The case for higher frequency inflation expectations

Guzman, Giselle C.

Economic Alchemy LLC

29 June 2010

The Case for Higher Frequency Inflation Expectations[1]

Giselle Guzmán, Ph.D., FRM

Economic Alchemy LLC

Draft, Preliminary and Incomplete

This Version: June 29, 2010

[1] A previous version of this paper circulated under the title "An Inflation Expectations Horserace" and was submitted in partial fulfillment of the Ph.D. degree in Finance and Economics at Columbia University. I would like to thank Lawrence Klein, Joseph Stiglitz, Marshall Blume, Marc Giannoni, Charles Jones, and Lawrence Glosten for their insights and guidance. Thanks also to Randy Moore, William Dunkelberg, Richard Curtin, Lynn Franco, and Dennis Jacobe for generously providing data. All equations estimated with EViews Version 6 for Windows PC by Quantitative Micro Software, LLC. Any errors are solely my own.

The Case for Higher Frequency Inflation Expectations

Giselle Guzmán, Ph.D.

Economic Alchemy LLC

Abstract:

I present evidence that higher frequency measures of inflation expectations outperform lower frequency measures of inflation expectations in tests of accuracy, predictive power, and rationality. For decades, the academic literature has focused on three survey measures of expected inflation: the Livingston Survey, the Survey of Professional Forecasters, and the Michigan Surveys of Consumers. While these measures have been useful in developing models of forecasting inflation, the data are low frequency measures that are anachronistic in the modern era of high frequency and real-time data. I present a collection of 37 different measures of inflation expectations, including many previously unexploited monthly and real-time measures of inflation expectations. These higher frequency measures tend to outperform the standard three low frequency survey measures in tests of accuracy, predictive power, and rationality, indicating that there are benefits to using higher frequency measures of inflation expectations. Out of sample forecasts confirm the findings.

The Case for Higher Frequency Inflation Expectations

I. Introduction

The importance of inflation expectations, for the real economy as well as for financial markets, cannot be overstated. Inflation expectations play a critical role in the Federal Reserve's determination of monetary policy and in establishing the Fed's credibility among market participants. Expectations of inflation are embedded in the investment and financing decisions of firms, the labor contract negotiations of managers and employees, and the consumption, investment, and savings decisions of individuals. For decades, economists have relied on a standard set of three survey measures of expected inflation, namely the semiannual Livingston Survey, the quarterly Survey of Professional Forecasters, and the quarterly Michigan Surveys of Consumers.[2] These three low frequency survey measures have been useful in developing models of inflation expectations formation, and in testing the rational expectations hypothesis.[3] Given the importance of inflation expectations, and the considerable attention the subject has received in the academic literature, it is somewhat surprising that economists have not endeavored to look beyond the standard set of three surveys to develop a more comprehensive set of measures to gauge inflation expectations. In particular, it seems odd that in a world driven by real-time information, economists are still relying on quarterly and semi-annual measures of inflation expectations, when higher frequency measures exist and are readily available.

In this paper, I introduce a collection of monthly and real-time measures of inflation expectations, and compare the performance of these higher frequency measures with the standard three quarterly and semi-annual surveys. I run a horserace between all the measures and compare their accuracy, predictive content, and rationality. The paper follows the spirit and methodology set forth in Thomas (1999), Grant and Thomas (1999), and Mehra (2002). I examine two types of measures – numerical forecasts of the inflation rate (survey-based and market-implied) and diffusion-style indexes (survey based) of the expected direction of inflation. The numerical forecasts are tested for accuracy by comparing summary statistics of the forecasting errors. A test of equal forecast accuracy is performed to evaluate competing numerical forecasts. The predictive content of both the numerical forecast and diffusion index inflation expectations measures are assessed with a test for Granger causality. Rationality for all measures is evaluated with tests for unbiasedness and efficiency. Out of sample tests of forecast accuracy are also conducted. By performing this analysis, I seek to answer the following questions:

[2] Data were quarterly prior to 1976 and monthly thereafter, but most researchers use only the quarterly data.
[3] Term-structure models, ARIMA time series models, and Phillips Curve motivated models of inflation expectations are important tools as well, but those are not emphasized here since the focus is mainly on survey expectations.

1) Are the higher frequency measures of inflation expectations accurate, predictive and rational compared to the standard three low-frequency surveys?

2) How does the out-of-sample forecasting performance of these higher frequency indicators compare to the low-frequency inflation expectations survey data?

3) Does it pay to venture beyond the status quo in terms of the economists' data set, or are economists correct in sticking with data that are tried and true?

The goal is to evaluate a set of unexploited measures of inflation expectations and determine if the academic literature has been correct in ignoring these measures, or if some of these measures could potentially replace or enhance the standard economists' data set on inflation expectations. The paper is organized as follows: Section II provides a brief review of related literature, Section III contains a description of the inflation expectations measures, Section IV describes the methodology, Section V presents the results, and Section VI concludes.

