

**PHILADELPHIA FED FORECASTING SURVEYS:  
THEIR VALUE FOR RESEARCH**

Dean Croushore

University of Richmond

Visiting Scholar, Federal Reserve Bank of Philadelphia

February 26, 2009

The two forecasting surveys conducted by the Federal Reserve Bank of Philadelphia, the Livingston Survey and the Survey of Professional Forecasters, provide a long track record of macroeconomic forecasts. Over many decades, they have provided not only timely information for policymakers and other economic analysts, but they have also been used to answer numerous research questions. This article describes the surveys and highlights some of the major ways the survey has been used by researchers.

## **PHILADELPHIA FED FORECASTING SURVEYS: THEIR VALUE FOR RESEARCH**

For almost twenty years, the Federal Reserve Bank of Philadelphia has conducted both the Survey of Professional Forecasters and the Livingston survey. Both surveys of private-sector forecasters provide researchers, central bankers, news media, and the public with detailed forecasts of major macroeconomic variables. The surveys have been made available to the public at no charge, reflecting the public education mission of the Federal Reserve. Based on the questions received by the Federal Reserve staff, the surveys have proved helpful for people who are planning for the future. They have also provided useful input into the decisions of policymakers at the Federal Reserve and elsewhere.

The Livingston survey is the oldest of the two Philadelphia Fed surveys, having started in 1946 when Joseph Livingston, a Philadelphia newspaper reporter, wanted to get a sense of what forecasters thought would happen to the economy in the next year. He continued to publish his survey every six months, gathering and reporting on the forecasts and tracking their evolution over time. His survey, which was the only collection of private-sector forecasts in the country, gained a national following. Economic researchers discovered the survey in the early 1970s, and began using it to test theories about people's expectations. By 1978, Livingston was having trouble keeping up with all the requests for the data and turned the data over to Research Department of the Federal Reserve Bank of Philadelphia, which organized the data in a computer database and made it available to researchers on request. Livingston still ran the survey but then had the Philadelphia Fed compile the results and maintain the database. Livingston provided the

first report of the survey's results in his column in the *Philadelphia Inquirer*. When Livingston died in 1989, the Fed took over the administration of the survey and carried on Livingston's legacy. Since the advent of the Internet, the Fed has made all the historical Livingston data available on its website.

The Survey of Professional Forecasters began as the idea of Victor Zarnowitz and others at the American Statistical Association and the National Bureau of Economic Research. They began the ASA/NBER Economic Outlook Survey in 1968 and carried it out successfully for 22 years. The survey was similar to the Livingston survey in asking private-sector forecasters for projections for the next year for major macroeconomic variables. But the ASA/NBER survey was conducted more frequently than the Livingston survey (quarterly instead of semi-annually), asked for quarterly forecasts (instead of Livingston's half-year forecasts), and included some unique questions asking forecasters about the probabilities of different outcomes, instead of just the point forecasts reported by the Livingston survey.

Over time, interest in the ASA/NBER survey waned, and the ASA and NBER agreed to fold the survey after the first quarter survey in 1990. However, I happened to be doing research on expectations in early 1990 and had obtained the data from the NBER. It seemed like a huge loss to let the survey disappear, so I began working to convince senior managers at the Philadelphia Fed that the Fed should take over the survey. They were at first reluctant because they had just begun to run the Livingston survey completely. But I was able to persuade them that the surveys were different enough from each other and that Federal Reserve policymakers could gain from the information in the survey. So, the Philadelphia Fed took on a second survey of forecasters, which we christened the Survey of Professional Forecasters (SPF).

Why do people need forecasts? In their personal budget planning, people need to forecast inflation; in planning production, firms need to forecast demand for their products; in buying and selling financial assets, investors need to forecast both inflation and future interest rates; in setting policy, government analysts need to know how the economy is likely to fare in the future. Forecasting surveys can help all these groups figure out the most likely outcomes for the variables they are most concerned about.

The Philadelphia Fed's surveys are not the only surveys of forecasters that exist. After the Livingston survey and ASA/NBER survey had been around for some time, the success of those surveys brought competitors. Perhaps the best known U.S. survey is the *Blue Chip Economic Indicators*, which was started by Robert Eggert in 1976. The Blue Chip concept was to publish forecasts monthly (compared with the quarterly SPF and the semi-annual Livingston survey), and to publish the names of each forecaster along with his or her forecast (whereas both the SPF and Livingston surveys were anonymous). In addition, the National Association of Business Economics (NABE) has produced a quarterly survey of forecasters since the early 1960s, and the *Wall Street Journal* also conducts a similar survey that is reported in great detail on its web site. Direct measurement of consumers' inflation expectations is gathered by the monthly University of Michigan survey of consumers, which asks a random sample of consumers for their forecast of inflation. For other countries there have been a number of surveys, most notably *Consensus Forecasts*, which gathers detailed forecasts for all major developed countries in the world and less detailed forecasts for numerous other countries. Also, the European Central Bank started a European version of the Survey of Professional Forecasters in 1999 after consulting with the Philadelphia Fed on its methods.

The SPF and Livingston surveys both forecast major macroeconomic variables, including:

- nominal GDP
- real GDP
- inflation (consumer price index)
- unemployment rate
- interest rate on 3-month Treasury bills
- interest rate on 10-year Treasury notes
- industrial production
- corporate profits after tax
- housing starts
- business fixed investment

Each survey also includes a number of other variables. The Livingston survey also asks forecasters for their projections of the producer price index, S&P 500 stock prices, average weekly earnings, retail trade sales, auto sales, and the prime interest rate. The SPF asks for forecasts of the GDP price index, payroll employment, the interest rate on AAA bonds, three additional inflation rates (consumer price index excluding food and energy prices, personal consumption expenditures price index, and personal consumption expenditures price index excluding food and energy prices), as well as all the remaining major components of real GDP (consumption, residential fixed investment, federal government spending, state and local government spending, net exports, and changes in private inventories). In addition, in every survey, the SPF asks a question about the probability of a decline in real GDP in each of the next five quarters and questions about the probability that real GDP, the inflation rate in the GDP price index, the CPI excluding food and energy prices, and the PCE price index excluding food and energy prices will fall into certain ranges. The latter questions are designed to get an idea of

the uncertainty that forecasters attach to their forecasts. Each SPF survey also asks about forecasts for inflation in both the CPI and PCE price index for the next 5 years and the next 10 years, while the Livingston survey asks about real GDP growth and CPI inflation for the next 10 years.

