No. 2499

INFLATION FORECAST UNCERTAINTY

Paolo Giordani and Paul Söderlind

INTERNATIONAL MACROECONOMICS

Centre for Economic Policy Research

ISSN 0265-8003

INFLATION FORECAST UNCERTAINTY

Paolo Giordani, Stockholm School of Economics
Paul Söderlind, Stockholm School of Economics and CEPR

Discussion Paper No.2499
July 2000

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: http://www.cepr.org

CEPR Discussion Paper No. 2499

July 2000

ABSTRACT

Inflation Forecast Uncertainty*

We study the inflation uncertainty reported by individual forecasters in the Survey of Professional Forecasters 1969–99. Three popular methods of estimating uncertainty from survey data are analysed in the context of models for forecasting and asset pricing. We find that inflation uncertainty fluctuates over time in a way that traditional time-series models fail to capture. Instead, uncertainty is highly correlated with the level of inflation, in particular with recent positive inflation surprises. We also find that disagreement among forecasters increases with the inflation rate and causes above-average fluctuations in individual uncertainty.

JEL Classification: C53, E31 and E37
Keywords: survey data, survey of professional forecasters, TGARCH and VAR

Paolo Giordani
Paul Söderlind
Stockholm School of Economics
PO Box 6513
Sveavagen 65
SE-113 83 Stockholm
SWEDEN
Tel: (46 8) 736 9252
(46 8) 736 9601
Fax: (46 8) 313207 (both)
Email: Paolo.Giordani@hhs.se
Paul.Soderlind@hhs.se

*We thank Tom Stark for much help with the data. This Paper is produced as part of a CEPR research network on New Approaches to the Study of Economic Fluctuations, funded by the European Commission under the Training and Mobility of Researchers Programme (Contract No ERBFMRX-CT98-0213).

Submitted 17 May 2000

NON-TECHNICAL SUMMARY

Modern economic theory predicts that agents' behaviour depends on their assessment of the probable distribution of the variables that they need to forecast. It is only under very restrictive assumptions that the point forecast is sufficient to characterize their choices. In general, higher moments also matter. This Paper focuses on inflation uncertainty, which is interesting since it is likely to affect the pricing of many financial assets, investment decisions, price-setting behaviour, and both the evaluation and conduct of monetary policy. For instance, inflation uncertainty can be used as an indicator of central bank credibility.

A large number of studies have investigated the forecasting properties (objectivity, efficiency) of point forecasts from surveys. This Paper has very little to say about that. We do use surveys, but we focus on uncertainty instead.

We study inflation uncertainty in the Survey of Professional Forecasters for the period 1969 to 1999. This survey is unique in asking the respondents about their subjective probability distribution of future inflation. We use this information to construct different measures of uncertainty. The correct definition of a measure of uncertainty from survey data is not self-evident. If the data set comprised a single forecaster, a quick agreement could be reached on the standard deviation of her subjective distribution as a good measure of uncertainty. The issue is more complicated when there are many forecasters. Should we take average individual uncertainty to be the most relevant measure of inflation uncertainty in the economy or should we also take the disagreement among forecasters into account? We try to answer these questions by using both theoretical models (of forecasting and asset pricing) and empirical estimates of the properties of different measures of uncertainty.

This Paper thus considers appropriate procedures to extract a measure of uncertainty from survey data, with an application to inflation uncertainty. We propose a new method, which allows decomposing uncertainty into the average individual uncertainty and a measure of disagreement. This interpretation implies that forecasters underestimate uncertainty on average, which is unappealing but supported by our sample. In any case, theoretical considerations suggest that this measure could be the most useful for asset-price analysis. Moreover, disagreement among forecasters is highly correlated with it. Since forecasters' disagreement is much more readily available than the average of individual variances (most surveys only ask for a point forecast), our analysis suggests that it is a good proxy for inflation uncertainty, for both theoretical and empirical reasons.

Having defined a small set of meaningful measures of uncertainty, we use the survey data to track how inflation uncertainty has changed over the last 30 years in the US. We study how these changes are correlated with macro variables like inflation, output growth and past macroeconomic volatility. Not surprisingly, inflation uncertainty is strongly correlated with the level of inflation.

We also compare the results from the survey data with results from several time-series models, with the aim of evaluating some commonly used forecasting models. A VAR model of US macro data estimated on a longer and longer sample (forecasts are made for the next out-of-sample period) fails to capture the decrease in inflation uncertainty since the early 1990s. The result is noteworthy since these VARs are commonly used as forecasting tools. We do not say anything about the quality of the point forecasts from the VAR, but, when combined with evidence that inflation is heteroskedastic in the data as well as in the survey, our results throw doubt onto the quality of the error bands constructed around the point forecasts.

The same VAR model estimated on a moving data window performs much better. We also study if these features could be due to ARCH/GARCH effects and find that positive inflation surprises seem to increase uncertainty for some time, while negative shocks do not. That is, an unanticipated increase in inflation increases uncertainty, while an unanticipated decrease in inflation does not. While the first applications of ARCH and GARCH models were on inflation, we suggest that a symmetric GARCH is not the most appropriate modelling tool for inflation.

All measures of uncertainty derived from the survey indicate that inflation is perceived as heteroskedastic, as is in fact the case in the data. In time-series forecasting this feature must be taken into account to construct accurate error bands. This is particularly relevant for the central banks that publish confidence inflation forecasts with confidence bands (‘fan charts’).

1 Introduction

Modern economic theory predicts that agents' behavior depends on their assessment of the probabilistic distribution of the variables that they need to forecast. It is only under very restrictive assumptions that the point forecast is sufficient to characterize their choices. In general, higher moments also matter. This paper focuses on inflation uncertainty, which is interesting since it is likely to affect the pricing of many financial assets, price setting behavior, and both the evaluation and conduct of monetary policy. For instance, inflation uncertainty can be used as an indicator of central bank credibility.

We study inflation uncertainty in the Survey of Professional Forecasters for the period 1969 to 1999. This survey is unique in asking the respondents about their subjective probability distribution of future inflation. We use this information to construct different measures of uncertainty.

A large number of studies have investigated the forecasting properties (unbiasedness, efficiency) of point forecasts from surveys (see, for instance, Thomas (1999) for a recent study of several surveys, including the Survey of Professional Forecasters). This paper has very little to say about that. Instead, we focus on uncertainty, which we take to be captured by the variance.

The survey data allows us track how inflation uncertainty has changed over the last 30 years. We study how these changes are correlated with macro variables like inflation, output growth, and past macro economic volatility. Not surprisingly, inflation uncertainty is strongly correlated with the level of inflation.

We also compare the results from the survey data with results from several time series models, with the aim of evaluating some commonly used forecasting models. A VAR model of US macro data estimated on a longer and longer sample (forecasts are made for the next out-of-sample period) fails to capture the decrease in inflation uncertainty since he early 1990s. The same VAR model estimated on a moving data window performs much better. We also study if these features could be due to ARCH/GARCH effects, and find that positive inflation surprises seem to increase uncertainty for some time, while negative shocks do not.

