Modeling Uncertainty: Predictive Accuracy as a Proxy for Predictive Confidence
Robert Rich and Joseph Tracy
Federal Reserve Bank of New York Staff Reports, no. 161
February 2003
JEL classification: C12, C22, E37

Abstract

This paper evaluates current strategies for the empirical modeling of forecast behavior. In particular, we focus on the reliability of using proxies from time series models of heteroskedasticity to describe changes in predictive confidence. We address this issue by examining the relationship between *ex post* forecast errors and *ex ante* measures of forecast uncertainty from data on inflation forecasts from the Survey of Professional Forecasters. The results provide little evidence of a strong link between observed heteroskedasticity in the consensus forecast errors and forecast uncertainty. Instead, the findings indicate a significant link between observed heteroskedasticity in the consensus forecast errors and forecast dispersion. We conclude that conventional model-based measures of uncertainty may be capturing not the degree of confidence that individuals attach to their forecasts but rather the degree of disagreement across individuals in their forecasts.

We thank Nathan Barczi and Rema Hanna for excellent research assistance. The authors have benefited from the suggestions of participants in the Federal Reserve System Committee on Macroeconomics Conference in Minneapolis. We are also grateful to Michelle Barnes and Dean Croushore for valuable comments on earlier versions of this paper. The views expressed in this paper are those of the individual authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System. Address correspondence to the authors at the Federal Reserve Bank of New York, Domestic Research Function, 33 Liberty Street, New York, NY 10045-0001. E-mail: robert.rich@ny.frb.org or joseph.tracy@ny.frb.org.

It is widely recognized that macroeconomic outcomes depend critically on both peoples' expectations and the confidence attached to those expectations. Because these magnitudes are largely unobservable, a considerable amount of work has focused on the empirical modeling of forecast behavior. As a result of this effort, there is now general agreement among applied researchers on this issue. Following the tenets of the rational expectations hypothesis, subjective predictions are assumed to be optimal forecasts given all available information and are equated to the objective conditional expectation from the specific model under consideration. With regard to modeling forecast uncertainty, the prevailing approach relies on time series models of heteroskedasticity in which the variance surrounding a prediction is allowed to change over time.[1] Temporal variation in subjective uncertainty is equated to the objective conditional variance of a series, with heightened (diminished) uncertainty associated with episodes of decreased (increased) predictability.

In spite of the important role played by uncertainty in economic behavior, there is a very limited understanding about the nature of this process. Consequently, the use of time series models of heteroskedasticity to generate estimates of uncertainty has largely been motivated by their characteristics and attractiveness for econometric applications, rather than by theoretical arguments or empirical verification. For example, it is not immediately obvious why the conditional variance of a time series, which relates to its *ex post* predictability, should be associated with forecast uncertainty, which relates to the *ex ante* confidence attached to a prediction. In addition, testing hypotheses about uncertainty is problematic because it requires information on individuals' assessments of possible outcomes of a predicted event.

---

[1]The most popular example of this modeling approach is the Autoregressive Conditional Heteroskedasticity (ARCH) model of Engle (1982) and its various extensions.

This paper is motivated by those situations in which a researcher requires a time-varying measure of uncertainty when no direct uncertainty measures (or potential proxies) are available. Specifically, the paper focuses on the reliability of using proxies from time series models of heteroskedasticity to describe changes in predictive confidence. Because these model-based measures of uncertainty are linked to the predictability of a series, we address this issue by using data from the Survey of Professional Forecasters (SPF) to examine the relationship between *ex post* squared inflation forecast errors and movements in *ex ante* inflation forecast uncertainty. The SPF inflation data is unique in that it provides direct observations on predicted outcomes as well as information on the distribution of possible outcomes, and thereby allows us to construct measures of expected inflation and inflation uncertainty for each individual respondent.

Our findings indicate only a weak relationship between observed heteroskedasticity in the consensus (average) forecast errors and forecast uncertainty. On the other hand, we document a significant and stable relationship between observed heteroskedasticity in the consensus forecast errors and forecast dispersion. Taken together, these results provide little support for current strategies used in the empirical modeling of forecast uncertainty and also raise questions of interpretation for proxies based on this approach. In particular, we would argue that the (estimated) conditional variance of a series likely is not identifying the degree of confidence that individuals attach to their forecasts of the variable, but instead is principally identifying the degree of disagreement across individuals in their forecasts.[2]

[2]Admittedly, our conclusions are based on the results from one empirical investigation. However, the lack of probabilistic forecast data for other time series will make it difficult to gain further insight into this issue.

The SPF data also allow us to examine the assumption that agents share identical beliefs and information about the structure of the economy. While the SPF data is based exclusively on professional forecasters, there is still evidence of systematic differences across respondents in their predictive confidence regarding future inflation. Consequently, changes in the set of SPF participants over time introduce important compositional effects into the measure of aggregate inflation uncertainty. Due to limited information about the respondents' characteristics, we can not undertake a detailed investigation into the possible source(s) for the observed heterogeneity in predictive confidence. Nevertheless, there is some evidence suggesting that the heterogeneity may be related to differences in either access to information or abilities to process information.

In the next section of the paper, we provide an overview of the SPF data. We also discuss the matched point and probabilistic forecasts of inflation as well as the construction of measures of forecast uncertainty. Section III describes our econometric methodology. We present the empirical results in Section IV. We conclude with a short summary of our findings.

II. Data

A. Background

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 conducted quarterly and originally asked respondents to provide point forecasts for 10 variables over a range of forecast horizons. Unlike other surveys, however, the questionnaire also asks respondents to provide probabilistic forecasts for aggregate output and inflation.[3] That is, respondents are asked to assign a probability to each of a number of 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 researchers with a unique basis from which to construct empirical measures of forecast uncertainty. 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.

There have also been changes to the SPF survey instrument over time. For example, in the early 1980s respondents switched from forecasting nominal output to real output. The lack of a homogenous sample makes it very difficult to undertake a historical evaluation of the output forecasts. Therefore, we will restrict our attention to data on the inflation forecasts.

There are several other features of the SPF survey that require careful consideration. Consequently, it is instructive to provide some details about the survey before turning to a discussion of our econometric methodology and empirical results. We begin by examining the structure of matched point and probabilistic forecasts of inflation and the construction of measures of expected inflation and inflation uncertainty. We then consider adjustments that account for changes over time in the number of intervals and their widths in the SPF's survey instrument. Because of turnover in the set of respondents, we also consider adjustments to the measures that control for compositional effects. Finally, our discussion touches upon additional issues related to errors in the conduct of the survey, changes in the base year of the price indexes, and the use of real-time versus final-revised data.

[3] The Livingston price expectations data and the Survey Research Center (SRC) expected price change data are two other major surveys of inflation expectations. However, neither survey series solicits probabilistic forecasts. Thomas (1999) contains a discussion of the Livingston series and SRC series.

B. Variable Definitions

The SPF 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. With regard to the probabilistic 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. This structure results in a target variable for the respondents that remains fixed for four consecutive surveys (from the fourth quarter of year t -1 through the third quarter of year t), with a corresponding forecast horizon (h) that declines from approximately 4½ quarters to 1½ quarters. For convenience, we shall refer to these horizons as h = 4, . . . , 1. Section III discusses the implications of this accordion pattern in the forecast horizon for the estimation.

For purposes of illustration and to provide a useful benchmark for the analysis, we will initially focus on the approach adopted by Zarnowitz and Lambros (1987) to construct measures of expected inflation and inflation uncertainty. Specifically, they assume a uniform distribution within each of the selected n intervals and apply the following formulas to calculate the mean (π^e) and variance (σ^2) of the individual histograms:

$$\pi^{e}_{jth} = \sum_{k=1}^{n} p_{jth}(k)\pi^{Mid}_{th}(k) \tag{1}$$

$$\sigma^{2}_{jth} = \sum_{k=1}^{n} p_{jth}(k)\left[\pi^{Mid}_{th}(k) - \pi^{e}_{jth}\right]^2 - \frac{w_t^2}{12} \tag{2}$$

where $p_{jth}(k)$ denotes the probability respondent *j* attaches to interval *k* in year *t* for horizon *h*, $\pi_{th}^{Mid}(k)$ denotes the midpoint of the corresponding interval, and the last term in the variance calculation is the Sheppard correction to take account of changes in the interval widths, w_t, over time.[4] The individual means and variances are then averaged across all N_t respondents for the same survey to obtain the consensus probabilistic forecast and consensus predictive uncertainty series given, respectively, by:

$$\bar{\pi}_{th}^{e} = (1/N_t)\sum_{j=1}^{N_t}\pi_{jth}^{e} \tag{3}$$

$$\bar{\sigma}_{th}^{2} = (1/N_t)\sum_{j=1}^{N_t}\sigma_{jth}^{2} \tag{4}$$

Turning to the point forecasts, the survey asks respondents for predictions of the price level for the current and the next four quarters. Because data is available on the price index in the preceding quarters, a point forecast f_{jth}^{e} can be constructed that matches each π_{jth}^{e}. For example, the point forecast for respondent *j* of the annual average percentage change in prices in the fourth quarter of year τ is given by:

$$f_{j\tau 4}^{e} = 100 * \left[\frac{P_{j\tau+1,1} + P_{j\tau+1,2} + P_{j\tau+1,3} + P_{j\tau+1,4}}{A_{\tau,1} + A_{\tau,2} + A_{\tau,3} + P_{j\tau,4}} - 1\right] \tag{5}$$

[4]See Kendall and Stuart (1963). Because the lowest and highest intervals are open-ended, our calculations treat them as closed intervals and set their width equal to that of the interior intervals. The sample period in Zarnowitz and Lambros (1987) is 1968:Q4-1981:Q2 which ended before there were switches in interval widths. Consequently, Zarnowitz and Lambros did not use the Sheppard correction.

where $P_{j\tau,q}$ is respondent *j*'s predicted value of the price level in quarter *q* of year τ and $A_{\tau,q}$ is the "actual" value of the price level in quarter *q* of year τ.[5] The subsequent point forecast in quarter 1 of year τ+1 is then given by:

$$f^{e}_{j\tau+13} = 100 * \left[\frac{P_{j\tau+1,1} + P_{j\tau+1,2} + P_{j\tau+1,3} + P_{j\tau+1,4}}{A_{\tau,1} + A_{\tau,2} + A_{\tau,3} + A_{\tau,4}} - 1 \right] \tag{6}$$

where the *P*'s and *A*'s reflect the new quarterly price level predictions and realizations, respectively. A similar updating would occur for $f^{e}_{j\tau+12}$ and $f^{e}_{j\tau+11}$.

