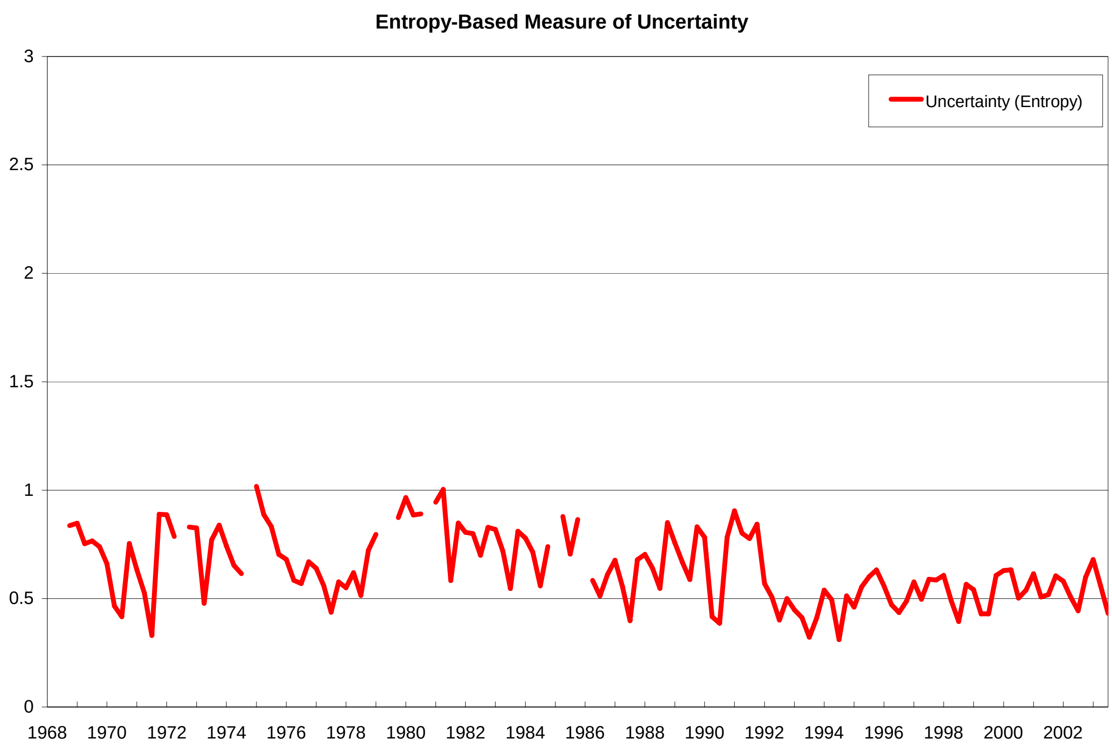


Figure 5:
Measures of Expected Inflation

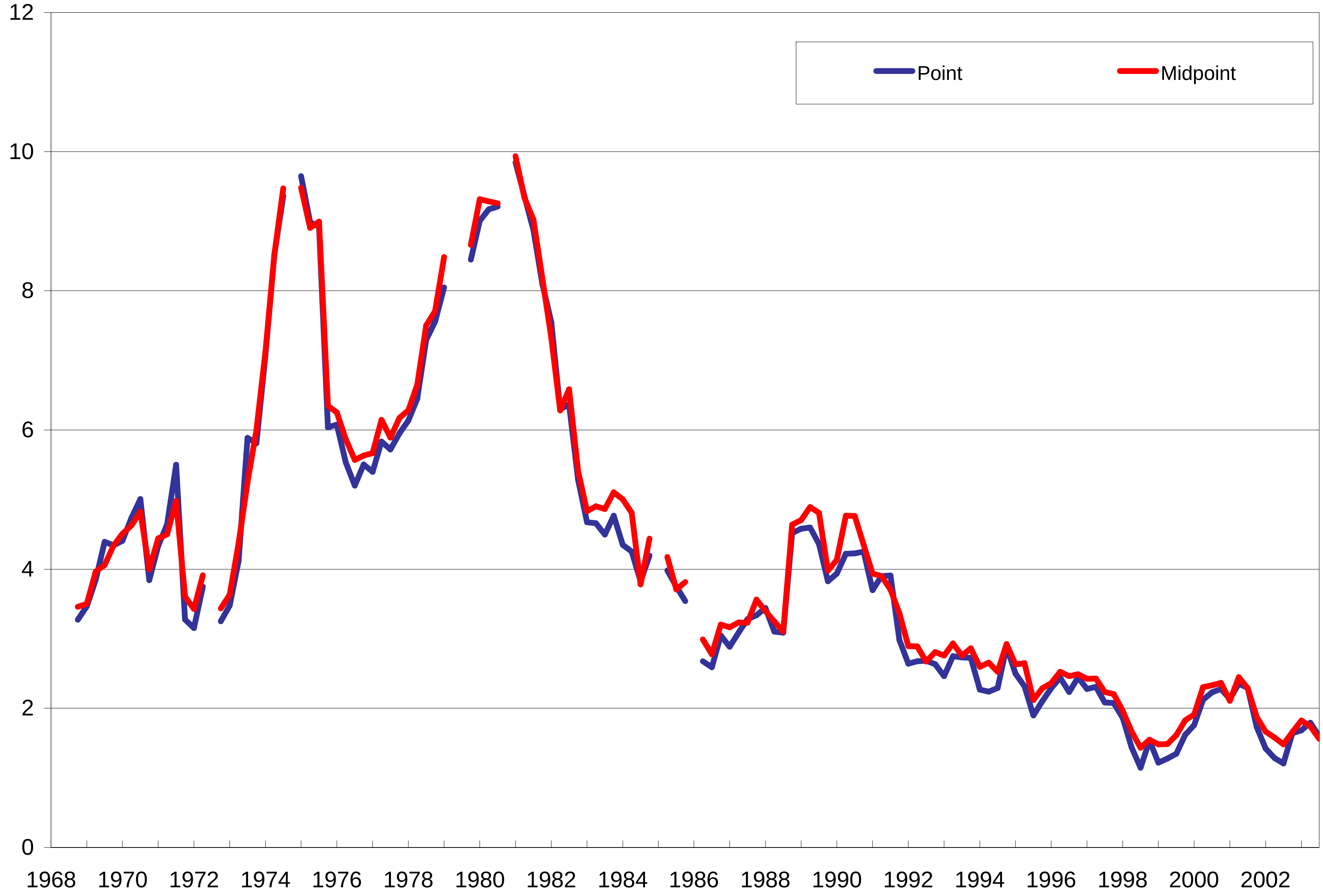


Figure 6:
Results from Fitted Distributions