Both the SPF and Livingston surveys provide anonymity for the forecasters. The survey news release lists the names of the forecasters, but a reader cannot tell which forecaster provided which forecast. The idea is that the forecasters are more likely to reveal their true forecasts if they know that their name will not be associated with a particular forecast. If they think that their forecast is very different from that of other forecasters, they have no incentive to hide it. However, if they were providing their forecasts in a non-anonymous survey (such as the Wall Street Journal survey or the Blue Chip survey), they might prefer to modify their forecast. Some forecasters might want to shade their forecasts closer to the consensus, out of fear that they will be seen as being out of the mainstream. Other forecasters might be looking for fame and intentionally make their forecasts stand out from the crowd. Since the forecasts are anonymous in the SPF and Livingston surveys, the surveys avoid these problems.

How do people use the surveys? From questions posed to the Philadelphia Fed staff over time, we can tell that people use these surveys in a variety of ways. Some labor unions and firms use the survey's forecasts as starting points for wage negotiations, especially the forecasts of inflation. Business firms, especially industrial firms, use the surveys to help them gauge the demand for their products. Investors use the surveys for guidance on the likely future path of interest rates and stock prices. Households, business firms, and Federal Reserve policymakers all use the surveys to inform them of inflation expectations over the next four quarters, and over the

next five years and ten years. Economic researchers find the survey useful for calculating expected real interest rates, which are nominal interest rates minus the expected inflation rate over a bond's time to maturity; expected real interest rates are the main variable driving many economic models of the economy.

The timing of the SPF and Livingston surveys differs, in part because the SPF is conducted four times each year, while the Livingston survey is conducted just twice a year. More importantly, the focus of the SPF is on the national income accounts, so the survey forms are mailed to participants immediately following the initial release of the GDP data for the preceding quarter, which occurs in late January, April, July, and October each year. The forecasters are given about 10 days to respond to the survey questions, and then email their responses to the Philadelphia Fed before the middle of the following month (when key data on consumer prices are released). The Livingston survey timing is based on the release of the consumer price index data in May and November, with the forecasts due prior to the next month's release of the consumer price index.

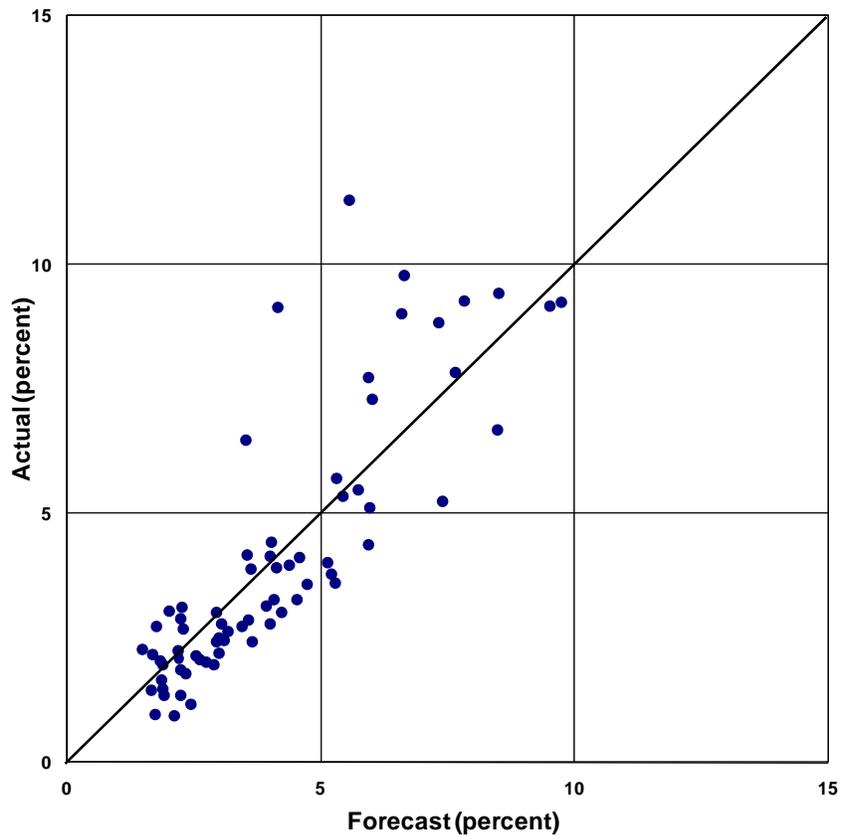
## **EVALUATING THE SURVEY FORECASTS**

Given the uses to which people, firms, and policymakers put the surveys, it is important that the forecasts be accurate. Of course, no forecast is going to be on the mark all the time. But we can hope that the SPF and Livingston surveys are at least as good as alternative methods of forecasting, so they provide some useful information to people who use them.

Economists have tested the surveys extensively. Simple tests examine the forecast errors over time to see if they are zero, on average, which would be the hallmark of an unbiased forecast. More sophisticated tests look at the correlation between forecast errors and information available to forecasters when they made their forecasts; if such a correlation exists, then the forecasters in the survey are not using that information efficiently.