As in the case of the probabilistic forecasts, the point forecasts can be averaged across individuals for the same survey to obtain a consensus point forecast series:

$$\overline{f}^{e}_{th} = (1/N_t)\sum_{j=1}^{N_t} f^{e}_{jth} \tag{7}$$

A measure of forecast dispersion can also be calculated based on the cross-sectional variance of the individual point forecasts:

$$s^2_{f_{th}} = (1/N_t)\sum_{j=1}^{N_t} \left(f_{jth} - \overline{f}^{e}_{th} \right)^2 \tag{8}$$

where equation (8) measures the degree of disagreement across individual predictions.

The measure of inflation uncertainty previously described in equation (4) reflects the

---

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

average subjective variance across individual respondents. However, we can consider an alternative measure of uncertainty based on the aggregate probability distribution from the survey. Specifically, the individual probability distributions can be combined into an aggregate probability distribution by calculating the mean probability for each inflation interval across respondents. We can then calculate the variance of the aggregate probability distribution (ϕ^2). A feature of this measure is the following decomposition:

$$\phi_{th}^2 = \overline{\sigma}_{th}^2 + \overline{s}_{th}^2 \tag{9}$$

where the variance of the aggregate probability distribution of inflation is equal to the mean of the respondents' variances plus the variance of the respondents' forecasted means (s^2). Thus, equation (9) motivates including dispersion measures as an additional control for uncertainty.

C. Data Considerations

The previous discussion abstracts from several important data issues. 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 aggregate probabilistic and point 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 probabilistic 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 probabilistic 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.

There are also issues pertaining to the calculation of summary measures using the raw probabilistic forecast data. One concern stems from 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.

Another concern with the SPF data is changes in the number and composition of the survey respondents. As shown in Figure 1, the number of respondents participating in the survey has changed over time and varied from a low of 14 to a high of 65. While new entrants to the survey were concentrated in the early 1970s and 1990s, there was a steady flow of permanent exits in the 1980s. Figure 2 provides further insight into the average annual turnover rate of respondents. In each year, there is a sizeable flow of entrants and exits from the survey. A comparison of Figures 1 and 2 illustrates that many of the exits in Figure 2 are temporary, with the respondent re-entering the survey at a later date. If there are systematic differences across individuals in their forecast behavior, then this churning of respondents could lead to problems in comparing over time aggregate measures of expected inflation and inflation uncertainty.

Evidence from two recent papers examining household-level survey data offers additional support for our concern about compositional effects. Souleles (2002) uses data from the Michigan Index of Consumer Sentiment and documents differences across demographic groups in their inflation expectations. Carroll (2001) develops a model in which there are differences across demographic groups in their propensity to pay attention to economic news.

Using data on inflation expectations and unemployment expectations from the Michigan household survey, he finds that the qualitative behavior of the cross-sectional standard deviation of inflation forecasts roughly matches the predictions of the model. Because of the greater similarity among SPF respondents as compared to the Michigan respondents (professional forecasters versus households), one might expect a closer correspondence in respondents' forecast behavior than that reported by Souleles (2002). Nevertheless, this remains an open empirical question.

In an attempt to gauge the importance of these various concerns, we will derive measures of expected inflation and inflation uncertainty from the raw probabilistic forecast data using a number of different approaches. The first approach relies on the formulas in equations (1) and (2) and uses a uniform distribution over the individual histograms and the original SPF intervals. As previously noted, the Sheppard correction is intended to control for the impact of varying interval widths on the estimated variance. Alternatively, we can redefine the intervals to impose a common 2% width throughout the whole sample period and drop the Sheppard correction from the variance calculation.[6] Using the 2% intervals, we can also reestimate the expected inflation rate and its variance by fitting a continuous distribution [for example, a Normal distribution as in Giordani and Söderlind (2000)] to the individual SPF's histograms.[7] This approach is likely to produce a lower estimate of the variance because the Normal distribution shifts the conditional

[6]Due to the odd number of intervals used over the subperiod 1968:Q4-1981:Q2, we use a unit interval length for the middle interval.

[7]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. Our approach differs from Giordani and Söderlind (2000) who fit normal distributions to the aggregate (or mean) histograms, rather than the individual histograms.

mean for each interval inward from the midpoint towards the overall mean.

We use the panel structure of the SPF to investigate how changes over time in the respondents may impact the estimates of expected inflation and inflation uncertainty.[8] For a given forecast horizon *h*, we regress the individual respondent's expected inflation (inflation uncertainty) on a set of year dummy variables and a set of respondent fixed-effects. The estimated respondent fixed effects reflect the extent to which a particular respondent's expected inflation (inflation uncertainty) systematically differs from the average adjusting for the years that the respondent participated in the survey. By subtracting out these fixed-effect estimates from the respondent's expected inflation (inflation uncertainty) estimates, we can control for changes in the composition of the survey.

There are a few remaining issues that merit discussion. First, 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).[9] Presently, the survey asks about inflation as measured by the chain-weighted GDP index. The analysis also accounts for periodic changes in the base year of the relevant price indexes.[10] Last, there is the question of whether to use real-time or final-revised data. To explore the importance of this issue, we examined both types of data in the empirical analysis. We generally found that the results were not sensitive to

---

[8]The survey assigns an identification number to each respondent. However, the SPF does not report any demographic information for the respondents.

[9]We do not analyze the SPF CPI inflation forecasts because they were not initiated until 1981:Q3.

[10]For surveys on or before 1975:Q4, the base year is 1958. From 1976:Q1-1985:4 the base year is 1972, while from 1986:Q1- 1991:Q4 the base year is 1982. Beginning in 1992:Q1, the base year is 1987.

this feature of the data. Consequently, we only report the results using the most recently revised values as our series for realized inflation.[11]

III. Econometric Methodology

The SPF inflation data have been used to investigate a wide range of issues.[12] For example, Zarnowitz (1984, 1985), Keane and Runkle (1990), and Bonham and Cohen (1995, 2001) evaluate the forecast performance and statistical properties of individual and consensus predictions. Zarnowitz and Lambros (1987) characterize the relationship between various summary measures such as the mean forecast, forecast dispersion, and forecast uncertainty. Other studies such as Lahiri, Teigland and Zaporowski (1988), Pennacchi (1991), and Neumark and Leonard (1993) quantify the effects of expected inflation and inflation uncertainty on key macroeconomic and financial variables.

While the SPF inflation data in principle can allow researchers to test competing theories of forecast behavior, there are only a few studies that have focused on *ex ante* measures of uncertainty and the ability of time series methods to approximate their behavior. Batchelor and Dua (1993, 1996) consider a host of proxies used in empirical studies such as forecast standard deviations from ARIMA, ARCH and structural models of inflation. They conclude that these proxies are not significantly correlated with the survey-based measures of uncertainty. Giordani

[11]We also considered alternative measures of the consensus inflation forecast and inflation uncertainty. The results were unaffected when we used the median forecast for the measure of expected inflation or inner quartile ranges for the measures of forecast dispersion and forecast uncertainty.

[12]The Federal Reserve Bank of Philadelphia maintains a web page listing academic articles that have used various SPF data series. See Croushore (1993) for additional details about the SPF and its history.

and Söderlind (2000) undertake a similar comparison of alternative time series measures to the survey-based measures of uncertainty. Their results largely corroborate the findings of Batchelor and Dua (1993, 1996).

While we are interested in answering the same general question as these other studies, our approach differs in a number of ways. One particularly important difference concerns the variables used for comparison to the *ex ante* measures of uncertainty. Batchelor and Dua (1993, 1996) and Giordani and Söderlind (2000) include proxies for inflation uncertainty derived from time series models of heteroskedasticity. However, the nature of the changing forecast horizon for the SPF inflation data presents difficulties for the estimation of conventional time series models of heteroskedasticity.[13] Abstracting from this issue, there is also the question of which model specification(s) should be selected. Finally, there is an apparent disconnect in these analyses. Specifically, Batchelor and Dua (1993, 1996) and Giordani and Söderlind (2000) construct model-based forecasts of inflation, rather than use data on expected inflation from the SPF, to generate their model-based measures of uncertainty. Because there is no matching of the inflation forecast series, there is no direct correspondence between the model-based measures of uncertainty and the survey-based measures of uncertainty. Taken together, these considerations raise concerns about the comparability of the measures employed in these studies as well the reliability of their conclusions.

