EXPECTATIONS FORMATION AND FORWARD INFORMATION

Nathan Goldstein
Yuriy Gorodnichenko

ABSTRACT

We propose a model where forecasters have access to noisy signals about the future (forward information). In this setting, information varies not only across agents but also across horizons. As a result, the estimated persistence of forecasts deviates from the persistence of fundamentals and the ability of forecasts at shorter horizons to explain forecasts at longer horizons is limited. These properties tend to diminish as the forecast horizon increases. We document that this novel pattern is consistent with survey data for professional forecasters. We provide further evidence that time-series and cross-sectional variation in professional forecasts are driven by forward information. We propose a simple method for extracting the forward information component from a survey and provide several applications of forward information.

Nathan Goldstein
Department of Economics
Building 504
Bar-Ilan University
Ramat-Gan 5290002
Israel
Nathan.Goldstein@biu.ac.il

Yuriy Gorodnichenko
Department of Economics
530 Evans Hall #3880
University of California, Berkeley
Berkeley, CA 94720-3880
and IZA
and also NBER
ygorodni@econ.berkeley.edu

1. Introduction

Yogi Berra famously summarized challenges in forecasting with "it's tough to make predictions, especially about the future." Yet economic agents routinely engage in forecasting because almost every decision they make depends on what they think inflation, income and other variables are going to be in a week, month, year, or longer. But how are expectations formed? What information do economic agents have about the current and future conditions? How useful are signals about the future? To shed more light on these questions, we propose a new approach that incorporates "forward information" in expectations formation.

We start with the observation that forecasts are more than just a model iterated forward. Indeed, using a special survey conducted among participants in the Survey of Professional Forecasters (SPF), Stark (2013) documents that a vast majority of professional forecasters (80%) describe their forecasting method as a combination of "model" and "subjective adjustment". For example, this practice means that when forecasters try to predict inflation, they use an econometric model to generate a first pass at the inflation projection, but then they adjust the prediction due to whatever information (e.g., news about future monetary/fiscal policy, technology) they have which is beyond their model.[1] This "forward" information is useful because it may reduce forecast errors. Since forecasters can face different and likely imperfect signals about the future, forward information can also be a key source of forecast dispersion and apparent deviations from full-information rational expectations (FIRE) documented in the literature. Our objective is to extract this forward information from forecasts and then study its properties and effects.

In a nutshell, consider the standard noisy information framework as a benchmark: agents face imperfect signals about realized fundamentals and use the Kalman filter or similar tools to filter out the noise. Once the forecast for the current period is formed, the forecasts for longer horizons obey the state equation and, hence, persistence in expectations should reflect persistence in the fundamental. For example, if the fundamental x_t is an AR(1) process with persistence ρ, the forecast for horizon h ($x_{t+h|t}$) is simply ρ^h times the belief about the current value of the fundamental ($x_{t|t}$): $x_{t+h|t} = \rho^h x_{t|t}$. As a result, when we regress the forecast for horizon $h+1$ ($x_{t+h+1|t}$) on the forecast for horizon h ($x_{t+h|t}$), the slope coefficient in this regression is ρ, i.e., persistence in the fundamental, for any horizon h. Furthermore, the R^2 in this regression should be one because there is no new information beyond the current estimate of the fundamental. We test these predictions using SPF forecasts for multiple horizons. In contrast to the predictions, we find that both the slope coefficient and the fit of the regression gradually increase and tend to "converge" as we apply the regression to forecasts with longer horizons. This pattern points to some additional forecasting component,

---

[1] A common technical tool for subjective adjustment is known in practice as "add factoring", where the forecaster adds some subjective factor to the estimated specification used for prediction. For an early discussion of add factoring, see Fair (1986).

beyond the application of the state equation, which produces a variation in correlation across the forecast horizons.

