Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

Understanding survey based inflation expectations

Travis J. Berge

2017-046

Please cite this paper as:
Berge, Travis J. (2017). "Understanding survey based inflation expectations," Finance and Economics Discussion Series 2017-046. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2017.046.

April 2017

Understanding survey based inflation expectations[*]

Abstract

Survey based measures of inflation expectations are not informationally efficient yet carry important information about future inflation. This paper explores the economic significance of informational inefficiencies of survey expectations. A model selection algorithm is applied to the inflation expectations of households and professionals using a large panel of macroeconomic data. The expectations of professionals are best described by different indicators than the expectations of households. A forecast experiment finds that it is difficult to exploit informational inefficiencies to improve inflation forecasts, suggesting that the economic cost of the surveys' deviation from rationality is not large.

- *Keywords*: Survey based inflation expectations; informational inefficiency; inflation forecasting; Phillips curve; boosting.
- *JEL Codes*: C53, E31, E37.

Travis J. Berge
Board of Governors of the Federal Reserve
20th and C St NW
Washington, DC 20551
e-mail: travis.j.berge@frb.gov

---

[*]The views expressed herein do not necessarily reflect those of the Federal Reserve Board or Federal Reserve System.

1 Introduction

Understanding expectations of the public is critically important to policymakers. Inflation expectations in particular carry important information about future realized inflation, provide real-time feedback about the prevailing real interest rate, and may elucidate the public's understanding of a central bank's inflation target. Yet despite their central importance, little is known about how inflation expectations are formed.

In the United States, several measures of inflation expectations are readily available. Market-based inflation expectations can be computed by comparing nominal interest rates to their inflation protected counterparts, or by modeling the yield curve (Gurkaynak, Sack & Wright, 2010; D'Amico, Kim & Wei, 2014). Perhaps the simplest strategy, however, is the use of surveys, which do not require additional modeling assumptions (in order to adjust for liquidity or risk premiums, e.g.), and provide timely real-time snapshots of the evolution of inflation expectations.

Survey based inflation expectations carry important economic information. Kiley (2009) and Clark & Davig (2011) argue that recent stability in inflation expectations is due to the systematic implementation of monetary policy, suggesting that surveys reflect, in part, the public's understanding of monetary policy. In addition, inflation expectations ought to play an important role in the price setting decisions of firms, and therefore realized inflation. A large empirical literature supports this supposition. Ball & Mazumder (2011) and Coibion & Gorodnichenko (2015*b*) argue that the behavior of inflation expectations during and immediately following the Great Recession can help explain the lack of a severe disinflation during this time. Ang, Bekaert & Wei (2007) and Faust & Wright (2013) find that raw surveys of inflation expectations often outperform empirical models when forecasting U.S. inflation; Grothe & Meyler (2015) find similar results for the Survey of Professional Forecasters in Europe. Groen, Paap & Ravazzolo (2013) use Bayesian Model Selection to select models to be used for forecasting purposes, and household inflation expectations are an important covariate. The forecast ability of survey based inflation expectations may exist since expectations serve as a proxy for the slow-moving trend in inflation, for example, Kozicki & Tinsley (2001, 2005), Cogley & Sbordone, (2008) and Clark & Doh (2014).

Surveys have also been used to help identify the New Keynesian Phillips curve. Roberts (1995), Roberts (1998), and Adam & Padula (2011) present estimates of the NKPC using survey measures as proxies for expected inflation. Del Negro & Eusepi (2011) show that inflation expectations from the SPF contain information about future inflation even after accounting for the model-consistent expectations that are embedded into a modern DSGE model. Similarly, Fuhrer (2012) finds that once one conditions on the SPF forecast of year-ahead CPI inflation, model-implied expectations carry little information regarding inflation dynamics.

While surveys are of critical importance, expectations as measured by surveys are not strictly rational. Thomas (1999), Mehra (2002) and Croushore (2010) document that survey measures often violate basic implications of rationality. Mankiw, Reis & Wolfers (2004) document notable disagreement across surveys and the persistence of forecast errors thereby. Forecast errors from respondents to the Michigan Survey of Consumers are also predictable, based on, for example, age cohort, sex, and movements in recent prices (Johannsen, 2014, Ehrmann, Pfajfar & Santoro, 2015; Malmendier & Nagel, 2016; Binder, 2016b; Schulhofer-Wohl & Kaplan, 2016). Coibion & Gorodnichenko (2012, 2015) document that forecast errors of the Survey of Professional Forecasters are predictable in the sense that they under react to incoming information, and utilize this predictability to differentiate models that deviate from pure rational expectations.

The economic literature has developed a number of models that account for imperfect information and the resulting deviations from rationality. In one class of models, the acquisition or processing of news is costly. As a result, agents update their expectations infrequently and expectations are slow-moving (Mankiw & Reis 2002). Alternatively, agents may update their beliefs continuously but receive imperfect information about the economy, as in Sims (2003). Since agents cannot perfectly extract the true state of the economy, expectations respond to news sluggishly. Carroll (2003) develops a different model, wherein households update their expectations towards those of professionals.

Yet important questions remain. Coibion & Gorodnichenko's results indicate that models of information rigidity cannot account for all of the deviation from rationality present in surveys. Further, the literature has not addressed whether deviations from rationality are economically meaningful. The presence of informational inefficiencies, alongside the fact that inflation expectations matter for future realized inflation, suggest that informational inefficiencies may be exploited to improve forecasts of realized inflation. Therefore, a model selection algorithm is used to produce models of inflation expectations for both households and professionals. With a model of expectations in-hand, a forecast experiment then attempts to exploit predictability the surveys to improve out-of-sample forecasts of realized inflation.

The approach outlined above produces several interesting results. After revisiting the evidence that survey expectations fail simple tests of informational efficiency, I find that the macroeconomic variables that are most highly correlated to survey based inflation expectations are different for households and professional forecasters. The inflation expectations of households correlate with particular subcomponents of the Consumer Price Index, food and energy prices. In contrast, inflation expectations from professionals depend importantly on interest rates, and to a lesser extent, other subcomponents of the CPI. In general, however, it is difficult to exploit any informational inefficiency when forecasting future inflation. Adjusting household expectations for biasedness improves their forecast ability by about 30 percent. In contrast, bias adjustment does not notably improve the forecast performance of the expectations of professionals. Once a forecast model includes additional covariates, for example, measures of slack as in a Phillips curve relationship, bias-adjusting expectations does not help to improve the forecast ability of the model. In any case, the surveys, whether used literally or bias-adjusted, do not outperform simple univariate time-series models: I find that an ARMA(1,1) produces the best forecasts, on average, from 1990 to 2015.

2 Revisiting the informational inefficiencies of survey based inflation expectations

Figure 1 shows three primary measures of inflation expectations: the Michigan Survey of Consumers year-ahead expectation (MSC), year-ahead CPI forecast from the Livingston Survey, and the year-ahead CPI forecast from the Survey of Professional Forecasters. Broadly speaking, the surveys measure the same object, expectations of average inflation for the next year, but they differ in their respondents, coverage and frequency. The Livingston survey is a survey of professional economists,

Figure 1: Survey based inflation expectations, 1960–2015. The panel shows the four-quarter change in the Total Consumer Price Index alongside the mean expectation for inflation in the next four quarters from each of the three surveys considered. See text for details.

has the longest history of the three measures, and is collected twice per year. The Survey of Professional Forecasters (SPF) is a quarterly survey of professional economists that began asking about total CPI inflation expectations in 1981. Finally, the Michigan Survey of Consumers (MSC) is a survey of households conducted each month, and began asking about year-ahead inflation expectations in 1978.[1]

Inflation expectations measured in the surveys tend to move in tandem. The SPF and Livingston surveys have a raw correlation of 0.98, while the expectations of households behave somewhat differently; their correlation with the expectations of professionals is about 0.65. As a result, the inflation expectations from the surveys can differ notably at any given point in time. Table 1 provides summary statistics for three periods, the surveys' entire sample, and the first and second halves of the period 1984–2015. The average behavior of each survey since 1984 (chosen so the surveys have a common sample) is quite similar. Each survey has a mean of about 3 percent and is quite stable, with ranges between just 1 and 5 percent. However, the behavior of the surveys diverges over the past 15 years. The average expected rate of inflation from the SPF and Livingston surveys is more than 1 percentage point lower in the second half of this period, and the forecasts of professionals have become less variable. In contrast, expectations from the Michigan survey are

[1] The Michigan Survey of Consumers does not specify a particular index. Instead, respondents are asked whether they expect prices to go up or down, followed by the question "how much?"

Table 1: Summary statistics of year-ahead survey based inflation expectations.

Survey	MSC	Liv.	SPF	MSC	Liv.	SPF	MSC	Liv.	SPF
Sample	Full sample			1984Q1–1999Q4			2001Q1–2015Q4		
N	152	112	138	64	32	64	64	32	64
Mean	3.6	3.4	3.1	3.1	3.5	3.6	3.0	2.1	2.2
Std dev.	1.7	2.0	1.2	0.5	0.8	0.8	0.6	0.4	0.3
Min	1.7	0.5	1.6	2.3	2.2	2.2	1.7	0.9	1.6
Max	10.2	10.2	7.8	4.7	5.6	5.3	5.1	2.8	2.7

largely unchanged from the first half of the sample.

Table 2 describes the behavior of the surveys' forecast errors. Panel 1 shows shows summary statistics for the *ex-post* forecast errors e_t, where $e_t = \pi^4_{t+4,t} - S_{j,t}\pi^4_{t+4,t}$ and $S_{j,t}$ denotes the 'survey operator,' the mean forecast from survey j, $j \in$\{MSC, SPF, Livingston\}. Throughout, a superscript on π denotes that it is an average rate of inflation across several quarters. Year-ahead inflation is denoted $\pi^4_{t+4,t}$, where $\pi^4_{t+4,t} = \frac{1}{4}\sum_{i=1}^{4}\pi_{t+i-1}$, since it is the average of realized quarterly inflation in the current and subsequent three quarters.[2]

The surveys are fairly accurate, on average, with mean errors of about 1 percentage point.[3] Over the full sample, professional forecasters are only slightly more accurate forecasters than households contacted by the Michigan Survey of Consumers. The average (signed) error from the MSC doubles when calculated since 2000 relative to the post-1984 period: from 1984 to 2000, the MSC did not over- or under-predict inflation on average but the MSC has overestimated realized inflation by nearly 1 percentage point since that time. In contrast, the forecast errors of professional forecasters tell the opposite story; professionals over-predicted inflation between 1984 and 2000 but had about mean-zero forecasts since that time. This simple evidence suggests that the expectations formulation process of households differs from that of professionals (Pfajfar & Santoro 2013).