II. Literature Review

For decades, the academic literature has devoted significant efforts to developing and evaluating methods of forecasting inflation. In addition to other methods of forecasting inflation, a large body of literature has evolved on the subject of survey-based inflation expectations, with researchers debating and discussing the rationality, accuracy, and predictive power of these measures. The vast majority of these studies focus on three surveys: the Livingston Survey, the Michigan Survey, and the Survey of Professional Forecasters (SPF). Thomas (1999) examines consensus forecasts of economists from the Livingston Survey and households from the Michigan Survey, and finds that these surveys outperform benchmark forecasts generated by a naïve model of lagged inflation and by the Fisher relation. In addition, households outperform economists in tests of accuracy and unbiasedness. Grant and Thomas (1999) provide evidence that the Livingston and Michigan survey measures of expected inflation are cointegrated with actual inflation realizations, supporting weak-form rationality of these survey respondents. Mehra (2002) examines the accuracy, predictive content, and rationality of the Livingston, Michigan, and SPF surveys, and reports that Michigan outperforms Livingston and SPF.

The Phillips curve has long been a standard tool for economists in forecasting inflation. Stock and Watson (1999) present an authoritative study of Phillips curve models and find that inflation forecasts generated by the Phillips curve produce the most accurate and reliable forecasts over the 1970-1996 period, compared with inflation forecasts using other macroeconomic variables and economic indicators. In addition, the authors find that the best-performing Phillips curve specification is one that uses a new composite index of aggregate economic activity comprising 168 individual activity measures, including surveys.

Indeed, an extensive literature has evolved on empirical factor models that exploit information from large data sets to predict key economic quantities such as inflation. Stock and Watson (2002) show that, when compared to standard benchmark models such as autoregressive, leading indicator, Phillips curve, and vector autoregressive models, the best forecast of inflation is obtained from a model employing lagged inflation and a single composite factor, constructed from a large set of indicators, including surveys. Other researchers, such as Guzmán (2003) have demonstrated that composite factors extracted from large data sets that include surveys along with other macroeconomic indicators can be effectively used to forecast aggregate stock returns. Guzmán (2008) shows how a composite factor constructed from a collection of surveys can improve both nowcasts and forecasts of aggregate stock returns as well as GDP growth. Similarly, Giannone, Reichlin, and Small (2008) show that composite factors obtained from high-frequency macroeconomic indicators and soft information such as surveys can significantly improve both nowcasts and forecasts of GDP growth. Surveys are gaining credibility as an important economic forecasting tool.

Economists currently rely on four primary methods of forecasting inflation: time series ARIMA models, forecasting regressions using variables motivated by the Phillips curve, term structure models, and inflation expectations derived from surveys of households and economists. Presumably, those economists participating as survey respondents are using some variation of the three non-survey methods to forecast inflation. Ang, Bekaert, and Wei (2007) compare and contrast these four methods of inflation forecasting and find that surveys outperform the other three methods. Adjustments to account for linear and non-linear bias in the survey data produce worse out-of-sample forecasting results than using the unadjusted survey median forecasts. In addition, the authors investigate models of combined forecasts and find that surveys outperform other model combinations, and when combined with other forecasts, the data tend to overweight survey forecasts and underweight the other forecasting methods.

However, Ang et. al. (2007), like others before them, examine only the three standard low frequency surveys – the quarterly Michigan Survey, the quarterly Survey of Professional Forecasters, and the semiannual Livingston Survey. Because of the long tradition these surveys have in the academic literature, many researchers are mistakenly under the impression that these three surveys are the *only available* surveys containing data on inflation expectations. In fact, there are at least 36 different survey measures of U.S. inflation expectations, available from a variety of sources, covering a wide range of respondent universes – households, businesses, economists, investors, manufacturers, retailers, and others. I examine a total of 37 different measures of inflation expectations, subjecting these measures to a battery of tests for accuracy, predictive content, and rationality, following the methods set forth in Thomas (1999), Grant and Thomas (1999), and Mehra (2002).[4]

[4] I examine 37 different measures of inflation expectations in this paper, but only 36 are survey-based measures, as the TIPS spread is a market-implied measure.

III. Description of Inflation Expectations Measures

In total, I examine 37 measures of inflation expectations from 14 sources. There are 36 survey measures and one market-implied measure. Table 1 contains the complete list and description of the indicators. A brief description of each data source follows, with an indication of the respondent universe and whether the measure is a diffusion-style index (D) or a numerical forecast (N).

[INSERT DATA DESCRIPTIONS]

III. Methodology

III.a. Accuracy

In order to test the inflation expectations measures for accuracy, I calculate and compare forecast errors over the full sample period for each numerical forecast. The forecast error e_i is calculated as the forecast inflation rate minus the actual inflation rate that subsequently occurred. One complication is that many of the surveys and other measures do not specify which rate of inflation is being forecast; they simply ask about changes in price levels, or about a general concept of inflation. In addition, the respondents are consumers, investors, economists, businesses, retailers, or manufacturers, and the definition of inflation surely varies depending on the profile of the respondent. Due to this vagueness, I calculate the errors comparing the forecasted rate with the actual rate of the Consumer Price Index (CPI), the Personal Consumption Expenditures deflator (PCE), and the Producer Price Index (PPI). A description of the inflation measures and data sources is contained in Appendix A. I use the forecast errors to identify the best actual rate of inflation that is being forecast by the inflation expectation measure, as well as the best horizon if no horizon is specified in the survey question.