We interpret this forecasting component as reflecting the "subjective adjustment." Specifically, forecasters adjust the state-equation prediction in light of forward information (i.e., signals about the future). Intuitively, when forecasters adjust their "econometric" predictions to incorporate forward information, they introduce an "error term" (which captures forward information) in the state equation: $x_{t+h+1|t} = \rho x_{t+h|t} + error$. Because forward information can influence not only beliefs about horizon $h + 1$ but also beliefs about horizon h, the error term is correlated with $x_{t+h|t}$ and hence can bias the estimate of the slope away from ρ and result in $R^2 < 1$. However, for long enough horizons, for which forward signals are no longer informative, the forecasts would only be based on the state equation without a biasing adjustment. Thus, the forward information setup should augment the standard noisy information framework in two important directions: First, there will be informative signals not only about the past but also about *future* fundamentals. Second, information will vary not only across agents, as in the standard framework, but also across future *horizons*.

We present a formal model of expectations formation with forward information along these lines. Agents receive multiple noisy signals referring to several periods in the future. They exploit these signals and optimally filter the noise when forming forecasts for multiple periods ahead. We show that forecast adjustment due to forward information can vary across forecasting horizons since agents need to optimally reweight the multiple forward signals when moving from one horizon to another. This reweighting mechanism diminishes and eventually stops when forecasting goes far enough into the future, at which point there are no further informative signals. We use simulations to show that this setup can qualitatively reproduce the patterns documented in the SPF data.

Our approach offers several additional insights. First, we can assess the importance of forward information. In particular, our generalized framework allows us to directly test the restricted form implied by the standard noisy information model, which includes only signals about the current and past fundamentals, against the unrestricted form, which includes forward information. We show that the coefficients in the augmented specification are directly mapped to the optimal weights placed on the multiple forward signals in our framework. Based on SPF data, our results point to a significant role of forward information and rule out a significant amount of noisy information in the sense of the standard model. These findings support the view that SPF forecasts are driven by dispersed information about the *future*, rather than about the past.

Second, our approach offers a straightforward way to extract forward information directly from expectations data. According to our forecast decomposition ($x_{t+h+1|t} = \rho x_{t+h|t} + error$), forward information is the component that leads forecasters to make an adjustment to their "econometric" prediction

(i.e., *error* in the state equation iterated forward). We apply this decomposition to SPF inflation forecasts. We show that forward information is helpful for predicting inflation at short horizons. Furthermore, forward information in the Greenbook forecasts from the Federal Reserve predicts policy rate, thus suggesting that the central bank acts on news about future inflation.

Finally, we apply our approach to other variables and survey data. When we examine SPF forecasts for additional macroeconomic variables, we find both indirect and direct evidence for the presence of forward information that varies across forecasting horizons. Similar evidence is found for the Greenbooks. We also apply inflation forecast data from the SPF run by the European Central Bank (ECB). We exploit the availability of both calendar-year and rolling-year forecasts in this survey and find a pattern of variation across calendar quarters, which is consistent with the utilization of forward information that varies across forecasting horizons.

Our study is related to several strands of literature. First, we contribute to the vast and growing literature on the expectations formation process (Coibion, Gorodnichenko and Kamdar (2018) provide a review of this literature). Survey forecasts are widely used to test and quantify models with information rigidities or behavioral frictions (e.g. Mankiw, Reis, and Wolfers, 2004; Capistran and Timmermann, 2009; Patton and Timmermann, 2010; Andrade et al., 2013; Coibion and Gorodnichenko, 2012, 2015; Fuhrer, 2018; Angeletos et al., 2020; Bordalo et al., 2020; Giacomini et al., 2020; Kohlhas and Walther, 2021; Goldstein, 2021). Building on this literature, we explicitly model the "subjective adjustment" of forecasts which we interpret as a practice driven by the availability of forward information. We emphasize that forward information plays a key role in how professional forecasters form their macroeconomic projections.