The correct definition of a measure of uncertainty from survey data is not self evident. If the data set comprised a single forecasters, a quick agreement could be reached on the standard deviation of her subjective distribution as a good measure of uncertainty. The issue is more complicated when there are many forecasters. Should we take average individual uncertainty to be the most relevant measure of inflation uncertainty in the economy or should we also take the disagreement among forecasters into account? We try to answer these questions by using both theoretical models (of forecasting and asset pricing) and empirical estimates of the properties of different measures of uncertainty.

The rest of the paper is organized as follows. Section 2 presents the data (mainly the Survey of Professional Forecasters). Section 3 discusses alternative measures of uncertainty from the survey data in the context of a forecasting model and also in a model of asset pricing. Section 4 discusses the estimation of uncertainty. Section 5 presents the empirical results, and Section 6 concludes.

2 The Survey of Professional Forecasters

The Survey of Professional Forecasters (SPF) is a quarterly survey of forecasters' views on key economic variables. The respondents, who supply anonymous answers, are professional forecasters from the business and financial community. The survey was started in 1968 by Victor Zarnowitz and others of the American Statistical Association and National Bureau of Economic Research. The number of forecasters was then around 60, but decreased in two major steps in the mid 1970s and mid 1980s to as low as 14 forecasters in 1990. The survey was then taken over by the Federal Reserve Bank of Philadelphia and the number of forecasters stabilized around 30. See Croushore (1993) for details.

A unique feature of the survey is that it asks for probabilities (on top of the usual point forecasts). In particular, it asks for probabilities of different intervals of (annual average) GDP deflator inflation, that is, the GDP deflator for year t divided by the GDP deflator for year $t-1$, minus one.[1] CPI inflation would perhaps have been better, but probabilities are not available for this variable. Besides, GDP deflator and CPI inflation are typically very similar (the correlation is 0.94 for the sample 1955Q1-1999Q3).[2]

[1]Before 1981Q3, the questions were about GNP and the GNP deflator instead.

[2]The major difference between the GDP deflator inflation and CPI inflation is that the latter is more volatile. In particular, it seems to temporarily react more to oil price shocks.

There are indications that the forecasting horizons for the probability values were not adjusted in a systematic way (from quarter to quarter) before 1981Q3. We therefore choose to focus on the first-quarter surveys, which appear to be correct: they refer to the growth rate of the deflator from the previous to the current year (annual-average). From 1981Q4 the survey includes also probabilities for GDP deflator inflation between the current and the next year. We will take a look at these numbers as well.[3]

Table 1 shows how the inflation intervals have changed over time. The general structure is that there is an open lower interval, a series of interior intervals of equal width, and an open upper interval. Unfortunately, the width of the intervals has changed over time (1% before 1981Q3 and after 1991Q4, 2% in the intermediate period). which may influence estimates of variance. We will take this into account.

Table 1: Inflation intervals in SPF

Period	No. of intervals	Intervals, %					No. of forecasters*
68Q4-73Q1	15	<-3	-3 – -2.1	...	9 – 9.9	10+	53 – 65
73Q2-74Q3	15	<-1	-1 – -0.1	...	11 – 11.9	12+	59 – 59
74Q4-81Q2	15	< 3	3 – 3.9	...	15 – 15.9	16+	26 – 47
81Q3-85Q1	6	< 4	4 – 5.9	...	10 – 11.9	12+	29 – 34
85Q2-91Q4	6	< 2	2 – 3.9	...	8 – 9.9	10+	14 – 34
92Q1-	10	< 0	0 – 0.9	...	7 – 7.9	8+	27 – 36

*Refers to the first quarter of each year.

The results from this survey are typically reported in three ways: the median point forecast, aggregate (or mean) histograms (built by averaging the probabilities from the individual histograms), and the dispersion of the individual point forecasts.

As a preview of the data, *Figure 1* shows the aggregate histogram for all first quarters 1969-1999 (disregard the vertical lines for the moment). The means of these "distributions" (vertical lines) follow the well-known story about US inflation, and the spread of the distributions seems to positively correlated with the mean. For most years, the histograms are reasonably symmetric with most of the probability mass in interior intervals. However, 1985 is a striking exception with 60% of the

[3]Errors in the surveys for 1985Q1, 1986Q1, and 1990Q1 mean that data for the next year is not available (see Federal Reserve Bank of Philadelphia (1999)).

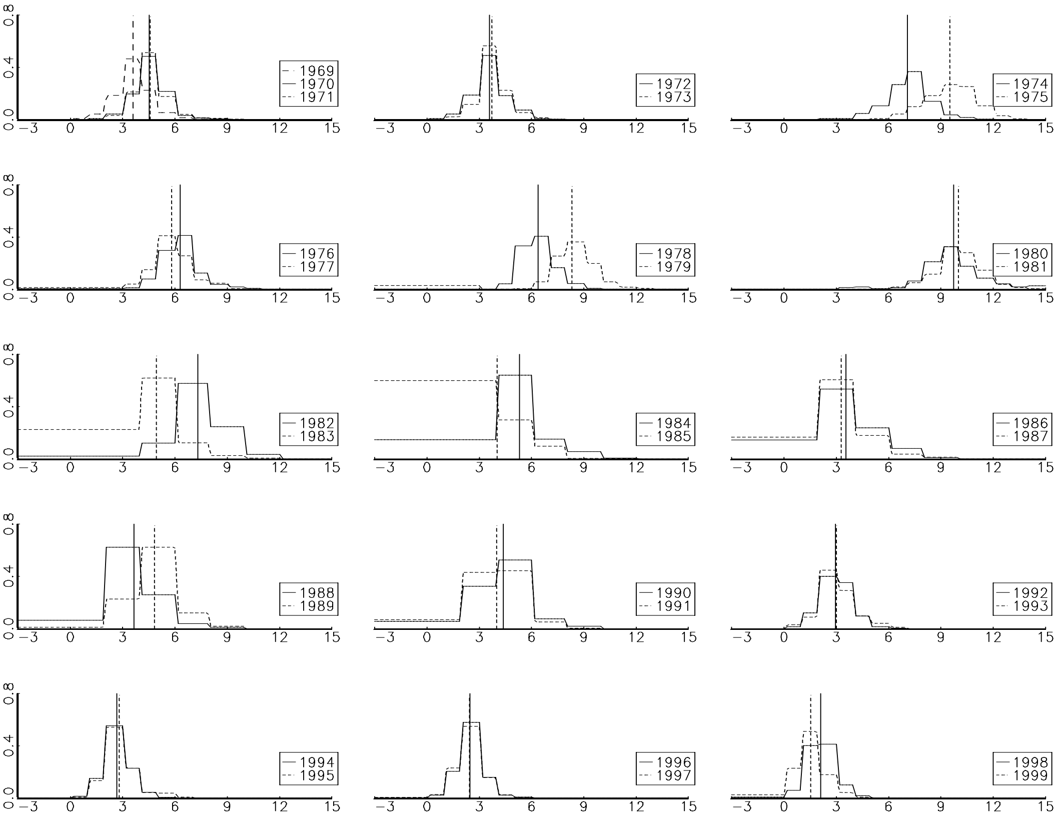

Figure 1: Aggregate probabilities in SPF 1969-1999. This figure shows the aggregate probabilities (vertical axis) for different inflation rates (horisontal axis). Estimated means are indicated with vertical lines. Each subfigure shows the probabilities for two (or three) years.

mass in the open lower interval (inflation lower than 4%). It is very difficult to say anything about the moments of the distribution for this period with so little information. We will therefore disregard this observation for the rest of the analysis.

