MEASURING SENTIMENT SHOCKS IN PROFESSIONAL SURVEY DATA AND THEIR ROLE IN AGGREGATE FLUCTUATIONS*

Rupal Kamdar† Christian Matthes‡ Anthony Vecchia§

May 16, 2026

Abstract

Professional forecasts contain common sentiment components that drive aggregate fluctuations but are not related to existing measures of consumer sentiment. Using Survey of Professional Forecasters data, we remove variation explained by macroeconomic factors and identified structural shocks, then extract factors from the residual forecast component. This approach reveals a substantial factor structure: the first three factors explain roughly a third of residual forecast variation. These factors can be interpreted as shocks that, in a VAR, explain 10 to 15% of medium-horizon aggregate fluctuations in real GDP growth, PCE inflation, the unemployment rate, and the federal funds rate.

JEL Codes: C38, E32, E70
Keywords: surveys, sentiment, factor analysis, business cycles

*We thank Rudi Bachmann and Jane Ryngaert for helpful comments.
†Department of Economics, Indiana University, rkamdar@iu.edu
‡Department of Economics, University of Notre Dame, cmatthes@nd.edu
§Department of Economics, University of Notre Dame, avecchia@nd.edu

1 Introduction

Surveys of professional forecasters provide key empirical measures of informed macroeconomic beliefs and, as such, are closely watched by markets and policymakers. Comovement in central tendencies of these forecasts is often interpreted as reflecting shared responses to economic fundamentals. However, common movements do not necessarily reflect only disciplined responses to macroeconomic fundamentals. Forecasters may also share changes in outlook that are not fully explained by observable conditions or standard structural shocks. In this paper, we ask whether professional survey forecasts contain such residual common movements, which we call "sentiment shocks".

We measure these shocks as common variation in professional forecasts that remains once we account for macroeconomic fundamentals and identified structural shocks. First, we extract common factors from a large set of macroeconomic variables. Second, we partial out these factors, as well as a large variety of structural shocks identified in the existing literature, from the mean Survey of Professional Forecasters (SPF) beliefs for a wide array of macroeconomic variables. Finally, we conduct principal component analysis on the residuals, which represent the remaining unexplained variation in the mean expectations of professional forecasters. We find that the residuals are low-dimensional, with the first three factors together explaining roughly a third of the variation in the residual SPF forecasts. The time series of the factors does not obviously covary with the business cycle; this may not be surprising because we purge macroeconomic fundamentals from the SPF data. Yet we still find a strong factor structure in the residuals. That is, there are common movements in SPF forecasts that cannot be explained by observables.

These sentiment factors are not merely a statistical summary of forecast residuals. When we include them in a VAR, they explain a sizable share of the variation in macroeconomic and financial variables. Their contribution is especially large for the federal funds rate, for which the three shocks together account for nearly 15 percent of the forecast error variance at medium horizons. More broadly, the three factors explain approximately 10 percent of the variation in inflation, unemployment, output growth, consumption, investment, housing starts, and term and credit spreads. In other words, our identified sentiment shocks affect professional forecasts and macroeconomic dynamics.

This link between beliefs and aggregates is economically meaningful because professional forecasts are not merely measured expectations; they are inputs into policy, financial-market, hiring, wage-setting, and investment decisions (Croushore, 1993). If professional forecasters share a residual change in outlook after conditioning on macroeconomic fundamentals and known shocks, that component can therefore matter for aggregate dynamics even when it is not captured by standard consumer sentiment measures. Firm-level studies provide further evidence that professional forecasts are used in firms' decision-making. Inflation-expectation treatments shift firms' prices, credit demand, employment, and capital; CFO earnings expectations explain both planned and realized investment; and firms' macro-forecasts predict subsequent employment, investment, and output growth (Coibion et al., 2018, 2020; Gennaioli et al., 2016; Tanaka et al., 2020).

The main empirical results—both the strong factor structure in residual forecasts and the importance of the resulting sentiment shocks for macroeconomic dynamics—are robust to a wide range of alternative specifications. The results are broadly similar under alternative choices for the number, timing, and lag length of macroeconomic factors used when we regress SPF beliefs on macro-factors and known shocks (the second step of the empirical approach). The results are also robust to including standard measures of sentiment (University of Michigan Sentiment Index, Conference Board Confidence Index, and the main business-cycle shock from Angeletos et al. (2020)) as regressors. These results suggest that we are not capturing known variation in expectations, but rather something new and economically important. We also conduct robustness checks around the VAR choices, including specifying variables in log levels rather than growth rates or using OLS for estimation. Again, we find no notable differences from the baseline.

This paper contributes to two areas of work. First, our measurement strategy builds on the literature that uses factor methods to summarize information in large macroeconomic datasets. Stock and Watson (2002) show how principal components can extract common information from many predictors, and McCracken and Ng (2020) create a standardized dataset that is commonly used to extract factors for the US economy (the FRED-QD dataset). Our approach to assessing the importance of the sentiment shocks draws on Bernanke et al. (2005), who use estimated factors in a VAR to account for information omitted from small macroeconomic systems. The contribution is to apply factor methods to residual professional forecasts, not to macroeconomic observables themselves. That is, we apply principal components to the *residual*, unexplained variation in professional forecasts. In this sense, we use factor methods to isolate a latent component of professional beliefs that is distinct from standard observable fundamentals.

Second, this paper contributes to the literature on sentiment- and belief-driven fluctuations. Angeletos and La'o (2013) provide a theory in which shifts in expectations about economic activity can drive fluctuations without changes in fundamentals, while Milani (2017) estimates the empirical contribution of sentiment shocks in a medium-scale model with observed expectations. Angeletos et al. (2020) identify a main business-cycle driver and provide a useful benchmark for comparing belief-driven shocks to other propagation mechanisms. Our paper also relates to recent work documenting low-dimensional structure in beliefs, including consumer beliefs (Kamdar and Ray, 2024; Ferreira and Pica, 2024), professional-forecast disagreement (Herbst and Winkler, 2025), and perceived shocks inferred from expectation revisions (Giacomini et al., 2024). Relative to this work, we extract factors from residualized average SPF forecasts, which lets us interpret the common component as professional sentiment not determined by current macroeconomic conditions or contemporaneous shocks identified elsewhere.

2 Construction of Sentiment Shocks as Factors of Residuals

We obtain sentiment shocks by extracting factors from survey forecasts after residualizing them on macroeconomic conditions in a three-step process. First, we extract factors from a large set of macroeconomic observables. Second, we include these first-step factors as regressors in predictive regressions for the survey data. Finally, we extract factors from the residuals of the second-step regressions. In this section, we describe this three-step process in further detail, then discuss the data, and finally present the baseline results.

2.1 THREE-STEP PROCESS

Let Y_t contain a large set of macroeconomic observables, which we demean and normalize to unit variance. In the first step, we extract N factors from Y_t using principal component analysis (PCA). We choose N using a standard information criterion.[1] Let the nth chosen factor be F_{nt}.

The goal of the second step is to orthogonalize mean (or median) survey forecasts to macroeconomic conditions represented by all factors F_{nt} as well as additional controls X_{mt}. We use these factors rather than the raw observables in Y_t to obtain a parsimonious regression specification. Let S_t^{ih} be the ith forecast variable at horizon h. We run the following regression via OLS separately for each variable i and horizon h:

$$S_t^{ih} = \alpha^{ih} + \sum_{n=1}^{N}\sum_{j=0}^{J}\beta_{nj}^{ih}F_{n,t-j} + \sum_{k=1}^{K}\gamma_k^{ih}S_{t-k}^{ih} + \sum_{m=1}^{M}\delta_m^{ih}X_{mt} + u_t^{ih}, \tag{1}$$

where J and K are lag lengths, both set to four as a baseline, and X_{mt} is the mth variable in a set of M controls, which are instruments for structural macroeconomic shocks in our baseline specification (described below). We demean F_{nt} and X_{mt} and normalize them to unit variance. We choose $N = 6$ as our benchmark, as suggested by the information criteria used by McCracken and Ng (2020). In total, there are $I \times H$ dependent variables (and residuals).