Our paper is also related to the literature about news shocks (Beaudry and Portier (2014) provide a survey of this literature and Chahrour and Jurado (2018) connect news shocks and noise in signals). Unlike the literature on expectations formation, this literature stresses the role of news about the future in driving the business cycle, but it largely ignores microfoundations of expectational assumptions (Chahrour and Jurado (2021) is a recent exception). Furthermore, with some exceptions (e.g., Barsky and Sims, 2012; Nguyen and Miamoto, 2019), it tends to extract news from realized macroeconomic series rather than expectations data. Our approach bridges these two literatures by linking forward information directly to the expectations formation process. Furthermore, we emphasize and exploit heterogeneity across both agents and forecast horizons. Our findings can be used to quantify forward information systematically using forecast data and to assess its role in predicting macroeconomic dynamics and affecting policy.

Our paper also contributes to the literature that studies the persistence of inflation and documents a notable decline in inflation persistence in the U.S. and other countries since the 2000s (Fuhrer (2011) surveys this literature). Jain (2019), which is the closest in spirit to the starting point of our analysis, utilizes revisions in SPF forecasts to estimate the size and time variation in the *perceived* persistence of inflation.

In contrast to Jain (2019), we allow persistence to vary with the forecast horizon. As explained above, this novel type of variation that we document motivates our idea of forward information. Based on our results, we argue that persistence estimated on long horizons can provide a better metric of perceived persistence of the fundamental because forward information is less likely to be quantitively important at long horizons. We also offer a real-time estimate of persistence based on cross-sectional/across-horizon variation in forecasts, which contrasts with the focus on time-series variation in actual data.

The remainder of the paper is organized as follows. Section 2 provides a simple noisy-information benchmark model to build intuition and testable predictions. We show that these predictions are rejected by stylized facts documented for SPF inflation forecasts. Section 3 presents our model of forward information and rationalizes the stylized facts. Section 4 reports further tests of forward information. Section 5 describes how one can extract forward information from forecast data and documents the properties of forward information, including the ability of forward information to predict policy choices and future inflation. Section 6 summarizes evidence from forecasts of other key macroeconomic variables in the U.S. SPF and the ECB SPF. Section 7 concludes.

2. Noisy information: The standard framework

To fix ideas, we use the standard noisy-information model as a starting point. Following Woodford (2002) and Coibion and Gorodnichenko (2015), assume that fundamental x_t follows an AR(1) process. While agents know the underlying process, they cannot perfectly observe x_t due to idiosyncratic noise in signals about x_t. Their forecasting problem can be described by a state-space representation:

State:

$$x_t = \rho x_{t-1} + \omega_t, \tag{1}$$

Measurement:

$$y_t^i = x_t + v_t^i, \tag{2}$$

where $\omega_t \sim iid\ N(0, \sigma_\omega^2)$ is a shock to the fundamental, $v_t^i \sim iid\ N(0, \sigma_v^2)$ is idiosyncratic noise, and ρ is the persistence of the fundamental. An optimal forecast $x_{t+h|t}^i$, for h steps ahead, is formed by each agent i with a Kalman filter:

$$x_{t+h|t}^i = x_{t+h|t-1}^i + G\big(\rho^h y_t^i - x_{t+h|t-1}^i\big), \tag{3}$$

where G is the Kalman gain. Thus, agents revise their former forecast, $x_{t+h|t-1}^i$, by placing some weight on the prediction $\rho^h y_t^i$, which utilizes the new signal.

An important implication of equation (3) is that forecasts made at time t for various horizons should obey the process of the state equation:

$$x^i_{t+h|t} = \rho x^i_{t+h-1|t}. \tag{4}$$

Equation (4) posits several predictions which can be tested with data on survey forecasts for multiple horizons. In particular:

When estimating a regression of the forecast $x^i_{t+h|t}$ on the forecast $x^i_{t+h-1|t}$:

    I.   The fit of the regression should be perfect ($R^2 = 1$) for any $h \geq 1$.

    II. The regression coefficient recovers ρ, the persistence parameter, for any $h \geq 1$.

Note that these predictions apply to both forecaster-level projections and consensus projections.