As an alternative approach, we will focus our attention on the relationship between the *ex post* predictive accuracy of the survey forecasts and *ex ante* measures of uncertainty. While we

[13]Batchelor and Dua appear to adopt a different definition for the target inflation rate forecasted by the survey respondents, while Giordani and Söderlind essentially ignore this feature of the data and generate conditional variance estimates that relate to a constant forecast horizon.

will not consider specific time series models of heteroskedasticity, our framework will be general enough to allow us to gauge the reliability of proxies for uncertainty derived from this modeling strategy. Moreover, the use of predictive accuracy maintains a consistency between the measures of expected inflation and inflation uncertainty in our investigation. Consequently, we will not need to introduce model-based measures of expected inflation or inflation uncertainty which will avoid problems of comparability to the SPF data.

There are several other aspects of our empirical analysis that are noteworthy. First, we pool the data across forecast horizons which allows for a substantial increase in degrees of freedom. This approach contrasts with other studies such as Batchelor and Dua (1993, 1996), Diebold, Tay and Wallis (1999), and Giordani and Söderlind (2000) who either restrict their attention to a single forecast horizon, or analyze the data over separate forecast horizons. Second, we focus on the consensus inflation forecast and consensus forecast error. This choice is motivated by our interest in the behavior of these measures at the aggregate level, as well as by consistent evidence indicating that an unweighted combination of forecasts is more accurate than individual forecasts.[14] Last, the analysis departs from these studies by incorporating forecast dispersion and exploring its relationship to heteroskedasticity in the consensus forecast errors.

With regard to this last point, our work is also related to Bomberger (1996). Specifically, he examines inflation forecasts from the Livingston price expectations data and finds a significant and stable relationship between the conditional variance of the inflation forecast errors and disagreement. Under the maintained assumption that the estimated conditional

[14]Zarnowitz and Braun (1993) provide such evidence in their analysis of the NBER-ASA Economic Outlook Survey. We recognize, however, that analysis at the aggregate level limits the conclusions that can be drawn about individual behavior. As Keane and Runkle (1990) note, the finding of unbiasedness in consensus regressions does not necessarily imply individual forecasts are unbiased.

variance process provides a reasonable measure of uncertainty, Bomberger concludes that disagreement is a potentially useful proxy for uncertainty.[15] Our analysis will provide an opportunity to evaluate the validity of Bomberger's maintained assumption and his conclusion.

There is one final aspect of the empirical analysis that merits discussion. Because the matched point and probabilstic forecast series each provide a measure of expected inflation, there is a question about which data should be used to construct the measures of predictive accuracy and forecast dispersion. Our belief is that the point forecast series is the more appropriate choice because it provides direct observations on expected inflation, and correspondingly on *ex post* predictive accuracy and disagreement. In contrast, the probabilistic forecast series requires the adoption of auxiliary assumptions to estimate a mean for each of the individual probability distributions. Nevertheless, we consider alternative constructs for some of the variables to gauge the robustness of the results.

We now turn to a discussion of the econometric methodology we use to examine the relationship between *ex post* predictive accuracy and *ex ante* forecast uncertainty as well as to test for unbiasedness of the inflation forecasts. We also describe a modified covariance matrix estimator that takes into account the unique correlation structure of the regression disturbances.

A. Aggregate Specifications

Time series models of heteroskedasticity use variation in second moments to measure uncertainty. Within the context of our study, we can describe this approach as:

---

[15]Recall that the Livingston survey does not provide probabilistic forecast data. It is worth noting that the Livingston survey data also involves forecast horizons that exceed unity, although Bomberger appears to ignore this feature of the data in his analysis.

$$
\begin{aligned}
\pi_{th}^{e} &= E\left[\pi_{th} \mid I_{th}\right] \\
V(\pi_{th}) &= E\left[\left(\pi_{th} - \pi_{th}^{e}\right)^{2} \mid I_{th}\right]
\end{aligned} \tag{10}
$$

where π_{th} denotes the SPF target inflation rate associated with the survey conducted in year *t* with a forecast horizon of *h* quarters, and π_{th}^{e} and $V(\pi_{th})$ denote, respectively, the mean and variance of π_{th} conditional on all information available at the time of the survey conducted in year *t* with a forecast horizon of *h* quarters, (I_{th}).

Under the assumption that respondents make efficient use of information, the system of equations in (10) implicitly defines regression models in which differences between the actual and expected value of variables reflects the influence of random disturbance terms. Thus, if the consensus point forecast is an unbiased estimator of expected inflation, then we can rewrite the system of equations in (10) as:[16]

$$
\begin{aligned}
\pi_{th} &= \bar{f}_{th}^{e} + \varepsilon_{t,h} \\
\left(\varepsilon_{t,h}\left[= \pi_{th} - \bar{f}_{th}^{e}\right]\right)^{2} &= V(\pi_{th}) + \eta_{t,h}
\end{aligned} \tag{11}
$$

where $E[\varepsilon_{t,h} \mid I_{th}] = E[\eta_{t,h} \mid I_{th}] = 0$. As shown, time series models of heteroskedasticity inherently link the conditional variance of a stochastic process (*V*) to changes in its predictability (ε^{2}). In practice, a regression function for $V(\pi_{th})$ can specified and estimated incorporating restrictions to ensure that the conditional variance process is well behaved. The value of the estimated conditional variance process at each time period can then be used as a proxy for

[16]The following analysis also holds if we used the estimated consensus mean of the probabilistic forecast series $(\bar{\pi}_{th}^{e})$ to measured expected inflation rather than the consensus point forecast $(\bar{f}_{th}^{e})$.

forecast uncertainty.[17]

The key question for our study centers on the reliability of using time series models of heteroskedasticity to construct empirical measures of uncertainty. If we assume that the conditional variance of the consensus inflation forecast errors is a linear function of the survey-based measures of inflation uncertainty, then we can use the system of equations in (11) to cast the analysis exclusively in terms of the relationship between the forecasting accuracy and forecast uncertainty of the SPF respondents:

$$\left(\varepsilon_{t,h}\right)^2 = \alpha + \beta\left(\bar{\sigma}_{th}^2\right) + \eta_{t,h} \tag{12}$$

Our testing procedure will focus on the hypothesis that $\beta > 0$. Because the magnitude and behavior of the inflation forecast errors reflect *ex post* outcomes, it is not immediately clear that they should provide a reliable basis to construct proxies for *ex ante* uncertainty.[18]

While equation (12) offers some insight into the relationship between observed heteroskedasticity in the forecast errors and survey-based measures of uncertainty, it is silent on whether the correlation is robust to the inclusion of other variables. Admittedly, the list of

[17]These formulations could include the autoregressive conditional heteroskedasticity (ARCH) model introduced by Engle (1982) in which the conditional variance of a series varies over time as a function of past squared forecast errors. Alternatively, one could postulate that the conditional variance is related to a set of observable variables. For the purpose of this study, however, the system of equations in (11) is only intended to illustrate the link between the squared forecast errors and the conditional variance of a time series. Consequently, we do not attempt to provide a detailed discussion of various specifications for the conditional variance of a time series process or of alternative methods for model estimation. See Bollerslev, Chou and Kroner (1992) for a survey of ARCH-type models and a review of this literature.

[18]This regression equation merely provides a convenient framework to address the issue of correlation and is not intended to imply a causal relation. We will also examine this relationship using the estimated mean of the probabilistic forecast series to construct the measure of predictive accuracy.

candidates for consideration is quite large. Drawing upon our earlier discussion of equation (9), we will limit our choice to one variable that can be easily motivated. Specifically, we will also consider the following regression equation:

$$\left(\varepsilon_{t,h}\right)^2 = \alpha + \beta\left(\overline{\sigma}^2_{th}\right) + \gamma\left(s^2_{f_{th}}\right) + \eta_{t,h} \tag{13}$$

where we include the cross-sectional variance of the point forecasts from equation (8) as an additional regressor.

Estimation of equation (13) is of additional interest because it bears directly upon the interpretation that researchers ascribe to the estimated conditional variance process of a time series. Specifically, we can test for the individual statistical significance of β and γ to determine if the variance of the forecast errors is reflecting effects associated with uncertainty or picking up additional effects related to disagreement. The regression equation also allows us to test the restriction $\beta = \gamma$ implied by the hypothesis that the variance of the forecast errors is more closely associated with the variance of the aggregate probability distribution of inflation.[19]

The previous analysis maintains the assumption that the consensus forecast is an unbiased estimator of expected inflation. Because the issue of unbiasedness bears upon the construction of the measure of predictive accuracy and is also of some interest in itself, we will consider the following regression equation to test for unbiasedness:

$$\pi_{th} = \alpha + \beta\,\overline{f}^{\,e}_{th} + \varepsilon_{t,h} \tag{14}$$

[19]The variance decomposition in (8) holds as an identity when we construct the two components using the probabilistic distributions. Equation (8) should hold approximately when we construct the dispersion measure using the point forecasts of inflation.

We test for unbiasedness in equation (14) by testing the joint restriction that $\alpha = 0$ and $\beta = 1$.[20] If the consensus forecasts were to reveal evidence of systematic bias, then the analysis could proceed by using a biased-adjusted measure of predictive accuracy. Estimation would then involve joint estimation of the system of equations for the mean and variance of inflation, with cross-equation restrictions imposed on the squared residual $\hat{\varepsilon}^2_{t,h}$.

B. Estimation

We use the method of ordinary least squares (OLS) to estimate the aggregate regression equations over pooled forecast horizons.[21] There is, however, one complication involved in the estimation. The regressions for equations (12)-(14) involve overlapping data due to the forecast horizon exceeding the sampling interval in each case.[22] A consequence of overlapping data is that the error terms in the regression may be autocorrelated and follow a moving average (MA) process. While OLS provides consistent parameter estimates, we need to compute the standard errors using a covariance matrix estimator that can account for this feature of the data.