In the third and final step, we assess whether a factor structure exists in the survey forecasts after controlling for observable conditions. Here, we collect all u_t^{ih} into U_t, normalize U_t to unit variance, and extract factors F_t^* using PCA. These F_t^* are the factors of interest. Any substantial structure in F_t^* summarizes relevant information contained in the set of forecasts beyond macroeconomic realizations. We call this information sentiment.

2.2 DATA

We use macroeconomic observables from the FRED-QD dataset and forecasts from the Survey of Professional Forecasters (SPF), both at the quarterly frequency. In the baseline specification, we end both samples in 2019Q4. Appendix A contains details on data sources and management not mentioned in the text below. FRED-QD includes 245 major series from the FRED database that are updated and maintained in real time.[2] Of these 245 series, we drop those that either contain substantial missing observations, are broadly considered asset prices (e.g., bond yields, stock indices, home prices), or are survey sentiment indices (e.g., University of Michigan consumer sentiment, banks tightening lending standards). Details on the dropped variables can be found in Appendix A. We later show that our results are robust to the inclusion of these variables. Each of the kept series is transformed to be stationary according to the choices provided in McCracken and Ng (2020). In the first step, we extract factors using data back to 1968Q1. In the second step, the sample starts in 1981Q3 such that we can include all the SPF variables listed below.

We use the mean forecasts[3] for all variables available beginning on or before 1981Q3 in the SPF: chain-

[1] Let Y be $T \times M$. The optimal number of factors N is chosen as $\arg\min_{n\in\{1,\dots,\bar{N}\}} \log(\frac{1}{M}\sum_{m=1}^{M}\frac{e_{nm}'e_{nm}}{T}) + \frac{M+T}{MT}n\log(\min(M,T))$ where e_{nm} is the prediction error for the mth variable when using n factors. We inherit this choice of factors from McCracken and Ng (2020).

[2] See McCracken and Ng (2020) for more details.

[3] Results are robust to using median forecasts, as we show later. Using individual-level forecasts is infeasible because we rarely observe long time series of consecutive forecasts from the same forecaster.

weighted GDP price index (PGDP), nominal corporate after-tax profits (CPROF), the unemployment rate (UNEMP), industrial production (INDPROD), housing starts (HOUSING), real GDP (RGDP), the three-month Treasury rate (TBILL), Moody's AAA corporate bond yield (BOND), real personal consumption expenditures (RCONSUM), real non-residential investment (RNRESIN), real residential investment (RRESINV), real federal government consumption and gross investment (RFEDGOV), real state and local government consumption and gross investment (RSLGOV), real change in private inventories (RCBI), real net exports (REXPORT), and CPI (CPI). We use forecasts for all variables from the current quarter (which the SPF indexes by 2) to four quarters ahead (indexed by 6). Here, we keep the naming and indexing convention in the SPF data for simplicity. Each mean SPF variable is transformed similarly to its corresponding FRED-QD variable.[4] In total, we use 80 SPF variables ($I = 16$ economic variables with $H = 5$ horizons each).

We also gather 48 instruments for structural shocks common in the macroeconomics literature, compiled by Adams and Barrett (2025), to use as contemporaneous controls in the baseline second-step regressions. We include these instruments to control for contemporaneous movements in variables that might not be captured by the macroeconomic factors but are also distinctly not caused by sentiment shocks. The compilation includes instruments for conventional and unconventional monetary policy (18), fiscal policy (16, e.g., consumption, investment, and taxes), technology (6), oil (5), inflation sentiment (1), severe weather (1), and uncertainty (1) shocks. The instruments are derived using a variety of methods, mostly structural vector autoregressions, high-frequency identification, or narratives. Monthly shocks are summed to quarterly. As a baseline, we harmonize the sample lengths of these shocks with the SPF variables by setting missing observations to zero. Results are robust to filling in missing values using a state-space model from McCracken and Ng (2020) (see Section 4).

Six macroeconomic factors are chosen in step one and included contemporaneously and with lags as independent variables alongside the structural shocks. These six factors cumulatively explain 48.6% of the variance in the FRED-QD data (21.7% from the first factor alone). We compute the F-statistics for the six macroeconomic factors and their lags (excluding coefficients on the lags of the dependent variable and structural shocks). For nearly all dependent variables in the second stage, the F-statistic is large and implies statistically significant dependence of the forecast variables on these factors and lags collectively. Exhibit 6 in Appendix B summarizes these F-statistics and R^2s from the second-step regressions. Importantly, the second-step regressions generally fit very well. This implies that professional forecasts are usually reasonably well explained by macroeconomic observables. However, as we show in Section 3, the resulting sentiment shocks based on the residuals of those regressions still substantially affect macroeconomic variables.

2.3 Baseline Specification Results

We now describe the results from the baseline specification described above. In Section 4, we discuss robustness to alternative data handling, lag lengths, factor

[4] To avoid inconsistencies of base years or scaling across surveys, we use the prior quarter realization posted *within* surveys when transforming SPF variables. For example, the log prior quarter realization of real GDP is subtracted from the log mean forecast at each horizon. This prior quarter realization is given to the forecasters in the survey and matches the units of their forecasts.

selection, and more.

Exhibit 1a summarizes the residual SPF factors F_t^* from the third step. The first three factors together explain roughly 31% of the variation in the residuals of the 80 SPF series; that is nearly one-third of the total variation contained in less than 4% of the factors (and nearly a quarter in just two factors). This result underscores substantial movements in forecasts *independent* of a large information set summarized with a small set of macroeconomic factors and instruments for structural shocks. We interpret the variation contained in the factors as pure forecaster sentiment. Importantly, since we use mean forecasts (and median forecasts in robustness checks), our notion of sentiment is one of average sentiment, not driven by outliers. We investigate the macroeconomic importance of these first three factors in Section 3. The variance explained by the next two factors trails off at roughly 5–6% each (not shown in the table).

Exhibit 1a also lists the dependent variables associated with the residuals with the largest factor loadings in absolute value for each factor (remember that all these variables are transformed to be stationary and standardized). Forecast horizons for the unemployment rate load most heavily onto the first factor, and housing starts and residential investment load most heavily onto the second. Thus, unexplained variation in labor and housing market forecasts is likely a key component of forecaster sentiment. Of particular relevance for the following section, the three-month Treasury yield loads most heavily onto the second factor (as opposed to the AAA corporate yield loading heavily onto the first). Horizons of net exports dominate the loadings of the third factor.

Exhibit 1b plots the three factors described above. These factors represent information independent of other major shocks and indices, as well as macroeconomic observables. The factors exhibit no significant correlation with the structural shocks included in the second-step regressions by construction, but also show no significant correlation with consumer sentiment as measured by the University of Michigan and Conference Board indices (as shown in Exhibit 1a). They also show no significant autocorrelation structure. Thus, we interpret these factors as sentiment shocks distinct from standard measures of consumer sentiment. Interestingly, the factors do not show unusual movements at the onset of recessions. So, while they explain a meaningful fraction of the variance of many macroeconomic variables (as shown below), we do not necessarily think of them as causing business cycles.

3 The Importance of Sentiment Shocks

The forecaster sentiment factors extracted from the process above demonstrate that information in forecasts beyond observables can be reasonably summarized in a low-dimensional vector. In this section, we ask whether this information is relevant for macroeconomic conditions. We include the first three sentiment factors in a VAR model with key macroeconomic and financial indicators, interpret them as structural shocks, and assess their importance via variance decompositions. The sentiment shocks explain a sizable share of important business cycle indicators, usually above ten percent for most horizons.