We test these predictions with CPI inflation forecasts from the Survey of Professional Forecasters. Quarterly forecasts are available since 1981Q3 for multiple horizons, running from the current quarter to four quarters ahead. In addition, there is a backcast provided in each quarter for the inflation rate in the previous quarter. Thus, we estimate five horizon-specific regressions for each h between 0 to 4.[2] The results for the full sample period (1981Q3-2017Q4) are described in Panel A of Figure 1. We observe several patterns, which deviate from the above predictions. First, the estimated persistence and R^2 vary with the forecast horizon significantly. Second, both the estimated persistence and R^2 increase in the forecast horizon. The estimated persistence rises from ≈0.4 for $h = 0$ to ≈0.9 for the longest horizon $h = 4$. Likewise, R^2 increases from ≈0.3 to ≈0.8. Finally, the fit of the regressions is far from perfect for the shorter horizons ($R^2 \ll 1$).

To explore the robustness of these patterns, we split the sample into four decades from the 1980s to the 2010s and repeat the estimation. We report results in Panel B of Figure 1. Overall, the coefficient estimates and R-squared show the same pattern of increase with the forecast horizon for each sub-period. The only difference is that the range of variation is getting narrower in recent decades. This could reflect the decline in inflation persistence in the US over the years, which was documented by the previous literature (e.g., Fuhrer, 2011). Nevertheless, differences across horizons are still highly significant as indicated by the p-values.[3]

---

[2] $h = 0$ in the data is equivalent to $h = 1$ in the above model, since SPF respondents receive a signal on realization in the previous quarter.

[3] We also preform the estimation with consensus forecasts and find the same pattern documented in Figure 1. See the first row in Panel A of Appendix Table G.3.

Because equation (4) holds for any period, we can estimate the corresponding regression quarter by quarter separately and examine time-series variation in the estimates.[4] Using a non-parametric smoother to remove high-frequency noise in the estimates, we report the results in Figure 2. We find that both the estimated persistence and R^2 are higher at longer horizons (Panels A and B, respectively). Furthermore, as the estimated persistence declines over the years, so does the R^2 of the cross-sectional regressions. This suggests that the dynamics of persistence and R^2 may be related phenomena.

Specification (4) reflects a simple AR(1) process, whereas agents may use more sophisticated processes. In Figure 3, we examine more general processes, such as AR(4), VAR(4), and AR(1) in terms of the inflation gap (the difference between inflation and trend inflation). Because we do not have enough forecast horizons to explore how $\hat{\rho}$ varies in h, we focus on the R^2 from a quarter-by-quarter estimation corresponding to each specification for $h = 3$. For the AR(4) process, we regress the forecast $x_{t+3|t}^i$ on the forecasts for the 4 preceding quarters. For the VAR(4), we add forecasts of "lagged" interest rate and unemployment from the SPF. The inflation gap specification utilizes long-run inflation forecasts, which are provided in the SPF more recently, and computes the forecast for the inflation gap as the difference between the CPI inflation forecast and the long-run forecast.[5] The figure describes a fit of the different regressions over the sample period which is very similar to the fit obtained with the simple AR(1). This is in line with Coibion and Gorodnichenko (2015), who find that a univariate AR(1) approximates well the expectations in the SPF data.[6]

Another potential explanation for the results in Figures 1 and 2 is that forecasters use different models, i.e., each forecaster applies her own (perceived) ρ^i, which may be different from the true persistence parameter ρ. Hence, the forecast of an agent i for time $t + h$, denoted by $x_{t+h|t}^i$, follows:

$$x_{t+h|t}^i = \rho^i x_{t+h-1|t}^i = \rho x_{t+h-1|t}^i + (\rho^i - \rho) x_{t+h-1|t}^i. \tag{5}$$