While there are several covariance matrix estimators that have been proposed for use with overlapping data, a common feature of their design is the assumption of a constant forecast horizon. That is, the data are assumed to contain observations on a *k*-step-ahead forecast, where the value of *k* does not change over the course of the sample period. As a result, the regression disturbances will follow a MA(k-1) process. Under this scenario, Hansen (1982) and White

[20]We carry out a similar test of the unbiasedness of the consensus estimated mean forecasts.

[21]We carried out Chow tests to verify that pooling the data was permissible.

[22]Recall the forecast horizons vary from 1½ - 4½ quarters, while the sampling interval is 1 quarter.

(1984, Chapter 6) have proposed the following covariance matrix estimator:

$$\left(\frac{\hat{\Omega}_T}{T}\right) = \left[\sum_{t=1}^{T} x_t x_t'\right]^{-1} \left[\sum_{t=1}^{T} \hat{\mu}_t^{\,2} x_t x_t' + \sum_{v=1}^{k-1}\left\{\sum_{t=v+1}^{T} x_t \hat{\mu}_t \hat{\mu}_{t-v} x_{t-v}' + x_{t-v}\hat{\mu}_{t-v}\mu_t x_t'\right\}\right]\left[\sum_{t=1}^{T} x_t x_t'\right]^{-1} \tag{15}$$

where x_t is a (n x 1) vector of explanatory variables and $\hat{\mu}_t$ is the estimated OLS sample residual for date *t*.[23]

In the case of the SPF data, however, the forecast interval displays an accordion pattern with a recurrent decline over the course of the sample period. While this pattern leads to a different correlation structure for the regression residuals from that typically observed with overlapping data, we can nevertheless design an appropriate covariance matrix estimator.[24] The details on this modified covariance estimator are provided in the Appendix.

IV. Empirical Results

Our sample includes the surveys conducted from 1968:Q4 through 2000:Q3, so that the values on the realized annual rate of inflation pertain to the periods 1968-69 through 2000-01. We begin by examining measures of expected inflation and inflation uncertainty, as well as the effect of various adjustment procedures on their behavior. We then present the results from estimation of the aggregate regression equations.[25]

---

[23]Another advantage of the estimator is that it also takes account of possible conditional heteroskedasticity of regression disturbances.

[24]Batchelor and Dua (1991) use of a similarly modified covariance matrix estimator in their investigation into the rationality of forecasts from the Blue Chip service.

[25]Preliminary estimation of the aggregate regression equations suggested the presence of an extreme outlier in 1973:Q4, resulting from the first oil price shock. Consequently, we exclude this observation from the empirical analysis.

A. Measuring Expected Inflation and Inflation Uncertainty

Figures 3-4 present measures of expected inflation and inflation uncertainty from the probabilistic forecast data. We do not plot the series individually, but rather depict selected pairings. The order of the pairings provides a sequential view of how each data adjustment affects the aggregate behavior of the series.

There are several important features of the data that emerge from inspection of the plots. The measures of expected inflation appear to be essentially unaffected by any of the data adjustments. That is, there is a strong similarity in the series after controlling for the switches in the number and widths of intervals as well as using the normal versus uniform distribution. With regard to the presence of compositional effects, we find evidence of statistically significant forecaster fixed effects which account for 2-25% of the residual variation across respondents (after controlling for year effects). However, the compositional effects do not appear to be economically significant in that controlling for their presence does not alter the behavior of the consensus inflation forecast series.

In contrast to the expected inflation series, there are much more noticeable differences across the measures of inflation uncertainty. As shown in the top panel of Figure 4, there is a large difference during the 1970s and 1990s between the estimates based on the original SPF intervals using the Sheppard correction and the 2%-adjusted SPF intervals. Moreover, there is clear evidence that the use of a uniform distribution leads to higher estimates of inflation uncertainty relative to those based on the normal distribution. When we fit normal distributions to the 2%-adjusted histograms, there is a downward shift on the order of 0.4 in the estimated inflation uncertainty series. Last, there are important systematic differences in predictive uncertainty across respondents. The estimated fixed effects are not only statistically significant, but also economically significant as they account for 35-50% of the residual variation across respondents (after controlling for year effects). As shown in the last panel, controlling for compositional effects raises the estimates of inflation uncertainty through the early 1990s and slightly lowers the estimates thereafter.

B. Aggregate Regression Results

We begin by examining issues related to the construction of the measures of predictive accuracy. Table 2 presents the results from testing for the unbiasedness property of the consensus point and consensus probabilistic mean forecasts. There is little difference in the results across the various measures of expected inflation. The forecasts can account for approximately 80% of the total variation in annual inflation rates over the last 30 years. In addition, we are unable to reject the unbiasedness property at conventional significance levels, indicating that we can measure predictive accuracy as the unadjusted (squared) difference between actual inflation and the consensus forecast. The robustness of the unbiasedness tests is not surprising in light of the very similar behavior of the series previously depicted in Figure 3.

We now turn our attention to Table 3 which presents the results from testing the relationship between predictive accuracy and predictive confidence.[26] The results are sensitive to whether we calculate the squared forecast errors using the consensus point forecasts or the consensus estimated mean forecasts. The upper panel of Table 3 reports the results based on the

[26]The regressions maintain a consistency between the adjusted inflation measures used to construct the consensus forecast errors and the corresponding adjusted measures of inflation uncertainty.

consensus point forecasts.[27] As shown, we only find evidence a positive and significant relationship between *ex post* accuracy and *ex ante* uncertainty in the case of the last uncertainty measure. This measure is derived from the normal approximations and has been adjusted to take account of composition changes over time in the survey.

The lower panel of Table 3 reports the results based on the consensus estimated mean forecasts. In contrast to the point forecast results, there is now consistent evidence of a statistically significant positive association between the *ex post* squared forecast errors and the *ex ante* measures of inflation uncertainty. There is also a noticeable improvement in the fit of the model and an increase in the statistical significance of the relationship as we consider additional adjustments to the measures of inflation uncertainty.[28]

Table 4 presents the results from estimation of equation (13) in which we augment the set of regressors to include a measure of forecast dispersion.[29] As in Table 3, the top panel of Table 4 uses an *ex post* accuracy measure based on the consensus point forecasts. The data indicate that periods in which there is more disagreement across forecasters are on average periods with larger

[27]As previously discussed, we view the *ex post* accuracy measure based on the consensus point forecasts as having an advantage relative to the *ex post* accuracy measure based on the consensus mean forecasts in that no estimation is involved in its construction.

[28]While the value of the $\overline{R}^2$ measure from the regression equations could be used as an additional metric to judge the reliability of proxies derived from time series models of heteroskedasticity, we do not focus on this issue. Because squared forecast errors can be highly noisy series, most practitioners estimating conditional variance processes are more concerned with statistical significance of the regressors than goodness of fit. While the reported values of $\overline{R}^2$ measure in the lower panels of Table 3 and Table 4 may appear low, they are nevertheless within the range of values obtained in studies such as Engle (1983).

[29]To maintain consistency with the corresponding uncertainty measure, we adjusted the forecast dispersion measure for compositional effects for the regression results reported in rows 5 and 9 of Table 4.

forecast errors. Controlling for the dispersion in point forecasts across forecasters, there is a positive but imprecisely estimated relationship between *ex post* accuracy and *ex ante* uncertainty.

The results change somewhat when we examine the lower panel of Table 4 and instead consider the *ex post* accuracy measure based on the consensus estimated mean forecasts. Specifically, predictive accuracy now displays a positive and significant relationship with average uncertainty using the Sheppard correction as well as the two uncertainty measures based on the normal approximation.

Irrespective of the approach used to compute predictive accuracy, there is consistent evidence of a significant positive relationship between the magnitude of the inflation forecast errors and forecast dispersion. A closer inspection of the results in Table 4 provides further insight into the nature of this relationship. When we measure the predictive accuracy of inflation using the consensus point forecasts, the coefficient on disagreement is larger and more precisely estimated than the coefficient on average uncertainty. When we measure the predictive accuracy of inflation using the consensus mean forecasts, we can not reject the hypothesis that the coefficient on disagreement and average uncertainty are the same. However, even in this case, disagreement plays a more important role than average uncertainty in explaining the variation in the predictive accuracy of inflation. Examining both individual and pooled horizons within our sample, we find that disagreement has a variance that is twice that for the average uncertainty. Thus, disagreement appears to be a more important contributor than average uncertainty to movements in predictive accuracy and consequently to conventional model-based measures of uncertainty.

There are two key results that emerge from the empirical analysis. First, the notion that the conditional variance of a time series provides a useful description of forecast uncertainty appears tenuous. While we are able to uncover a link between observed heteroskedasticity in forecast errors and *ex ante* uncertainty, the evidence in support of this relationship is not robust across different measures of predictive accuracy. Moreover, we would argue that the more favorable evidence relies on the less preferred construct for predictive accuracy. Secondly, the findings speak much more directly to a separate and stable link between observed heteroskedasticity in the consensus forecast errors and disagreement.

These findings have a number of important implications. For example, they suggest the need to develop theoretical models that clearly differentiate between the effects of uncertainty and disagreement on the decision-making process of agents.[30] This research might then be able to offer some insight or explanation for the observed relationship between predictive accuracy and disagreement. For the moment, however, our results bear upon the common practice of estimating variances from time series models of heteroskedasticity as well as the general acceptance of these measures as valid proxies for uncertainty. At best, these measures may not be capturing effects solely related to the confidence that individuals attach to their forecasts. At worst, these measures may instead be primarily describing changes in the degree of disagreement across individuals in their forecasts.