The remainder of the table tests for simple deviations from rational expectations. The second panel of the table presents two measures of the persistence of survey errors, the estimated coefficient from an AR(1) process and the sum of autoregressive coefficients (SARC).[4] *Ex-post*, it is always

[2] The quarterly inflation rate is by $\pi_t = 400 \times log(P_t/P_{t-1})$, where P_t is the quarterly value of the price index.

[3] Of course, the forecast accuracy of the surveys may not not entirely surprising. Inflation is highly persistent, so future inflation may be highly predictable given recent inflation. A more formal comparison of the relative forecast accuracy of the surveys is performed in section 4.

[4] The lag length of the autoregressive process is determined by the AIC.

possible to decompose forecast errors into two portions, a bias and an error consistent with rational expectations. Under the null of rationality, this bias should be zero; i.e., forecast errors should not be predictable. The third panel, labeled *Mincer-Zarnowitz*, presents the Mincer-Zarnowitz regression wherein a forecast error is regressed onto the forecast itself. The final panel presents regressions as in Nordhaus (1987), which regress forecast errors onto forecast revisions. The sets of regressions include two important macroeconomic variables as covariates, the unemployment rate and the yield on the 10-year Treasury Bond.

It is difficult to reconcile tables 1 and 2 with fully rational, informationally efficient expectations formulation. In contrast to full-information rational expectations, it is clear that the forecast errors resulting from the surveys are highly autocorrelated. Panel 3 indicates that the surveys are, at times, biased. The 10-year Treasury yield often predicts forecast errors, and at times the unemployment rate also contains useful information for errors. Although not shown, tests of the null hypothesis that all coefficients in the regressions are zero are very strongly rejected. In that sense, the results conform to the prior literature, which concludes that expectations are not informationally efficient (Thomas, 1999; Mehra, 2002; Fuhrer, 2015).

Coibion & Gorodnichenko (2015*a*) use regressions of the type presented in the third and fourth panels to discern amongst models of expectations formulation that allow for deviations from full rationality. Specifically, they show that both sticky and noisy information models imply reduced-form relationships between forecast errors and forecast revisions as in Nordhaus (1987). Coibion & Gorodnichenko (2015*a*) find strong evidence of positive correlation between forecast errors and revisions, as predicted by models of sticky and imperfect information. Further, Coibion & Gorodnichenko (2015*a*) find that macroeconomic variables enter significantly in their Mincer-Zarnowitz regressions but not in the Nordhaus regressions. They interpret these results as supportive of models of informational rigidities: once one controls for the revision to the forecast, macroeconomic data no longer predict forecast errors. The results here are less clear cut, but strongly indicate deviations from rationality. The coefficient on forecast revisions is consistently positive, if not always statistically significant. Further, macroeconomic indicators continue to enter the Nordhaus regression in a statistically significant manner, even after conditioning on the forecast revision.

Table 2: Behavior of forecast errors from survey based inflation expectations.

Survey	MSC	Liv.	SPF	MSC	Liv.	SPF	MSC	Liv.	SPF
Sample	Full sample			1984Q1–2015Q4			2000Q1–2015Q4		
N	149	110	135	128	62	128	64	32	64
Mean and absolute forecast error									
Mean error	-0.15	0.35	-0.36	-0.40	-0.18	-0.25	-0.83	0.12	0.03
Mean abs error	1.11	1.14	0.96	0.99	0.87	0.89	1.25	1.00	1.03
Persistence of forecast errors									
AR(1)	0.74	0.46	0.68	0.73	0.49	0.69	0.67	0.39	0.65
SARC	0.75	0.60	0.56	0.68	0.32	0.51	0.57	0.06	0.39
Mincer-Zarnowitz regressions									
Const.	-1.06	3.09**	1.35*	0.56	1.50*	1.14	-1.27	-2.85	1.32
	(0.73)	(0.65)	(0.65)	(0.65)	(0.96)	(0.65)	(0.73)	(1.89)	(1.48)
$S_t\pi^4_{t+4,t}$	0.35	0.54**	-0.22	-0.66	-0.21	-0.54	-0.99**	-0.24	-1.91
	(0.15)	(0.14)	(0.17)	(0.36)	(0.38)	(0.36)	(0.43)	(0.79)	(0.90)
u_{t-1}	-0.14	-0.41**	-0.11	-0.06	-0.13	-0.08	0.14	0.19	0.03
	(0.09)	(0.11)	(0.09)	(0.09)	(0.09)	(0.09)	(0.15)	(0.20)	(0.19)
i^{10y}_{t-1}	0.14*	-0.33**	-0.05	0.24**	-0.04	0.12	0.67**	0.59*	0.70**
	(0.07)	(0.09)	(0.11)	(0.07)	(0.15)	(0.14)	(0.198)	(0.21)	(0.20)
R^2	0.22	0.30	0.16	0.26	0.12	0.07	0.40	0.16	0.26
Nordhaus regressions									
Const.	-0.62	1.82**	1.16	-1.23	1.36*	0.72	-4.99**	-2.81	-3.43*
	(0.72)	(0.65)	(0.67)	(0.76)	(0.63)	(0.72)	(1.72)	(1.92)	(1.71)
$S_t\pi^4_{t+4,t} - S_{t-1}\pi^4_{t+4,t}$	0.36	1.20**	0.56	0.24	0.52	0.47	0.27	0.66*	0.47
	(0.23)	(0.32)	(0.34)	(0.20)	(0.34)	(0.40)	(0.27)	(0.27)	(0.91)
u_{t-1}	-0.14	-0.25**	-0.11	-0.07	-0.13	-0.07	0.20	0.17	0.23
	(0.09)	(0.08)	(0.09)	(0.08)	(0.09)	(0.09)	(0.15)	(0.18)	(0.16)
i^{10y}_{t-1}	0.21**	0.01	-0.13*	0.21**	-0.11*	-0.08	0.74**	0.49*	0.52**
	(0.06)	(0.06)	(0.05)	(0.06)	(0.05)	(0.05)	(0.19)	(0.21)	(0.20)
R^2	0.19	0.30	0.17	0.20	0.15	0.06	0.22	0.21	0.12

Notes: Panel labeled *Mincer-Zarnowitz* displays results from the regression $(\pi^4_{t+4,t} - S_t(\pi^4_{t+4,t})) = \alpha + \beta_1 S_t(\pi^4_{t+4,t}) + \beta_2 u_{t-1} + \beta_3 i^{i10y}_{t-1} + \varepsilon_t$. The final panel presents a Nordhaus (1987) regression: $(\pi^4_{t+4,t} - S_t(\pi^4_{t+4,t})) = \alpha + \beta_1(S_t\pi^4_{t+4,t} - S_{t-1}\pi^4_{t+4,t}) + \beta_2 u_{t-1} + \beta_3 i^{i10y}_{t-1} + \varepsilon_t$. Newey-West standard errors in parentheses.

These results suggest that sticky or noisy information models cannot fully explain deviations from rationality observed in the surveys.

3 Modeling inflation expectations

This section uses a machine learning algorithm to uncover the drivers of survey based inflation expectations. Inflation expectations are regressed onto a large number of variables that may or may not influence inflation expectations. The set of covariates includes a large number of variables that may influence inflation expectations, such as employment and prices. The machine learning algorithm selects covariates important for explaining movements in inflation expectations, ignoring those that do not.

3.1 Method

Model selection is done with the *boosting* algorithm, initially introduced by Schapire (1990). Friedman, Hastie & Tibshirani (2000) later provided statistical foundations for the method (see also Friedman (2001) and Hastie, Tibshirani & Friedman (2001)). The algorithm is a shrinkage estimator; it builds a statistical model from a potentially very large set of regressors, shrinking the influence of regressors that do not co-vary with the dependent variable. The idea of the algorithm is to minimize the expected value of a user-specified loss function. In contrast to other linear shrinkage methods, boosting allows for a much more general relationship between a vector of covariates and the target variable. The algorithm is outlined below.

Let y_t, the outcome of interest, be a function of a $K \times 1$ vector of potential explanatory covariates, x_t (x includes a constant). Boosting estimates the function $F : \mathbb{R}^K \to \mathbb{R}$ that minimizes the expected loss $\mathcal{L}(y, F)$, i.e.,

$$\hat{F}(x) \equiv \arg\min_{F(x)} E\left[\mathcal{L}(y, F(x))\right]. \quad (1)$$

The setup as described is very flexible. The strategy below will be to specify $\mathcal{L}(y, F(x))$ as squared-error loss and to assume that $F(x)$ is linear in each covariate. These assumptions reduce the problem to one that is analogous to a standard OLS regression, and the result is an algorithm known as the $L2$-boost (Buhlmann & Yu 2003).

In practice, the algorithm searches over the covariates in x_t for those that have the highest correlation with the outcome variable y_t. Given the setup described here, let $y_t = \beta' x_t + u_t$. Boosting is an iterative procedure, so let m denote iterations and $\beta_{m,k}$ denote the k^{th} element of β at iteration m. The algorithm is as follows:

1. Initialize $m = 0$ and let $\beta_{0,k} = 0$ for $k = 1, ..., K$.

2. For m = 1,...,M

    (a) For $t = 1, ..., T$, produce the current residuals $u_t = y_t - \beta'_{m-1} x_t$.

    (b) For each $k = 1, ..., K$, regress the current residuals u onto each regressor, x_k, to find $\hat{b}_{m,k}$. Calculate the residuals $\hat{e}_k = u - \hat{b}'_k x_k$ and the sum of squared residuals $\hat{e}'_k \hat{e}_k$.

    (c) Choose the regressor with the minimum sum squared residuals: $I_m = \min_{k=1,..,K} \hat{e}'_k \hat{e}_k$. Let $\hat{b}_m$ denote the $K \times 1$ vector that consists of zeros in all elements except for element I_m, which is equal to $\hat{b}_{I_m}$.

    (d) Update slope vector $\hat{\beta}_m = \hat{\beta}_{m-1} + \rho \hat{b}_m$, where $\rho \in (0, 1)$ is the shrinkage factor or step-size.

$L2$-boosting is repeated least-squares regression, where at each iteration the variable that is most highly correlated with the outcome is included in the model. Each covariate may be selected many times throughout the course of the boosting algorithm, or not at all. Buhlmann (2006) shows that $L2$-boosting has appealing consistency properties when the true model is sparse.