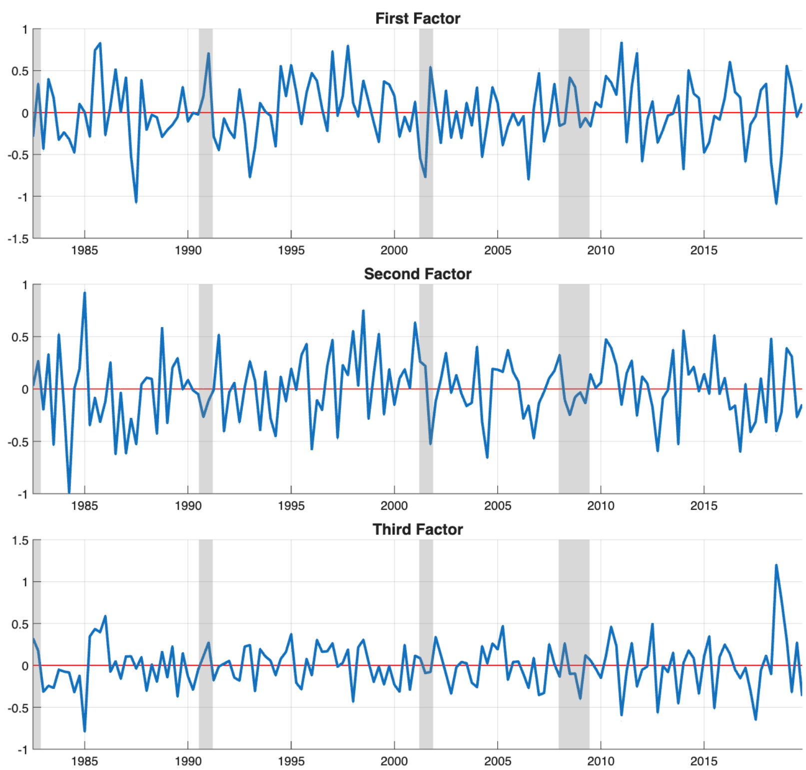
We identify structural sentiment shocks using a recursive identification assumption. While we directly interpret the extracted sentiments as shocks, the derived impulse responses are also valid if we use a weaker assumption: Plagborg-Møller and Wolf (2021) show that recursive identification with an instrument ordered first is a valid method for structural VAR estimation with a proxy. Fusari et al. (2026) extend

Exhibit 1: Residual SPF Factors

(a) Explanatory Power & Loadings

Factor	1	2	3
Variance Explained			
Marginal (%)	13.72	10.02	7.31
Cumulative (%)	13.72	23.74	31.05
Correlations			
Conference Board US Consumer Confidence	-0.07 [-0.20, 0.07]	0.06 [-0.07, 0.20]	0.02 [-0.11, 0.16]
Michigan Consumer Sentiment	-0.06 [-0.19, 0.08]	0.04 [-0.10, 0.17]	-0.00 [-0.14, 0.13]
Top Loadings			
1.	UNEMP4 (1.83)	TBILL3 (-1.77)	REXPORT2 (-3.21)
2.	UNEMP5 (1.82)	TBILL2 (-1.76)	REXPORT3 (-3.17)
3.	UNEMP3 (1.77)	HOUSING4 (1.72)	REXPORT4 (-3.15)
4.	UNEMP6 (1.75)	RRESINV3 (1.69)	REXPORT5 (-3.12)
5.	BOND4 (-1.56)	HOUSING5 (1.66)	REXPORT6 (-3.08)

(b) Time Series Plots



Notes: Panel (a) summarizes the factors extracted from the final step of the three-step process. The top portion displays the percent of the variance of the second-step residuals explained by each of the first three factors individually and cumulatively, which is computed as the corresponding eigenvalue divided by the sum of all eigenvalues. The middle portion reports correlations (with 90% confidence intervals) of each factor with consumer confidence indices. The bottom portion displays the dependent variables associated with the residuals with the largest loadings in absolute value for each factor (remember that we standardize the regression residuals before computing principal components, so these coefficients are not determined by different scales of the variables). UNEMPL denotes the forecasts for the unemployment rate; BOND Moody's AAA corporate bond yield; TBILL the three-month Treasury rate; HOUSING housing starts; RRESINV real residential investment; and REXPORT real net exports. Current-quarter forecasts are indexed by 2 and four-quarter-ahead forecasts by 6. Panel (b) plots the corresponding factors extracted in the final step of the three-step process, each normalized to unit variance. Shaded areas indicate official NBER recession dates.

this result to multiple instruments together ordered first.

We argue that our three factors collectively act as instruments for forecaster sentiment. Note that they are independent of macroeconomic observables and other shocks, as previously discussed. Thus, we order our three factors first (in their respective order).[5] They are followed by PCE inflation, the unemployment rate, real GDP growth, durable consumption growth, investment growth, housing starts growth, the BAA/AAA corporate bond spread, the 10-year/3-month Treasury spread, and the Federal Funds Effective Rate (FFR). All growth rates are quarter-over-quarter. Appendix A contains more details about the data.

We use the MATLAB package from Canova and Ferroni (2024) to estimate the VAR model via Bayesian methods. Specifically, we use a Gibbs sampler over the posterior of the reduced-form parameters. We employ the Minnesota prior with the package's default hyperparameters. The baseline estimation includes four lags of the dependent variables. All factors are normalized to unit variance before estimation.

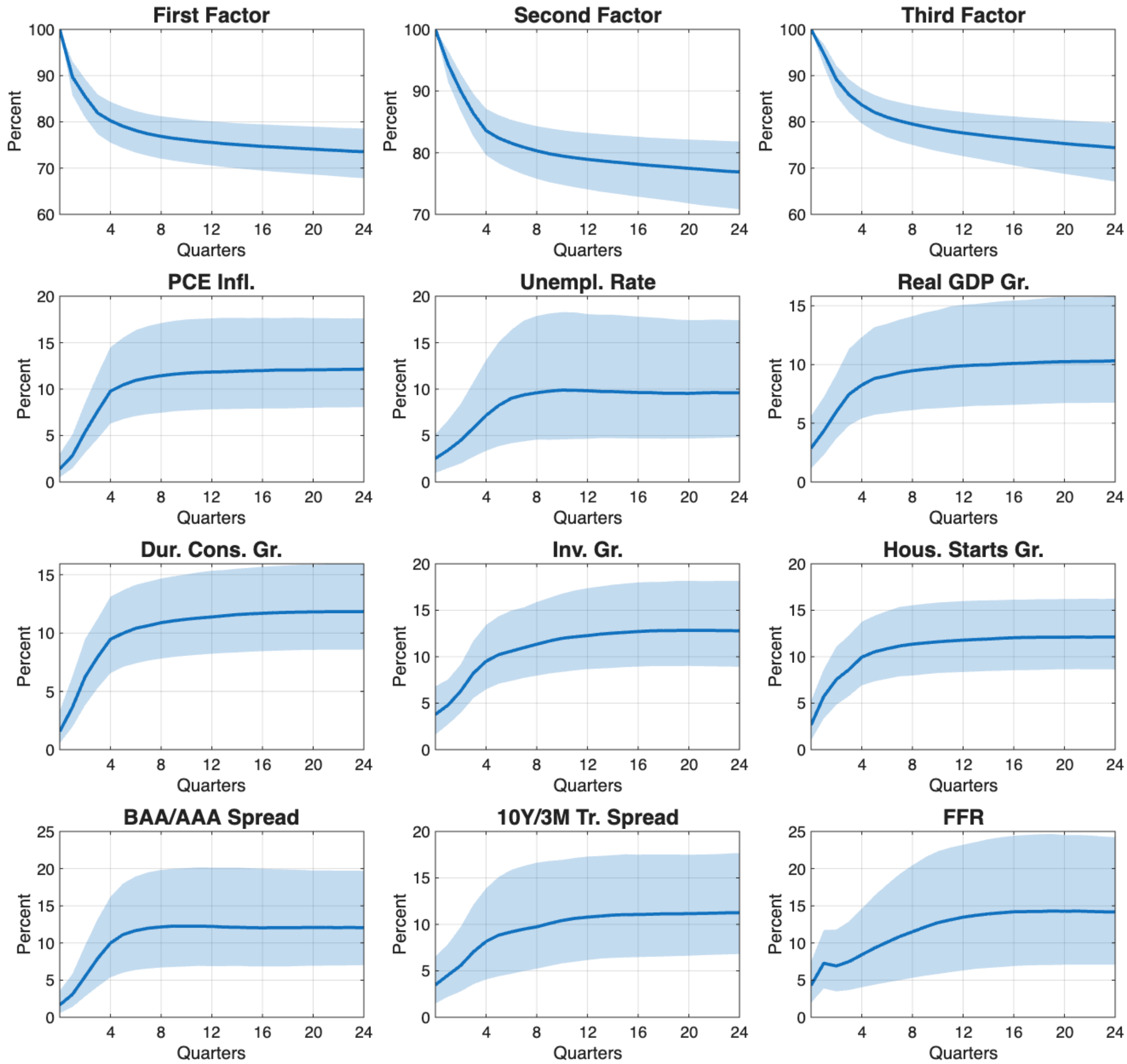
Exhibit 2 shows that the three sentiment shocks together explain a large share of variation in economic aggregates, particularly the FFR. This is the paper's main empirical result. It plots the posterior median path of the percentage of variation attributed to the three sentiment shocks collectively along with 68% probability bands. We emphasize the collective explanatory power of the sentiment shocks because the factors used for their identification together contain information about forecaster sentiment for a sizable set of forecast variables. Overall, the three shocks explain at least 10% of the variation in each economic indicator in the system after four to eight quarters. The three shocks propagate most quickly to investment, durable consumption, inflation, real GDP, and housing starts. The shocks are most important for the FFR, explaining nearly 15% of its variation after roughly twelve quarters. This suggests that forecaster sentiment is an important input to monetary policy decisions.