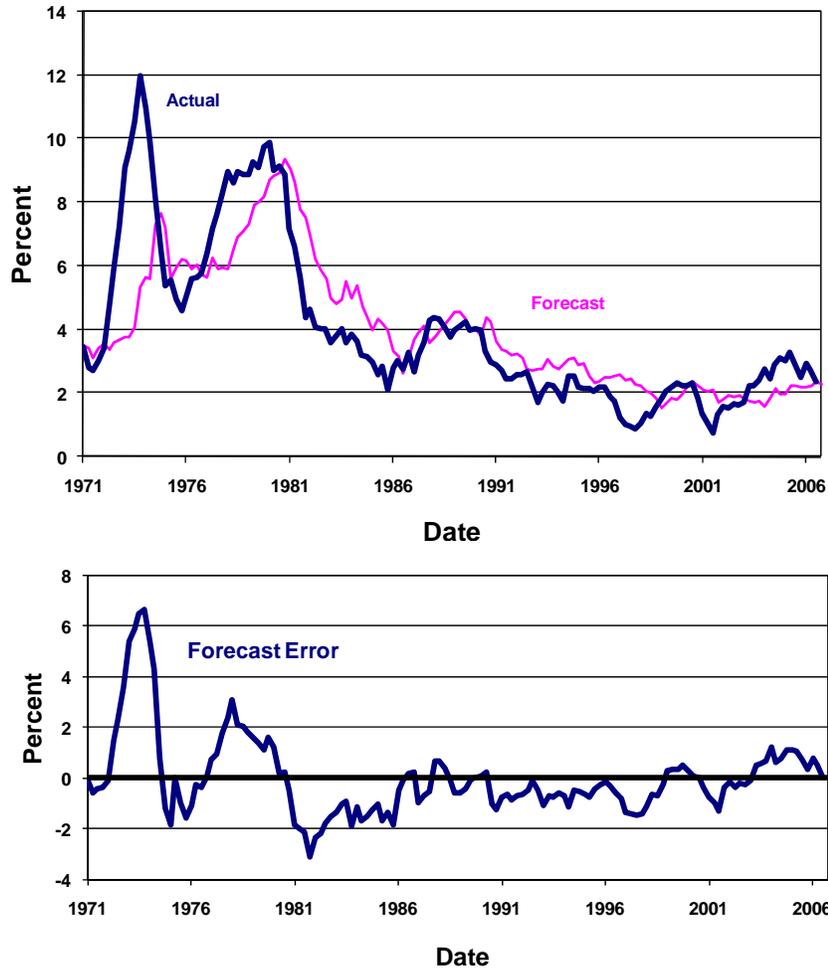
A visual inspection of the data sometimes suffices to see whether a particular forecast has forecast errors that are zero on average. Figure 1 shows a scatter plot in which the value of the inflation rate (based on the GDP deflator) is plotted on the vertical axis and the forecasts from the Livingston survey are plotted on the horizontal axis. The 45-degree line in the figure helps interpret the figure because if the forecasts were completely accurate, every point in the diagram would be on that line. The fact that most of the points in the graph are close to the 45-degree line suggests that the forecasts are fairly accurate. Formal statistical tests confirm that the mean forecast error in this series is not statistically significantly different from zero (Croushore, 2008).

**Figure 1**  
**Forecasts Versus Actuals**  
**Livingston Survey: 1971:H1 to 2006:H1**



Despite the unbiasedness of the survey forecasts over the entire period from the early 1970s to the mid-2000s, there are numerous periods in which the survey forecasts appear to have performed poorly. Figure 2 shows the SPF forecasts and actual values of inflation (measured using the GDP price index over the coming year from the date of the forecast, where the actuals are measured using the data as measured one quarter after the end of the forecast period).

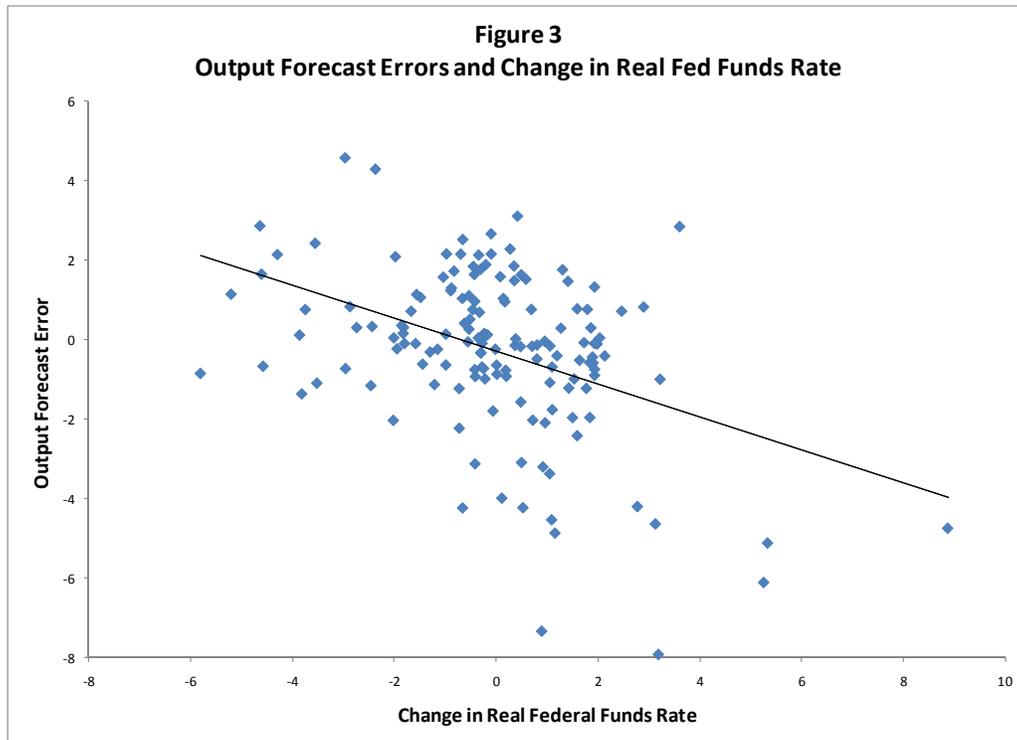
**Figure 2**  
**Median Inflation Forecast from SPF**