The parameters ρ and M jointly determine the fit of the algorithm. Small values of ρ prevent overfitting, at the cost of needing to employ a higher number of iterations. Although one can estimate an additional minimization problem to find an optimal step-size, the empirical application below uses a fixed value of 0.1 (a commonly used value). The algorithm must be stopped early if any substantive model selection is to take place. To that end, M is chosen to minimize the Schwarz information criterion; i.e. $M \equiv \arg\min_m BIC(m)$, where $BIC(m)$ the value of the BIC of the model at iteration m (Buhlmann 2006; Buhlmann & Hothorn 2007).

3.2 Data

The analysis focuses on the quarterly frequency. The SPF is collected once per quarter, with the survey undertaken in the middle month of the quarter. In order to ensure that respondents to the MSC have approximately the same information set as respondents to the SPF, quarterly observations of the Michigan Survey of Consumers are constructed using the monthly survey from the month in the middle of each quarter. Monthly data are converted to quarterly by computing the average level of the indicator within each quarter, then applying any relevant transformations.

Covariates fall into two groups. Since individuals may benchmark inflation expectations to recent movements in the overall price level, or may focus on the prices of the goods and services that they most frequently purchase or observe (Schulhofer-Wohl & Kaplan 2016), the dataset includes several subcomponents of the Consumer Price Index. The second group of regressors are common macroeconomic indicators, the well-known Stock-Watson macroeconomic dataset. The data are a wide range of macroeconomic indicators regarding the labor market, housing market, financial indicators such as equity price indices and interest rate spreads, industrial production, manufacturing surveys, exchange rates, and consumer sentiment.[5] Finally, d lags of each candidate predictor is included in the model search. Time-t predictors are gathered into an $n+1$-dimensional vector of data, x_t, where x includes a constant and n denotes the number of predictors. In total, $n = 104$ and $d = 3$, so K, the number of possible covariates is $n \times d + 1 = 313$.

3.3 Full-sample results

Table 3 shows the ten covariates most frequently selected by the model search, as well as summary statistics describing model fit, for the models fit to the Michigan Survey of Consumers and the

[5] All variables are transformed to stationarity, see appendix table A1 for details. The data can be found via FRED at: https://research.stlouisfed.org/pdl/788. Producer price indices from the standard Stock-Watson dataset are not included in the data, since I include the measures from the CPI instead. Inclusion of the PPI subcomponents does not qualitatively alter the results.

Survey of Professional Forecasters.[6] Recall that the boosting algorithm estimates

$$S_{jt}(\pi^4_{t+4,t}) = \sum_{k=1}^{K} \beta_k X^{(k)}_{t-1} + \varepsilon_t,$$

where j denotes the survey and X_{t-1} contains a large number of covariates $k = 1, ..., K$.[7] The estimated regression coefficient is given in the second column of the table. Since many of the price indices included can be very volatile on a quarter-to-quarter basis, regression coefficients for price indicators have been standardized. Variables are sorted by their importance in explaining movements in inflation expectations, as measured by the marginal R^2 of each variable. Specifically, let $I(.)$ be the function selects a variable at each model iteration, and let κ denote the selected variable. Then the frequency that variable k is selected by the algorithm is $\phi_k = \frac{1}{M}\sum_{m=1}^{M} I(k = \kappa)$. Column 3 of each panel weights each iteration by the contribution to the model's overall fit: $\Phi_k = \sum_{m=1}^{M} w_m I_m$, where $w_m = \frac{R^2_m - R^2_{m-1}}{R^2_M}$.

The left panel shows the ten indicators with the highest correlation with household expectations. The final model includes 78 of the possible covariates, although many of the covariates explain only minute portions of the total variance, with coefficients that are heavily shrunk towards zero. The algorithm produces a model that nearly perfectly fits the path of inflation expectations, in sample, the R^2 of the model is nearly one. The regression coefficients are shown in the second column of the panel.[8] Household inflation expectations appear heavily influenced by movements of particular prices, especially food and energy prices. Measures of real activity are also included in the model, such as average hourly earnings in manufacturing, consumer sentiment and personal income. The signs of each indicator are as one would have expected; higher prices and wages move inflation expectations higher. Sentiment is negatively correlated with inflation expectations, conditional on the other macroeconomic variables. In contrast, the covariates selected for the model of expectations of professional forecasters depends heavily on macroeconomic data. Interest rates are responsible

[6] In the interest of concision, and because they tend to be similar to those from the SPF, I do not report results for the Livingston survey.

[7] The covariates in X are dated with subscript *t-1* in order to account for the lag of the macroeconomic data-flow.

[8] Coefficients that correspond to indicators that are volatile on a quarter-to-quarter basis have been standardized to show the estimated response to a one standard deviation change in that variable (and are denoted with a †). It is worth emphasizing that the regression coefficients should be interpreted only as suggestive correlations; the predictors are highly correlated among themselves, and no attempt has been made to produce an identified model.

Table 3: Covariates with highest explanatory power for year-ahead inflation expectation, MSC and SPF surveys.

Michigan Survey of Consumers			Survey of Professional Forecasters		
Covariate	$\beta \times 100$	$\Phi \times 100$	Covariate	$\beta \times 100$	$\Phi \times 100$
CPI (food away)$_{t-1}$	47†	37	i^{ff}_{t-1}	9	35
CPI (food away)$_{t-2}$	38†	25	i^{10y}_{t-1}	8	28
CPI (food away)$_{t-3}$	18†	10	CPI (med. serv.)$_{t-1}$	17†	9
CPI (comm. lfe)$_{t-1}$	15†	5	CPI (comm. lfe)$_{t-1}$	16†	7
AHE, goods$_{t-1}$	5	4	CPI (OER)$_{t-1}$	17†	7
Personal income$_{t-1}$	4	4	CPI (comm. lfe)$_{t-3}$	18†	3
MSC sentiment$_{t-1}$	-2	2	Personal income$_{t-1}$	12†	2
CPI (food at home)$_{t-1}$	6†	1	Empl. (whole.)$_{t-1}$	5	1
i^{BAA}_{t-1}	-5	1	Empl. (nondur.)$_{t-1}$	1	1
CPI (gasoline)$_{t-1}$	7†	1	CPI (food away)$_{t-2}$	4†	1
R^2	96%		R^2	98%	
N (included)	78		N (included)	74	
Estimation period	1978Q4 – 2015Q4		Estimation period	1982Q3 – 2015Q4	

Notes: Stopping criteria chosen with BIC.
†Regression coefficient is standardized and gives estimated response of expectations to a one-standard deviation move in that subcomponent (×100).

for a bulk of the total explained variance of the model fit to the SPF. Other important covariates include measures of the labor market and certain subcomponents of the CPI.

Overall, it appears that consumers and professional forecasters have very different models of future inflation. Consumer-based inflation expectations appear mostly correlated with prices that are salient, such as food, energy and wages. In contrast, expectations of professional forecasters are more heavily correlated with macroeconomic indicators. Finally, it is important to explicitly note that these models do not attempt to uncover causal relationships between the covariates and survey expectations. The remainder of the analysis focuses on whether the algorithm can be used to anticipate movements in inflation expectations and realized inflation itself.

3.4 An evaluation of recent movements in inflation expectations

The past 15 years has seen a divergence in the inflation expectations of households and professional forecasters. It is also a time of large macroeconomic shocks, and is thus an important episode that can be used to understand the relationship between macroeconomic fundamentals and inflation expectations. The models of inflation expectations developed above are a natural tool to understand the behavior of inflation expectations throughout this period. To do so, the boosting algorithm is estimated with data that ends in the fourth quarter of 2005 to build a model for both the MSC and SPF.[9] These models produce an estimate of inflation expectations, conditional on the realized macroeconomic data. A second set of models are fit to the entire sample, and in-sample estimates of expectations are produced. The comparison of the two models may shed light on the ability of the boosting model to fit out-of-sample and whether the behavior of inflation expectations was abnormal through this period.

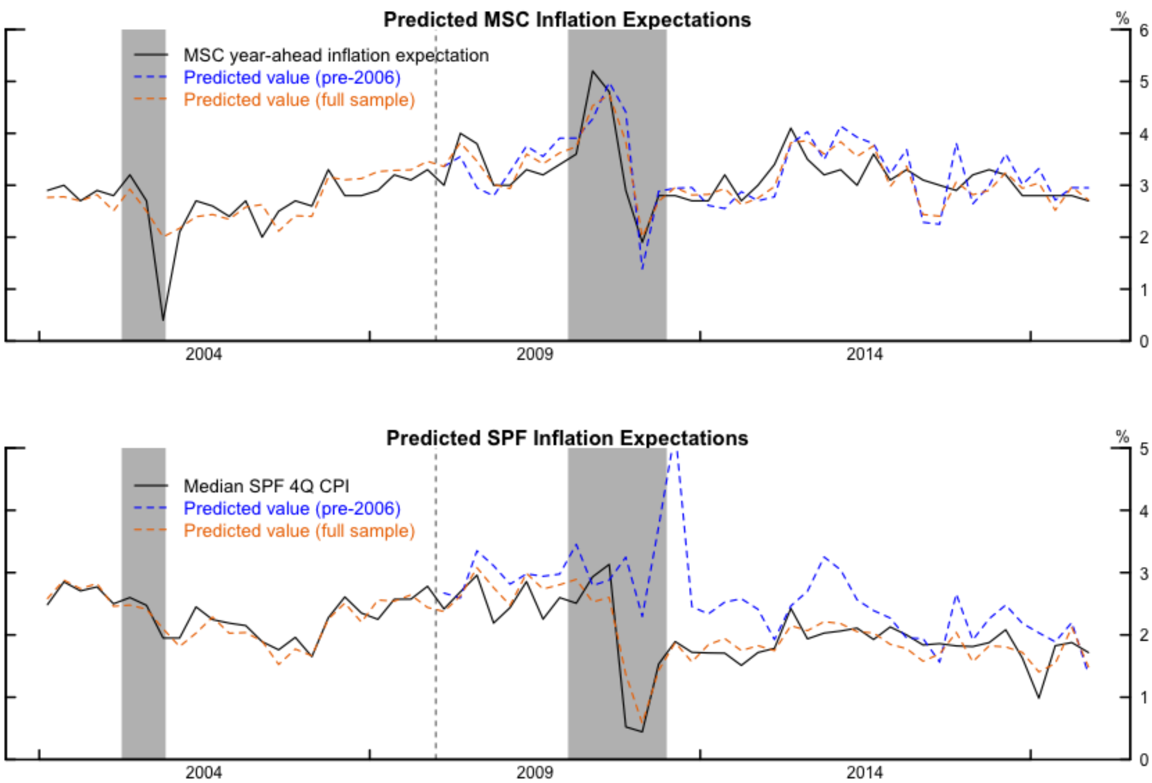
The solid black lines in figure 2 show the path realized year-ahead inflation expectations for the MSC (top panel) and SPF (bottom panel). Blue dashed lines show the model's predicted path for inflation expectations when models are fit using data only prior to 2006, and orange lines indicate in-sample projections from models fit to the entire sample.