I calculate three summary statistics of the forecast errors for each indicator: the mean error (ME), mean absolute error (MAE), and the root mean square error (RMSE). The mean error can be interpreted as a basic measure of forecasting bias, and represents the average magnitude of the forecast error over the n periods being forecast. A positive mean error indicates a propensity to overestimate inflation; whereas a negative mean error indicates a propensity to underestimate inflation. The mean absolute error measures the accuracy of forecasts, as does the root mean square error. However, the RMSE amplifies the effect of large forecast errors.

The ME, MAE, and RMSE are calculated in the standard fashion, as follows:

[INSERT EQUATION]

III.b. Forecast Comparison Tests

With measures of inflation expectations from so many different sources, it is inevitable that there will be apparent differences in forecast accuracy within the sample. This raises the question as to whether the outcome is due to pure chance. A test of equal predictive accuracy is performed to determine whether these observed differences are statistically significant or not.

Since the information set is limited, i.e., available data only include a set of forecasts and actual values of the predictand, a model-free test is appropriate. I employ a variant of the Morgan-Granger-Newbold (1977) (MGN) test, proposed by Harvey, Leybourne, and Newbold (1997) (HLN). The test will allow an objective evaluation of the forecast accuracy of each of the numerical forecasts and determine whether the observed differences are due to chance or due to superior forecasting ability. The methodology is described as follows.

[INSERT EQUATION]

III.c. Predictive Power

Predictive content is measured by a test of Granger Causality. This test evaluates the possibility that inflation expectations and inflation realizations may be co-integrated, in the sense of Engle and Granger (1987). One would expect that actual inflation rates may influence inflation expectations. But, if inflation expectations influence actual future rates of inflation, this would be of significant interest to policymakers, as it implies a bilateral feedback effect between inflation and inflation expectations.

The tests for Granger Causality are specified as follows:

[INSERT EQUATION]

where π_t is the actual rate of inflation and π^e_t is the expected rate of inflation, and ε_{it} is a white noise error. The null hypothesis is $\lambda_\pi = 0$ and $\lambda_{\pi e}=0$. If both λ_π and $\lambda_{\pi e}$ are significantly different from zero then forecasters respond to the behavior of inflation, and in addition, inflation responds to the behavior of forecasters. This is a fundamental proposition of the rational expectations paradigm.[5]

III.d. Rationality – Unbiasedness and Efficiency

According to Thomas (1999), "If inflation expectations are fully rational, they should exhibit two fundamental characteristics. First, they should be unbiased – that is, agents should forecast inflation correctly on average. Second, forecasts should be efficient – that is, agents should

[5] Grant and Thomas (1999).

employ all relevant information for which the marginal benefit of gathering and utilizing the information exceeds the marginal cost."

The test for bias is estimated by OLS and specified as follows:[6]

[INSERT EQUATION]

The equation is estimated by regressing the actual inflation rate π_t on the previously made forecast of inflation π^e_t and testing the joint null hypothesis that $\alpha=0$ and $\beta=1$. Forecasts are considered unbiased if the null hypothesis cannot be rejected. The joint null hypothesis is tested with a Chi-squared test. The Chi-squared test only applies to the numerical forecasts, since the hypothesis that $\beta=1$ would be meaningless for a diffusion index.

The test for efficiency is estimated by OLS and specified as follows:[7]

[INSERT EQUATION]

The equation is estimated by regressing the forecast error e_t on the information set I_t either individually or jointly. The information set includes those variables that are pertinent to a comprehensive model of inflation. The variables are tested for significance first individually, then jointly. If any or all of the variables in the information set are significantly negatively correlated with the forecast error, this implies that agents failed to take all relevant information into account when forming their inflation expectations. Weak-form efficiency implies that agents have taken into consideration only the information contained in past inflation rates, while strong-form efficiency implies that agents have considered information about all variables that are germane to forecasting inflation.

Following Thomas (1999) and Mehra (2002), the variables employed in the information set I_t are: the lagged 12-month rate of CPI inflation, a measure for the output gap, M1 and M2 growth, and a measure for oil price inflation. Since most of the data have a monthly frequency, the unemployment rate is used as a proxy for the output gap, with this substitution following Gramlich (1983). The measure for oil price inflation is the lagged 12-month rate of change for the producer price index for fuels. A description of the variables and data sources is contained in Appendix A.

III.e. Out of Sample Forecasts

[6] Model is estimated by OLS with Newey-West HAC standard errors with lag truncation parameter set to equal forecast horizon in order to avoid overlapping standard errors.

[7] Model is estimated by OLS with Newey-West HAC standard errors with lag truncation parameter set to equal forecast horizon in order to avoid overlapping standard errors.