Exhibit 7 in Appendix B plots the paths of the percentage of variation attributed to each sentiment shock individually. These results are robust to different orderings within the factor block because the factors are not contemporaneously correlated with each other. The second sentiment shock is the largest contributor (or close to it) to the variation in every variable except durable consumption, investment, and the 10-year/3-month Treasury spread. Thus, variation in the forecast residuals does not automatically translate to high economic importance. Otherwise, we would expect the first factor to contain the highest explanatory power. The second shock is most important for the FFR, which is likely a result of second-step residuals for horizons of the three-month Treasury yield forecasts loading most heavily onto the second factor. The three shocks propagate similarly to the economy, as shown by their variance decompositions, which follow similar paths for most variables. Only the magnitudes differ.

Exhibit 3 displays the impulse responses to each sentiment shock. All three shocks reduce inflation but differ in their unemployment response: the first factor does not move the unemployment rate significantly (in the sense that zero is always contained in the 68% posterior bands), whereas the second shock moves the unemployment rate up and the third factor decreases the unemployment rate. All shocks, when normalized to increase on impact, decrease the federal funds rate, except for the impact response

[5] The order of the factors relative to each other should not impact the results under recursive identification because they are not contemporaneously correlated with each other by construction and are not substantially autocorrelated.

Exhibit 2: Variance Explained by the Three Factors Collectively

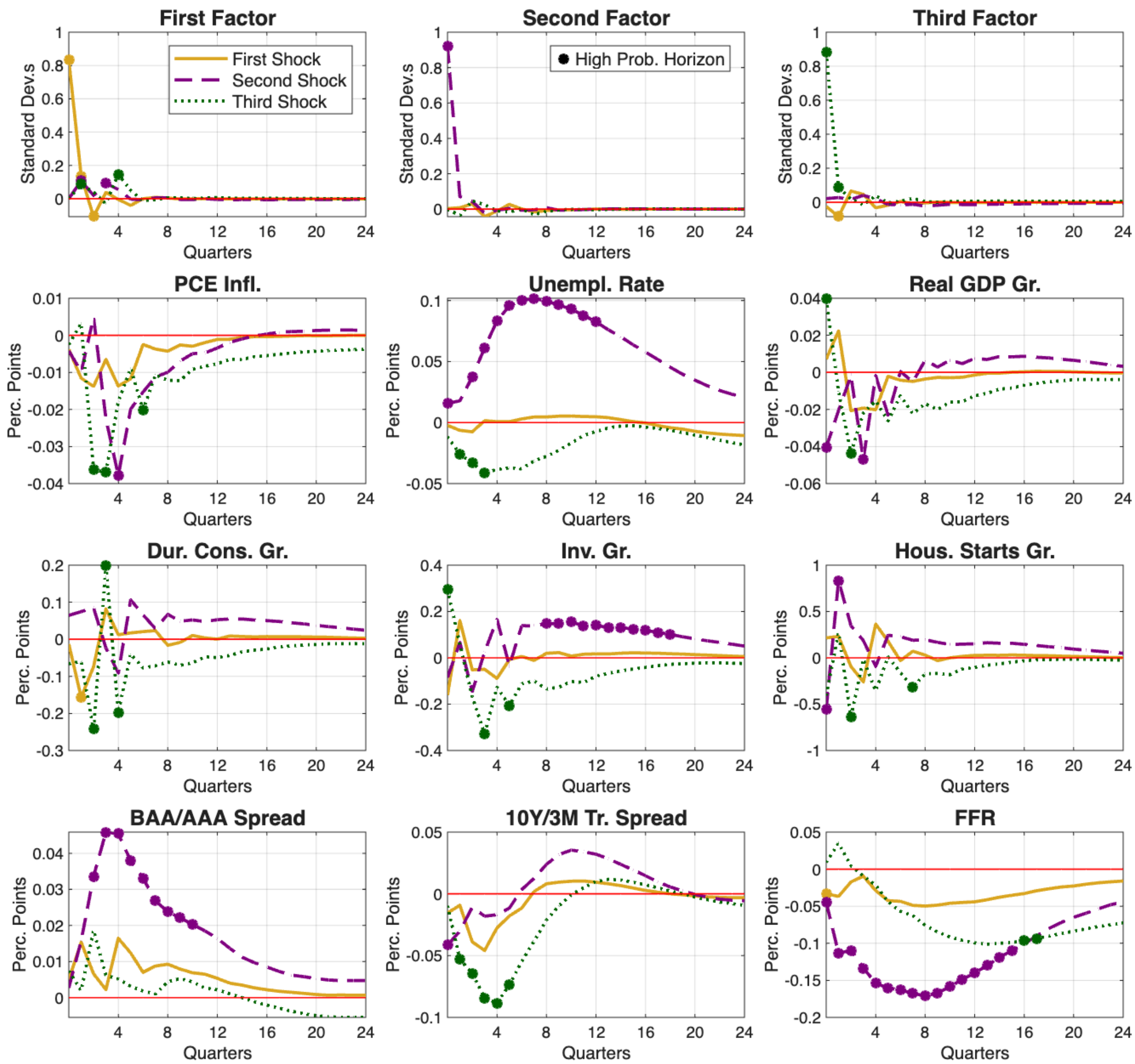


Notes: The solid blue line represents the pointwise posterior median path of the sum of the variance explained by each of the three residual SPF factor shocks in the forecast error variance decomposition. The system is identified recursively using the order in which the variables appear. The shaded area indicates the 68% probability bands based on five thousand draws from the posterior distribution.

to the third factor.