Focusing first on MSC expectations, the model-implied expectations based on pre-crisis data are broadly consistent with the realized changes in inflation expectations data. The most important model coefficients are presented in table 4 for the model fit to the two periods of data. Comparing the left-side panel to the one on the right, the models are quite similar. The most important impact of an additional 10 years of data appears to be slight changes in the model coefficients. The response of inflation expectations to measures of CPI, especially food, are smaller in absolute value when estimated with the full sample. Notably, the most important indicators are largely the same between the two models, a regime change does not appear to have occurred, despite the large macroeconomic shocks during this period.

Expected inflation from the SPF was remarkably stable through this period. In contrast, the model fit to pre-crisis data would have expected a sharp uptick in inflation expectations in 2009, as both interest rates and commodity prices moved higher. After that, the fitted value drifts lower, and the model fit to pre-crisis data would have expected an inflation expectation of about 1.4 percent at the end of the sample. As seen in table 5, estimating the model on the full sample

[9] The end date was chosen arbitrarily, but leaves nearly two years of data prior to the onset of the Great Recession.

Figure 2: Event study of survey based inflation expectations, 2007-2015.



Notes: Solid black lines show MSC (top) and SPF (bottom) year-ahead expected inflation. Colored dashed lines indicate model projections, with blue lines showing predicted values using boosting algorithm fit to data ending in 2005Q4 and orange line fit to data through 2015Q4. Vertical line indicates 2006Q1. See text for details.

Table 4: Covariates with highest explanatory power for MSC year-ahead inflation expectation.

Covariate	$\beta \times 100$	$\Phi \times 100$	Covariate	$\beta \times 100$	$\Phi \times 100$
CPI (food away)$_{t-1}$	56†	36	CPI (food away)$_{t-1}$	47†	37
CPI (food away)$_{t-3}$	43†	25	CPI (food away)$_{t-2}$	38†	25
CPI (food away)$_{t-2}$	27†	16	CPI (food away)$_{t-3}$	18†	10
AHE, goods.$_{t-1}$	6	4	CPI (comm. lfe)$_{t-1}$	15†	5
MSC sentiment$_{t-1}$	-2	2	AHE, goods$_{t-1}$	4	4
CPI (comm. lfe)$_{t-1}$	13†	3	Personal income$_{t-1}$	4	4
Personal income$_{t-1}$	5	3	MSC sentiment$_{t-1}$	-2	2
CPI (food at home)$_{t-1}$	11†	2	CPI (food at home)$_{t-1}$	6†	1
AHE, manu.$_{t-1}$	8	2	i^{BAA}_{t-1}	-5	1
i^{BAA}_{t-1}	-6	1	CPI (gasoline)$_{t-1}$	7†	1
In-sample R^2	97%		In-sample R^2	96%	
N (included)	63		N (included)	74	
Estimation period	1978Q4–2005Q4		Estimation period	1978Q4–2015Q4	

Notes: Stopping criteria chosen with BIC. See text for details.
†Regression coefficient is standardized and gives estimated response of expectations to a one-standard deviation move in that subcomponent.

of data dampened the estimated response of expectations to the covariates, as would be expected given how flat realized expectations had been during the crisis.

There are several conclusions that can be drawn from this simple event study. The boosting algorithm is able to nearly perfectly fit data in-sample, suggesting that it may be overfit. Secondly, whereas recent literature has tied inflation expectations to movements in energy prices, e.g. Coibion & Gorodnichenko (2015*b*), Binder (2016), this work suggests that food prices also importantly affect inflation expectations, especially for households. The models do not view the fall in realized inflation expectations from 2012 to 2015 as abnormal. Finally, the models tended to become less sensitive, especially to movements in food and energy prices, through this period. Although since the costs associated with infrequently updating beliefs increase during times of uncertainty, that the estimated reaction of expectations changed during this period is not necessarily surprising.

Table 5: Covariates with highest explanatory power for SPF year-ahead inflation expectation.

Covariate	$\beta \times 100$	$\Phi \times 100$	Covariate	$\beta \times 100$	$\Phi \times 100$
i^{5y}_{t-1}	7	25	i^{ff}_{t-1}	9	35
i^{ff}_{t-1}	5	20	i^{10y}_{t-1}	8	28
CPI (comm. lfe)$_{t-1}$	$19^{\dagger}$	14	CPI (med. serve)$_{t-1}$	$17^{\dagger}$	8
CPI (food away)$_{t-1}$	$19^{\dagger}$	11	CPI (comm. lfe)$_{t-1}$	$16^{\dagger}$	8
CPI (med. serve)$_{t-1}$	12	5	CPI (OER)$_{t-1}$	$17^{\dagger}$	7
CPI (OER)$_{t-1}$	15	5	CPI (comm. lfe)$_{t-3}$	$8^{\dagger}$	3
i^{10y}_{t-1}	2	5	Personal income$_{t-1}$	12	2
CPI (commodities lfe)$_{t-2}$	$6^{\dagger}$	3	Employment, whole.$_{t-1}$	5	1
Employment (non-dur)$_{t-1}$	4	2	Employment, non-dur.$_{t-1}$	3	1
MSC sentiment$_{t-2}$	-1	2	CPI (food away)$_{t-1}$	3	1
In-sample R^2	97%		In-sample R^2	95%	
N (included)	72		N (included)	72	
Estimation period	1982Q3–2005Q4		Estimation period	1982Q3–2015Q4	

Notes: Stopping criteria chosen with BIC. See text for details.
†Regression coefficient is standardized and gives estimated response of expectations to a one-standard deviation move in that subcomponent.

4 Forecasting realized inflation

Inflation expectations are not informationally efficient, yet one reason for the interest that surrounds the surveys is that they carry important information for future realized inflation. This section attempts to uncover whether deviations from rationality are economically important when viewed through the lens of a forecasting horserace. If deviations from informational efficiency are economically meaningful, then a procedure that bias-adjusts the surveys ought to result in notable improvements to forecast ability. For completeness, this section produces forecasts with several broad methods: surveys, time-series models, Phillips curve models, and by applying model selection to inflation itself.

Forecasts are made at the one- and two-year horizon,

$$E_t(\pi^h_{t+h,t}) = E_t\left(\frac{1}{h}\sum_{i=1}^{h}\pi_{t+i-1}|\Omega_{t-1}\right), \tag{2}$$

where $h \in \{4, 8\}$ denotes the forecast horizon and Ω represents the information set available to the model. Three price indices are forecasted: the total consumer price index, the consumer price index less food and energy, and the consumer price index less shelter. Since real-time data for the Stock-Watson dataset does not have a long history, all data is of the June 2016 vintage. I reserve the period 1990Q1-2015Q4 as an out-of-sample forecast period, for a total of 104 forecasts per model, price index and forecast horizon.

4.1 The forecast models

4.1.1 Survey based forecasts

Three sets of forecasts are produced using primarily the surveys. The first set is the surveys themselves. In this case,

$$E_t(\pi^h_{i,t+h,t}) = S_{j,t}(\pi^h_{t+h,t}), \tag{3}$$

for $j \in \{\text{SPF, Michigan}\}$ and price index i. Since neither survey elicits an explicit 8 quarter ahead forecast, I use the four-quarter ahead expectation to forecast both one and two years ahead. Forecasts are denoted as MSC and SPF, respectively.

The surveys are adjusted for biasedness in two ways. First, Mincer & Zarnowitz (1969) regressions are used to bias-adjust the surveys, so that predictions from the regression

$$(\pi^h_{i,t+h,t} - S_{jt}\pi^h_{t+h,t}) = \alpha + \beta S_{jt}\pi^h_{t+h,t} + \varepsilon_t \tag{4}$$

are used to adjust survey j for its estimated bias with respect to price index i, with α and β estimated via OLS. Forecasts of $\pi^h_{t+h,t}$ given $t-1$ are denoted with a superscript MZ.

Thirdly, boosting models are used to estimate per-period bias in the survey measures. Specifically, the boosting algorithm estimates

$$(\pi^h_{i,t+h,t} - S_{jt}\pi^h_{i,t+h,t}) = \sum_{k=1}^{K} \beta_k X^{(k)}_{t-1} + \varepsilon_t. \tag{5}$$

The estimated models are used to produce bias-adjusted inflation projections. Note that this model is the one used in section 4, but that we model bias instead of the survey itself. Hyperparameters of the algorithm are set as above. These forecasts are denoted with a dagger superscript (†).

4.1.2 Time-series models

As a point of comparison, four univariate time-series models are used to produce inflation forecasts: the random walk model, an $ARMA(1,1)$ model, an AR(p) model, and the UC-SV model of Stock & Watson (2007).

Since Atkeson & Ohanian (2001), the random walk model has served as the benchmark model of inflation forecasting. Thus, I produce forecasts of h-quarter ahead inflation using

$$E(\pi^h_{t+h,t}|\Omega_{t-1}) = \pi^h_{t-1}, \tag{6}$$

where π^h_{t-1} is the previous period's h-quarter averaged rate of inflation. This model is denoted AO.

An ARMA(1,1) model serves as a slightly more sophisticated time-series model. As noted by Ang et al. (2007), an $ARMA(1,1)$ model can be motivated by rational expectations models: if expected inflation follows an AR process, then a reduced-form model for inflation is ARMA:

$$\pi_t = \mu + \phi\pi_{t-1} + \psi\varepsilon_{t-1} + \varepsilon_t. \tag{7}$$

The $AR(p)$ model is:

$$\pi_t = \mu + \phi_1\pi_{t-1} + \phi_2\pi_{t-2} + ... + \phi_p\pi_{t-p} + \varepsilon_t, \tag{8}$$

where p denotes the lag-length for the autoregressive process. The AIC criterion is used to select the lag length with a maximum of 6 possible lags.