As mentioned, the SPF combines the individual histograms into aggregate (or mean) probabilities, by taking the average (across forecasters) probability for each inflation interval. One way to formalize this approach is to think of both future inflation and forecaster i's "signal" as random variables. Let $\text{pdf}(\pi|i)$ be the probability density function of inflation conditional on receiving the signal of forecaster i; this corresponds to the individual histograms in a given period. Also, let $\text{pdf}(i)$ be the density function of receiving the signal of forecaster i in the same period. We then see that the aggregate distribution in that period, $\text{pdf}_A(\pi)$, which averages pdf($\pi|i$) across forecasters, amounts to calculating the "marginal" distribution of π

$$\text{pdf}_A(\pi) = \int_{-\infty}^{\infty} \text{pdf}(\pi|i)\text{pdf}(i)di. \tag{1}$$

Let μ_i and σ_i^2 be the mean and variance in forecaster i's distribution, pdf($\pi|i$). We know that (if the moments exist) the variance of this distribution is[4]

$$\text{Var}_A(\pi) = \text{Var}(\mu_i) + \text{E}(\sigma_i^2). \tag{2}$$

This equation shows that the variance of the aggregate distribution of π can be decomposed into the variance of the forecasters' means (disagreement) and the average of the forecasters' variances (average individual uncertainty). Note that this equation refers to the distribution in a particular period, so the distributions and moments should carry time subscripts. To economize on notation we suppress all such time subscripts, with the hope that the context makes it clear that we are discussing separate distributions for each period.

3 Which Measure of Uncertainty?

Equation (2) contains three possible measures of uncertainty: disagreement on the most likely outcome, Var(μ_i), average standard deviation of individual histograms, E(σ_i^2), and standard deviation of the aggregate histogram, $\text{Var}_A(\pi)$. This section discusses which of these that makes most sense.

Disagreement on the most likely outcome has the advantages of being readily available and easy to compute. The disadvantages are also clear. This measure becomes meaningless as the number of agents goes to one or when agents have the same information and agree on the model to use in forecasting. In this case, the measure of uncertainty is zero as if the economy was deterministic. One of the questions tackled in this paper is the degree to which disagreement mirrors measure of uncertainty that are theoretically more appealing, but less easily available.

The average standard deviation of individual histograms does not have the drawbacks of the first measure. It is attractive because it is easy to associate with the uncertainty of a representative agent. On the other hand, it sweeps disagreement

[4]For any random variables y and x we have $\text{Var}(y) = \text{E}[\text{Var}(y|\,x)]+ \text{Var}[\text{E}(y|\,x)]$.

under the rug, and disagreement must reflect some type of uncertainty.

The third measure of aggregate uncertainty is computed as the standard deviation of the aggregate histogram. In this case aggregate uncertainty is higher than the average standard deviation of the individual distributions, which is to say that individual forecasters underestimate uncertainty. This last result does cause some perplexity. However, the possibility cannot be excluded a priori: if someone consults two doctors, obtaining different diagnoses, she will be uncertain to some degree, no matter how certain the doctors are.

3.1 Combining Forecasts

It is well established, both in theory and practice, that an unweighted combination of several different methods/forecasters typically reduces the forecast uncertainty (see, for instance, Granger and Ramanathan (1984) for a general discussion of optimal combinations; and Zarnowitz (1967) and Figlewski (1983) for applications to inflation data).

The theory says that if there are n different unbiased forecasters whose average forecast error variance is $\mathrm{E}(\sigma_i^2)$, then the (expected) forecast error variance of the unweighted forecast is

$$\mathrm{Var}\left(\pi - \textstyle\sum_{i=1}^{n}\mu_i/n\right) = \mathrm{E}\sigma_i^2/n + \mathrm{E}\gamma_{ij}\left(1 - 1/n\right), \tag{3}$$

where $\mathrm{E}\gamma_{ij}$ is the average covariance of two individual forecast errors.

The gain from using the combined forecast instead of a (randomly chosen) individual forecast, $\left(\mathrm{E}(\sigma_i^2) - \mathrm{E}\gamma_{ij}\right)(1 - 1/n)$, is positive as long as individual forecasts are not perfectly correlated. This theoretical argument suggests that the mean or median forecast in the SPF should be assigned a smaller uncertainty than the average individual uncertainty.

3.2 A Simple Model of Forecasting

It is not possible to say very much about the relation between individual and "aggregate" uncertainty unless we are very clear about the information sets and forecasting procedures. To gain some insights, this section sets up a simple forecasting model.

Individual forecasters face different, but correlated, information sets which they use in an optimal way to make the best possible inflation forecast. Suppose the information set of forecaster i can be summarized by a scalar "signal" z_i, which is useful for forecasting since it is correlated with actual inflation, π. Assume that the unconditional distribution of π and z_i, is multivariate normal with zero means (to simplify the algebra), variances $s_{\pi\pi}$ and s_{ii}, and covariance $s_{\pi i}$.

Forecaster i calculates the mathematical expectation of π conditional on z_i. The resulting distribution of inflation conditional on z_i is the standard result

$$\begin{aligned} &\pi|z_i \sim N\left(\mu_i, \sigma_i^2\right), \text{ with} \\ &\quad \mu_i = \left(s_{\pi i}/s_{ii}\right) z_i \text{ and } \sigma_i^2 = s_{\pi\pi} - s_{\pi i}^2/s_{ii}. \end{aligned} \tag{4}$$

Anyone who has access to the survey data can use the cross sectional average of the individual forecasts, $\mathrm{E}\mu_i$, as a "combined forecast." This wipes out any individual noise, so the distribution of the error of the combined forecast is normal with zero mean and variance equal to the covariance between individual forecast errors, as in (3). For simplicity, we assume that individual signals have the same variance and covariances with each other and actual inflation. It can then be shown (see Appendix B) that the forecast error variance of the combined forecast equals $\sigma_i^2 - \mathrm{Var}(\mu_i)$, which is the individual forecast error variance minus the cross sectional variance of point forecasts. The combined forecast is thus better than individual forecasts, especially if forecasters disagree.