Exhibit 3 also clarifies the relation between our sentiment shocks and the main business-cycle (MBC) shock in Angeletos et al. (2020). For their shock, unemployment falls, real activity rises, inflation barely moves, and the nominal interest rate rises. In contrast, for our sentiment shocks, PCE inflation falls after all three shocks, and most responses fade within a few years. The strong real comovement present in Angeletos et al. (2020) does not appear. The first shock has little effect on unemployment and small real effects. The second shock is closest to the recessionary MBC sign in its unemployment, credit-spread, and funds-rate responses: unemployment and the BAA/AAA spread rise, while the federal funds rate

Exhibit 3: Impulse Responses to the Three Factor Shocks



Notes: IRFs to a one-standard-deviation shock to each of the three residual SPF factors identified recursively using the order in which the variables appear. The lines represent the pointwise posterior median paths. The solid yellow line corresponds to the first factor shock, the dashed purple to the second, and the dotted green to the third. The asterisks indicate horizons at which the 68% probability bands lie entirely above or below zero. All IRFs are based on five thousand draws from the posterior distribution.

falls. However, investment and housing starts rise at medium horizons. The third shock lowers unemployment, but it also lowers the federal funds rate and is followed by weaker GDP, durable consumption, investment, and housing growth. The propagation of our shock series is more heterogeneous and less tied to the positive real comovement that defines the MBC shock.

4 ALTERNATIVE CHOICES IN THE CONSTRUCTION OF SENTIMENT SHOCKS

The sentiment shocks remain economically relevant across a wide range of specification choices. Exhibit 4 summarizes these checks by reporting the average share of forecast-error variance explained by the three sentiment shocks, across horizons up to 24 quarters, for inflation, unemployment, real GDP growth, and the federal funds rate. The checks vary three parts of the procedure: the information set used in the first-step FRED-QD factor extraction, the controls used in the second-stage regressions, and the specification and estimation of the VAR.

Expanding the first-step information set to include asset prices and the Michigan Consumer Sentiment Index leaves the variance shares close to the baseline values. This check addresses whether the baseline FRED-QD factors omit financial or survey-sentiment information that could change the residual forecast factors.

The second set of checks changes the second-stage residualization. Adding the Michigan Survey Consumer Sentiment Index, the Conference Board Consumer Confidence Index, or the Angeletos et al. (2020) sentiment shock as controls has little effect on the variance shares. These results suggest that the SPF factors are not simply reproducing existing sentiment measures. The same conclusion holds when we use median SPF forecasts instead of mean forecasts, which reduces the influence of extreme survey responses.

We also vary the macroeconomic controls used in the second stage. Filling in missing values for the shock controls, doubling the number of FRED factors, and changing the timing of the FRED factors all leave a substantial role for the sentiment shocks. The lead-FRED specification is the most conservative of these checks: it controls for one-quarter-ahead movements in the macroeconomic factors, even though the baseline results imply that some of those movements may themselves respond to sentiment shocks. The sentiment shocks still explain meaningful forecast-error variance under this specification. Conversely, using only lagged FRED factors strengthens the results. This specification is no less consistent with the timing of the SPF than the benchmark choice: surveys are collected in the middle of quarter t, so forecasters cannot respond to macroeconomic data releases that arrive in the second half of the same quarter.

Finally, the results are not driven by the VAR implementation. Estimating the VAR in log levels rather than growth rates, using OLS instead of the Bayesian specification, or extending the sample through 2025Q1 does not eliminate the role of the sentiment shocks.

Exhibit 4: Summary of Robustness

	PCE Infl.	Unempl. Rate	Real GDP Gr.	FFR
Baseline	10.6	8.8	9.2	12.0
	(7.1, 15.4)	(4.7, 15.1)	(6.1, 13.6)	(6.4, 19.9)
Include Asset Prices and Sentiment (Step 1)	9.9	9.4	9.6	10.6
	(6.6, 14.4)	(4.9, 16.1)	(6.5, 14.1)	(5.6, 17.9)
Include Michigan Consumer Sentiment (Step 2)	10.7	8.2	9.1	12.8
	(7.2, 15.6)	(4.3, 14.5)	(6.1, 13.4)	(6.8, 20.9)
Include Conference Board US Consumer Confidence (Step 2)	10.9	8.2	9.2	12.6
	(7.3, 15.8)	(4.3, 14.4)	(6.1, 13.5)	(6.8, 20.6)
Include Angeletos et al. (2020) Shock (Step 2)	10.3	9.4	12.4	12.4
	(6.9, 15.1)	(4.8, 16.1)	(8.0, 18.4)	(6.8, 20.5)
Use Median SPF Forecasts (Step 2)	10.9	8.4	8.9	12.3
	(7.3, 15.8)	(4.5, 14.6)	(6.0, 13.1)	(6.6, 20.2)
Fill in Missing Shock Values (Step 2)	11.7	9.3	9.8	13.1
	(8.0, 16.5)	(4.9, 15.4)	(6.5, 14.4)	(7.0, 21.3)
Include 12 FRED Factors (Step 2)	9.5	5.8	10.7	9.5
	(6.4, 13.6)	(3.1, 10.2)	(7.4, 15.2)	(5.2, 16.3)
Include FRED Factors Lead (Step 2)	7.0	7.9	6.5	7.6
	(3.4, 11.4)	(2.4, 15.9)	(3.1, 10.6)	(2.4, 15.7)
Include Only FRED Factors Lags (Step 2)	12.3	6.9	9.7	15.2
	(8.3, 17.6)	(3.7, 12.2)	(6.4, 14.6)	(8.5, 24.1)
Include Lag Length from 4 to 8 (Step 2)	11.5	6.9	9.8	15.8
	(8.0, 15.9)	(3.7, 12.3)	(6.8, 13.7)	(9.4, 24.6)
OLS VAR Estimation	10.4	6.2	5.8	12.2
	(6.3, 15.3)	(1.1, 13.5)	(2.4, 9.6)	(5.2, 21.2)
Log Levels (VAR)	9.7	8.2	8.4	12.9
	(4.5, 17.6)	(4.1, 14.9)	(3.9, 15.4)	(6.8, 21.6)
Extend Sample to End in 2025Q1	13.3	13.4	21.6	26.8
	(9.3, 18.5)	(9.0, 19.0)	(17.6, 25.9)	(19.6, 35.2)

Notes: The rows list the specific robustness check (see text for more details). The first summarizes the corresponding baseline results in Exhibit 2. The columns correspond to the selected variables included in the VAR system. For each robustness check and variable, we report posterior percentiles of the average variance contribution of the three shocks across horizons up to horizon 24, as in Exhibit 2. We report the median and 16th and 84th percentiles, shown underneath in parentheses. For OLS estimation, we construct error bands using a bootstrap approximation. For the check corresponding to "Log Levels (VAR)" only, the variables included in the VAR system are in fact log levels instead of growth rates implied by the column labels. For the sample ending in 2025Q1, we control for the pandemic using dummy observations.

Most robustness checks recover factors that are close to the baseline factors. Exhibit 5 compares each baseline factor with the closest factor from each robustness check, using the maximum absolute correlation. We report this comparison for robustness checks that alter the second-stage residuals. In 23 of 33 cases, the correlation is at least 0.8, indicating that the baseline factor structure is not tied to a narrow specification choice.

The main departures occur in specifications that either estimate many more second-stage parameters or extend the sample through the pandemic. The 12-factor and 8-lag specifications are harder to interpret as clean alternatives to the baseline because both increase the risk of overfitting. The pandemic extension raises a different issue: dummy observations may not fully capture nonlinear forecast behavior during that period. Thus, the extended-sample factors are less comparable to the baseline factors, even though the sentiment block continues to account for a substantial share of macroeconomic variation.

Exhibit 5: Correlations of Baseline and Robustness Factors

	1	2	3
Include Asset Prices and Sentiment (Step 1)	0.91	0.86	0.92
Include Michigan Consumer Sentiment (Step 2)	0.99	0.99	1.00
Include Conference Board US Consumer Confidence (Step 2)	0.98	0.98	0.98
Include Angeletos et al. (2020) Shock (Step 2)	0.92	0.88	0.64
Use Median SPF Forecasts (Step 2)	0.96	0.93	0.95
Fill in Missing Shock Values (Step 2)	0.91	0.93	0.87
Include 12 FRED Factors (Step 2)	0.69	0.61	0.39
Include FRED Factors Lead (Step 2)	0.92	0.89	0.94
Include Only FRED Factors Lags (Step 2)	0.81	0.86	0.89
Include Lag Length from 4 to 8 (Step 2)	0.53	0.54	0.40
Extend Sample to End in 2025Q1	0.33	0.45	0.15

Notes: The rows list the specific robustness check (see text for more details). The columns correspond to the baseline factors. For each baseline factor, we report the maximum absolute correlation with each of the three robustness factors.