Out of sample forecasts are performed using a basic predictive model for actual inflation regressed on expected inflation and past inflation. Due to the high serial correlation in the rate of inflation, the model is specified to test whether the survey forecasts have any predictive power for the future rate of inflation beyond the information contained in past inflation data. The model is estimated by OLS as follows:[8]

[INSERT EQUATION]

A static forecast is produced by estimating parameters using data available through December 2005. The estimated parameters are then used to fit the equation over the out-of-sample period, January 2006 to October 2008. For the five-year inflation forecasts, parameters are estimated with data through September 2003 and the out-of-sample period is October 2003 to October 2008. The forecasts are then evaluated by comparing the Root Mean Squared Errors to determine the accuracy of the forecasts.

IV. Results

IV.a. Accuracy

Accuracy is evaluated for numerical forecasts only. Table 2 presents the results for the accuracy test using the CPI as the actual inflation rate. The inflation forecast with the lowest RMSE is the Michigan Median 5-year inflation forecast, with RMSE = 0.7939. The inflation forecast with the highest RMSE is the Livingston Mean PPI forecast, with RMSE = 2.9680.

Table 3 presents the results for the accuracy test using the PCE as the actual inflation rate. The inflation forecast with the lowest RMSE is again the Michigan Median 5-year inflation forecast, with RMSE = 0.7668. The inflation forecast with the highest RMSE is the Conference Board inflation forecast, with RMSE = 2.6679. The PCE results are of particular interest given that the PCE deflator is frequently the preferred inflation indicator used by the Federal Reserve in conducting monetary policy.

Table 4 presents the results for the accuracy test using the PPI as the actual inflation rate. The inflation forecast with the lowest RMSE is once again the Michigan Median 5-year inflation forecast, with RMSE = 1.8169. The inflation forecast with the highest RMSE is again the Conference Board inflation forecast, with RMSE = 3.4189. Notice that the RMSE of the Livingston Survey forecasts for mean and median PPI are 3.1096 and 3.1110, respectively, when forecasting the PPI. However, the same Livingston Survey forecasts for mean and median PPI have RMSEs of 2.9680 and 2.9566, respectively, when forecasting the CPI, and 2.0144 and

[8] Model is estimated by OLS with Newey-West HAC standard errors with lag truncation parameter set to equal forecast horizon in order to avoid overlapping standard errors.

2.0405, respectively, when used to forecast the PCE. Thus, the Livingston Survey forecasts of the PPI are better predictors of the CPI and the PCE than they are for the PPI.

IV.b. Forecast Comparison

The HLN (1997) variation of the MGN (1977) test is performed to test for equal forecasting accuracy, i.e., equality of forecast error variances. Table 5 presents results of the HLN test of numerical forecasts for the CPI over a 12-month horizon. The benchmark selection rule indicates that the Michigan Survey's 1-year median inflation forecast is the benchmark measure. The null hypothesis of $\beta = 0$ can be decisively rejected for all the measures of inflation expectations except for the Survey of Professional Forecasters. This means that the null hypothesis of equal forecasting accuracy is rejected for the majority of the measures.

Table 6 presents results of the HLN test of numerical forecasts for the PCE over a 12-month horizon. The benchmark selection rule indicates that the Michigan Survey's 1-year median inflation forecast is again the benchmark measure. The null hypothesis of $\beta = 0$ is rejected for the Blue Chip GDP Deflator, the Blue Chip CPI, the TIPS Spread, Michigan 1-year mean, and the Conference Board 1-year inflation forecasts. The null hypothesis fails to be rejected for the SPF 1-year CPI forecast and all of the Livingston forecasts.

Table 7 presents results of the HLN test of numerical forecasts for the PPI over a 12-month horizon. The benchmark selection rule indicates that the Blue Chip 1-year CPI inflation forecast is the benchmark measure. The null hypothesis of $\beta = 0$ is rejected for the Blue Chip GDP Deflator, the Michigan 1-year median, Michigan 1-year mean, and the Conference Board 1-year inflation forecasts. The null hypothesis fails to be rejected for the SPF 1-year CPI forecast, the TIPS Spread, and all of the Livingston forecasts.

IV.c. Predictive Power

It is natural to expect that inflation expectations would be influenced by the past actual inflation rate. However, if inflation expectations influence the future actual inflation rate, then this would be of interest to policymakers and investors alike. Tables 8 and 9 present results for the test of predictive power, using a Granger Causality test at 3 and 12 lags, respectively. The null hypothesis for Equations (4) and (5) is that the actual inflation rate does not Granger Cause inflation expectations and inflation expectations do not Granger Cause the actual inflation rate.

Table 8 shows that, at 3 lags, the actual inflation rate influences most of the measures of inflation expectations, and this is not a surprise, as one would expect agents to form expectations based in part on recent past experience. What is intriguing is that several of the measures of inflation expectations influence the future actual inflation rate. In this case the null hypothesis for the absence of Granger Causality is rejected. Significant predictive power is demonstrated by the following measures of inflation expectations: Small Business 3-month Price Plans, Richmond Fed 6-month Retail Prices, Michigan Vehicles Price Conditions, Blue Chip 1-year CPI forecast, Survey of Professional Forecasters 1-year CPI forecast, Livingston 1-year Median CPI forecast and the Michigan 5-year mean inflation forecast. Since many of these indicators are available at a monthly frequency, there is a clear advantage to using them instead of or in addition to the quarterly and semi-annual frequency measures.