The aggregate distribution, calculated as (1), is characterized by

$$\mathrm{pdf}_A(\pi) \text{ is from } N\left[\mathrm{E}\mu_i, \mathrm{Var}_A(\pi)\right], \text{ with } \mathrm{Var}_A(\pi) = \sigma_i^2 + \mathrm{Var}\left(\mu_i\right). \tag{5}$$

It is not obvious what this distribution represents. It is more "informed" than the unconditional distribution of π, but less informed than the individual conditional distributions. An appealing interpretation is that a reader of forecasts faces two sources of uncertainty: which forecast to trust and then that forecast's uncertainty. This is wrong, however, if the individual forecaster understands that he could make a more precise forecast if he had all the information of the other forecasters, that is, if he incorporates his individual uncertainty in the forecasting error variance he reports. The forecasters in (4) do.

This formal, but simple, model of forecasting suggests a few things. First, combined forecasts are likely to be much better than individual forecasts only if disagreement is large compared to individual uncertainty. Second, it seems hard to justify the aggregate distribution unless we believe that individual forecasters underestimate uncertainty. Whether this is the case empirically is discussed in Section 5.2.

3.3 Individual Beliefs and Asset Pricing

It has so far been difficult to motivate why we should care about the variance of SPF's aggregate distribution, $\text{Var}_A(\pi)$. This section presents a simple economic model, where this aggregate distribution is crucial for asset pricing, even if may be of no particular importance for forecasting. This model argues that taking the average across forecasters' distributions can make a lot of sense—from an economic point of view. The basic idea is the beliefs of an agent (forecaster) will influence his consumption and investment decisions, so the aggregate economy is likely to be affected by some kind of average beliefs.

To demonstrate this point, we consider the pricing of Arrow-Debreu (A-D) assets when agents have different beliefs (see, for instance, Varian (1985) and Benninga and Mayshar (1997)); A-D asset s has a payoff of one in state s and zero in all other states.

This is an endowment economy with n agents. Endowment in period 1 is known, but endowment in period 2 is random and may take any of S different values ("states") for each of the n agents. The agents have identical preferences, but may have different endowments and also different beliefs about the state probabilities. Agent i maximizes expected utility subject to the budget restrictions

$$\max C_{i1}^{1-\gamma}/(1-\gamma) + \beta \text{E} C_{i2}^{1-\gamma}/(1-\gamma)\text{, subject to} \tag{6}$$

$$Y_{i1} = C_{i1} + \textstyle\sum_{s=1}^{S} p(s)\, B_i(s)\text{, and} \tag{7}$$

$$C_{i2}(s) = Y_{i2}(s) + B_i(s) \text{ for } s = 1,..,S. \tag{8}$$

The period utility function has constant relative risk aversion γ and the discount factor is β. The budget restriction in period 1, (7), says that agent i's endowment is spent on consumption and purchases of the S different A-D assets at the prices $p(s)$. The budget restriction in period 2, (8), says that i's consumption in state s

equals his endowment in that state and the number of A-D asset s bought in the previous period.

Agent i's first order condition for A-D asset s is

$$p\left(s\right)=\beta\Pr_{i}\left(s\right)C_{i1}^{\gamma}/C_{i2}^{\gamma}\left(s\right), \tag{9}$$

where $\Pr_i(s)$ is agent i's subjective probability assessment of state s, and $C_{i2}(s)$ his consumption in that state. Solve this equation for $C_{i2}(s)$, sum over all n agents and use the fact that aggregate consumption equals aggregate endowment, $Y_2(s)$, in every state. This gives the price of A-D asset s

$$p\left(s\right)=\beta\left(\textstyle\sum_{i=1}^{n}C_{i1}\Pr_{i}\left(s\right)^{1/\gamma}\right)^{\gamma}/Y_{2}^{\gamma}\left(s\right). \tag{10}$$

When $\gamma = 1$ (logarithmic utility), then this simplifies to $\beta/Y_2^{\gamma}(s)$ times a weighted average of the agents' probabilities that state s will occur. The weights are proportional to the agents' consumption in period 1. If these are roughly equal, then average is an unweighted average similar to the aggregate distribution in the SPF. This can then be interpreted as the appropriate belief of a "representative agent."

Any asset can be seen as a combination of A-D assets, so the results in this section suggests that the "aggregate distribution" may be an important factor in the pricing of all assets.[5] For instance, a real bond pays one unit in all states, so its price must be $\Sigma_{s=1}^{S}p(s)$. A nominal bond, will require a compensation for inflation. For simplicity, suppose there is no inflation in states $s = 1, \ldots, S-1$, but inflation is π in state S (this makes the inflation risk "systematic"). The price of a nominal bond is then equal to the price of a real bond minus $\pi p(S)$, so difference between the real and nominal interest rate is affected by the weighted average of individual beliefs as in (10).

4 Estimating Uncertainty from Survey Probabilities

We want to estimate the variance of both the aggregate distribution and of each individual distribution. A straightforward, but crude, way to estimate the mean and

[5]When $\gamma \neq 1$, then the asset price depends not only on the "aggregate distribution" but also on how $\Pr_i(s)$ is distributed among agents. In particular, if $\gamma > 1$, if the probabilities for state s is made more different between agents but with $\Sigma\Pr_i(s)$ unchanged, then $p(s)$ decreases.

variance from a histogram (and used in a similar context by Lahiri and Teigland (1987)) is

$$\widetilde{\mathrm{E}}\pi = \sum_{k=1}^{K} \bar{\pi}(k) \Pr(k) \text{ and } \widetilde{\mathrm{Var}}(\pi) = \sum_{k=1}^{K} \left(\bar{\pi}(k) - \widetilde{\mathrm{E}}\pi\right)^2 \Pr(k), \tag{11}$$

where $\bar{\pi}(k)$ and $\Pr(k)$ are the midpoint and probability of interval k, respectively. The lowest and highest intervals, which are open (see Table 1), are assumed to be closed intervals of the same width as the interior intervals. This approach essentially assumes that all probability mass are located at the interval midpoints.[6] However, the shape of the histograms in *Figure 1*, which often look fairly bell shaped, suggests that this approach overestimates the variance. It rather seems plausible that relatively more of the probability mass within an interval is located closer to the mean. One very convenient possibility is that the histograms are generated from normal distributions.

The mean and variance of the normal distribution are estimated by minimizing the sum of the squared difference between the survey probabilities and the probabilities for the same intervals implied by the normal distribution. This can be thought of as a non-linear least squares approach where the survey probability is the dependent variable and the interval boundaries the regressors.

Figure 2.a shows the aggregate survey probabilities once again, and *Figure 2.b* shows the difference between the aggregate survey probabilities and the probabilities implied by the normal distribution. The normal distribution (with two parameters) seems to be able to fit most of the intervals (6, 10, or 15 depending on period) most of the time.