---

[4] Notice that this estimation is also in line with the well-known unobserved-component modelling of inflation. Specifically, suppose that inflation follows $x_t = \mu_t + x_t^{GAP}$, where μ_t is a trend component following a random walk and x_t^{GAP} is a transitory component following AR(1) with persistence $0 < \rho \le 1$. Then, the process of x_t can be written as AR(1), $x_t = c_t + \rho x_t + \varepsilon_t$, where ε_t captures shocks to both components and $c_t = (1 - \rho)\mu_{t-1}$ behaves like a "constant" that varies over time due to the trend component. When applying our AR(1) regression to inflation forecasts quarter by quarter, the regression constant can vary over time in a similar way.

[5] From the previous footnote, it is clear that the inflation-gap autoregression arises from the same unobserved-component modelling of inflation, since $x_t^{GAP} = x_t - \mu_t$. The difference in the current estimation is that we allow heterogeneity in forecasters' views on trend inflation, by using their long-run forecasts from the survey.

[6] Another robustness check for our stylized pattern is reported in Appendix Figure G.1. As for the AR(1), we applied estimations across different horizons for higher order auto-regressions. Each additional "lag" implies one less regression (i.e., instead of 5 regressions for the different horizons with AR(1), we can ran 4 regressions with AR(2) and so on). The figure reports persistence estimates as the sum of coefficients and the R^2 values. In all the panels they clearly increase with the horizon, consistent with the pattern documented for AR(1).

If we regress $x^i_{t+h|t}$ on $x^i_{t+h-1|t}$ as in specification (4), the error term corresponding to $(\rho^i - \rho)x^i_{t+h-1|t}$ is clearly correlated with the regressor $x^i_{t+h-1|t}$. Furthermore, the cross-sectional estimate of average (perceived) persistence could depend on the distribution of $(\rho^i)^h$, thus creating variation in h. Finally, differences in ρ^i generate disagreement in forecasts.[7]

To assess this alternative explanation, we estimate the persistence in expectations across horizons for a *specific* forecaster i, using OLS:

$$x^i_{t+h|t} = \rho^i_h x^i_{t+h-1|t} + error^i_t. \tag{6}$$

For each horizon h, we compute the average (across forecasters) value of the estimated ρ^i_h. If heterogeneity in ρ^i were driving our results in Figures 1 and 2, we should observe that the average value of $\hat{\rho}^i_h$ should not increase in horizon h. Panel A of Figure 4 documents that forecasters indeed have different perceived persistence of inflation: the standard deviation of $\hat{\rho}^i_{h=3}$ is above zero and it is generally increasing over time. However, this heterogeneity does not have a materially important effect on how inflation persistence varies across horizons and over time. Similar to Panel A in Figure 2, Panel B in Figure 4 documents that persistence (average of $\hat{\rho}^i_h$ across i) is larger for longer horizons and there is a broad decline in persistence over time. Thus, model heterogeneity does not explain why persistence increases in horizon h.

In summary, our evidence suggests that the standard noisy information framework misses some important forecasting component, which varies both cross-sectionally and across horizons. In what follows, we argue that this component is associated with the practice of "subjective adjustment" of forecasts, or "add-factoring". This subjective adjustment can capture forecasters' information about future fundamentals.

3. A model of forward noisy information

To build intuition, we begin with a tractable example using two forward signals. The example is used to derive some analytical results that illustrate how forward information augments the forecast in the standard noisy information framework and introduces an additional key component to equation (4). We will then present a more general framework and use simulations to show how one can rationalize the patterns documented in the previous section.

3.1. A case of two forward signals

---

[7] The correlation between the regressor and the error term can arise for other, potentially behavioral reasons. In Appendix D, we show how such correlation can arise in the model of asymmetric loss function by Elliott, Komunjer and Timmermann (2008) and Capistran and Timmermann (2009). However, this model does not explain why estimated persistence should increase with the forecast horizon.