Table 9 shows that, at 12 lags, the actual inflation rate once again influences many of the measures of inflation expectations. In addition, several measures of inflation expectations demonstrate predictive power over the actual future inflation rate. The Livingston 6-month mean PPI forecast, the Blue Chip 1-year CPI forecast, the Survey of Professional Forecasters 1-year CPI forecast and the Michigan median 1-year inflation forecast all demonstrate a statistically significant ability to anticipate the future actual inflation rate.

The results of the Granger Causality tests lend support to some alternative theoretical macroeconomic models. For instance, the finding that inflation expectations of businesses and retailers Granger cause future inflation rates makes sense to the extent that there may exist strategic complementarities between the price-setting decisions of manufacturers or suppliers of different goods, in the sense suggested by Calvo (1983). This theory of pricing can justify an aggregate supply relation that takes the form of an expectations-augmented Phillips curve relation, where the location of the short-run Phillips curve is determined by expectations regarding future inflation. Indeed, in many macroeconomic models of the New Keynesian variety, current inflation is mainly determined by current expectations of future inflation. This is because price-setters will optimally adjust their prices such that current prices reflect a mark-up above their expected average nominal marginal costs for the duration that prices are expected to remain fixed. Therefore, expected future inflation will affect current inflation because current prices are aligned with average expected future nominal marginal costs. Thus, inflation expectations can lead to self-fulfilling deflations or inflations, i.e., there is convergence to a rational-expectations equilibrium as a result of adaptive learning dynamics.[9]

Alternatively, the results could be explained by a sticky information model as proposed by Mankiw and Reis (2002), rather than the sticky prices underlying the New Keynesian models. In the sticky information model, current inflation is determined by past expectations of current

[9] Woodford (2003)

inflation.[10] Some researchers, most notably Carroll (2003) and Lanne, Luoma, and Luoto (2009), argue that the inflation expectations data from the Michigan Survey is consistent with a sticky information model and that agents are slow to update their beliefs, thus providing the microfoundations for the model proposed by Mankiw and Reis.

Finally, another alternative for the Granger Causality results could be that the apparent cointegration could be explained by a common shock affecting both current and future inflation. For example, even if actual inflation and expected inflation are unrelated, a commodity price shock could induce a revision of today's expectations of inflation one year from now, and also affect inflation every month from now on. While this explanation is possible, it is not probable due to the fact that many of the sample periods occur over a time span during which there was no major commodity price shock.

IV.d. Rationality – Unbiasedness and Efficiency

Table 10 contains the results of the test for unbiasedness, where the joint null hypothesis $\alpha=0$ and $\beta=1$ is tested for Equation (6), for the 17 numerical forecasts of inflation expectations.[11] The results of the Chi-squared tests indicate that the null hypothesis is decisively rejected for 16 of the 17 numerical forecasts. This means that each of these 16 indicators systematically either overestimate or underestimate the actual inflation rate. The only measure of inflation expectations where the null hypothesis fails to be rejected is the Blue Chip Indicators Survey one-year forecast for CPI inflation. In this case, the Chi-squared p-value is 0.3093, and we fail to reject the joint null hypothesis $\alpha=0$ and $\beta=1$.

Tables 11 through 16 present results for the tests for efficiency, to find out if agents employed relevant information in forming inflation expectations. In this test, forecast errors are regressed on inflation-related variables to determine if there is a correlation. The variables are first tested separately and then together in a joint specification. Testing whether agents used knowledge of lagged inflation in forming expectations is a test of weak-form efficiency. To test for strong-form efficiency, four variables were tested: the unemployment rate (a substitute measure for the output gap), the lagged 12-month growth rate of the narrow (M1) and broad (M2) monetary aggregates, and a measure for energy price inflation (the 12-month rate of change of the producer price index for fuels). In each case, the independent variable is defined so that failure of agents to take account of the variable in the manner suggested by conventional economic theory would result in a negative and statistically significant coefficient on the variable.[12] That is, if agents fail to account for past inflation, money growth, etc., they would underestimate inflation and have a

[10] In a sense, the sticky information model is like a Phillips curve with backward-looking expectations instead or forward-looking expectations.

[11] The Chi-squared test is not applicable to the 20 diffusion indexes of inflation expectations.

[12] Thomas (1999)

negative forecasting error, resulting in a negative sign on the coefficient for the variable. Conversely, if agents take too much account for past inflation, money growth, etc., they would overestimate inflation and have a positive forecasting error, resulting in a positive sign on the coefficient for the variable.

Table 11 contains the results for the efficiency test with respect to the most recent 12-month rate of CPI inflation known to agents at the time the inflation expectations are measured. The table indicates that most agents adequately took into account the past rate of CPI inflation, but respondents to some surveys did not. Specifically, the Small Business Price Plans, the Philadelphia Fed's, Dallas Fed's, New York Fed's, and Kansas City Fed's expectations for Prices Paid, and the Livingston Mean and Median CPI forecasts all failed to consider adequately the lagged inflation rate in forming inflation expectations. Due to the insufficient use of information concerning the past inflation rate, weak-form efficiency can be rejected for these measures of inflation expectations.