The SPF survey forecasts for inflation were clearly far from the mark in the early and late 1970s, with very large forecast errors. Perhaps these forecast errors were understandable, given the unprecedented increase in money-supply growth that occurred in the decade, which caught forecasters by surprise. In addition, in most of the 1990s, the forecasters made a string of forecast errors, with inflation continually coming in lower than the forecasters had projected. In that period, productivity growth surged and it took some time before the forecasters realized that the

economy was not overheating, but rather that potential output was increasing more rapidly than before, so inflation would not be rising significantly. (These concepts are explored in more detail in Croushore, 2008). Thus, the forecasters clearly go through periods in which they make persistent forecast errors.

In addition to periods in which the forecasters make persistent forecast errors, the forecasters in the surveys may be inefficient in their use of other information. Economists test this idea by examining the relationship between the survey's forecast errors and data that was known when the forecasters made their forecasts. For example, Ball and Croushore (2003) found that output forecast errors were associated with changes in the real federal funds rate, which means that the forecasters did not accurately modify their forecasts in response to a change in monetary policy. This can be seen in Figure 3, which plots the output forecast error from the SPF against the lagged change in the real federal funds rate. The negative relationship between these two variables implies that the output forecasts from the SPF are not efficient with respect to changes in monetary policy.



A little-explored aspect of the SPF is the probability distribution forecasts that it provides. Each forecaster is asked to list the probability that real GDP growth and inflation in the GDP price index will fall into certain ranges. In the most recent surveys, the forecasters are asked to state the probability that real GDP growth in the next year will be 6% or more, 5.0 to 5.9%, 4.0 to 4.9%, 3.0 to 3.9%, 2.0 to 2.9%, 1.0 to 1.9%, 0.0 to 0.9%, -1.0 to -0.1%, -2.0 to -1.1%, and -2.0% or less. The same question is also asked for real GDP growth in the following year. For the percent change in the GDP price index, the ranges are two percentage points higher, so the top range is 8% or more, and so on.

Diebold-Tay-Wallis (1999) analyzed these probability distribution forecasts from the SPF using new methods. Their goal was to test the accuracy of the distribution forecasts, and for the most part they found that the forecasts were reasonably accurate. However, the forecasts failed to pass some tests: (1) they place too large a probability on a large negative shock to inflation; and

(2) they make persistent inflation forecast errors, though the forecasters eventually adapt and the errors disappear. They also found that when inflation is low, uncertainty about inflation is also low.

Overall, recent research on the accuracy of the SPF and Livingston forecasts has found that they are reasonable, even if there are a few areas in which they are imperfect. However, as the literature using the surveys for research evolved over time, the accuracy of the surveys was often called into question.

## **USING THE SURVEYS TO ANSWER RESEARCH QUESTIONS**

The economics literature has been served well by both the Livingston survey and the SPF. Major research papers have been written by economists using both surveys, beginning with Turnovsky (1970). Turnovsky tried to show how forecasters formed their expectations and developed an early test of rationality of the forecasts using the Livingston survey. The first paper to use the SPF was Su-Su (1975), which evaluated the accuracy of the survey forecasts using only a few years of data.

None of the earliest papers to use the Livingston survey is reliable, however, because Carlson (1977) discovered a major flaw in the data. Because the original purpose of the survey was for journalism, Livingston did not report the actual forecasts of the forecasters in his newspaper column. Instead, he modified the forecast data if a data release occurred after the forecasters had submitted their forecasts but before his newspaper column was posted and if the data release would have changed the overall nature of the forecasts.

Carlson gives the following example. Suppose the CPI was released in September and October with a value of 121.1 and the forecasters have an average forecast for the following June of 121.2. Then, if the November data release (which came out after the forecasters had answered the survey but before the survey results were reported) for the CPI is 121.1, the June forecast is reasonable and Livingston would not adjust the forecast. But suppose the November CPI data was released as 121.6. Then if Livingston reported the November number and the June forecast, it would appear that the forecasters thought there would be deflation, even though they were really forecasting a small amount of inflation. So, Livingston would instead report a forecast of 121.7, which maintains the 0.1 increase in the CPI that the forecasters thought would happen. But this means that the forecasts that were reported were fictional and depended on Livingston's personal judgment. Carlson remedied this situation by obtaining the true forecast values from Livingston and thus restoring the integrity of the data set.