5 Conclusion

Professional forecasts contain a meaningful residual factor structure even after we control for macroeconomic conditions and known economic shocks. Because these factors are not persistent, we interpret them as unforeseeable shocks. These shocks are important determinants of volatility in many key macroeconomic variables. Interestingly, our shock series do not show any unusual behavior at the onset of NBER recessions, implying that they should likely not be thought of as causing recessions.

REFERENCES

Adams, J. J. and Barrett, P. (2025). What Are Empirical Monetary Policy Shocks? Estimating the Term Structure of Policy News . IMF Working Paper.

Angeletos, G.-M., Collard, F., and Dellas, H. (2020). Business-Cycle Anatomy. American Economic Review, 110(10):3030–3070.

Angeletos, G.-M. and La'o, J. (2013). Sentiments. Econometrica, 81(2):739–779.

Bernanke, B. S., Boivin, J., and Eliasz, P. (2005). Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach. The Quarterly Journal of Economics, 120(1):387–422.

Canova, F. and Ferroni, F. (2024). A Hitchhiker's Guide to Empirical Macro Models. Working Paper.

Coibion, O., Gorodnichenko, Y., and Kumar, S. (2018). How Do Firms Form Their Expectations? New Survey Evidence. American Economic Review, 108(9):2671–2713.

Coibion, O., Gorodnichenko, Y., and Ropele, T. (2020). Inflation Expectations and Firm Decisions: New Causal Evidence. The Quarterly Journal of Economics, 135(1):165–219.

Croushore, D. D. (1993). Introducing: The Survey of Professional Forecasters. Business Review-Federal Reserve Bank of Philadelphia, 6:3.

Ferreira, C. and Pica, S. (2024). Households' Subjective Expectations: Disagreement, Common Drivers and Reaction to Monetary Policy. Working Paper.

Fusari, F., Marlow, J., and Volpicella, A. (2026). More on VARs and Local Projections Equivalence: Unit Roots and Multiple Instruments. Working Paper.

Gennaioli, N., Ma, Y., and Shleifer, A. (2016). Expectations and Investment. NBER Macroeconomics Annual, 30(1):379–431.

Giacomini, R., Lu, J., and Smetanina, K. (2024). Perceived Shocks and Impulse Responses. Technical Report 21/24, Institute for Fiscal Studies.

Herbst, E. and Winkler, F. (2025). The Factor Structure of Disagreement. Journal of Business & Economic Statistics, pages 1–14. Published online December 2, 2025.

Kamdar, R. and Ray, W. (2024). Attention-Driven Sentiment and the Business Cycle. Technical report, Center for Applied Economics and Policy Research.

McCracken, M. W. and Ng, S. (2020). FRED-QD: A Quarterly Database for Macroeconomic Research . Federal Reserve Bank of St. Louis Working Paper, 2020-005B.

Milani, F. (2017). Sentiment and the US Business Cycle. Journal of Economic Dynamics and Control, 82:289–311.

Plagborg-Møller, M. and Wolf, C. K. (2021). Local Projections and VARs Estimate the Same Impulse Responses. Econometrica, 89(2).

Stock, J. H. and Watson, M. W. (2002). Forecasting Using Principal Components from a Large Number of Predictors. Journal of the American Statistical Association, 97(460):1167–1179.

Tanaka, M., Bloom, N., David, J. M., and Koga, M. (2020). Firm Performance and Macro Forecast Accuracy. Journal of Monetary Economics, 114:26–41.

SUPPLEMENTAL APPENDIX FOR MEASURING SENTIMENT SHOCKS IN PROFESSIONAL SURVEY DATA AND THEIR ROLE IN AGGREGATE FLUCTUATIONS

Rupal Kamdar Christian Matthes Anthony Vecchia

A DATA

A.1 FRED-QD FRED-QD is a dataset of 245 macroeconomic indicators maintained on the FRED website and originally compiled by McCracken and Ng (2020). The dataset contains series from 14 "groups": NIPA; industrial production; employment and unemployment; housing; inventories, orders, and sales; prices; earnings and productivity; interest rates; money and credit; household balance sheets; exchange rates; other; stock markets; and non-household balance sheets. The source series are observed at different frequencies. For FRED-QD, each series is aggregated to the quarterly frequency. Series are seasonally adjusted when applicable. Any data updates or revisions are automatically incorporated in each release of FRED-QD. We use the *"2025-07-QD"* vintage.

We use FRED-QD data in the first step of the three-step process and in the VAR system in Section 3. We begin the sample in 1968Q1, which corresponds to the first observation of SPF data. Before the first step, we drop 47 series that either contain substantial missing observations, can broadly be considered asset prices, or are sentiment indicators, all of which are listed below. The kept series are transformed to be stationary according to codes included with the dataset. Transformations are most often first differences of logs for series in dollars or indices, and first differences for percents (with no transformation for spreads). Finally, each transformed series is demeaned and normalized to unit variance for factor extraction.

For use in the VAR system, we select from FRED-QD PCE inflation (first difference of log 'PCECTPI'), the unemployment rate ('UNRATE'), real GDP growth (first difference of log 'GDPC1'), durable consumption growth (first difference of log 'PCDGx'), gross private domestic investment growth (first difference of log 'GPDIC1'), housing starts growth (first difference of log 'HOUST'), the BAA/AAA corporate bond spread ('BAA' minus 'AAA'), the 10-year/3-month Treasury spread ('GS10' minus 'TB3MS'), and the Federal Funds Effective Rate ('FEDFUNDS'). The sample length is harmonized to that of the residual SPF factor extracted from the third step.

The following 22 series are dropped for containing missing values.

- 'ACOGNO': manufacturers' new orders (consumer goods)
- 'COMPAPFF': 3-month commercial paper minus Federal Funds Rate
- 'COMPRMS': manufacturing sector (real hourly compensation for all workers)
- 'CP3M': 3-month AA financial commercial paper rate

- 'CPF3MTB3Mx': 3-month commercial paper minus 3-month Treasury bill (secondary market)
- 'CUSR0000SEHC': consumer price index for all urban consumers (owners' equivalent rent of residences in U.S. city average)
- 'DRIWCIL': net percentage of domestic banks reporting increased willingness to make consumer installment loans
- 'EXUSEU': U.S. dollars to euro spot exchange rate
- 'HOAMS': manufacturing sector (hours worked for all workers)
- 'MORTGAGE30US': 30-year fixed rate mortgage average in the U.S.
- 'MORTG10YRx': 30-year conventional mortgage rate relative to 10-year Treasury constant maturity
- 'NASDAQCOM': NASDAQ composite index
- 'OILPRICE': crude oil prices (West Texas Intermediate, Cushing, Oklahoma)
- 'OPHMFG': manufacturing sector (labor productivity, output per hour for all workers)
- 'OUTMS': manufacturing sector (real sectoral output for all workers)
- 'RSAFS': advance retail sales (retail trade and food services)
- 'SPCS10RSA': S&P CoreLogic Case-Shiller 10-city composite home price index
- 'SPCS20RSA': S&P CoreLogic Case-Shiller 20-city composite home price index
- 'TWEXAFEGSMTH': nominal advanced foreign economies U.S. dollar index
- 'ULCMFG': manufacturing sector (unit labor costs for all workers)
- 'USEPUINDXM': economic policy uncertainty index for United States
- 'USSTHPI': all-transactions house price index for the U.S.

The following 24 *additional* series are dropped for being classified as asset prices. We classify asset prices broadly.