Figure 3 shows the aggregate standard deviation estimated in three ways: (*i*) as in the crude method in (11) using SPF's intervals; (*ii*) from a fitted normal distribution using SPF's intervals; and (*iii*) from a fitted normal distribution, but using 2% intervals for the whole sample. The general pattern is the same for all three methods, but there are some differences. As expected, the crude method is consistently above the estimated normal and the difference is particularly large

[6]An alternative, which gives a continuous distribution, is to assume a flat (uniform) distribution within each interval. This gives the same estimator of the mean as in (11), but the variance is increased by one twelfth of the squared interval width, that is, by 1/12 for most of the sample, but 1/3 for 74Q1-91Q3 (see Table 1).

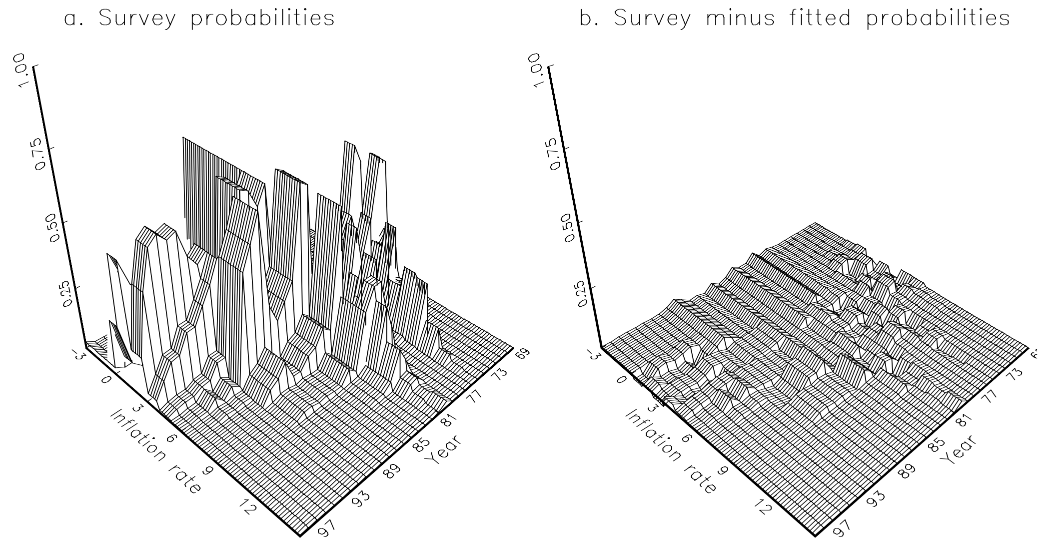

Figure 2: Aggregate probabilities and estimation error. Subfigure a. shows the aggregate survey probabilities for every year 1969-1999. Subfigure b. shows the difference between the survey probabilities and the implied probabilities from fitted normal distributions.

during the 1980s when there were few and wide intervals (see Table 1).

The two estimates of normal distributions are fairly similar, but wide and few intervals give somewhat higher standard deviations. This suggests that the high standard deviations during the 1980s are partly due to the 2% intervals used by SPF at that time (this period is marked by vertical lines). However, the difference between the two estimates during the 1970s and 1990s, when SPF used 1% intervals, suggest that this effect is small. In fact, none of our main findings are affected by using 2% intervals throughout the sample.

To mitigate the effects of a few extreme outliers (forecasters), $\mathrm{Var}(\mu_i)$ is approximated with the square of half of the distance between the 84th and 16th percentiles of μ_i. If μ_i were normally distributed, then this would be the same as the variance.[7]

[7] The square of the interquartile range divided by 1.35, which is another common robust measure of dispersion, gives very similar result except for 1980 and 1986 when it gives lower values.

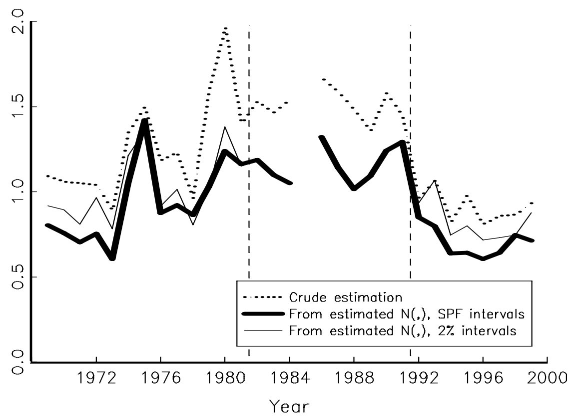

Figure 3: Aggregate standard deviation from different methods. This figure compares the result from three different ways of estimating the aggregate standard deviation: (*i*) the crude method in (11) using SPF's intervals; (*ii*) by fitting a normal distribution, using SPF's intervals; and (*iii*) by fitting a normal distribution, but using 2% intervals for the whole sample.

Similarly, the median of σ_i^2 is used instead of the mean.[8,9]

5 Empirical Results

5.1 A Decomposition of Aggregate Uncertainty

Figure 4.a shows approximations to the decomposition in (2). Figure 4.a also shows that the general uncertainty of forecasters typically contribute more to the aggregate variance than the disagreement among forecasters. *Figure 4.b* shows that Median(σ_i^2) and Var(μ_i) indeed sum (approximately) to the variance estimated from the aggregate probabilities, Var$_A(\pi)$, which supports out interpretation of the ag-

[8]Var(μ_i) and the quasi-variance evolve quite diffently over 1969-1999. The average is much higher for the variance (0.66 and 0.35, respectively), as is the standard deviation (0.55 and 0.27, respectively), even if the correlation is as high as 0.81. E(σ_i^2) and Median(σ_i^2) also evolve differently. The average has a higher average (0.81 and 0.52, respectively) and standard deviation (0.37 and 0.28, respectively), and the correlation is 0.78. It can be noted that Zarnowitz and Lambros (1987) use a traditional estimator of the cross sectional variance, which probably explains the very large variation they get.

[9]An alternative way of estimating Eσ_i^2, which is not used in this paper, is to use the decomposition in (2). It is typically easier to estimate the variance of the aggregate distribution than of the individual distributions, since some forecasters fill in probabilities for only one or two intervals. It is also relatively straightforward to estimate the dispersion of individual point forecasts.