Conversely, Table 11 indicates that some measures of inflation expectations attributed too much influence to the past CPI inflation rate, resulting in a positive forecasting error. The Richmond Fed's survey expectations for 6-month prices paid, prices received, retail prices, non-retail prices, and services prices all have a positive and statistically significant coefficient on lagged CPI. Similarly, the Blue Chip 1-year GDP deflator and 1-year CPI forecasts, the Michigan 1-year median and 5-year mean and median inflation forecasts, the Survey of Professional Forecasters 1-year CPI, and the Livingston survey's 6-month mean and median CPI and PPI forecasts all have forecast errors that are positively correlated with the lagged inflation rate. This indicates that respondents to these surveys overestimated the impact of past inflation when forming their expectations for future inflation.

Table 11 indicates that weak-form efficiency is supported for the majority of the inflation expectations measures. The Philadelphia Fed 6-month prices received, Richmond Fed 6-month wages, Kansas City Fed 6-months prices received, New York Fed 6-months prices received, Dallas Fed 6-months prices received, and Dallas Fed 6-month wages, are all measures of inflation expectations that adequately took account of the lagged CPI inflation rate. The same is true for the Michigan Survey's price conditions for durable goods, vehicles, and housing. In addition, the UBS/Gallup 1-year inflation forecast, the Michigan 1-year mean inflation forecasts, the TIPS spread, the Conference Board 1-year inflation forecast, and the Livingston survey's 1-year mean and median PPI forecasts are also weak-form efficient measures of inflation expectations.

Table 12 presents the results of the efficiency test with respect to the lagged 12-month growth rate of the narrow monetary aggregate (M1). Expectations for the Blue Chip 1-year GDP deflator, Michigan 1-year median inflation and Livingston 1-year mean and median CPI forecasts all have negative and statistically significant coefficients, meaning that they fail to take sufficient account of M1 growth. Because these survey measures failed to take adequate account of M1 growth, strong-form efficiency can be rejected for these measures of inflation expectations. Conversely, the Kansas City Fed's 6-month prices received, New York Fed's 6-month prices paid, and the Michigan 5-year mean and median inflation forecasts all have forecast errors that are positively correlated with lagged M1 growth, indicating that respondents to these surveys overestimated the influence of lagged M1 growth when forming their inflation expectations.

Strong-form efficiency with respect to lagged M1 growth is supported for several measures of inflation expectations. The Small Business 3-month price plans, the Philadelphia Fed's 6-month prices paid and prices received, the Richmond Fed's 6-month prices paid, prices received, retail prices, non-retail prices, services prices, and wages all take into account the lagged growth rate of the narrow monetary aggregate. The Kansas City Fed's 6-month prices paid, the New York Fed's 6-month prices received, the Dallas Fed's 6-month prices paid, prices received, and wages, and the Livingston Survey's 6-month mean and median CPI and PPI also efficiently incorporate information about lagged M1 growth, as do the Michigan Survey's price conditions for durable goods, vehicles, and housing. The UBS/Gallup 1-year inflation forecast, the Michigan 1-year mean inflation, the Blue Chip 1-year CPI forecast, the TIPS spread, the Conference Board 1-year inflation forecast, the SPF 1-year CPI forecast, and the Livingston survey mean and median 1-year PPI forecasts adequately take M1 growth into account as well. Strong-form efficiency is supported for all these measures of inflation expectations.

Table 13 contains the results of the efficiency test with respect to the lagged 12-month growth rate of the broad monetary aggregate (M2). The Richmond Fed Survey's expectations for 6-month prices paid and prices received, the Livingston Survey's 1-year mean and median CPI expectations, and the Michigan 5-year mean and median inflation forecasts all fail to take proper account of the lagged 12-month growth rate of the broad monetary aggregate, as indicated by the significant negative correlation between the forecasting error and lagged M2 growth. Because these survey measures failed to take adequate account of M2 growth, strong-form efficiency can be rejected for these measures of inflation expectations. Conversely, the Philadelphia Federal Reserve's 6-month prices received and the Richmond Fed's 6-month Retail prices, as well as the Michigan Survey's durable goods and housing price conditions, and the UBS/Gallup 1-year inflation forecast are measures of inflation expectations with forecast errors that are positively significantly correlated with M2 growth, suggesting that forecasters attributed too much influence of M2 growth on the future inflation rate.