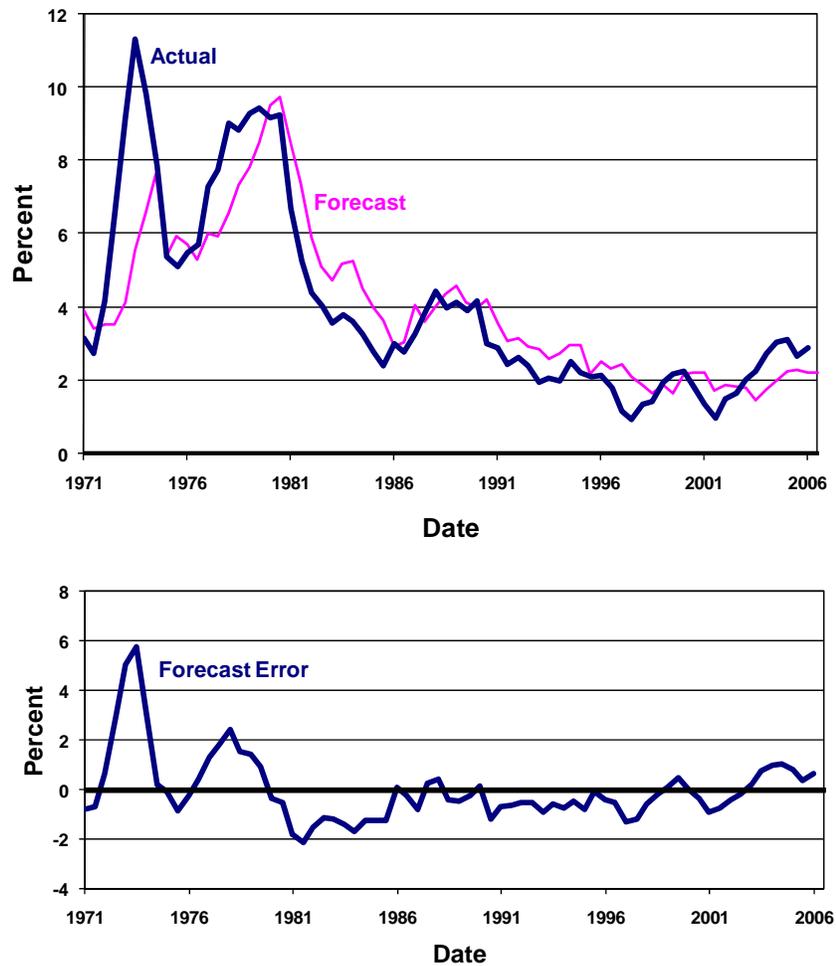
Carlson then provided a great additional public service by questioning the information set that the forecasters had at the time the survey was made. Researchers prior to Carlson had assumed that the forecasts made by the forecasters in the survey had a horizon of 6 or 12 months. For example, in the December survey, the survey form clearly asks the forecasters to forecast the June value and the December value of the CPI in the next year. The survey form usually listed the current year's September value of the CPI that was reported by the BLS. But a substantial amount of time had passed between the mailing of the survey form and the return of the form by each forecaster. Carlson suspected that many of the forecasters already knew the October value of the CPI when they sent in their forecasts. So, he sent out a survey to all the forecasters asking them what they knew when they sent in their forecasts, discovering that most of them indeed based their forecasts on the October CPI. Knowing this, the forecast horizon from October to

June is 8 months, and the horizon from October to December of the following year is 14 months. So, Carlson changed the implicit forecast horizons in the Livingston survey to 8 and 14 months instead of the commonly accepted 6 and 12 months. This made statistical work based on the survey far more accurate.

*Rational Expectations.* The Philadelphia Fed's surveys of forecasters were initially used by researchers in the early 1970s to investigate the concept of rational expectations. A number of early papers had used the Livingston survey forecasts of inflation and rejected the rational-expectations hypothesis because the survey forecasts were biased (with a non-zero mean forecast error) and inefficient (because the forecast errors were correlated with data known when the survey was taken). But Mullineaux (1978) found a major flaw in previous studies that had tested for and rejected rationality of expected inflation in the Livingston survey. The flaw is that the Chow test used in much previous work assumes identically and independently distributed (i.i.d.) errors in a framework where that is unlikely to hold. He showed that the assumption can be rejected. Mullineaux then proposed a new test that is not subject to the same statistical problem and found that the properly specified data are consistent with rationality.

Economic researchers were not convinced by Mullineaux's work, however, because a glance at the forecast errors suggested a problem. As Figure 4 shows, there were clearly some very large forecast errors in the 1970s in the Livingston survey.

**Figure 4**  
**Median Inflation Forecast from Livingston Survey**



Empirical work claiming irrationality in the Livingston survey was still not convincing until Pearce (1979) did a simple experiment to show how irrational the survey forecasts were. Pearce did a simple exercise: suppose someone just used the past data on inflation and ignored data on other variables that a forecaster might use to forecast, including the money supply and the strength of the economy. Pearce then constructed a simple forecast of inflation in which the change in the inflation rate from one period to the next depended only on the unexpected change

in the inflation rate in the previous period. Formally, economists refer to this model as an IMA(1,1) model of inflation. Pearce correctly used only the data that the participants in the Livingston survey had available to them at the time when they made their forecasts (see Croushore-Stark 2000 for more on this concept of real-time data analysis). Pearce compared his simple model's forecasts with the forecasts from the Livingston survey, and found that his model had much better forecasts for inflation than the survey. He also showed that the rise in interest rates in the 1970s was better explained by his simple IMA(1,1) model than by the Livingston survey. So if a very simple model can provide better forecasts than the survey, there does not appear to be much value in the survey, and professional forecasters are irrational because they could have used Pearce's model and made better forecasts.

The early literature on rational expectations often ran tests for unbiasedness and inefficiency of the survey forecasts. But those tests were flawed in an important way, because they failed to account for the fact that a forecast error in one survey forecast carried over to other surveys because the length of the forecast horizon (8 or 14 months) was longer than the interval between surveys (6 months). Thus a shock to inflation in one period could lead to forecast errors in two or three consecutive surveys—a situation that has come to be known as the overlapping observations problem. By failing to account for this correlation in the forecast errors, the researchers' tests for unbiasedness and inefficiency were overstating the case against the surveys. Brown-Maital (1981) finally remedied this situation, making a key methodological contribution: recognizing the overlapping-observations problem and showing how to adjust the statistical tests so they gave the correct inference. Brown-Maital then tested the Livingston survey data for unbiasedness and efficiency. They generally found no bias, unlike many earlier research papers.

But they did find some evidence that the Livingston survey forecast errors were correlated with monetary policy.

Another challenge to rational expectations using the surveys came from Fama-Gibbons (1984). They created alternative inflation forecasts based on levels and changes in nominal and real interest rates. They showed that the inflation forecasts based on interest rates outperformed the Livingston survey forecasts of inflation from 1977 to 1982. Interest-rate-based models also did better at forecasting than univariate time-series models in this period.