Finally, forecasts are produced from the unobserved component model of Stock & Watson (2007), since the model is able to capture important changes in the time-series behavior of inflation. Inflation is modeled as the sum of a trend and a serially uncorrelated error, each of which has stochastic volatility:

$$\pi_t = \tau_t + e^{h_{\pi t}/2}\varepsilon_{\pi t} \tag{9}$$

$$\tau_t = \tau_{t-1} + e^{h_{\tau t}/2}\varepsilon_{\tau t} \tag{10}$$

$$h_{\pi t} = h_{\pi t-1} + \varepsilon_{h_\pi,t} \tag{11}$$

$$h_{\tau t} = h_{\tau t-1} + \varepsilon_{h_\tau,t} \tag{12}$$

where $\varepsilon_t = [\varepsilon_\pi, \varepsilon_\tau, \varepsilon_{h_\pi}, \varepsilon_{h_\tau}]' \sim N(0, \Sigma)$; $\Sigma = diag(I_2, \gamma I_2)$. The parameter γ controls the degree to which the stochastic volatility can vary over time, and is set to 0.04. The model's forecast of $\pi_t^h|\Omega_{t-1}$ is $\tau_t|\Omega_{t-1}$.

4.1.3 Phillips-curve models

I consider several Phillips-curve models of inflation, which combine a measure of economic slack with a measure of expected inflation.

$$\pi_{t+h,t}^h = \alpha + \gamma(L)E_t\pi + \beta(L)X_{t-1} + \varepsilon_t, \tag{13}$$

where $E_t\pi$ denotes some measure of expected inflation and the vector X denotes a measure of economic slack. Since the literature using Phillips curves to forecast inflation has not coalesced around a single econometric specification, I produce forecasts from a number of specifications.

Three different proxies are used for inflation expectations. The first proxy is recent realized inflation, known as a 'backwards-looking' Phillips curve. Secondly, I consider models that use surveys to proxy for expectations. I use both the SPF and the Michigan Survey in these expectations-augmented Phillips curves. Finally, as above, I consider a Phillips curve that uses the bias-adjusted measure of inflation expectations produced above. In these models, boosting is used to extract a bias-adjusted measure of inflation expectations. This measure of expected inflation is used in the Phillips curve model in 13. These models are again denoted with a superscript †.

Four measures of real activity are used within the Phillips curve specification: the CBO produced GDP gap, the CBO produced unemployment rate gap, the Stock-Watson unemployment rate gap

(Stock & Watson 2010), and labor's share of income. The first three indicators are measures of economic activity commonly utilized in the empirical Phillips curve literature; while the fourth is motivated by the New Keynesian Phillips curve, as in Galì & Gertler (1999).[10]

Thus, all told, I produce forecasts from 20 different Phillips curve specifications—five different measures of expected inflation (backwards-looking, SPF, MSC, SPF†, MSC†), each with four different measures of economic activity. The lag length of equation (13) is determined via the AIC, and I impose that the number of lags of inflation is the same as the number of lags in the indicator of economic activity.

4.1.4 Model selection

The final model applies boosting to inflation itself. The model is similar in spirit to Wright (2009), who uses BMA to weight forecasts produced by a suite of univariate inflation models. Wright finds that BMA outperforms forecasts produced by equally weighting forecasts produced by the same suite of univariate models. The covariates in this model are the same used in section 3 to produce models of inflation expectations.

4.2 Results

Tables 6 and 7 contain the main results. The tables treat the Atkeson & Ohanian (2001) random walk as the null model, and show the root mean squared forecast error for that model, in percentage points. For all other models, the table gives the RMSE relative to the AO null, so that a value less than one indicates that the model outperforms the random walk null in the sense that it has a smaller root mean-squared forecast error. Asterisks indicate statistical significance according to a Diebold-Mariano-West test of predictive accuracy, and bolded entries indicate the best performer for each price index.

Beginning with the surveys, the unadjusted MSC and SPF projections outperform the AO random walk when forecasting total CPI and the CPI excluding shelter. The improvement of the MSC over the AO null is not usually statistically significant, and the MSC performs worse than the

[10] Details on the slack measures for the Phillips curves can be found in table A3.

null when forecasting CPI less food and energy. The SPF outperforms the null for all three indices at both horizons, and usually in a statistically significant manner. In addition, the SPF always outperforms MSC, in contrast evaluations that used earlier samples (Thomas, 1999; Ang Bekaert & Wei, 2007).

Bias-adjustment does improve the forecast ability of the surveys, especially for MSC forecasts. Focusing at the four-quarter horizon, bias-adjusting the MSC with the Mincer-Zarnowitz regressions reduces the RMSFE by 0.3-0.4 percentage point; for example, the RMSFE of the survey with respect to total CPI falls from 1.35 to 0.95 pp. These differences are statistically meaningful when tested with a DMW test (not shown). The boosting algorithm cannot consistently improve the forecast ability of the survey beyond the simple MZ-regressions. Bias-adjustment of the SPF produces forecasts that are less dramatically improved. MZ adjusting the SPF improves forecasts on average, but not in a statistically significant way. Bias-adjusting with the boosting algorithm also improves the forecasts over the raw survey, but usually not in a statistically significant way. Further, the improvement is not consistent; bias-adjustment the two year ahead forecast of core CPI with boosting actually causes a deterioration in forecast performance.

The various survey based forecasts are presented in figures 3 and 4. The figure reveal why the boosted surveys perform poorly when forecasting CPI less food and energy; the boosted forecasts reacted strongly to the turmoil that occurred during the most recent recession. For the models of MSC, these movements largely reflect movements in commodity prices, as we have seen, the SPF forecasts also react to large movements in interest rates. In contrast, simple MZ bias-adjustment served to smooth through these large shocks.

Turning to the Phillips curve models, conditioning on surveys almost always improves the forecast ability of Phillips curve models, relative to their backwards-looking Phillips curve counterpart, although the improvements are not statistically significant (not shown). Interestingly, Phillips curve models that condition on MSC outperform their SPF counterparts, suggesting that once one accounts for slack in some way, the MSC can carry additional information for forecasting future inflation while the SPF does not. No single measure of slack provides clearly superior inflation projections. The Stock & Watson (2010) unemployment rate gap performs well. Given that the

Table 6: Relative out-of-sample root mean squared forecast errors, four quarter ahead forecasts, 1990-2015.

Model	Total CPI	CPI ex. food & energy	CPI ex. shelter
AO random walk	1.52	0.56	2.24
MSC	0.90	1.92	0.81*
SPF	0.69**	0.91	0.71**
MSC^{MZ}	0.63**	1.28	0.61**
SPF^{MZ}	0.65**	0.83*	0.67**
MSC†	0.74**	1.89	0.53**
SPF†	0.61**	0.88	**0.52****
UCSV	0.92	1.07	0.79**
ARMA(1,1)	0.56**	**0.57****	0.52**
AR(p)	**0.52****	1.11	**0.52****
Acc. Phillips curve (CBO GDP gap)	0.69**	1.28	0.65**
Acc. Phillips curve (CBO un. gap)	0.80	1.58	0.70**
Acc. Phillips curve (S&W un. gap)	0.80	1.58	0.70**
Acc. Phillips curve (labor's share)	0.72**	1.34	0.64**
MSC Phillips curve (CBO GDP gap)	0.68**	1.45	0.64**
MSC Phillips curve (CBO un. gap)	0.69**	1.45	0.65**
MSC Phillips curve (S&W un. gap)	0.66**	1.37	0.59**
MSC Phillips curve (labor's share)	0.68**	1.29	0.67**
SPF Phillips curve (CBO GDP gap)	0.75*	1.09	0.74**
SPF Phillips curve (CBO un. gap)	0.74*	1.23	0.71*
SPF Phillips curve (S&W un. gap)	0.70**	0.97	0.67**
SPF Phillips curve (labor's share)	0.79**	0.98	0.78**
MSC† Phillips curve (CBO GDP gap)	0.74**	1.06	0.71**
MSC† Phillips curve (CBO un. gap)	0.71**	1.15	0.67**
MSC† Phillips curve (S&W un. gap)	0.68**	1.03	0.66**
MSC† Phillips curve (labor's share)	0.66**	0.91	0.67**
SPF† Phillips curve (CBO GDP gap)	0.79	1.19	0.79*
SPF† Phillips curve (CBO un. gap)	0.83	1.27	0.81
SPF† Phillips curve (S&W un. gap)	0.85	1.16	0.83
SPF† Phillips curve (labor's share)	0.95	1.05	0.91
Boosted inflation model	1.14	1.54	1.10