- 'AAA': Moody's seasoned AAA corporate bond yield
- 'AAAFFM': Moody's seasoned AAA corporate bond minus Federal Funds Rate
- 'BAA': Moody's seasoned BAA corporate bond yield
- 'BAA10YM': Moody's seasoned BAA corporate bond yield relative to yield on 10-year Treasury constant maturity

- ‘EXCAUSx’: Canada/U.S. foreign exchange rate
- ‘EXJPUSx’: Japan/U.S. foreign exchange rate
- ‘EXSZUSx’: Switzerland/U.S. foreign exchange rate
- ‘EXUSUKx’: U.S./U.K. foreign exchange rate
- ‘FEDFUNDS’: effective Federal Funds Rate
- ‘GS1’: 1-year Treasury constant maturity rate
- ‘GS5’: 5-year Treasury constant maturity rate
- ‘GS10’: 10-year Treasury constant maturity rate
- ‘GS1TB3Mx’: 1-year Treasury constant maturity minus 3-month Treasury bill (secondary market)
- ‘GS10TB3Mx’: 10-year Treasury constant maturity minus 3-month Treasury bill (secondary market)
- ‘NIKKEI225’: Nikkei stock average
- ‘S&P 500’: S&P’s common stock price index (composite)
- ‘S&P div yield’: S&P’s composite common stock (dividend yield)
- ‘S&P PE ratio’: S&P’s composite common stock (price-earnings ratio)
- ‘T5YFFM’: 5-year Treasury constant maturity minus Federal Funds Rate
- ‘TB6M3Mx’: 6-month Treasury bill minus 3-month Treasury bill (secondary market)
- ‘TB3MS’: 3-month Treasury bill (secondary market)
- ‘TB6MS’: 6-month Treasury bill (secondary market)
- ‘TB3SMFFM’: 3-month Treasury constant maturity minus Federal Funds Rate
- ‘VIXCLSx’: CBOE volatility index (VIX)

The following *additional* series is dropped for being classified as a sentiment index.

- ‘UMCSENTx’: University of Michigan consumer sentiment index (1996Q1 base)

A.2 SPF The Survey of Professional Forecasters (SPF) is a quarterly survey of roughly 30–50 forecasters conducted by the Federal Reserve Bank of Philadelphia. The survey is sent to forecasters after the release of the Bureau of Economic Analysis's advance NIPA report around the end of the first month of the quarter. The survey posts the realizations from this report along with the most recent realizations of other forecast variables. The forecasters report their forecasts in the same units (e.g., scale, seasonal adjustment, base years) as these posted realizations. According to the Philadelphia Fed, the deadline for forecaster responses is the second or third week of the second month of the quarter. Often, other data, such as the BLS Employment Situation Summary, are released after the survey is sent out but before it is due. The Philadelphia Fed took over survey administration in 1990Q2 and acknowledges that many details about the survey's prior administration may have differed and are unknown to them. We downloaded the most recent survey data from the Philadelphia Fed as of September 6, 2025.

Individual forecaster identity is only observable using a confidential identifier number. The Philadelphia Fed uses its own discretion (and does not disclose) whether this identifier is attached to a firm or individual for specific forecasters. Consecutive forecasts from the same forecaster for long periods are rarely observed. Therefore, we use mean within-survey forecasts. In Section 4, we show that results are robust to using median forecasts instead.

We use the SPF data only as dependent variables in the second-step regressions of the three-step process. Each dependent variable corresponds to a specific horizon of a specific forecast variable. We use five forecast horizons (current quarter to four quarters ahead) of the 16 variables available beginning on or before 1981Q3 (for 80 total dependent variables): chain-weighted GDP price index (PGDP), nominal corporate after-tax profits (CPROF), the unemployment rate (UNEMP), industrial production (INDPROD), housing starts (HOUSING), real GDP (RGDP), the three-month Treasury rate (TBILL), Moody's AAA corporate bond yield (BOND), real personal consumption expenditures (RCONSUM), real non-residential investment (RNRESIN), real residential investment (RRESINV), real federal government consumption and gross investment (RFEDGOV), real state and local government consumption and gross investment (RSLGOV), real change in private inventories (RCBI), real net exports (REXPORT), and CPI (CPI). Thus, we exclude forecast variables like the ten-year Treasury yield, nonfarm payroll employment, core CPI, PCE, and Moody's BAA corporate bond yield that are only available for shorter samples. Note that forecasters report the quarterly average for variables released at a monthly frequency.

The SPF indexes the current-quarter forecast by 2 and the four-quarter-ahead forecast by 6. The most recent realization is indexed by 1. We use this realization to transform the SPF variables similarly to their corresponding FRED-QD variables. For example, the prior quarter realization is logged and subtracted from the log mean forecast of real GDP at each horizon.

A.3 STRUCTURAL SHOCKS Adams and Barrett (2025) compile instruments for common structural macroeconomic shocks as provided by their original sources. We use these instruments only as contemporaneous independent variables in the second-step regressions of the three-step process. The table below lists all 48 instruments included in the compilation along with relevant information reproduced from Adams and Barrett (2025). We downloaded the compilation from the authors' webpage on September 25, 2025. Monthly shocks are summed to the quarterly frequency. The compilation does not distinguish between zeros and missing observations. Thus, we set all missing observations to zero in our baseline specification to harmonize the sample length of these instruments. In this process, it is almost never necessary to impute a zero in the middle of an instrument's observed sample (after summing monthly shocks to quarterly). As a robustness check, we instead use an expectations-maximization algorithm from McCracken and Ng (2020) to fill in missing observations. See Section 4 for more details.

Complete List of Instruments Used

Source	**Freq.**	**Derivation**	**Description**
Monetary Policy			
Romer and Romer (2004)	Monthly	Narrative	
Gertler and Karadi (2015)	Monthly	High-Frequency Identification (HFI)	
Jarociński and Karadi (2020)	Monthly	HFI	
Jarociński and Karadi (2020)	Monthly	HFI	Fed information effect
Miranda-Agrippino and Ricco (2021)	Monthly	HFI	
Bu, Rogers, and Wu (2021)	Monthly	HFI	
Bauer and Swanson (2023)	Monthly	HFI	
Aruoba and Drechsel (2024)	Monthly	Narrative	
Bundick, Herriford, and Smith (2024)	Monthly	HFI	Term structure uncertainty level
Bundick, Herriford, and Smith (2024)	Monthly	HFI	Term structure uncertainty slope
Drechsel (2024)	Quarterly	Narrative	Political pressure
Jarociński (2024)	Monthly	HFI	FFR
Jarociński (2024)	Monthly	HFI	Forward guidance
Jarociński (2024)	Monthly	HFI	Large-scale asset purchases
Jarociński (2024)	Monthly	HFI	Information effect
Swanson (2024)	Monthly	HFI	FFR
Swanson (2024)	Monthly	HFI	Forward guidance
Swanson (2024)	Monthly	HFI	Large-scale asset purchases
Fiscal Policy			
Ramey (2011)	Quarterly	Narrative	Government spending from military news
Leeper, Richter, and Walker (2012)	Quarterly	Other	Fiscal news implied by bond markets and forecasts
Mertens and Ravn (2012)	Quarterly	Narrative	Tax surprise
Mertens and Ravn (2012)	Quarterly	Narrative	Tax news
Fisher and Peters (2010)	Quarterly	Other	Government spending from excess returns of defense contractors
Romer and Romer (2016)	Monthly	Narrative	Permanent Social Security expansion
Romer and Romer (2016)	Monthly	Narrative	Temporary Social Security expansion
Ben Zeev and Pappa (2017)	Quarterly	Narrative	Government spending from defense news
Fieldhouse, Mertens, and Ravn (2018)	Monthly	Narrative	Non-cyclical federal housing purchases
Fieldhouse and Mertens (2024)	Quarterly	Other	R&D defense spending