After much literature finding fault with the forecasting surveys, many economists began to believe that either people did not have rational expectations, or that the surveys were not representative of people's true forecasts, or both. Keane-Runkle (1990) sought to disprove both hypotheses, arguing that much of the literature on testing survey forecasts for rationality suffered from four flaws: (1) the use of the median or mean forecast across forecasters is wrong since forecasters may have different information sets; (2) past research studies failed to adjust properly for data revisions; (3) research papers that had used the Michigan survey of consumers were bad because consumers don't necessarily have an incentive to be rational in responding to the survey; and (4) past research studies failed to account for the correlation of forecast errors across forecasters. Keane-Runkle avoided these problems by using individual forecasts on the GNP deflator, basing their analysis on real-time data (the first revision of the national income data, which comes out one month after the initial release), basing their research using data from the SPF professional forecasters who have the incentive to be rational, and developing a statistical method that accounts for the correlation of forecast errors across forecasters. They evaluated current-quarter inflation forecasts from the SPF, finding that they were unbiased and efficient.

Overall, the literature on rational expectations has benefited tremendously from the existence of the SPF and Livingston survey. Though the results of tests for rationality have been mixed over time, more recent evaluations generally suggest that the survey forecasts are fairly accurate and pass most, though not all, tests for rationality.

*Expectations formation.* Research on how people form expectations has used the Philadelphia Fed forecasting surveys extensively. This line of research has a slightly different goal than the literature on testing rational expectations; it uses the surveys to investigate what information forecasters use to form their forecasts and the properties of their forecasts.

Cukierman-Wachtel (1979) introduced the idea that inflation expectations differ across individuals because of differing information sets. In their model, people observe their own prices but not other prices or the general price level, a model consistent with the theory of rational expectations. Under this model, an increase in people's uncertainty about inflation leads to more variability in their inflation expectations over time than when inflation is more stable. Cukierman-Wachtel used the Livingston survey forecasts on CPI inflation to examine the variance of inflation expectations across forecasters. They found that the variance of expected inflation across forecasters is positively related to the variance of the GNP inflation rate and the growth rate of nominal GNP. Thus volatility in the economy translates into uncertainty in people's forecasts.

One branch of this literature is devoted to finding variables that are correlated to the survey forecasts, thus revealing the data that forecasters find important in forming their forecasts.

Mullineaux (1980) used the Livingston survey forecasts to examine how forecasters form

inflation expectations, using real-time data on the money supply. He found that the forecasters used money-growth data in forming their forecasts, not just lagged inflation data, so that inflation forecasting models (like Pearce's) that are just based on past inflation rates are not efficient. Mullineaux found evidence that the expectations formation process had structural shifts over time, perhaps in response to policy changes. This is an important finding, as it provides evidence that is consistent with the Lucas critique of econometric policy evaluation, which argues that when the Federal Reserve changes the process by which it sets monetary policy (a process that clearly changed in the 1970s), then equations such as those describing the formation of inflation expectations will undergo structural shifts. Mullineaux also found evidence that the same model determining expectations also determines actual inflation, so that survey forecasts are rational.

One of the most important papers in this literature is that of Zarnowitz-Lambros (1987), who were the first to combine and compare the SPF point forecasts with the probability distribution forecasts, showing that consensus (the degree of agreement of point forecasts across forecasters) is very different from uncertainty (the degree of uncertainty about each point forecast, measured by the probability distribution questions in the survey). While the mean of the point forecasts across forecasters is close to the mean of the probability distribution forecasts across forecasters, the probability forecasts allow researchers to measure uncertainty properly. The previous method of using the standard deviation of the point forecasts across forecasters understates the degree of uncertainty. Zarnowitz-Lambros also found that higher inflation rates were associated with greater uncertainty about inflation, and showed that that increased inflation uncertainty was associated with lower real output growth.

Evans-Wachtel (1993) used the nascent literature on Markov-switching models of the economy to explore the formation of expectations. A Markov-switching model is one in which the economy shifts back and forth between two different alternative regimes, for example, between a high and a low inflation rate. In their model, Evans-Wachtel used a Markov switching model of inflation regimes to account for changes in inflation uncertainty. Changes in inflation regimes led to persistent forecast errors because it took forecasters some time to learn of the regime switch. The model's forecasts looked very much like Livingston forecasts, in the sense that the Markov switching probabilities were similar to those implied by the Livingston forecasts. A vector autoregression model (VAR), in which each variable in the model is a function of lagged values of all the other variables in the model, shows that uncertainty about the inflation regime has affected employment.

Recently, numerous researchers have begun focusing on how households form their own inflation expectations. Mankiw-Reis-Wolfers (2003) noted that professional forecasters disagree with each other in their forecasts of inflation, as do consumers. They showed that the degree of disagreement changes over time. To explain these disagreements, they developed a "sticky-information" model. The basic idea of sticky information is that there are costs of collecting and analyzing information, so that agents update their expectations infrequently. They then used the Michigan survey of consumers, the SPF, and the Livingston survey to verify their model. They found that their model helps to explain the irrationality of inflation expectations, including why forecast errors are persistent and why it takes some time before news is incorporated into the forecasts.

A related paper is that of Carroll (2003), who developed an interesting hypothesis—that households may not have rational expectations, but rather form their expectations by reading professional forecasts, which are rational. Households' expectations may not be rational because they only occasionally read the forecasts of professional forecasters and don't always pay attention to them. To test this view, he uses a statistical method known as an encompassing test to see whether the forecasts in the Michigan survey of consumers incorporates information in the SPF, or vice versa. By examining the relationship between the actual inflation rate, the Michigan consumer survey forecasts, and the SPF forecasts of inflation, he was able to show that the Michigan forecast contains no additional information that is not already in the SPF, but the SPF does contain additional information that is not in the Michigan survey. He also found evidence that SPF survey forecasts affected later Michigan surveys, but that the Michigan survey did not affect later SPF forecasts. This result suggests that over time households come to incorporate the SPF forecasts. Carroll's results are also supported by the fact that when news coverage of inflation is high, Michigan forecasts get closer to SPF forecasts. Similar results occur when Carroll uses the unemployment rate in his empirical work, rather than the inflation rate.