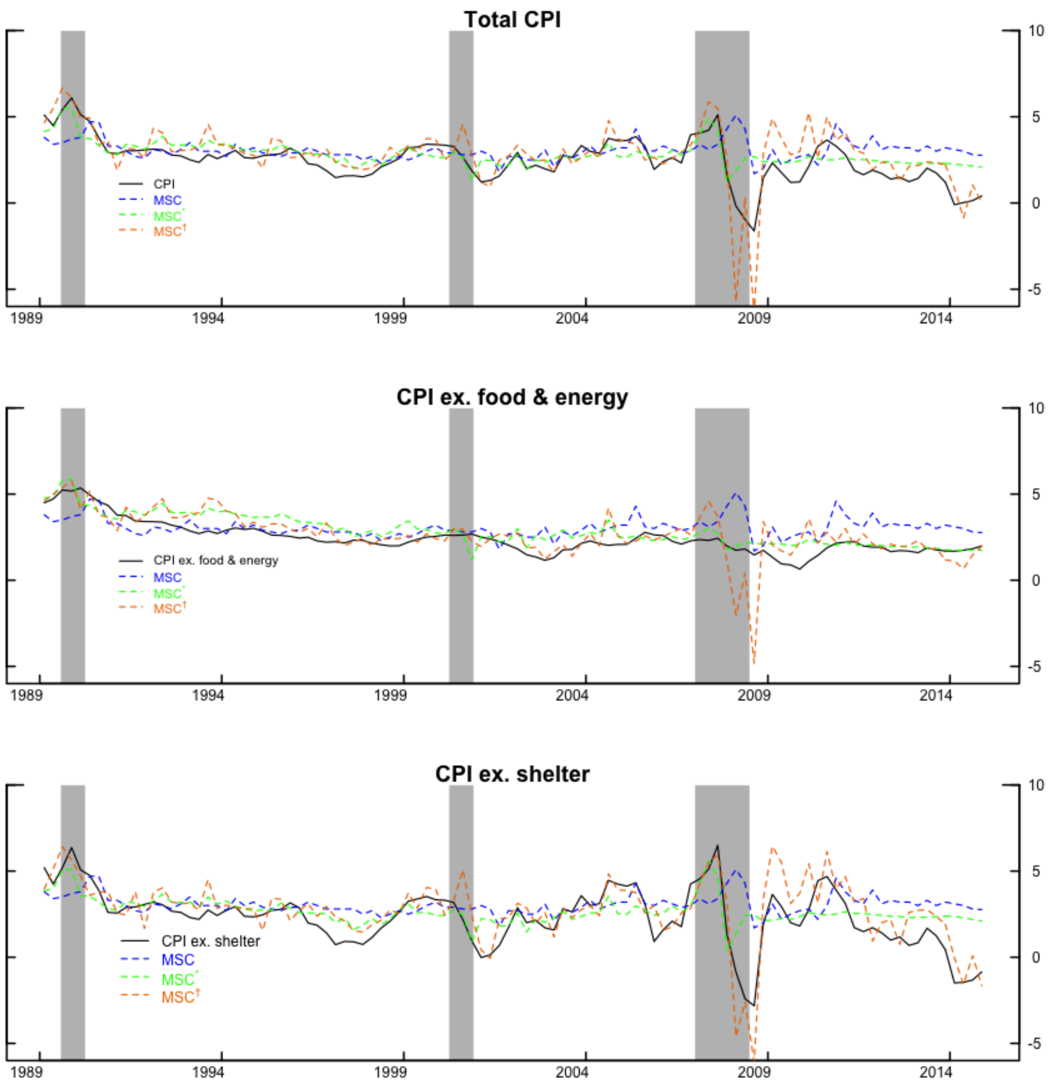
Notes: Root mean squared forecast error shown for AO model, all other models show RMSE relative to *AO* model, $\frac{RMSE^j}{RMSE^{AO}}$ for model j. Asterisks indicate statistical significance of Diebold-Mariano-West test of comparative predictive accuracy at 10 (*) and 5 (**) percent, relative to AO model. Bolded entry indicates best performing model for that price index. See text for details.

Table 7: Relative out-of-sample root mean squared forecast errors, eight quarter ahead forecasts, 1990-2015.

Model	Total CPI	CPI ex. food & energy	CPI ex. shelter
AO random walk	1.32	0.59	1.86
MSC	0.86	1.77	0.73**
SPF	0.60**	0.80**	0.58**
MSC^{MZ}	0.58**	1.20	0.53**
SPF^{MZ}	0.61**	0.68**	0.62**
$\text{MSC}^{\dagger}$	0.62**	1.49	0.45**
$\text{SPF}^{\dagger}$	0.55**	1.13	0.49**
UCSV	0.64**	0.58**	0.55**
ARMA(1,1)	**0.32****	**0.31****	**0.31****
AR(p)	0.47**	0.87	0.40**
Acc. Phillips curve (CBO GDP gap)	0.86	1.54	0.67**
Acc. Phillips curve (CBO un. gap)	0.95	1.74	0.71**
Acc. Phillips curve (S&W un. gap)	0.90	1.77	0.69**
Acc. Phillips curve (labor's share)	1.07	2.09	0.75*
MSC Phillips curve (CBO GDP gap)	0.85	1.54	0.77**
MSC Phillips curve (CBO un. gap)	0.91	1.47	0.84
MSC Phillips curve (S&W un. gap)	0.76**	1.34	0.67**
MSC Phillips curve (labor's share)	0.68**	1.21	0.65**
SPF Phillips curve (CBO GDP gap)	0.70**	0.98	0.69**
SPF Phillips curve (CBO un. gap)	0.83*	1.10	0.76**
SPF Phillips curve (S&W un. gap)	0.84*	0.99	0.76**
SPF Phillips curve (labor's share)	0.77**	0.98	0.71**
$\text{MSC}^{\dagger}$ Phillips curve (CBO GDP gap)	0.75**	0.90	0.75**
$\text{MSC}^{\dagger}$ Phillips curve (CBO un. gap)	0.90	1.15	0.88
$\text{MSC}^{\dagger}$ Phillips curve (S&W un. gap)	0.78**	1.00	0.75**
$\text{MSC}^{\dagger}$ Phillips curve (labor's share)	0.70**	0.80	0.70**
$\text{SPF}^{\dagger}$ Phillips curve (CBO GDP gap)	0.67**	0.87	0.74**
$\text{SPF}^{\dagger}$ Phillips curve (CBO un. gap)	0.95	1.01	0.92
$\text{SPF}^{\dagger}$ Phillips curve (S&W un. gap)	1.41	1.37	1.44
$\text{SPF}^{\dagger}$ Phillips curve (labor's share)	1.16	1.11	1.00
Boosted inflation model	0.91	1.66	0.83

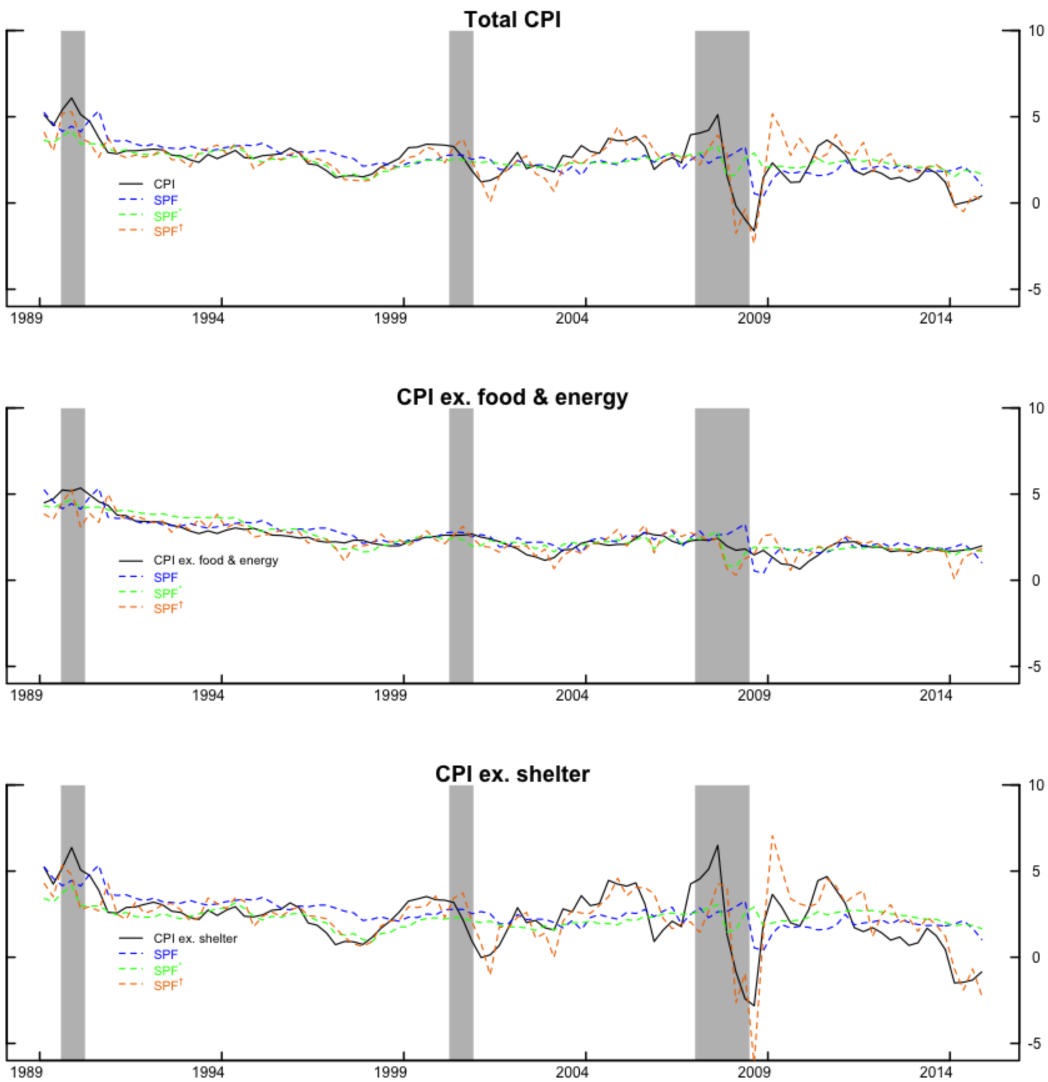
Notes: Root mean squared forecast error shown for AO model, all other models show RMSE relative to *AO* model, $\frac{RMSE^j}{RMSE^{AO}}$ for model j. Asterisks indicate statistical significance of Diebold-Mariano-West test of comparative predictive accuracy at 10 (*) and 5 (**) percent, relative to AO model. Bolded entry indicates best performing model for that price index. See text for details.

Figure 3: Out-of-sample four quarter ahead prediction from MSC and bias-adjusted MSC forecasts.



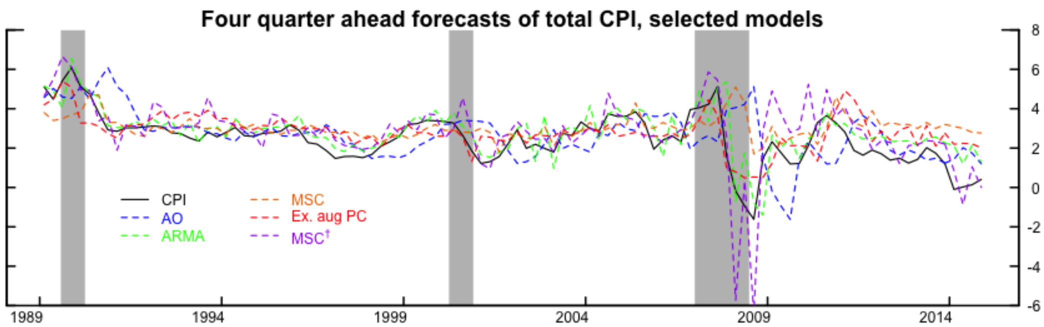
Notes: Out-of-sample predictions produced beginning in 1990Q1. See text for details.