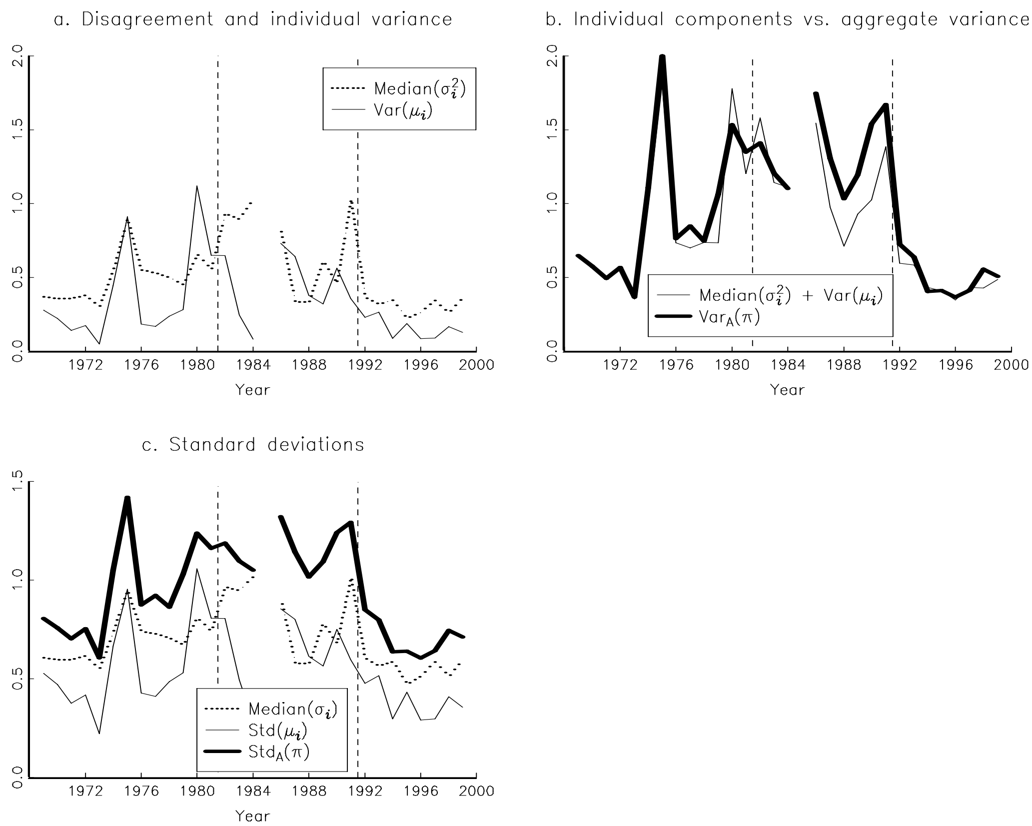

Figure 4: Individual and aggregate uncertainty. Subfigure a. shows the median individual uncertainty and the variance of individual point forecasts. Subfigure b. compares the sum of the two series in subfigure a. with the variance of the aggregate distribution obtained from fitting a normal distribution. Subfigure c. compares the standard deviations of the aggregate distribution with the median individual standard deviation and also with the standard deviation of individual point forecasts.

gregate distribution.

The implication of the finding in Figure 4.a is that the aggregate standard deviation and the median individual standard deviation evolve quite similarly in *Figure 4.c.* However, the disagreement among forecasters often moves in the same direction as the other two measures of uncertainty. The general pattern is that uncertainty was low before 1973 and after 1992 and fairly high in between with peaks in the early 1970s, late 1970s, and early 1990s.

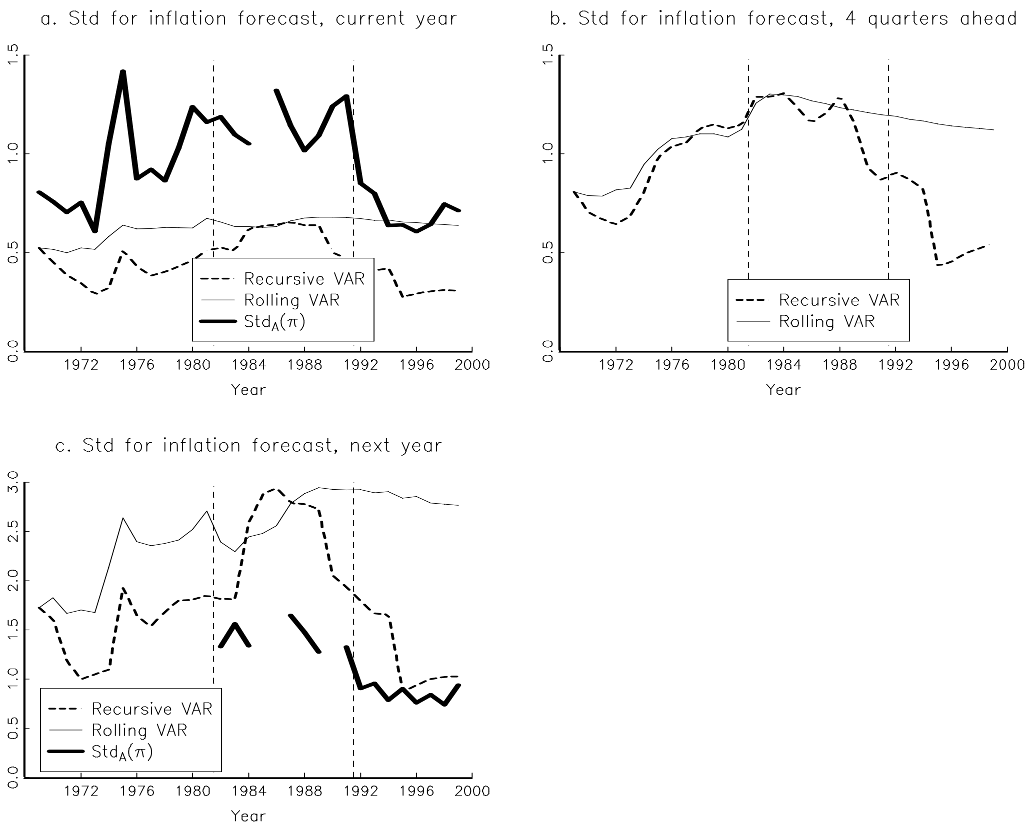

Figure 5: Uncertainty in survey and VAR models. This figure compares the aggregate standard deviation, Std_A, from SFP with estimated standard deviations for out-of-sample forecasts with different VAR models. Subfigure a. shows the standard deviations for inflation during the current year from recursive and rolling VARs and from the survey. Subfigure b. shows the std for 4-quarter inflation from the resursive and rolling VARs. Subfigure c. is similar to subfigure a., except that it shows std for inflation during the next year.

5.2 Comparison with Other Measures of Uncertainty

We will compare the uncertainty estimated from the survey data with that from a VAR model of quarterly US data on real GDP, GDP deflator inflation, the federal funds rate, and a 3-year interest rate for the period 1954Q3 to 1999Q3. The VAR is estimated in two different ways: on a rolling data window of 15 years ("rolling VAR") and with more and more data ("recursive VAR"). We then use the implied inflation forecast error standard deviation for the first and second year out-of-sample as our measure of VAR uncertainty of inflation.

Figure 5.a shows three estimates of the standard deviation of inflation during

the current year (forecasts are made in the first quarter of each year). The first thing to note is that the VAR uncertainty is much lower than the aggregate survey uncertainty. We shortly will return to this issue, and in the meantime focus on how uncertainty changes over time.