Strong-form efficiency with respect to M2 growth is supported for several of the measures of inflation expectations. The Small Business price plans, Philadelphia Fed's 6-month prices paid, Richmond Fed's 6-month non-retail prices, services prices, and wages, Kansas City Fed's 6-month prices paid and prices received, New York Fed's 6-month prices paid and prices received, and the Dallas Fed's 6-month prices paid, prices received, and wages all efficiently incorporate information about M2 growth,, thus exhibiting strong-form efficiency. Similarly, the Michigan Survey's price conditions for vehicles, the Blue Chip Survey's 1-year forecast for the GDP deflator and CPI, the Michigan Survey 1-year mean and median inflation forecast, the TIPS spread, the Conference Board Survey's 1-year inflation forecast, the Survey of Professional Forecasters 1-year CPI forecast, and the Livingston Survey's mean and median 6-month CPI and PPI, and mean and median 1-year PPI forecasts are also strong-form efficient with respect to M2 growth.

The results for the efficiency test with respect to oil price inflation are displayed in Table 14. Survey expectations for 6-month prices paid from neither the Philadelphia Fed, nor the Kansas City Fed, nor the Dallas Fed adequately took into account the lagged oil price inflation, as indicated by the negative and statistically significant coefficient. Due to the inadequate use of information concerning energy price inflation, strong-form efficiency can be rejected for these survey measures of inflation expectations. Conversely, a positive correlation between oil price inflation and the forecast error is noted for the Richmond Fed's 6-month prices paid, prices received, retail, non-retail, and services prices, the New York Fed's 6-month prices received, and the Livingston Survey's 6-month mean and median CPI forecasts, indicating that these measures of inflation expectations attributed too much importance to oil price inflation in forming expectations for future inflation.

Several of the measures demonstrate strong-form efficiency with respect to oil price inflation. The Small Business price plans, Philadelphia Fed's 6-month prices received, Richmond Fed's 6-month wage expectations, Kansas City Fed's 6-month prices received, New York Fed's 6-month prices paid, and the Dallas Fed's 6-month prices received and wage expectations all adequately took account of oil price inflation in forming expectations for future inflation. The Michigan Survey's price conditions for durable goods, vehicles, and housing prices, as well as the mean and median 1-year and 5-year inflation forecasts, also sufficiently incorporate information regarding oil price inflation, thereby exhibiting strong-form efficiency. The Blue Chip 1-year GDP deflator and CPI forecasts, the UBS/Gallup 1-year inflation forecast, the TIPS spread, the Conference Board 1-year inflation forecast, the SPF 1-year CPI forecast, and the Livingston Survey's mean and median 6-month PPI, and 1-year PPI and CPI forecasts are also strong-form efficient with respect to oil price inflation.

Table 15 presents the results for the efficiency test with respect to the lagged unemployment rate, a proxy for the output gap. The table indicates that none of the Richmond Fed's measures of inflation expectations for the services sector effectively incorporates information about the unemployment rate. The Richmond Federal Reserve Surveys of expectations for retail prices, non-retail prices, and service sector prices all have a negative and statistically significant coefficient on the lagged unemployment rate. Due to the inadequate use of information concerning the unemployment rate, strong-form efficiency can be rejected for the Richmond Fed's service sector surveys. Conversely, Small Business price plans, the Blue Chip 1-year CPI forecast, the Survey of Professional Forecasters 1-year CPI forecast, and the Michigan 5-year mean and median inflation forecasts, all have forecast errors that are positively correlated with the unemployment rate, suggesting that forecasters attributed too much influence from the unemployment rate on their forecasts of future inflation.

Strong-form efficiency with respect to the unemployment rate is indicated for several of the measures. The Philadelphia, Kansas City, and New York Fed's 6-month prices paid and prices received, and the Dallas and Richmond Fed's 6-month prices paid, prices received, and wages all efficiently incorporated information about the unemployment rate in forming inflation expectations. Additionally, the Michigan survey's price conditions for durable goods, vehicles, housing, 1-year mean and median inflation forecasts, the Blue Chip 1-year GDP deflator forecast, the UBS/Gallup 1-year inflation forecast, the TIPS spread, the Conference Board 1-year inflation forecast, and the Livingston Survey's mean and median 6-month and 1-year CPI and PPI forecasts also display strong-form efficiency with respect to the lagged unemployment rate.

Table 16 presents the results for the efficiency test using the joint specification, with the lagged CPI, M1 and M2 growth, oil price inflation, and unemployment rate tested together. The table indicates that most of the measures do not efficiently incorporate information from all of these variables simultaneously, refuting strong-form efficiency. Note that only the Conference Board Survey 1-year inflation expectations and the Michigan Survey median 1-year inflation expectations pass the joint specification test with statistical significance, indicating strong-form efficiency for these two survey measures.

IV.e. Out of Sample Forecasts

Table 17 presents results for out-of-sample forecasts using a basic predictive model for actual inflation regressed on expected inflation and past inflation.[13] The table reveals that the most accurate out of sample forecast is given by the Philadelphia Fed's 6-month Prices Received index, with RMSE = 1.1989. The standard economists' data set does not perform as well, with the SPF 1-year CPI forecast registering a RMSE of 1.5430, the Michigan Mean 1-year inflation forecast registering a RMSE of 2.5405 and the Livingston mean 1-year CPI forecast registering a RMSE of 3.4935. Most of the monthly measures of inflation expectations outperform the standard quarterly and semiannual survey measures, indicating that there are benefits to using higher frequency data.