C. Heterogeneity in Predictive Uncertainty

Before concluding, we think it is instructive to try and gain some insight into the issue of

[30]There are theoretical studies that have examined the relationship between uncertainty and forecast dispersion. For example, Cukierman and Wachtel (1979) investigate this relationship within the Lucas (1972) 'island' model of imperfect information. However, their analysis focuses on the response of forecast dispersion to exogenous changes in uncertainty, rather than the nature of these two effects on the behavior of economic agents.

heterogeneity in predictive uncertainty. Our interest is partly attributable to the central role of predictive uncertainty in this paper. In addition, the literature has focused much more of its attention on differences in individual forecasts rather than individual uncertainty. Because the SPF data does not provide information about any of the characteristics of the respondents, we recognize that there are limitations to the depth of our analysis.

The first hypothesis we test is whether the forecaster fixed-effects for inflation uncertainty reflect learning as respondents gain experience with the SPF survey instrument. As discussed earlier, there is considerable variability across respondents in the length of their participation in the SPF. Given the uniqueness of the SPF survey in asking respondents to fill in subjective probabilities for intervals around their point forecast, new respondents may systematically answer this question differently from seasoned respondents.[31]

We can test for learning effects by correlating the forecaster specific fixed-effects with the total number of times that the forecaster has participated in the SPF survey (where we do this separately for each forecast horizon). The data indicate that at the 1, 2 and 4 quarter horizons there is a negative and significant correlation between these fixed-effects and the respondent's overall experience with the survey. Each additional year of survey experience reduces the respondent's *ex ante* inflation uncertainty by around 0.026 to 0.04. This represents approximately a 5% change relative to the inner quartile range (IQR) of these fixed-effects. The IQR for the number of times a respondent participates in the survey for a given forecast horizon is 21. Scaling the marginal effect given above by this variation in length of participation roughly accounts for the IQR in the fixed-effects.

[31]Due to data limitations, these comparisons only relate to the relative experience of participants in the SPF, and not to their relative historical experience with forecasting.

The negative relationship between the respondent's *ex ante* uncertainty fixed effects and his/her overall experience with the SPF survey is consistent with learning effects. Alternatively, it is possible that there is non-random attrition in the SPF survey. For example, Figure 5 illustrates a situation in which respondent *j* has a lower *ex ante* uncertainty fixed effect and longer participation in the survey relative to respondent *i*. In this figure, there is no learning for either respondent, but a regression of participant fixed effects on overall experience will generate a negative relationship. We test for this possibility in the SPF data by regressing each participants' *ex ante* uncertainty on a set of year effects and a measure of the cumulative (not overall) experience to date with the SPF survey. At each horizon, we find positive but insignificant coefficients on cumulative experience. The data, then, are consistent with the situation depicted in Figure 5 and do not support the hypothesis of learning about the survey instrument.

The second hypothesis that we test is whether the forecaster fixed-effects for inflation uncertainty reflect differential access to information among the respondents. If this is the underlying factor generating the *ex ante* inflation uncertainty fixed-effects, then we would expect to see a positive correlation between these *ex ante* fixed effects and *ex post* inflation precision fixed-effects. To carry out this test, for each forecast horizon we regress measures of the *ex post* precision of each forecast (defined again as the squared difference between the actual inflation rate and the respondent's point forecast) on a set of year effects and a set of forecaster fixed-effects. For each respondent, we match up his/her *ex ante* and *ex post* fixed-effects for a given forecast horizon.

Figure 6 shows the scatter plots of these forecaster *ex ante* and *ex post* fixed-effects for each forecast horizon. At each forecast horizon there is a positive and significant correlation between these *ex ante* and *ex post* fixed-effects. This indicates that SPF respondents who are on average more certain of their inflation forecasts also on average have more precise forecasts.[32] This is consistent with these respondents having access to superior information on which to base their forecasts (or having the same information but possessing a superior ability to process the information).

V. Conclusion

Because of the general absence of data on forecast uncertainty, time series models of heteroskedasticity have gained widespread popularity as a technique for generating such measures. Our paper focuses on this class of models and the reliability of using the estimated conditional variance of a time series to construct proxies for predictive uncertainty. While Batchelor and Dua (1993, 1996) and Giordani and Söderlind (2000) undertake a similar investigation, we attempt to remedy a number of shortcomings in their analyses. In addition, we extend these previous studies by developing a general framework that incorporates forecast dispersion into the analysis. This latter issue is particularly important for purposes of interpreting conventional model-based measures of uncertainty.

We explore these issues by using data on matched point and probabilistic forecasts of inflation from the Survey of Professional Forecasters. The results offer little evidence of a systematic relationship between observed heteroskedasticity in the consensus forecast errors and forecast uncertainty. In particular, we only find evidence of a consistent significant correlation

[32]These correlations range from a low of 0.14 (probability value of 0.08) for the two quarter forecast horizon to a high 0.31 (probability value of 0.0002) for the four quarter forecast horizon.

between *ex post* predictive accuracy and *ex ante* uncertainty when we use the probabilistic mean forecast series to measure expected inflation. On the other hand, the findings document an important and robust link between observed heteroskedasticity in the forecast errors and forecast dispersion. Consequently, these results lead us to conclude that conventional model-based measures of uncertainty may be describing the degree of disagreement across forecasters in their predictions rather than the (average) confidence associated with the predictions. Moreover, the results lead us to cast doubt on the validity of inferences and conclusions drawn from studies that have employed these measures.

The results also lead us to question the conclusion of Bomberger (1996) that disagreement is a potentially useful proxy for uncertainty. Our analysis demonstrates that forecast uncertainty and forecast disagreement do not need to display a parallel relationship to the conditional variance of a time series. Consequently, we would argue that Bomberger's previous finding of a significant link between observed heteroskedasticity in forecast errors and forecast dispersion for the Livingston survey has no relevance for judging the validity of disagreement as a proxy for uncertainty. Rather, any attempt to address this issue requires empirical measures of each variable and the application of direct testing procedures.

Our study also supports the recent findings of Carroll (2001) and Souleles (2002) that cast doubt on the assumption that agents share identical beliefs and information about the economy. We find the heterogeneity in forecast behavior in conjunction with changes in panel composition over time introduces important compositional effects in the measure of inflation uncertainty. We also find evidence across individuals of a strong positive association between their average *ex post* predictive accuracy and their average *ex ante* uncertainty, although it is important to note that this relationship is quite distinct and has very different implications from the analysis conducted at the aggregate level. The results from comparing the forecast behavior across respondents may be important and useful in developing alternative modeling strategies that can provide a better description of uncertainty and its empirical features.

References

Batchelor, Roy, and Pami Dua. "Blue Chip Rationality Tests." *Journal of Money, Credit and Banking* 23 (November 1991): 692-705.

------, and Pami Dua. "Survey vs ARCH Measures of Inflation Uncertainty." *Oxford Bulletin of Economic and Statistics* 55 (August 1993): 341-353.

------, and Pami Dua. "Empirical Measures of Inflation Uncertainty: A Cautionary Note." *Applied Economics* 28 (March 1996): 333-341.

Bollerslev, Tim, Ray Y. Chou, and Kenneth F. Kroner. "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence." *Journal of Econometrics* 52 (April-May 1992): 5-59.

Bomberger, William A. "Disagreement as a Measure of Uncertainty." *Journal of Money, Credit and Banking* 28 (August 1996): 381-392.

Bonham, Carl S., and Richard H. Cohen. "Testing the Rationality of Price Forecasts: Comment." *American Economic Review* 85 (March 1995): 284-289.

------, and Richard H. Cohen. "To Aggregate, Pool, or Neither: Testing the Rational-Expectations Hypothesis Using Survey Data." *Journal of Business and Economic Statistics* 19 (July 2001): 278-291.

Carroll, Christopher D. "The Epidemiology of Macroeconomic Expectations." Working Paper No 8695. National Bureau of Economic Research, December, 2001.

Croushore, Dean. "Introducing: The Survey of Professional Forecasters." *Federal Reserve Bank of Philadelphia Business Review*, November/December 1993, 3-15.

Cukierman, Alex, and Paul Wachtel. "Differential Inflationary Expectations and the Variability of the Rate of Inflation: Theory and Evidence." *American Economic Review* 69 (September 1979): 595-609.

Diebold, Francis X., Anthony S. Tay, and Kenneth F. Wallis. "Evaluating Density Forecasts of Inflation: The Survey of Professional Forecasters." In *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W. J. Granger*, edited by Robert F. Engle and Halbert White, 76-90. Oxford and New York, Oxford University Press, 1999.

Engle, Robert F. "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* 50 (July 1982): 987-1007.

------. "Estimates of the Variance of U.S. Inflation Based Upon the ARCH Model." *Journal of Money, Credit, and Banking* 15 (August 1983): 286-301.

Giordani, Paolo, and Paul Soderlind. "Inflation Forecast Uncertainty." Working Paper. Stockholm School of Economics, October, 2000.

Hansen, Lars Peter. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50 (July 1982): 1029-1054.

Keane, Michael F., and David E. Runkle. "Testing the Rationality of Price Forecasts: New Evidence from Panel Data." *American Economic Review* 80 (September 1990): 714-735.

Kendall, Maurice.G., and Alan Stuart. *The Advanced Theory of Statistics, Vol. 1*. London, Griffin, 1963.

Lahiri, Kajal, Christie Teigland, and Mark Zaporowski. "Interest Rates and the Subjective Probability Distribution of Inflation Forecasts." *Journal of Money, Credit and Banking* 20 (May 1988): 233-248.

Lucas, Robert E. Jr. "Expectations and the Neutrality of Money." *Journal of Economic Theory* 4 (April 1972): 103-124.