The survey and the rolling VAR changes in similar ways over time. Both have relatively low uncertainty before 1974 and after 1992, but relatively high uncertainty in between. There are differences, however, and the correlation between the two series is only 0.55 for 1968-1981, but as high as 0.80 for 1982-1999. In particular, the rolling VAR seems unable to capture the extent to which uncertainty increased directly after the two oil price shocks (in 1974 and 1979). The recursive VAR (where more and more data is used) has very different changes over time: uncertainty increases after the first oil price shock, but is then almost flat for the rest of the sample. In particular, the recursive VAR fails to capture the decrease in inflation uncertainty since the early 1990s.

We now return to the issue of why the level of VAR uncertainty is so much lower. The first possibility is that the aggregate standard deviation is the wrong benchmark, and that the VAR uncertainty should instead be compared with the median individual uncertainty. However, by comparing with Figure 4.c we see that the VAR uncertainty is still too low, although somewhat less so. For instance, the median individual uncertainty peaks at almost one, whereas the standard deviation of rolling VAR peaks at 0.6.

Another possibility is that the forecasters misinterpret the survey questions (as the authors of this paper did for some time). The survey asks for the distribution of the GDP deflator for year t divided by the GDP deflator for year $t-1$. This is uncertainty for the average inflation in year t. For instance, if $\pi_t(q)$ is the inflation during the first q quarters of year t, then the average inflation in year t is close to $\pi_t(1)+\pi_t(2)+\pi_t(3)+\pi_t(4)$, which is the way we have calculated the VAR uncertainty. Suppose the survey respondents instead thought about inflation over year t, that is, $\pi_t(4)$. *Figure 5.b* shows the standard deviations from the VARs using this definition of inflation—and they are now much more in line with the survey uncertainty.

We are uncertain about how we should think of the level of uncertainty in the survey and VARs, but it is comforting that our conclusions regarding the changes in uncertainty are the same in Figures 5.a and b: the survey and rolling VAR seem to reflect the same reality, whereas the recursive VAR fails to capture the decline in uncertainty during the 1990s.

Figure 5.c shows the values for one-year inflation for the next year. There is survey data only from 1981Q3, and some years are missing because of errors in the survey (see footnote 3). Still, the figure suggests the same pattern as for the current year inflation.

Table 2 shows the fraction of years that actual deflator inflation was within the $x\%$ confidence interval calculated from the different alternative measures of the standard deviation. The calculations uses the mean of the aggregate distribution as the point forecast and assumes a normal distribution. The correct definition of the standard deviation implies that actual inflation should be within an $x\%$ confidence interval $x\%$ of the time (in a large sample). We note from Table 2 that the aggregate standard deviation, $\text{Std}_A(\pi)$, satisfies this requirement fairly well. However, the median individual uncertainty, $\text{Median}(\sigma_i)$, generates too narrow confidence bands. This does not square well with the theoretical argument in Section 3 where we found that the unweighted forecast (which is the one used in Table 2) should have a smaller forecast uncertainty than the average individual uncertainty. This may simply due to the effect of a small sample, or that forecasters underestimate uncertainty.

5.3 Inflation Uncertainty and Macro Data

The upper panel of *Table 3* shows how the survey measures of uncertainty are correlated with the point forecasts as well as the level and recent volatility (the standard deviation over the last 5 years) of inflation and GDP growth.

The aggregate standard deviation, $\text{Std}_A(\pi)$, is strongly correlated with the point forecasts as well as last quarter's actual inflation rate. The median standard deviation across forecasters, $\text{Median}(\sigma_i)$, is also strongly correlated with the other variables.[10] The standard deviation of means across forecasters, $\text{Std}(\mu_i)$, which measures disagreement among the forecasters, is very highly correlated with both the point forecasts and inflation.

The lower panel of Table 3 shows (the lower half of) a correlation matrix of the different measures of survey uncertainty. The aggregate standard deviation,

[10] Zarnowitz and Lambros (1987) also found that average expected inflation and average individual uncertainty are positively correlated.

Table 2: Comparison of confidence bands and ex post inflation. This table shows the fraction of years when actual inflation is inside the $x\%$ confidence bands from different measures of forecast uncertainty.

Confidence band from:	Confidence level ($x\%$): 95%	90%	80%	66%
$\text{Std}_A(\pi)$	0.90	0.90	0.84	0.58
Median(σ_i)	0.84	0.81	0.68	0.35
Std(μ_i)	0.61	0.55	0.39	0.29

The confidence bands are calculated assuming a normal distribution and are calculated as: mean inflation forecast $\pm$ the critical value times the standard deviation. Actual inflation is measured as the percentage change in the GDP deflator (annual-average).

$\text{Std}_A(\pi)$, is positively correlated with both its components, but more so with the degree of disagreement among forecasters, Std(μ_i), which therefore tends to dominate the movements in aggregate uncertainty.

It is also interesting to see if the positive correlation of uncertainty and level of inflation holds also on the "micro level" in the sense that forecasters with a high mean in a certain year also tend to report high uncertainty (here measured as standard deviation of the distribution) in the same year. This correlation fluctuates substantially over time (between -0.32 and 0.52), but has only a slightly positive mean (0.12). Another possibility is that forecasters whose point forecasts are far from the medium also are more uncertain. However, this hypothesis is not supported by data either. The correlation between the standard deviation and the absolute deviation of the point forecast from the median fluctuates over time around zero (between -0.29 and 0.41 with a mean of 0.02).

Table 3: Correlation of uncertainty and macro series

	$\text{Std}_A(\pi)$	$\text{Median}(\sigma_i)$	$\text{Std}(\mu_i)$	$\frac{\text{Var}(\mu_i)}{\text{Var}(\pi)}$
Point forecast	0.60	0.53	0.61	0.31
Inflation	0.50	0.51	0.53	0.22
GDP growth	-0.36	-0.32	-0.43	-0.24
Std(inflation)	0.48	0.62	0.27	-0.04
Std(GDP growth)	0.41	0.48	0.29	0.03
$\text{Std}_A(\pi)$	1.00			
$\text{Median}(\sigma_i)$	0.78	1.00		
$\text{Std}(\mu_i)$	0.85	0.46	1.00	
$\text{Var}(\mu_i)/\text{Var}(\pi)$	0.50	-0.09	0.83	1.00

The sample is 1969-1999 (excluding 1985). Inflation is measured by the annual GDP deflator inflation rate. Both inflation and GDP are lagged one quarter relative to the measure of uncertainty. Std(x) is the standard deviation of x over the previous 5 years.

5.3.1 Time Series Results

The high correlation of all measures of uncertainty with inflation and the failure of the recursive VAR to capture forecasters' uncertainty suggest that models of inflation are likely to have heteroskedastic errors. This conjecture is strongly supported by the data. A first implication is that inflation forecast uncertainty should be positively correlated with the level of inflation in the data, just as it is in expectations. This means that in modelling inflation as a GARCH process (this was indeed the first application of both ARCH and GARCH processes; see Engle (1982) and Bollerslev (1986)) one should expect to find an asymmetric component, such that a positive error leads to higher variance. The data on quarterly US inflation strongly support this view, with the coefficient of the asymmetric shock in the TGARCH(1,1) (see Glosen, Jagannathan, and Runkle (1993)) having p-value below 2% (the mean was modelled as an AR(3)). When a recursive TGARCH is used to construct error bands for the inflation forecasts, the resulting standard deviations perform quite well in approximating the aggregate uncertainty from the survey, at least somewhat better than the recursive VAR (which is a fair comparison since the TGARCH we estimate is also recursive).