Consider a state-space representation of a fundamental x_t, which follows an AR(1) process. However, unlike the one-signal state-space in equations (1) and (2), here agents receive three informative signals at time t: A perfect signal about realization (the current value of the fundamental) and two forward signals referring to two subsequent periods, $t + 1$ and $t + 2$. Accordingly, the state-space representation is

State:

$$x_t \equiv \begin{bmatrix} x_{t+2} \\ x_{t+1} \\ x_t \end{bmatrix} = \begin{bmatrix} \rho & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} x_{t-1} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \omega_{t+2} = \mathrm{P}x_{t-1} + S'\omega_{t+2} \tag{7}$$

where $\omega_t \sim iid\ N(0, \sigma_\omega^2)$.

Measurement:

$$y_t^i \equiv \begin{bmatrix} y_{t,t+2}^i \\ y_{t,t+1}^i \\ y_{t,t}^i \end{bmatrix} = \begin{bmatrix} x_{t+2} \\ x_{t+1} \\ x_t \end{bmatrix} + \begin{bmatrix} v_{t,t+2}^i \\ v_{t,t+1}^i \\ 0 \end{bmatrix} = x_t + v_t^i \tag{8}$$

where $v_{t,t+1}^i \sim iid\ N(0, \sigma_1^2)$ and $v_{t,t+2}^i \sim iid\ N(0, \sigma_2^2)$ are idiosyncratic noise in the two forward signals, $y_{t,t+1}^i$ and $y_{t,t+2}^i$, respectively. Notice that signal $y_{t,t}^i$ is perfect (it does not contain noise) so that x_t is perfectly known at time t. Only the forward signals $y_{t,t+1}^i$ and $y_{t,t+2}^i$ are imperfect. We use this example to demonstrate how the empirical patterns in the previous section can be driven only by forward signals, even in the absence of noise about the realized fundamental, as in equation (2).

Because $x_{t|t}^i = x_t$, we can simplify the analytical derivation of the weights in the optimal forecasts for the subsequent horizons. One can show (Appendix B) that the one-step ahead forecast is given by:

$$x_{t+1|t}^i = W_1\rho x_t + W_2 y_{t-1,t+1}^i + W_3 y_{t,t+1}^i + W_4\rho^{-1} y_{t,t+2}^i, \tag{9}$$

where W_1, W_2, W_3 and W_4 are the optimal weights placed on each informative signal, which minimizes $E_t\left(x_{t+1} - x_{t+1|t}^i\right)^2$ and obeys $W_1 + W_2 + W_3 + W_4 = 1$. As derived in Appendix B.1., the optimal weights depend on the noise-to-signal ratios ($\sigma_1^2/\sigma_\omega^2$ and $\sigma_2^2/\sigma_\omega^2$) and the persistence parameter ρ.

In a similar way, the two-step-ahead forecast is a weighted sum of the same four signals, which are shifted by one period to the future:

$$x_{t+2|t}^i = w_1\rho^2 x_t + w_2\rho y_{t-1,t+1}^i + w_3\rho y_{t,t+1}^i + w_4 y_{t,t+2}^i. \tag{10}$$

If the corresponding optimal weights in equations (9) and (10) are equal to each other, that is, if $W_k = w_k$ for each $k = 1,2,3,4$, then the relationship between the forecasts would simply obey $x_{t+2|t}^i = \rho x_{t+1|t}^i$, which follows the AR(1) process in the state equation. Hence, our setup can nest the standard noisy information model. However, this simple relationship $x_{t+2|t}^i = \rho x_{t+1|t}^i$ does not hold when we introduce forward signals. As shown in Appendix B.1., the optimal weights, when minimizing the two-step-ahead squared forecast error, are smaller than the corresponding weights in equation (9), except for $w_4 \geq W_4$.