Neumark, David, and Jonathan S. Leonard. "Inflation Expectations and the Structural Shift in Aggregate Labor-Cost Determination in the 1980s." *Journal of Money, Credit and Banking* 25 (November 1993): 786-800.

Pennacchi, George G. "Identifying the Dynamics of Real Interest Rates and Inflation: Evidence Using Survey Data." *Review of Financial Studies* 4, no. 1 (1991): 53-86.

Souleles, Nicholas S. "Consumer Sentiment: Its Rationality and Usefulness in Forecasting Expenditure? Evidence from the Michigan Micro Survey." *Journal of Money, Credit and Banking* (2002).

Thomas, Lloyd B Jr. "Survey Measures of Expected U.S. Inflation." *Journal of Economic Perspectives* 13 (Fall 1999): 125-144.

White, Halbert. *Asymptotic Theory for Econometricians*. Orlando, FL, Academic Press, 1984.

Zarnowitz, Victor. "Business Cycle Analysis and Expectational Survey Data." Working Paper No. 1378. National Bureau of Economic Research, June, 1984.

------. "Rational Expectations and Macroeconomic Forecasts." *Journal of Business and Economic Statistics* 3 (October 1985): 293-311.

------, and Phillip A. Braun. "Twenty-two Years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Comparisons of Forecasting Performance." In *Business Cycles, Indicators, and Forecasting*, edited by James H. Stock and Mark W. Watson, 11-84. NBER Studies in Business Cycles, vol. 28. Chicago and London, University of Chicago Press, 1993.

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

Appendix

This Appendix describes our modification to the covariance matrix estimators proposed by Hansen (1982) and White (1984, Chapter 6). Let $\varepsilon_{t,h}$ represent the inflation forecast error in equation (14) such that:[33]

$$\pi_{th} = \alpha + \beta \overline{f}_{th}^{e} + \varepsilon_{t,h} \tag{A1}$$

As previously noted, the target inflation rate (π_{th}) will remain fixed for four consecutive surveys. For example, we will have:

$$\pi_{t4} = 100 * \left[\frac{P_{t+1,1} + P_{t+1,2} + P_{t+1,3} + P_{t+1,4}}{P_{t,1} + P_{t,2} + P_{t,3} + P_{t,4}} - 1 \right] = \pi_{t+13} = \pi_{t+12} = \pi_{t+11} \tag{A2}$$

Consequently, the surveys undertaken from the last quarter in a year through the third quarter of the subsequent year result in forecast intervals of 4½, 3½, 2½ and 1½ quarters, respectively.[34] This implies that the corresponding forecast errors will follow, respectively, an MA process of order 4, 3, 2 and 1.

We will now consider the following sequence of inflation forecast errors that relates to the surveys conducted in years τ-1 through year τ+1:

$$\left[\varepsilon_{\tau-1,3},\ \varepsilon_{\tau-1,2},\ \varepsilon_{\tau-1,1},\ \varepsilon_{\tau-1,4},\ \varepsilon_{\tau,3},\ \varepsilon_{\tau,2},\ \varepsilon_{\tau,1},\ \varepsilon_{\tau,4},\ \varepsilon_{\tau+1,3},\ \varepsilon_{\tau+1,2},\ \varepsilon_{\tau+1,1},\ \varepsilon_{\tau+1,4} \right] \tag{A3}$$

Relative to the case in which there were a fixed forecast interval of 4½ quarters and the forecast errors followed an MA(4) process, the declining forecast intervals now imply that the correlation between certain elements in equation (A3) will be zero. These elements are:

$$\begin{aligned} &E\left[\varepsilon_{t,3}\ \varepsilon_{t-1,1}\right] = E\left[\varepsilon_{t,3}\ \varepsilon_{t-1,2}\right] = E\left[\varepsilon_{t,2}\ \varepsilon_{t-1,1}\right] = E\left[\varepsilon_{t,3}\ \varepsilon_{t-1,3}\right] \\ &= E\left[\varepsilon_{t,2}\ \varepsilon_{t-1,2}\right] = E\left[\varepsilon_{t,1}\ \varepsilon_{t-1,1}\right] = 0 \end{aligned} \tag{A4}$$

---

[33]This same discussion applies to the innovations to the process for the squared inflation forecast errors in equations (12) and (13).

[34]As shown in equation (5), this timing convention takes into account the existence of publication lags in the price index and the fact that the survey respondents' inflation forecasts are constructed using a prediction of the price level for the current quarter.

Therefore, we can obtain the appropriate covariance matrix estimator by excluding these terms from equation (15).

To construct our modified covariance matrix estimator, let $\hat{\mu}$ denote the (T x 1) vector of ordered quarterly OLS residuals from estimation of equation (A1) using pooled forecast horizons such that:

$$\hat{\mu} = \left[\hat{\mu}_1,\ \hat{\mu}_2,\ \ldots,\ \hat{\mu}_T\right] = \left[\ldots,\ \hat{\varepsilon}_{\tau-1,4},\ \hat{\varepsilon}_{\tau,3},\ \hat{\varepsilon}_{\tau,2},\ \hat{\varepsilon}_{\tau,1},\ \hat{\varepsilon}_{\tau,4},\ \ldots\right] \tag{A5}$$

We will now consider the following variant of equation (15):

$$\left(\frac{\hat{\Omega}_T}{T}\right) = \left[\sum_{t=1}^{T} x_t x_t'\right]^{-1} \left[\sum_{t=1}^{T} \hat{\mu}_t^2 x_t x_t' + \sum_{v=1}^{4}\left\{\sum_{t=v+1}^{T} D_t(v)\cdot(x_t \hat{\mu}_t \hat{\mu}_{t-v} x_{t-v}' + x_{t-v}\hat{\mu}_{t-v}\hat{\mu}_t x_t')\right\}\right]\left[\sum_{t=1}^{T} x_t x_t'\right]^{-1} \tag{A6}$$

where D_t is an indicator function that takes on the value of 0 or 1 according to the following rule:

$$\begin{aligned} &D_t(1) = 1 \ \forall\ t \\ &D_t(2) = 1 \ if\ \ \hat{\mu}_t = \ quarter\ 2, quarter\ 3\ or\ quarter\ 4 \\ &D_t(3) = 1 \ if\ \ \hat{\mu}_t = \ quarter\ 3\ or\ quarter\ 4 \\ &D_t(4) = 1 \ if\ \ \hat{\mu}_t = \ quarter\ 4 \end{aligned} \tag{A7}$$

The indicator function essentially defines a set of dummy variables that allows us to include or exclude the relevant autocorrelations of the forecast errors from the covariance matrix estimator.

Table 1

Intervals for probabilistic 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. Test for unbiasedness of inflation forecasts, 1968:Q4 - 2000:Q3

a) Consensus point forecast

$$\pi_{th} = \alpha + \beta \bar{f}^{e}_{th} + \varepsilon_{t,h}$$

Variable	α	β	$\bar{R}^2$	$\chi^2(2)$ [a]
Point Forecast	0.113 (0.353)	0.946** (0.090)	0.804	0.555

b) Consensus probability mean forecast

$$\pi_{th} = \alpha + \beta \bar{\pi}^{e}_{th} + \varepsilon_{t,h}$$

Variable	α	β	$\bar{R}^2$	$\chi^2(2)$ [a]
SPF Histogram	-0.173 (0.360)	0.957** (0.087)	0.797	3.692
2-unit Histogram	-0.216 (0.372)	0.962** (0.089)	0.796	4.158
Normal Approximation	-0.163 (0.364)	0.959** (0.088)	0.800	3.352
FE Adjusted Normal Approx.	-0.154 (0.369)	0.957** (0.088)	0.793	3.242

Note: Standard errors are calculated using the modified covariance matrix estimator described by equation (A6) in the Appendix.

[a] $\chi^2(2)$ H_0: α=0 and β=1

*Significant at the 5 percent level

**Significant at the 1 percent level

Table 3. Test of *ex ante* and *ex post* inflation uncertainty relationship, 1968:Q4 - 2000:Q3

a) Squared forecast errors based on consensus point forecasts.

$$\pi_{th} = \overline{f}^{e}_{th} + \varepsilon_{t,h}$$
$$(\varepsilon_{t,h})^2 = \alpha + \beta(\overline{\sigma}^2_{t,h}) + \eta_{t,h}$$

Variable	α	β	$\overline{R}^2$
Histogram, Sheppard adj.	0.555 (0.581)	0.607 (0.573)	0.011
2-unit Histogram	0.149 (1.032)	0.712 (0.741)	0.008
Normal Approximation	0.351 (0.646)	1.252 (0.990)	0.022
FE Adjusted Normal Approx.	0.126 (0.450)	1.319** (0.616)	0.044

b) Squared forecast errors based on consensus probability mean forecasts.

$$\pi_{th} = \overline{\pi}^{e}_{th} + \varepsilon_{t,h}$$
$$(\varepsilon_{t,h})^2 = \alpha + \beta(\overline{\sigma}^2_{t,h}) + \eta_{t,h}$$

Variable	α	β	$\overline{R}^2$
Histogram, Sheppard adj.	0.075 (0.504)	1.366* (0.595)	0.104
2-unit Histogram	-0.668 (0.995)	1.484* (0.801)	0.075
Normal Approximation	-0.037 (0.553)	2.158* (0.981)	0.093
FE Adjusted Normal Approx.	-0.107 (0.408)	1.874** (0.664)	0.113

Note: Standard errors are calculated using the modified covariance matrix estimator described by equation (A6) in the Appendix.
One-tailed test for statistical significance of β .
H_0: β=0, H_1: β>0
*Significant at the 5 percent level
**Significant at the 1 percent level