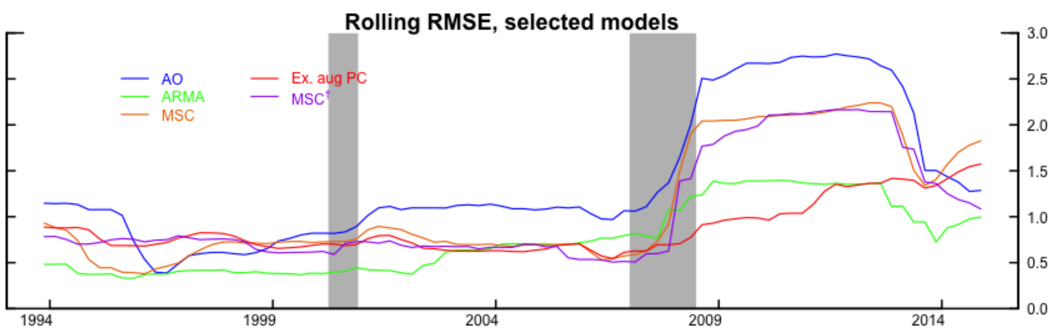
Figure 4: Out-of-sample four quarter ahead prediction from SPF and bias-adjusted SPF forecasts.



Notes: Out-of-sample predictions produced beginning in 1990Q1. See text for details.

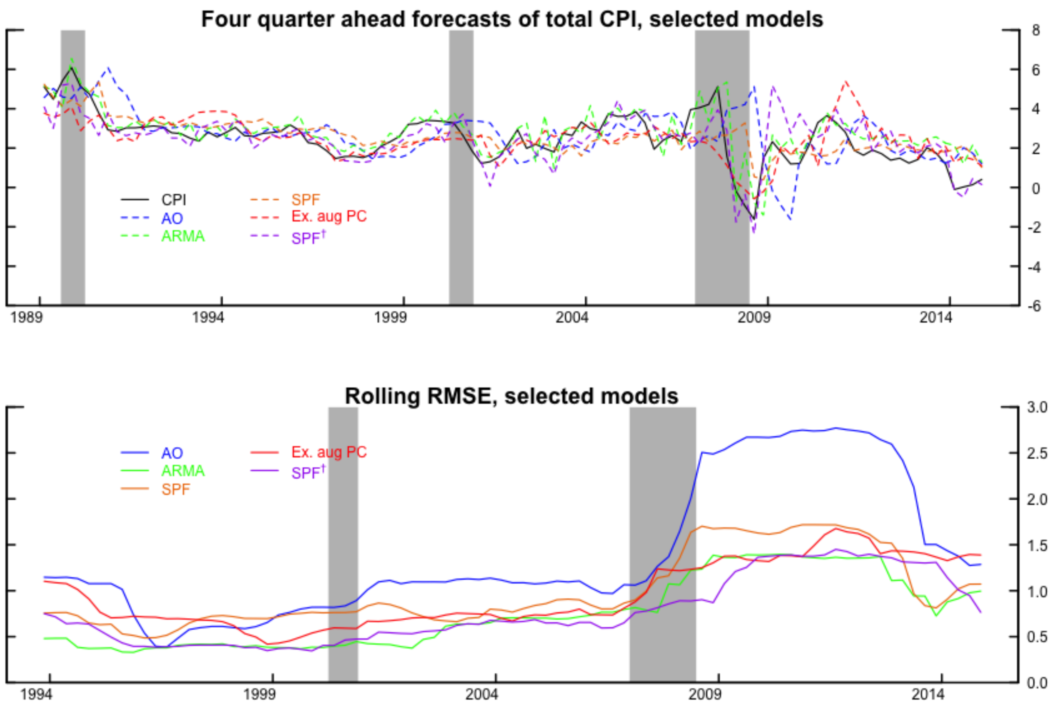
Figure 5: Out-of-sample four-quarter ahead prediction from selected models.





Notes: Top panel shows out-of-sample predictions beginning in 1990Q1. Bottom panel shows 20-quarter rolling root mean squared forecast errors of the selected models. See text for details.

Figure 6: Out-of-sample four-quarter ahead prediction from selected models.



Notes: Top panel shows out-of-sample predictions beginning in 1990Q1. Bottom panel shows 20-quarter rolling root mean squared forecast errors of the selected models. See text for details.

unemployment rate is not heavily revised, compared to the CBO measures and labor's share of income, the SW unemployment rate gap may be a useful covariate when producing inflation projections.

The univariate time-series models are consistently among the best forecasters. The ARMA(1,1) model is among the best performers when forecasting four quarters ahead and is the best performer when forecasting eight quarters ahead for all three price indices. The four quarter ahead RMSFEs for the ARMA(1,1) model are 0.95, 0.39 and 1.30 percentage point for total CPI, CPI less food and energy, and the CPI excluding shelter, respectively.

Although the AO random walk model is outperformed over the 1990-2015 sample by alternative time-series models, the Atkeson & Ohanian (2001) results that it is difficult to outperform a random walk forecast can be seen in figures 5 and 6. The bottom panels of these figures show, for a selection of models, the root mean squared forecast error, calculated with a 20-quarter rolling window. The random walk (the blue lines) performs well early in the sample—the sample at which Atkeson & Ohanian (2001) analyze—but poorly during the 2001 and 2008-2009 recessions. In contrast, the deterioration in performance of the ARMA(1,1) model is much less pronounced. Models that condition on a measure of slack perform best, relative to other models, during these periods of turmoil. This can be most clearly seen in figures 5 and 6, where the red line denotes the Phillips curve that conditions on MSC and the Stock-Watson unemployment rate gap. It is notable that during the relatively tranquil periods of the 1990s and mid-2000s, this Phillips curve model does not stand out from the others. However, during the most recent recession the model does best; although the RMSFE of the model becomes larger during this period, it does much better than the other models shown in the figures.

5 Discussion

While the analysis performed here is purely of reduced form, the results have clear implications for models that deviate from the rational expectations paradigm. The evidence presented here suggests that households and professionals have different models with which they build inflation expectations. Inflation expectations of households are influenced by recent movements in the prices salient to them. Although many recent studies have emphasized the effect of energy price movements on expectations (Binder, 2016b; Cao & Shapiro, 2016), the results here suggest that food prices also play an important role in the inflation expectations of households, consistent with the recent findings of Schulhofer-Wohl & Kaplan, 2016. Household expectations are also influenced by macroeconomic conditions, most notably income and their own sentiment. In contrast, professional forecasters are much more heavily influenced by macroeconomic indicators, especially interest rates. This result consistent with the findings of, for example, Kiley (2009), who suggests that the stance of monetary policy is an important determinant of expectations, as well as Fuhrer (2015), who shows that expectations have important effects on future macroeconomic conditions. At the least, the results show that expectations are highly endogenous to many macroeconomic variables, underlining the difficulties faced when structurally modeling expectations in a macroeconomic model.

The analysis indicates that deviations from full rationality are common, but not economically meaningful, at least for professional forecasters. A pseudo real-time forecast experiment confirmed that surveys contain important information about future realized inflation. Bias-adjusting the expectations of households improved the forecast performance of the survey, reducing the root mean-squared forecast error by about 30 percent. In contrast, bias-adjusting the SPF did not result in large improvements in forecast ability. Further, the improvement in forecast ability from the household survey disappears once the forecast model also conditions on an indicator of macroeconomic slack. On the whole, these results suggest that although the surveys do not strictly adhere to rationality, the economic significance of the deviation therefrom is minor.

A few other results from the forecast experiment are worth mentioning. Simple univariate time series models proved difficult to outperform in terms of forecast ability. In contrast to Ang et al. (2007), the ARMA(1,1) model was the best-performing model, producing more accurate forecasts than surveys. Conditioning on slack improved the forecasts of Phillips curve models relative to other forecasts, but only during periods of economic distress. No measure of slack considered here produced forecasts that were clearly superior to other Phillips curve specifications, although given its relative robustness to vintage revisions, the Stock & Watson unemployment rate gap appears to be a useful covariate in Phillips curve models that forecast inflation.

References

Adam, K. & Padula, M. (2011), 'Inflation Dynamics and Subjective Expectations in the United States', *Economic Inquiry* **49**(1), 13–25.

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

Atkeson, A. & Ohanian, L. E. (2001), 'Are phillips curves useful for forecasting inflation?', *Federal Reserve Bank of Minneapolis Quarterly Review* **25**(2).

Ball, L. & Mazumder, S. (2011), 'Inflation dynamics and the great recession', *Brookings Papers on Economic Activity* **42**(1 (Spring)), 337–405.

Binder, C. (2016), Inflation expectations and the price at the pump. Available at SSRN: https://ssrn.com/abstract=2784793.

Buhlmann, P. (2006), 'Boosting for high-dimensional linear models', *The Annals of Statistics* **34**(2), 559–583.

Buhlmann, P. & Hothorn, T. (2007), 'Boosting algorithms: regularization, prediction and model fitting', *Statistical Science* **22**(4), 477–505.

Buhlmann, P. & Yu, B. (2003), 'Boosting with the l2 loss: Regression and classification', *Journal of the American Statistical Association* **98**, 324–340.

Cao, Y. & Shapiro, A. H. (2016), 'Energy's impact on inflation expectations', *Federal Reserve Bank of San Francisco Economic Letter* (19).

Carroll, C. D. (2003), 'Macroeconomic expectations of households and professional forecasters', *The Quarterly Journal of Economics* **118**(1), 269–298.

Clark, T. E. & Davig, T. (2011), 'Decomposing the declining volatility of long-term inflation expectations', *Journal of Economic Dynamics and Control* **35**(7), 981–999.

Clark, T. E. & Doh, T. (2014), 'Evaluating alternative models of trend inflation', *International Journal of Forecasting* **30**(3), 426–448.

Cogley, T. & Sargent, T. J. (2005), 'Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S', *Review of Economic Dynamics* **8**(2), 262–302.

Cogley, T. & Sbordone, A. M. (2008), 'Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve', *American Economic Review* **98**(5), 2101–2126.

Coibion, O. & Gorodnichenko, Y. (2012), 'What can survey forecasts tell us about informational rigidities?', *Journal of Political Economy* **120**(1), 116–159.

Coibion, O. & Gorodnichenko, Y. (2015*a*), 'Information rigidity and the expectations formation process: A simple framework and new facts', *American Economic Review* **105**(8), 2644–2678.

Coibion, O. & Gorodnichenko, Y. (2015*b*), 'Is the phillips curve alive and well after all? inflation expectations and the missing disinflation', *American Economic Journal: Macroeconomics* **7**(1), 197–232.