Intuitively, because the fourth signal ($y_{t,t+2}^i$) refers directly to x_{t+2}, it is given an extra weight in the two-step-ahead forecast.

As a result, the relationship between forecasts at different horizons is more complex:

$$x_{t+2|t}^i = \rho x_{t+1|t}^i + \sum_{k=1}^{4}(w_k - W_k)\, Signal_{k,t+2|t}^i \tag{11}$$

where $Signal_{k,t+2|t}^i$ corresponds to each of the four signals used in equation (10) ($Signal_{1,t+2|t}^i \equiv \rho^2 x_t$, $Signal_{2,t+2|t}^i \equiv \rho y_{t-1,t+1}^i$, etc.). Equation (11) illustrates that forward signals induce an adjustment to the "standard" forecast $\rho x_{t+1|t}^i$, and that this adjustment depends on how the optimal weights change across forecasting horizons. From an econometric perspective, this adjustment introduces an "error term" in specification (4), which can account for the empirical patterns from the previous section. For example, the signals in this error term are correlated with the regressor $x_{t+1|t}$, thus, potentially explaining why estimated persistence can vary with the horizon.

Importantly, equation (11) could also be interpreted as a decomposition of the forecast $x_{t+2|t}^i$ into two components. The first component, $\rho x_{t+1|t}^i$, is a standard prediction based on the state process, while the second component is an adjustment due to forward information, which is beyond the information already included in the forecast for the previous step. In Section 5, we use this interpretation to quantify the second component as a measure of news. We further note that the decomposition in equation (11) provides a formal description of the common forecasting practice, that applies some "subjective adjustment" to the model-based forecast (Stark, 2013).

We now examine what happens to the adjustment component when increasing the horizon of the forecast to $h + 3$ (i.e., $x_{t+3|t}^i$), for which the forecaster does not have any forward information. As shown in Appendix B.1., the optimal weights do not change and remain w_k as in equation (10) so that:

$$x_{t+3|t}^i = w_1\rho^3 x_t + w_2\rho^2 y_{t-1,t+1}^i + w_3\rho^2 y_{t,t+1}^i + w_4\rho y_{t,t+2}^i = \rho x_{t+2|t}^i. \tag{12}$$

Since the forward signals do not refer to future periods beyond $t + 2$, $x_{t+2|t}$ is sufficient for an optimal forecast for $t + 3$ and there is no adjustment component as we have in equation (11). More generally, this result illustrates that the relationship between forecasts with two consecutive horizons would obey the simple relation of $x_{t+h|t}^i = \rho x_{t+h-1|t}^i$ for any horizon beyond the horizons of the forward signals ($h \geq 3$ in our example).

We could use $x_{t|t}^i = x_t$ and equation (9) to express the relation between the forecasts with the shortest horizons, $x_{t+1|t}^i$ and $x_{t|t}^i$:

$$x_{t+1|t}^i = \rho x_{t|t}^i + (W_1 - 1)\rho x_t + W_2 y_{t-1,t+1}^i + W_3 y_{t,t+1}^i + W_4 \rho^{-1} y_{t,t+2}^i. \tag{13}$$

The interpretation is similar to equation (11), where the simple relation $x_{t+1|t}^i = \rho x_{t|t}^i$, based on the state process, is modified due to variations in the optimal weights placed on the signal when the forecast horizon changes from $h = 0$ to $h = 1$. The optimal weights in $x_{t+1|t}^i$ are W_k, while for $x_{t|t}^i$ the whole weight is placed on the first signal, which is the perfectly observed realization of x_t (i.e., $x_{t|t}^i = x_t$). These weight differentials are multiplied by the corresponding signals as in equation (11).