Table 4. Augmented test of *ex ante* and *ex post* inflation uncertainty relationship, 1968:Q4 - 2000:Q3

a) Squared forecast errors based on consensus point forecasts.

$$\pi_{th} = \overline{f}^{e}_{th} + \varepsilon_{t,h}$$
$$(\varepsilon_{t,h})^2 = \alpha + \beta(\overline{\sigma}^2_{th}) + \gamma(s^2_{f_{th}}) + \eta_{t,h}$$

Variable	α	β	γ	$\overline{R}^2$	$\chi^2(1)$ [b]
Histogram, Sheppard adj.	0.547 (0.519)	0.128 (0.477)	1.298** (0.476)	0.074	2.12
2-unit Histogram	0.491 (0.938)	0.123 (0.642)	1.312** (0.467)	0.074	1.68
Normal Approximation	0.448 (0.589)	0.392 (0.860)	1.245** (0.473)	0.076	0.58
FE Adjusted Normal Approx.	0.115 (0.391)	0.722 (0.507)	1.341** (0.555)	0.090	0.44

b) Squared forecast errors based on consensus probability mean forecasts.

$$\pi_{th} = \overline{\pi}^{e}_{th} + \varepsilon_{t,h}$$
$$(\varepsilon_{t,h})^2 = \alpha + \beta(\overline{\sigma}^2_{th}) + \gamma(s^2_{f_{th}}) + \eta_{t,h}$$

Variable	α	β	γ	$\overline{R}^2$	$\chi^2(1)$ [b]
Histogram, Sheppard adj.	0.069 (0.455)	1.019* (0.502)	0.941* (0.465)	0.138	0.01
2-unit Histogram	−0.405 (0.889)	1.031 (0.679)	1.009* (0.512)	0.117	0.00
Normal Approximation	0.034 (0.503)	1.527* (0.827)	0.912* (0.488)	0.124	0.34
FE Adjusted Normal Approx.	−0.116 (0.353)	1.405** (0.549)	1.055* (0.594)	0.135	0.14

Note: Standard errors are calculated using the modified covariance matrix estimator described by equation (A6) in the Appendix. One-tailed test for statistical significance of β and γ.
H_0: β=0, H_1: β>0
H_1: γ=0, H_1: γ>0
[a]$\chi^2(1)$ H_0: β=γ
*Significant at the 5 percent level
**Significant at the 1 percent level

Figure 1. SPF Survey Size, New Entrants and Permanent Exits

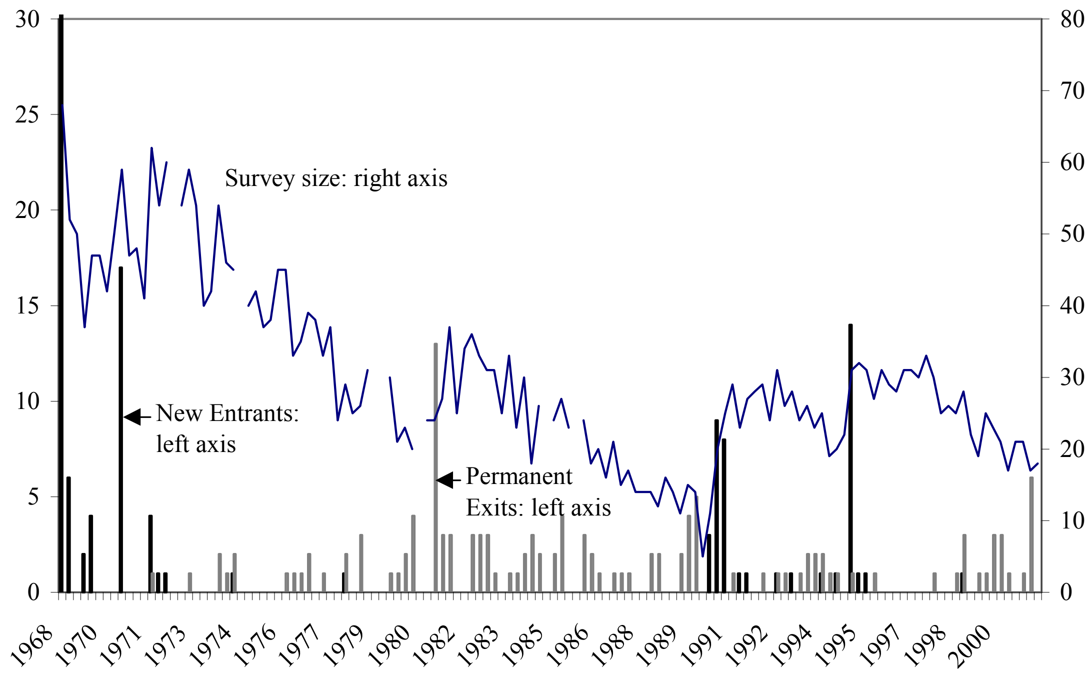


Figure 2. SPF Turnover of Participants

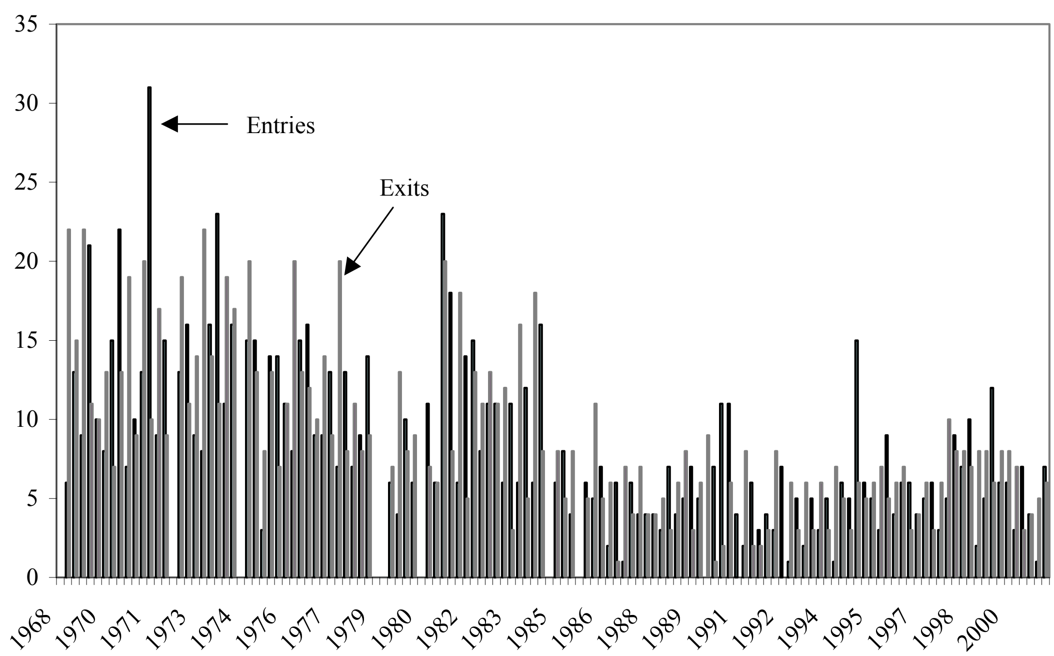
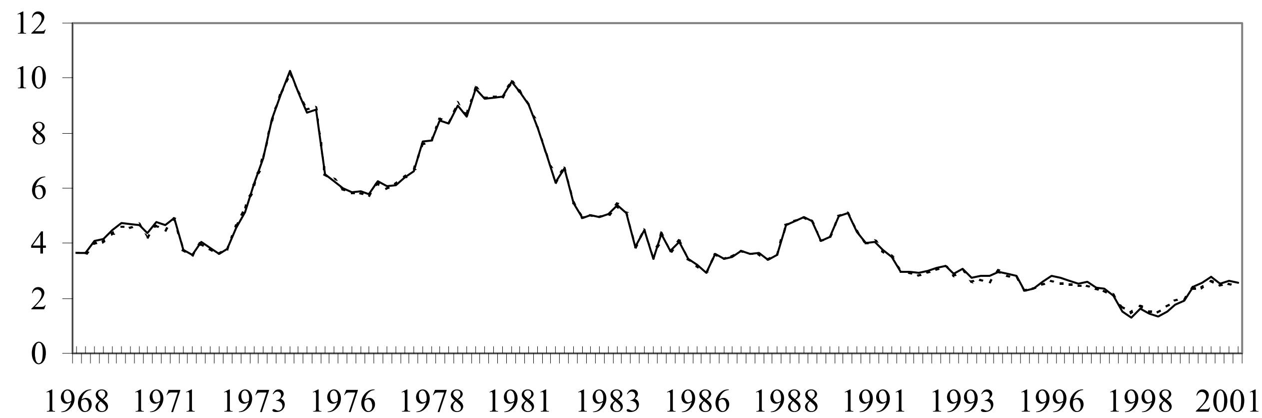
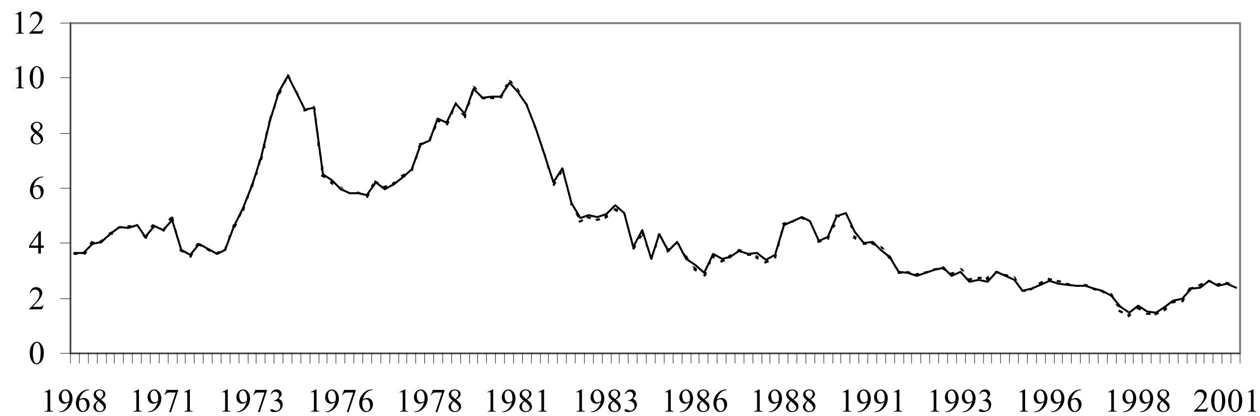