Continued on next page

Source	**Freq.**	**Derivation**	**Description**
Fieldhouse and Mertens (2024)	Quarterly	Other	R&D non-defense spending
Lieb et al. (2024)	Quarterly	Narrative	Tax news from presidential speeches
Phillot (2025)	Monthly	HFI	Tax using Treasury auction announcements (2-year yields)
Phillot (2025)	Monthly	HFI	Tax using Treasury auction announcements (5-year yields)
Phillot (2025)	Monthly	HFI	Tax using Treasury auction announcements (10-year yields)
Phillot (2025)	Monthly	HFI	Tax using Treasury auction announcements (30-year yields)
Technology			
Barsky and Sims (2011)	Quarterly	Structural Vector Autoregression (SVAR)	TFP news
Fernald (2014)	Quarterly	Other	Utilization-adjusted TFP growth
Ben Zeev and Kahn (2015)	Quarterly	SVAR	IST surprise
Ben Zeev and Kahn (2015)	Quarterly	SVAR	IST news
Ben Zeev and Kahn (2015)	Quarterly	SVAR	TFP
Miranda-Agrippino, Hacioglu-Hoke, and Bluwstein (2025)	Monthly	Other	Patent filing
Oil			
Kilian (2008)	Quarterly	Narrative	Conflicts in OPEC countries
Baumeister and Hamilton (2019)	Monthly	SVAR	Oil supply
Baumeister and Hamilton (2019)	Monthly	SVAR	Consumption demand
Baumeister and Hamilton (2019)	Monthly	SVAR	Inventory demand
Känzig (2021)	Monthly	HFI	Oil supply news
Inflation Sentiment			
Adams and Barrett (2024)	Monthly	SVAR	
Severe Weather			
Kim, Matthes, and Phan (2024)	Monthly	SVAR	
Uncertainty			
Piffer and Podstawski (2018)	Monthly	HFI	Uncertainty from gold prices
Information taken directly from Adams and Barrett (2025).			

A.4 ADDITIONAL DATA We gather the following two series, which are not included in the datasets described above, to use as independent variables in separate robustness checks in Section 4.

- Monthly Conference Board Consumer Confidence Index for the United States (1985 base), downloaded on March 11, 2026 and averaged to quarterly frequency
- Quarterly "Main Business-Cycle" shock from Angeletos et al. (2020), downloaded from Angeletos' webpage on March 11, 2026 (sample ends in 2017Q4)

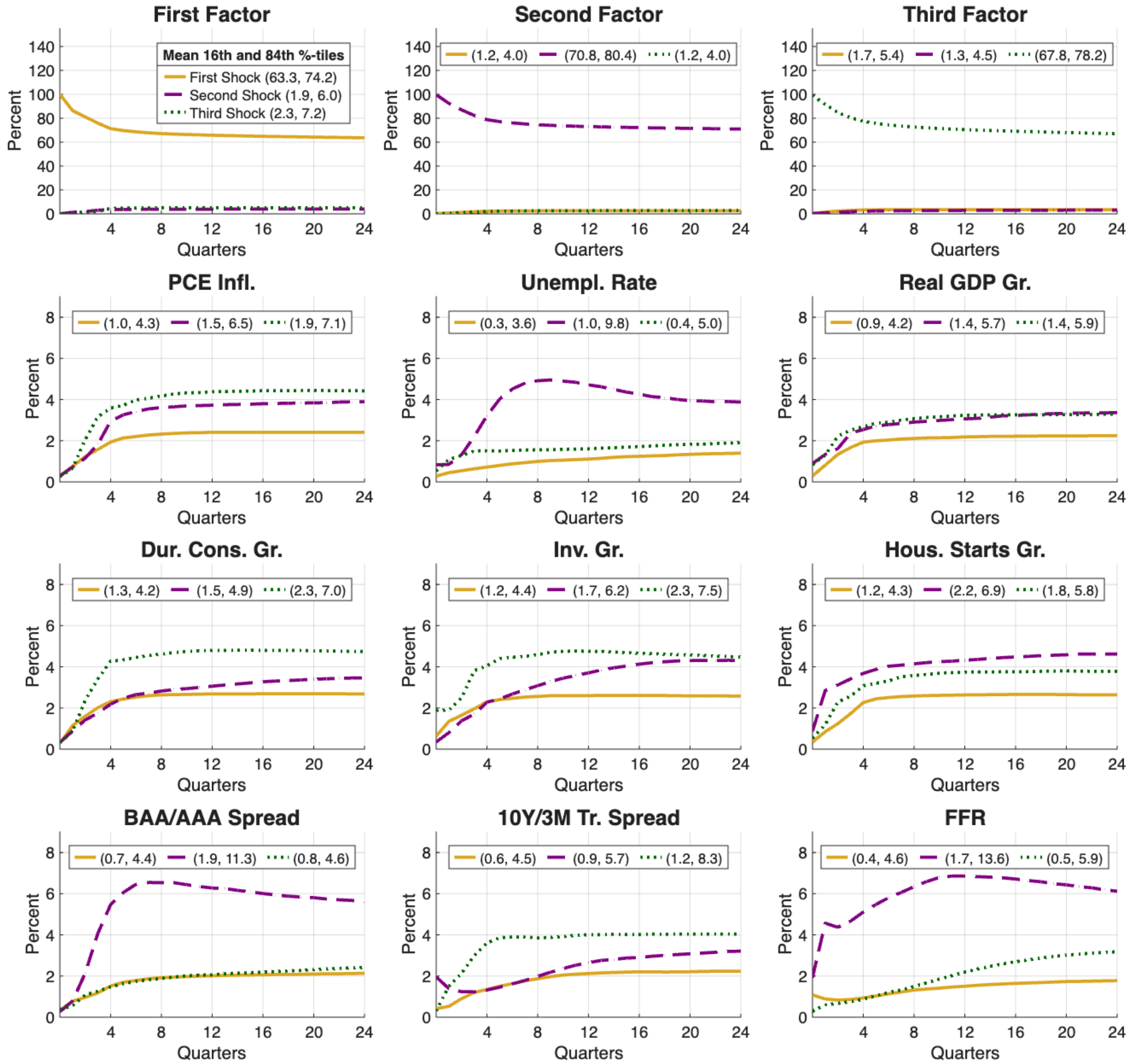
B ADDITIONAL RESULTS

Exhibit 6: Second-Step Regression Results

Dependent Variables	Mean F-Stat	Mean R^2	Mean Adj. R^2
Horizons of...			
PGDP	267.64	0.97	0.93
UNEMP	5889.36	1.00	0.99
INDPROD	44.57	0.92	0.83
HOUSING	4030.68	0.99	0.99
RGDP	99.43	0.90	0.77
TBILL	796.18	1.00	0.99
BOND	1701.80	0.99	0.99
RCONSUM	59.36	0.83	0.62
RNRESIN	46.41	0.94	0.86
RRESINV	23.28	0.95	0.89
RFEDGOV	7.02	0.82	0.59
RSLGOV	18.83	0.89	0.76
RCBI	47.33	0.95	0.89
CPI	323.27	0.97	0.94
REXPORT	480.65	0.99	0.98
CPROF	5.87	0.80	0.55
Number of dep. var.s: 80 (16 SPF var.s, 5 hor.s)			
Number of ind. var.s: 84			
Number of obs.: 150 (1982Q3–2019Q4)			

Notes: Summary of estimation output from Equation 1 with dependent variables for each horizon of each forecast variable. The rows correspond to groups of regressions by the SPF dependent variable (across horizons). The columns correspond to the mean of the respective statistics within each group. The F-statistic is computed for a test of joint significance of the coefficients on the first-step macroeconomic factors and their lags (excluding coefficients on other controls).

Exhibit 7: Variance Explained by the Three Factors Separately



Notes: The lines represent the pointwise posterior median path of the variance explained by each of the three residual SPF factor shocks separately in the forecast error variance decomposition. The solid yellow line corresponds to the first factor shock, the dashed purple to the second, and the dotted green to the third. The legend in each panel displays the mean of the 16th and 84th percentiles across horizons for each shock. The shocks are identified recursively using the order in which the variables appear.