V. Conclusion

I have shown that the higher frequency survey measures of inflation expectations tend to outperform the standard three low frequency surveys – the quarterly Michigan Survey, the quarterly Surveys of Professional Forecasters, and the semiannual Livingston Survey – in terms

[13] Out-of-sample tests for the Michigan 5-year forecasts cannot be analyzed due to insufficient data.

of accuracy, predictive power, rationality, and out-of-sample forecasts. While there is no single winner that consistently outperforms all of the other measures on the complete battery of tests, the results indicate that several of the surveys conducted by the regional Federal Reserve banks perform well, as do the Small Business Survey, the Conference Board Survey, the Blue Chip Survey and the TIPS spread. It is worth noting that the Blue Chip survey is the only indicator that passes the test for unbiasedness.

What is interesting is that many of the surveys that are not typically used in the academic literature perform better relative to those that are typically used. In particular, given that other authors have found that inflation forecasts from the standard three low frequency surveys outperform inflation forecasts generated by time series ARIMA models, regression models using Phillips curve-derived real activity measures, and term-structure models, then by the transitive property, since the higher frequency surveys examined in this paper outperform the standard three low frequency surveys, we can surmise that the higher frequency surveys would likely outperform inflation forecasts generated from the aforementioned other methods as well.

More research is needed to understand better the efficacy of these higher frequency measures of inflation expectations to determine if they should replace or enhance the standard three low frequency survey measures. There are many obvious benefits to using monthly or real-time measures versus quarterly or semiannual data for forecasters who wish to have their models reflect the most up-to-date information possible. The academic literature has been myopic in ignoring the availability of these higher frequency measures of inflation expectations.

References

Ang, A., Bekaert, G., and Wei, M. (2007). Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics*, 54 (2007), pp. 1163-1212.

Calvo, G. (1983). Staggered Prices in a Utility-Maximizing Framework. *Journal of Monetary Economics*, 12 (1983), pp. 383-398.

Carroll, C. (2003). Macroeconomic Expectations of Households and Professional Forecasters. *The Quarterly Journal of Economics*, Vol. 118, No. 1 (Feb. 2003), pp. 269-298.

Engle, Robert F., and Granger, C.W.J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, Vol. 55, March, pp. 251-276.

Giannone, D., Reichlin, L., and Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55 (2008), pp. 665-676.

Gramlich, E.M. (1983). Models of Inflation Expectations Formation: A Comparison of Household and Economist Forecasts. *Journal of Money, Credit, and Banking*, Vol. 15, No. 2, May 1983, pp. 155-173.

Granger, C.W.J., and Newbold, P. (1977). *Forecasting Economic Time Series*, Orlando, FL. Academic Press.

Grant, A.P. and Thomas, L.B. (1999). Inflationary expectations and rationality revisited. *Economics Letters*, Vol. 62, pp. 331-338.

Guzmán, Giselle (2003). GDP Growth Signals, Investor Psychology, and Hedging Pressure: A Multivariate Approach to Forecasting Returns on the S&P500 Index. *Manuscript,* The Wharton School, University of Pennsylvania, April, 2003.

Guzmán, Giselle (2008). Using Sentiment Surveys to Predict GDP Growth and Stock Returns. *The Making of National Economic Forecasts*, edited by Lawrence R. Klein. Edward Elgar Publishing, Ltd. London, 2009.

Harvey, D., Leybourne, S., and Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13, (1997), pp. 281-291.

Lanne, M., Luoma, A. and Luoto, J. (2009). A naïve sticky information model of households' inflation expectations. *Journal of Economic Dynamics & Control*, 33 (2009), pp. 1332-1344.

Mankiw, G. and Reis, R. (2002). Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. *The Quarterly Journal of Economics*, Vol. 117, No. 4 (Nov. 2002), pp. 1295-1328.

Mariano, R. and Preve, D. (2008). *Statistical Tests for Multiple Forecast Comparison.* Presentation, T.W. Anderson Conference, Stanford University, 2008.

Mehra, Y.P. (2002). Survey Measures of Expected Inflation: Revisiting the Issues of Predictive Content and Rationality. *Federal Reserve Bank of Richmond Economic Quarterly*, Vol. 88/3, Summer 2002, pp. 17-35.

Morgan, W.A. (1939). A Test for Significance of the Difference Between the Two Variances in a Sample From a Normal Bivariate Population. *Biometrika*, 31, 13-19.

Thomas, L.B. (1999). Survey Measures of Expected U.S. Inflation. *The Journal of Economic Perspectives*, Vol. 13, No. 4, Autumn 1999, pp. 125-144.

Stock, J.H. and Watson, M.W. (1999). Forecasting inflation. *Journal of Monetary Economics*, 44 (1999), pp. 293-335.

Stock, J.H. and Watson, M.W. (2002). Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business & Economic Statistics.* Vol. 20, No. 2 (April 2002), pp. 147-162.

Woodford, M. (2003). *Interest and Prices.* Princeton University Press, Princeton, NJ.

Appendix A. Data List

[INSERT DATA LIST]

[INSERT TABLES]