Thus the SPF and Livingston surveys have both been helpful to researchers interested in how professional forecasters and households form their expectations.

*Empirical Macroeconomics.* The Philadelphia Fed surveys of forecasters have been used in some of the most important empirical research papers in the past 30 years, ranging from tests of theory to explaining movements in macroeconomic data.

One puzzle that survey forecasts helped solve was the issue of why real interest rates declined so much in the 1970s. In a major research result estimating an empirical IS-LM model, Wilcox (1983) used the Livingston survey forecasts of inflation in an attempt to determine the main factors affecting nominal and real interest rates. He discovered that much of the decline in real interest rates in the 1970s (though not all) was due to increases in expected inflation rates. He argued that previous models were mis-specified mainly because they failed to include a supply-shock variable representing the prices on inputs. Once he included such a variable and used the Livingston survey forecasts as a proxy for expected inflation in calculating the real interest rate, his model fit the data well.

One of the most famous papers that empirically tests macroeconomic theory was Hall (1988), who found evidence supporting the life-cycle/permanent-income hypothesis using data on U.S. consumption spending. Hall used the Livingston survey to calculate the expected inflation rate and the expected return to the stock market. He also found that the intertemporal elasticity of substitution is close to zero, which means that changes in the real interest rate have little effect on consumption spending.

Economists have used movements in unemployment to help predict changes in wages and prices, based on a Phillips curve relationship. Neumark-Leonard (1993) explored why Phillips curve forecasts of inflation in the 1980s were much higher than inflation turned out to be. One reason why may be a shift in the dynamic specification of inflation, which would lead the Phillips curves to overpredict inflation. But suppose instead of using the dynamic specification of inflation in the Phillips curve, we were to substitute survey forecasts in their stead; do the Phillips curve fare any better? In fact, the Phillips curve inflation forecasts perform worse using

the Livingston survey than using the usual time-series process built into the Phillips curve. Neumark-Leonard found the same result using SPF forecasts of inflation. Their results suggest that wages grew less rapidly than prices because of a structural shift in the labor market, consistent with a decline in real wage growth.

This survey only touches on some of the main studies in the empirical macroeconomic literature. It is clear that the Philadelphia Fed's surveys have been instrumental in helping economists test theories and investigate structural changes in the economy.

*Monetary Policy.* One of the main mechanisms by which monetary policy affects the economy is by affecting people's inflation expectations. The Philadelphia Fed surveys have helped researchers investigate the impact of monetary policy and how policymakers determine monetary policy.

Orphanides (2003) used the SPF to provide forecasts of inflation and output to use in Taylor rule, which is a reaction function for monetary policy that suggests that the Fed sets its target for the federal funds interest rate as a function of expected inflation and the difference between output and potential output. Orphanides suggested that basing a Taylor rule on forecasts produces better estimates of what the Fed did than basing the rule on past data. Thus the Fed appears to follow a forward-looking Taylor rule rather than a backward-looking rule.

Economists have also tried to determine exactly how monetary policy affects the economy and why money is not neutral. Ball-Croushore (2003) investigated how monetary policy shocks affect actual output and inflation relative to the forecasts of output and inflation from the SPF

and Livingston survey. A monetary tightening reduces actual output, but forecasters do not reduce their forecasts by as much as output actually falls; thus output forecast errors are predictable by looking at changes in the real federal funds rate. For inflation forecasts, a change in monetary policy leads both actual and forecasted inflation to decline by about the same amount; so inflation forecast errors are not predictable from monetary policy measures. The output finding could help explain the non-neutrality of monetary policy.

How does the Fed respond to changes in expected inflation? Leduc-Sill-Stark (2007) investigated this issue, using the Livingston survey to represent expected inflation. They found that monetary policy accommodated shocks to expected inflation before 1979 by allowing real interest rates to decline, so that monetary policy contributed to the Great Inflation of the 1970s. However, after 1979, the Fed did not accommodate inflation, and the real interest rate increased in response to a shock to expected inflation.

These papers have all helped researchers investigate the formation and impact of monetary policy, and the use of the Philadelphia Fed surveys has been helpful in this empirical work.

*Other Important Research Results.* The Philadelphia Fed surveys have been used in key papers concerned with data production, optimal forecasting, and financial analysis.

One key question about the data that are issued by government statistical agencies is whether data revisions are forecastable or not. Mork (1987) sought to answer that question using the SPF survey as a measure of information known at the time the BEA releases its initial release of GDP data. Results show that the use of lagged data and the SPF forecast of GDP are

significantly correlated with the flash release and advance release of GDP. Later estimates are better, and the final (3 month later) estimate is efficient with respect to lagged data and the SPF forecast.

Some economists have also used the Philadelphia Fed surveys to investigate hypothesis in financial economics. For example, Sharpe (2002) related the SPF forecasts of one-year inflation rates and ten-year inflation rates to stock returns. Importantly, he found that a one-percentage-point rise in the long-term expected inflation rate implies a 20 percent reduction in stock prices.

The surveys have also been used to investigate optimal methods of forecasting. Fair-Shiller (1989) compared the output growth forecasts of different forecasters, using the SPF as a benchmark and seeing how well forecasts from forecasting firms DRI and WEFA did compared with the SPF. They also compared forecasts based on time-series (ARIMA) models that used only lagged values of output growth, as well as vector autoregressions and a structural model (the Fair model). For one-quarter-ahead output forecasts, the SPF dominates all the other models. But for four-quarter-ahead forecasts, DRI, WEFA, and Fair do better than the SPF and the SPF only fares well in comparison to the time-series models.