Croushore, D. (2010), 'An evaluation of inflation forecasts from surveys using real-time data', *The B.E. Journal of Macroeconomics* **10**(1), 1–32.

D'Amico, S., Kim, D. H. & Wei, M. (2014), Tips from tips: The informational content of treasury inflation-protected security prices, Finance and Economics Discussion Series 2014-24, Board of Governors of the Federal Reserve.

Del Negro, M. & Eusepi, S. (2011), 'Fitting observed inflation expectations', *Journal of Economic Dynamics and Control* **35**(12), 2105–2131.

Diebold, F. X. & Mariano, R. S. (1995), 'Comparing Predictive Accuracy', *Journal of Business & Economic Statistics* **13**(3), 253–263.

Ehrmann, M., Pfajfar, D. & Santoro, E. (2015), Consumers' Attitudes and Their Inflation Expectations, Finance and Economics Discussion Series 2015-15, Board of Governors of the Federal Reserve System (U.S.).

Faust, J. & Wright, J. H. (2013), Forecasting inflation, *in* 'Handbook of Economic Forecasting', Vol. 2, Elsevier, chapter Chapter 1, pp. 2–56.

Friedman, J. (2001), 'Greedy function approximation: A gradient boosting machine', *The Annals of Statistics* **29**(5).

Friedman, J., Hastie, T. & Tibshirani, R. (2000), 'Additive logistic regression: A statistical view of boosting', *The Annals of Statistics* **28**, 337–374.

Fuhrer, J. (2012), 'The role of expectations in inflation dynamics', *International Journal of Central Banking* **8**(S1), 137–174.

Fuhrer, J. (2015), Expectations as a source of macroeconomic persistence: An exploration of firms' and households' expectation formation. Federal Reserve Bank of Boston Working Paper 2015-5.

Galì, J. & Gertler, M. (1999), 'Inflation dynamics: A structural econometric analysis', *Journal of Monetary Economics* **44**, 195–222.

Groen, J., Paap, R. & Ravazzolo, F. (2013), 'Real-Time Inflation Forecasting in a Changing World', *Journal of Business & Economic Statistics* **31**(1).

Grothe, M. & Meyler, A. (2015), Inflation forecasts: Are market- based and survey-based measures informative?, Working Paper 1865, European Central Bank.

Gurkaynak, R. S., Sack, B. & Wright, J. H. (2010), 'The tips yield curve and inflation compensation', *American Economic Journal: Macroeconomics* **2**(1), 70–92.

Hastie, T., Tibshirani, R. & Friedman, J. (2001), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer.

Johannsen, B. K. (2014), Inflation experience and inflation expectations: Dispersion and disagreement within demographic groups, Finance and Economics Discussion Series 2014-89, Federal Reserve Board.

Kiley, M. T. (2009), 'Inflation Expectations, Uncertainty, the Phillips Curve, and Monetary Policy', Finance and Economics Discussion Series.

Kozicki, S. & Tinsley, P. (2005), 'What do you expect? Imperfect policy credibility and tests of the expectations hypothesis', *Journal of Monetary Economics* **52**(2), 421–447.

Malmendier, U. & Nagel, S. (2016), 'Learning from inflation experiences', *Quarterly Journal of Economics* .

Mankiw, N. G. & Reis, R. (2002), 'Sticky Information Versus Sticky Prices: A Proposal To Replace The New Keynesian Phillips Curve', *The Quarterly Journal of Economics* **117**(4), 1295–1328.

Mankiw, N. G., Reis, R. & Wolfers, J. (2004), Disagreement about inflation expectations, *in* 'NBER Macroeconomics Annual 2003, Volume 18', NBER Chapters, National Bureau of Economic Research, Inc, pp. 209–270.

Mehra, Y. P. (2002), 'Survey measures of expected inflation: Revisiting the issues of predictive content and rationality', *Federal Reserve Bank of Richmond, Economic Quarterly* **88**(3).

Mincer, J. A. & Zarnowitz, V. (1969), The Evaluation of Economic Forecasts, *in* 'Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance', NBER Chapters, National Bureau of Economic Research, Inc, pp. 3–46.

Nordhaus, W. D. (1987), 'Forecasting efficiency: Concepts and applications', *The Review of Economics and Statistics* **69**(4), 667–674.

Pfajfar, D. & Santoro, E. (2013), 'News on Inflation and the Epidemiology of Inflation Expectations', *Journal of Money, Credit and Banking* **45**(6), 1045–1067.

Roberts, J. (1995), 'New keynesian economics and the phillips curve', *Journal of Money, Credit, and Banking* **27**(4).

Roberts, J. (1998), Inflation expectations and the transmission of monetary policy, Finance and Economics Discussion Series 1998-43, Board of Governors of the Federal Reserve System.

Schapire, R. E. (1990), 'The strength of weak learnability', *Machine Learning* **5**, 197–227.

Schulhofer-Wohl, S. & Kaplan, G. (2016), Inflation at the Household Level, 2016 Meeting Papers 529, Society for Economic Dynamics.

Sims, C. A. (2003), 'Implications of rational inattention', *Journal of Monetary Economics* **50**, 665–690.

Stock, J. H. & Watson, M. W. (2007), 'Why has u.s. inflation become harder to forecast?', *Journal of Money, Credit and Banking* **39**(1).

Stock, J. H. & Watson, M. W. (2010), Modeling inflation after the crisis, NBER Working Paper Series 16488, National Bureau of Economic Research.

Thomas, L. B. (1999), 'Survey measures of expected u.s. inflation', *Journal of Economic Perspectives* **13**(4), 125–144.

Wright, J. H. (2009), 'Forecasting u.s. inflation by bayesian model averaging', *Journal of Forecasting* **28**, 131–144.

Appendix

	Michigan Survey of Consumers	**Livingston Survey**	**Survey of Professional Forecasters**
Organization	University of Michigan	FRB-Philadelphia	FRB-Philadelphia
Survey population	Households	Academic and professional economists	Professional economists
Number of respondents	500-700	14-63 (48 on average)	9-38 (34 on average)
Survey availability	Quantitative responses: January 1978-present	1946-present	Total CPI: 1981Q3
Frequency of survey	Quarterly from 1947 to 1977, monthly 1978-present	Semi-annual	Quarterly
Timing of survey	Telephone survey throughout month	Month of June, December	Responses due second to third week of the middle month of quarter.
Inflation measure	Expected change in prices during the next 12 months	CPI level (inflation rate imputed)	GDP deflator; more recently CPI and PCE
Inflation horizon	One year ahead	Several	Quarterly up to 6 quarters ahead
Website	`www.um.edu`	`www.frbphilly.org`	`www.frbphilly.org`

Table A1: Survey measures of inflation expectations

Price measures		
Name	Transformation	Source
Total CPI price index	$400\times\log(P_t/P_{t-1})$	BLS
CPI excluding food and energy	$400\times\log(P_t/P_{t-1})$	BLS
CPI excluding shelter	$400\times\log(P_t/P_{t-1})$	BLS
Energy services	$400\times\log(P_t/P_{t-1})$	BLS
Energy commodities	$400\times\log(P_t/P_{t-1})$	BLS
Food away from home	$400\times\log(P_t/P_{t-1})$	BLS
Food at home	$400\times\log(P_t/P_{t-1})$	BLS
Medical commodities	$400\times\log(P_t/P_{t-1})$	BLS
Medical services	$400\times\log(P_t/P_{t-1})$	BLS
Used vehicles	$400\times\log(P_t/P_{t-1})$	BLS
New vehicles	$400\times\log(P_t/P_{t-1})$	BLS
Apparel	$400\times\log(P_t/P_{t-1})$	BLS
OER	$400\times\log(P_t/P_{t-1})$	BLS
Rent	$400\times\log(P_t/P_{t-1})$	BLS
Tobacco	$400\times\log(P_t/P_{t-1})$	BLS
Alcohol	$400\times\log(P_t/P_{t-1})$	BLS
Transportation services	$400\times\log(P_t/P_{t-1})$	BLS
Fuel	$400\times\log(P_t/P_{t-1})$	BLS
Macroeconomic indicators		
IP by sector	$400\times\log(x_t/x_{t-1})$	FRB
Housing starts by region	$400\times\log(x_t/x_{t-1})$	U.S. Census
Housing permits by region	$400\times\log(x_t/x_{t-1})$	U.S. Census
Personal consumption expenditures: services	$400\times\log(x_t/x_{t-1})$	BEA
Personal consumption expenditures: non-durables	$400\times\log(x_t/x_{t-1})$	BEA
Personal consumption expenditures: durables	$400\times\log(x_t/x_{t-1})$	BEA
Nominal and real personal income	$400\times\log(x_t/x_{t-1})$	BEA
Initial claims	$400\times\log(x_t/x_{t-1})$	BEA
Unemployment by duration	N.A.	BLS
Employment by sector	$400\times\log(x_t/x_{t-1})$	BLS
Labor force participation rate	N.A.	BLS
Average hourly earnings by sector	$400\times\log(x_t/x_{t-1})$	BLS
Average weekly hours	N.A.	BLS
ISM manufacturing surveys by sector	N.A.	ISM
Financial indicators		
Dollar exchange rate against major currencies	$400\times\log(x_t/x_{t-1})$	FRED
Fed funds rate and Treasury yields (3-mo., 2y, 5y, 10y)	N.A.	FRED
Treasury yield-fed funds rate spreads (3-mo., 2y, 5y, 10y)	N.A.	FRED
Moody's corporate interest rates (AAA, BAA)	N.A.	FRED
Moody's corporate interest rate spreads (AAA, BAA)	N.A.	FRED
Commercial and industrial loans	$400\times\log(x_t/x_{t-1})$	FRED
Real estate loans	$400\times\log(x_t/x_{t-1})$	FRED
Total reserves of depository institutions	$400\times\log(x_t/x_{t-1})$	FRED
Consumer Sentiment	N.A.	UM/Reuters

Table A2: Data description. An complete listing can be found at the FRED website, https://research.stlouisfed.org/pdl/788.

Slack measures used in Phillips curve models		
Name	Definition	Source
CBO output gap	Real GDP less real potential GDP	CBO
CBO unemployment rate gap	Unemployment rate less natural rate of unemployment	CBO
Stock-Watson unemployment rate gap	Unemployment rate less minimum unemployment rate from previous three years	Stock & Watson (2010)
Labor's share of income	Log labor's share of income, nonfarm business sector	BLS

Table A3: Description of measures of slack used in Phillips curve models.