Figure 3. Comparison of Aggregate Expected Inflation Measures by Method of Construction

a) Uniform Distribution: Original SPF intervals vs. adjusted intervals



b) Uniform distribution vs. normal distribution



c) Normal distribution with and without adjustment for composition shifts

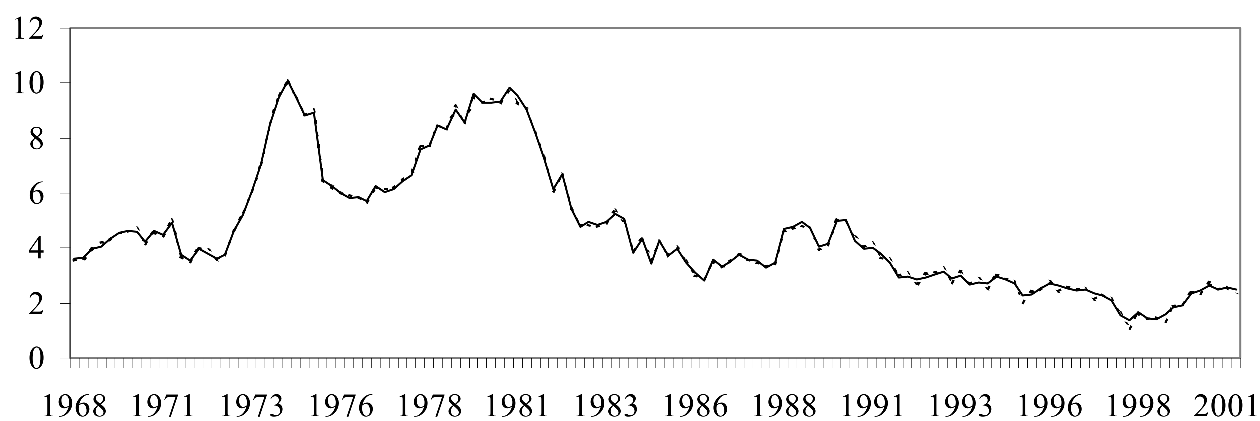
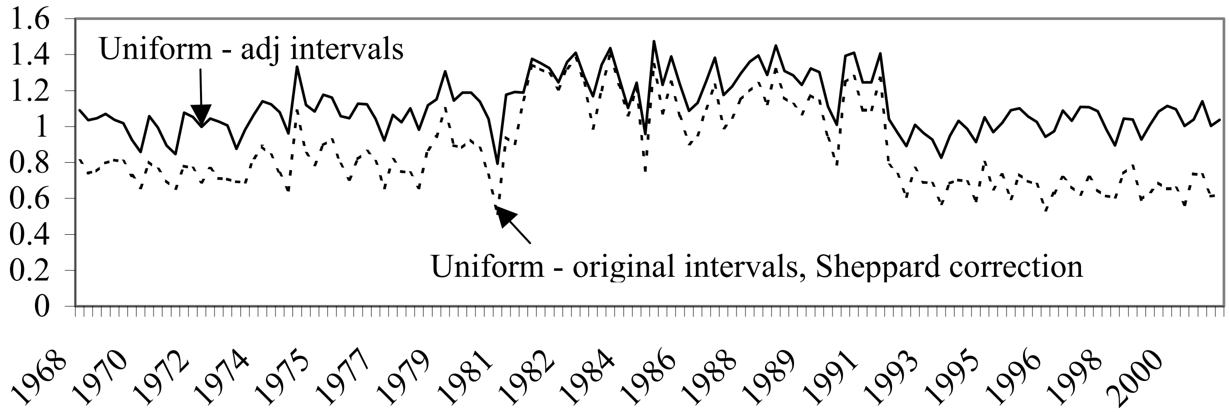
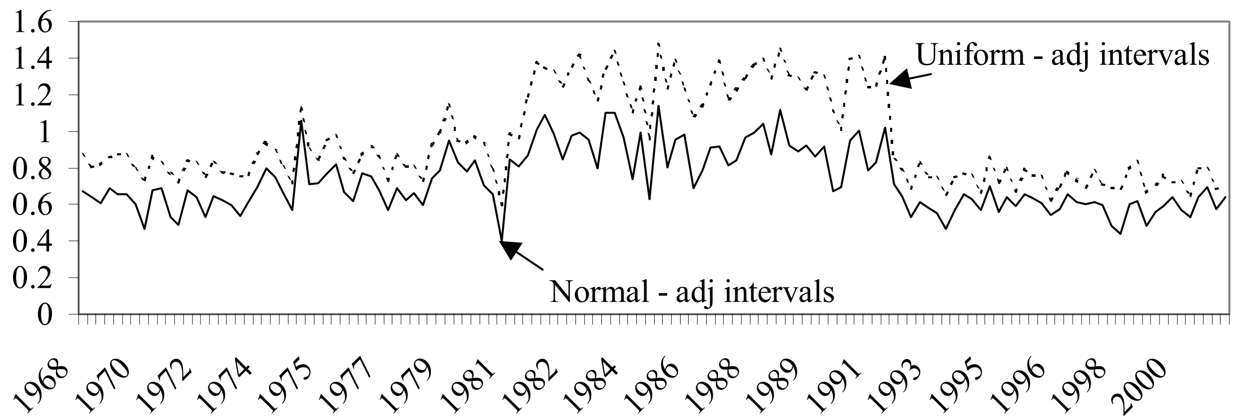


Figure 4. Comparison of Aggregate Uncertainty Measures by Method of Construction

a) Uniform Distribution: Original intervals w. Sheppard correction vs. adjusted intervals



b) Uniform distribution vs. normal distribution



c) Normal distribution with and without adjustment for composition shifts

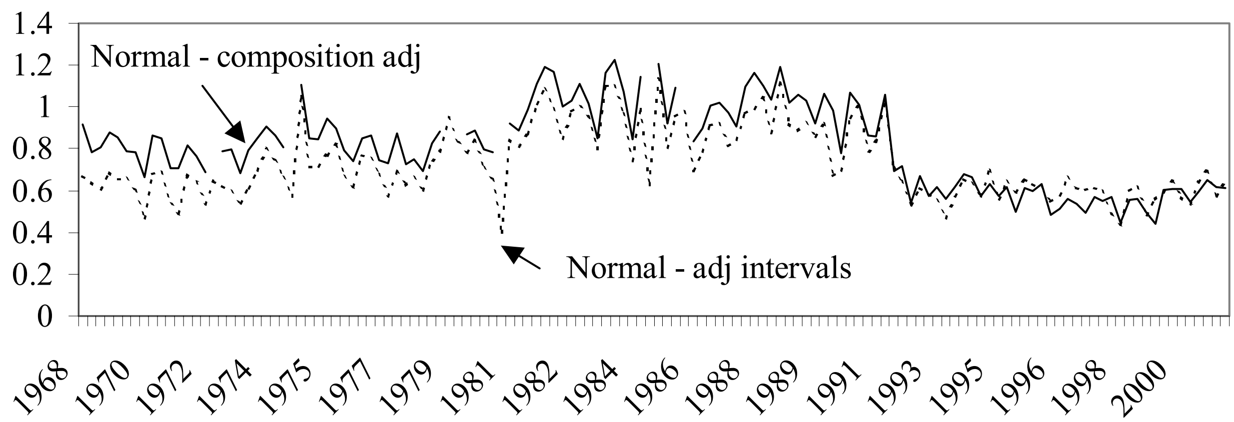


Figure 5. Learning versus Non-random Attrition

Ex Ante Uncertainty

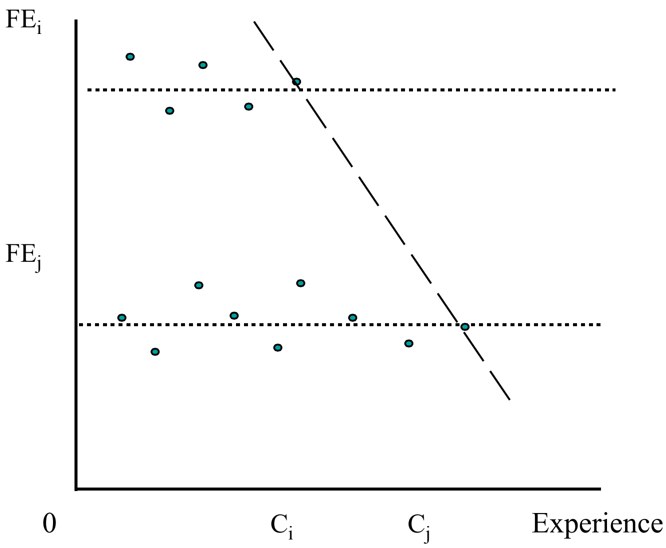


Figure 6. Ex Ante Uncertainty and Ex Post Accuracy Fixed Effects