In a broader and more recent comparison, Ang-Bekaert-Wei (2007) compared inflation forecasts from the Livingston survey, SPF, and the Michigan survey of consumers. They found that the surveys forecast inflation better out of sample than do ARIMA models, term-structure models, or Phillips-curve models. Even combining forecasts from all the methods does not perform significantly better than just using the survey forecasts. They also found that the Michigan forecasts are only slightly worse than the SPF and Livingston forecasts, but still do better than the other forecasting methods.

## **SUMMARY AND CONCLUSIONS**

There can be little doubt that the Philadelphia Fed's surveys of forecasters have played an instrumental role in economic research in the past 40 years. The surveys have been used to test rational-expectations theory, to analyze the formation of inflation expectations, to use in empirical research in macroeconomics, to investigate the formation and impact of monetary policy, and to use in a variety of other studies.

## REFERENCES

- Ang, Andrew, Geert Bekaert, and Min Wei. "Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?" *Journal of Monetary Economics* 54 (May 2007), pp. 1163–1212.
- Ball, Laurence, and Dean Croushore. "Expectations and the Effects of Monetary Policy." *Journal of Money, Credit, and Banking* 35 (August 2003), pp. 473–484.
- Brown, Bryan W., and Shlomo Maital. "What Do Economists Know? An Empirical Study of Experts' Expectations." *Econometrica* 49 (March 1981), pp. 491–504.
- Carlson, John A. "A Study of Price Forecasts." *Annals of Economic and Social Measurement* 6 (Winter 1977), pp. 27–56.
- Carroll, Christopher D. "Macroeconomic Expectations of Households and Professional Forecasters." *Quarterly Journal of Economics* 118 (February 2003), pp. 269–298.
- Croushore, Dean. "An Evaluation of Inflation Forecasts from Surveys using Real-Time Data." Working paper, University of Richmond, December 2008.
- Croushore, Dean, and Tom Stark. "A Funny Thing Happened on the Way to the Data Bank: A Real-Time Data Set for Macroeconomists," Federal Reserve Bank of Philadelphia *Business Review* (Sept./Oct. 2000), pp. 15–27.
- 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), pp. 595–609.
- Diebold, Francis X., Anthony S. Tay, and Kenneth F. Wallis. "Evaluating Density Forecasts of Inflation: The Survey of Professional Forecasters." In: Robert F. Engle and Halbert White, eds., *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W.J. Granger*. (Oxford: Oxford University Press, 1999), pp. 76–90.
- Evans, Martin, and Paul Wachtel. "Inflation Regimes and the Sources of Inflation Uncertainty." *Journal of Money, Credit and Banking* 25 (August 1993, part 2), pp. 475–511.
- Fair, Ray C., and Robert J. Shiller. "The Informational Content of Ex Ante Forecasts." *Review of Economics and Statistics* 71 (1989), pp. 325–331.
- Fama, Eugene R., and Michael R. Gibbons. "A Comparison of Inflation Forecasts." *Journal of Monetary Economics* 13 (1984), pp. 327–348.
- Hall, Robert E. "Intertemporal Substitution in Consumption." *Journal of Political Economy* 96 (April 1988), pp. 339–357.

- Keane, Michael P., and David E. Runkle. "Testing the Rationality of Price Forecasts: New Evidence From Panel Data." *American Economic Review* 80 (1990), pp. 714–735.
- Leduc, Sylvain, Keith Sill, and Tom Stark. "Self-Fulfilling Expectations and the Inflation of the 1970s: Evidence from the Livingston Survey." *Journal of Monetary Economics* 54 (March 2007), pp. 433–459.
- Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers. "Disagreement about Inflation Expectations." NBER *Macroeconomics Annual* 2003, pp. 209–248.
- Mork, Knut Anton. "Ain't Behavin': Forecast Errors and Measurement Errors in Early GNP Estimates." *Journal of Business and Economic Statistics* 5 (April 1987), pp. 165–175.
- Mullineaux, Donald J. "On Testing for Rationality: Another Look at the Livingston Price Expectations Data." *Journal of Political Economy* 86 (April 1978), pp. 329–336.
- Mullineaux, Donald J. "Inflation Expectations and Money Growth in the United States." *American Economic Review* 70 (March 1980), pp. 149–161.
- Neumark, David, and Jonathan S. Leonard. "Inflation Expectations and the Structural Shift in Aggregate Labor-Cost Determination in the 1980's." *Journal of Money, Credit, and Banking* 25 (1993), pp. 786–800.
- Orphanides, Athanasios. "Historical Monetary Policy Analysis and the Taylor Rule." *Journal of Monetary Economics* 50 (July 2003), pp. 983–1022.
- Pearce, Douglas K. "Comparing Survey and Rational Measures of Expected Inflation: Forecast Performance and Interest Rate Effects." *Journal of Money, Credit, and Banking* 11 (November 1979), pp. 447–456.
- Sharpe, Steven A. "Reexamining Stock Valuation and Inflation: The Implications of Analysts' Earnings Forecasts," *Review of Economics and Statistics* 84 (November 2002), pp. 632–648.
- Su, Vincent, and Josephine Su. "An Evaluation of ASA/NBER Business Outlook Survey Forecasts." *Explorations in Economic Research* 2 (Fall 1975), pp. 588–618.
- Turnovsky, Stephen J. "Empirical Evidence on the Formation of Price Expectations." *Journal of the American Statistical Association* 65 (December 1970), pp. 1441–1459.
- Wilcox, James A. "Why Real Interest Rates Were So Low in the 1970's." *American Economic Review* 73 (March 1983), pp. 44–53.
- Zarnowitz, Victor, and Louis A. Lambros. "Consensus and Uncertainty in Economic Prediction." *Journal of Political Economy* 95 (June 1987), pp. 591–621.