Yearly data seem well modelled with a more traditional form of heteroskedasticity, where the variance of inflation depends on the past level of inflation. The point is shown with a simple LM test in two regressions. We model inflation as a AR model. Then we run a regression of the squared residual on a constant and on the level of inflation in the previous period: the residuals are heteroskedastic at the 1% significance level. A recursive AR model with standard deviations built according to this approach produces a series for standard deviations which is as closely correlated with the aggregate survey standard deviations as the results from the recursive VAR.

The implication is that the assumption of iid errors gives rise to misleading error bands for inflation forecasts. This problem is particularly severe if time series models are estimated on period of high (low) inflation and used for prediction in periods of low (high) inflation. In policy application this point is important, as more central banks are publishing inflation forecasts and associated error bands, often produced by time series models.

6 Summary

This paper considers appropriate procedures to extract a measure of uncertainty from survey data, with an application to inflation uncertainty. We propose a new method, which allows decomposing uncertainty into the average individual uncertainty and a measure of disagreement. This interpretation implies that forecasters underestimate uncertainty on average, which is unappealing but supported by our sample. In any instance, theoretical considerations suggest that this measure could be the most useful for asset price analysis. Moreover, disagreement among forecasters is highly correlated with it. Since forecasters' disagreement is much more readily available, our analysis suggests that it is a good proxy for inflation uncertainty, for both theoretical and empirical reasons.

All measures of uncertainty indicate that inflation is perceived as heteroskedastic, as is in fact the case. In time series forecasting this feature must be taken into account to construct accurate error bands. This is particularly relevant for the central banks that publish confidence inflation forecasts with confidence bands ("fan charts").

A Appendix: Data

This appendix presents the data sources.

Real GDP and the GDP deflator series (1955Q1to 1999Q3) are from the Bureau of Economic Analysis (1959Q1-1999Q3 is available at http://www.bea.doc.gov/bea/dn1.htm which are spliced in 1959Q1 with the data for 1955Q1-1959Q1 available at http://www.bea.doc.gov/bea/dn/0898nip3/maintext.htm). The Federal funds rate and the 3-years T-bill rate (constant maturity) are aggregated to quarterly from monthly data by taking the average of the data at FRED (available at www.stls.frb.org/fred/data/irates.html).

B Appendix: Proof of Result in Section 3

By taking the limit of (3) as $n \to \infty$, we note that the forecast error variance of the combined forecast equals the (average) covariance of individual forecast errors. From (4) this covariance (the covariance forecast errors of agent i and j) is

$$\begin{aligned} \mathrm{E}(\pi - \mu_i)(\pi - \mu_j) &= \mathrm{E}\left[\pi - (s_{\pi i}/s_{ii})z_i\right]\left[\pi - (s_{\pi j}/s_{jj})z_j\right] \\ &= s_{\pi\pi} - s_{\pi j}^2/s_{jj} - s_{\pi i}^2/s_{ii} + (s_{\pi i}/s_{ii})(s_{\pi j}/s_{jj})s_{ij}, \end{aligned}$$

where s_{ji} is covariance of z_i and z_j. We assume $s_{\pi j} = s_{\pi i}$ and $s_{jj} = s_{ii}$, which simplifies the expression to

$$\begin{aligned} \mathrm{E}(\pi - \mu_i)(\pi - \mu_j) &= s_{\pi\pi} - 2s_{\pi i}^2/s_{ii} + (s_{\pi i}/s_{ii})^2 s_{ij} \\ &= \sigma_i^2 - (s_{\pi i}/s_{ii})^2(s_{ii} - s_{ij}), \end{aligned}$$

where σ_i^2 is the individual forecast error variance from (4). We therefore have to show that the last term is the cross sectional (across agents $i = 1, 2, \ldots$) variance of μ_i, $\mathrm{Var}(\mu_i)$. Note that $(s_{\pi i}/s_{ii})^2$ is the square of the individual projection coefficient; see (4). We therefore only have to show that the cross sectional variance of z_t is $s_{ii} - s_{ij}$.

Let $\bar{z}$ be the cross sectional sample mean of z_i. It is clear that $\mathrm{E}z_i|\bar{z} = \bar{z}$, and that $\mathrm{Var}(z_i|\bar{z}) = s_{ii} - \mathrm{Var}(\bar{z})$; see footnote 4. The variance of $\bar{z}$ is

$$\mathrm{Var}\left(\sum_{j=1}^{n} z_j/n\right) = \frac{1}{n^2}\left[n s_{ii} + n\left(n-1\right)s_{ij}\right],$$

so the limiting value, as $n \to \infty$, is s_{ij}, and the proof is done.

References

Benninga, S., and J. Mayshar, 1997, "Heterogeneity and Option Pricing," mimeo Tel-Aviv University.

Bollerslev, T. P., 1986, "Generalized Autoregressive Conditional Heteroscedasticity," *Journal of Econometrics*, 31, 309–328.

Croushore, D., 1993, "Introducing: The Survey of Professional Forecasters," *Federal Reserve Bank of Philadelphia Business Review*, pp. 3–13.

Engle, R. F., 1982, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation," *Econometrica*, 50, 987–1008.

Federal Reserve Bank of Philadelphia, 1999, "Suvey of Professional Forecasters: Data Information," Documentation of historical data series. http://www.phil.frb.org/econ/spf/spfdoc.html.

Figlewski, S., 1983, "Optimal Price Forecasting Using Survey Data," *Review of Economics and Statistics*, 65, 13–21.

Glosen, L. R., R. Jagannathan, and D. Runkle, 1993, "On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance*, 48, 1779–1801.

Granger, C. J., and R. Ramanathan, 1984, "Improved Methods of Combining Forecasts," *Journal of Forecasting*, 3, 197–204.

Lahiri, K., and C. Teigland, 1987, "On the Normality of Probability Distributions of Inflation and GNP Forecasts," *International Journal of Forecasting*, 3, 269–279.

Thomas, L. B., 1999, "Survey Measures of Expected U.S. Inflation," *Journal of Economic Perspectives*, 13, 125–144.

Varian, H. R., 1985, "Divergence of Opinion in Complete Markets: A Note," *Journal of Finance*, 40, 309–317.

Zarnowitz, V., 1967, *Ann Appraisal of Short-Term Economic Forecasts* . No. 3, NBER, New York.

Zarnowitz, V., and L. A. Lambros, 1987, "Consensus and Uncertainty in Economic Prediction," *Journal of Political Economy*, 95, 591–620.