This simple model with forward information can thus shed light on the empirical patterns when we examine regressions of $x_{t+h|t}^i$ on $x_{t+h-1|t}^i$. Specifically, if we estimate a cross-sectional regression of $x_{t+1|t}^i$ and $x_{t|t}^i$ generated by our model, the estimated persistence and R^2 would be biased towards zero, since $x_{t|t}^i$ has no cross-sectional variation. At the other extreme, if we estimate a cross-sectional regression of $x_{t+3|t}^i$ on $x_{t+2|t}^i$, the coefficient will be exactly ρ with a perfect fit. The "middle"-horizon regression of $x_{t+2|t}^i$ and $x_{t+1|t}^i$ will provide intermediate estimates. More generally, changes in the optimal weights across forecast horizons induce a "deviation" from the state process, manifested as a regression error which is also correlated with the regressor $x_{t+h-1|t}^i$. Consequently, the OLS coefficient estimate could be biased away from the underlying persistence parameter ρ and the fit of the regression could be poor. However, as we move to the longer horizons and estimate the regression, the variation in the optimal weights across horizons would diminish, thus shrinking the error term of the regression. As a result, the relation between forecasts horizons converges to $x_{t+h|t}^i = \rho x_{t+h-1|t}^i$ as h increases.

Our example can provide further intuition for this pattern of gradual convergence. Specifically, instead of increasing h, we can focus on the middle horizon $h = 2$, and examine how the relation between $x_{t+2|t}^i$ and $x_{t+1|t}$, as specified in equation (11), converges to $x_{t+2|t}^i = \rho x_{t+1|t}^i$ when increasing the noise in the more forward-looking signal, namely, when increasing σ_2^2. We show in Appendix B.3. how the properties of the regression (estimated ρ and R^2) change when increasing σ_2^2, in line with the empirical pattern across horizons. We also analyze how changes in ρ affect the regression properties and we find that one can generate co-movement of the estimated persistence coefficient and R^2 that mimics the empirical pattern. A summary of the results is presented in Table 1. More generally, our key insight from this analysis is that the empirical patterns are driven by the gap between optimal weights at different horizons ($w_k - W_k$), which tends to diminish in both the forecast horizon (or σ_2^2) and the underlying persistence ρ.

Finally, similar to the standard noisy information model, the augmented model with forward information can generate predictability of forecast errors, in the spirit of Coibion and Gorodnichenko

(2015). Consider the one step-ahead forecast error. After plugging equations (7) and (8) into equation (9) and taking the average across agents (and thus eliminating all the idiosyncratic terms), we get:

$$x_{t+1} - x_{t+1|t} = W_1 \omega_{t+1} + W_4(-\rho^{-1}\omega_{t+2}), \tag{14}$$

where $x_{t+1|t}$ (without superscript i) denotes the cross-sectional average (consensus) forecast. Using equations (9) and (10), and averaging across agents, we find (see Appendix B.5 for derivations) that the revision of the average forecast is:

$$x_{t+1|t} - x_{t+1|t-1} = w_1 \rho \omega_t + (w_1 + w_2 + w_3 - W_1)\omega_{t+1} + W_4 \rho^{-1}\omega_{t+2}. \tag{15}$$

Notice that both the forecast error and forecast revision contain the shocks ω_{t+1} and ω_{t+2}, which produces a correlation between the forecast error and forecast revision. In the standard noisy information setup, the correlation arises from shocks up to time t, due to the gradual processing of imperfect information about the past. By contrast, in the forward information setup, the correlation between forecast errors and revisions arises from future shocks, due to the gradual processing of *forward* information. In Appendix B.5., we derive the explicit expression for the OLS coefficient in a regression of the (mean-level) forecast error on forecast revision and show that it should be positive, in line with Coibion and Gorodnichenko (2015). Hence, the forward information model brings a different interpretation to the well-documented predictability of forecast errors.[8]

3.2. General Framework

At time t, agents receive multiple signals denoted by $y^i_{t,t+h}$, which refers to a time $t + h$ in the future, where h is the horizon of the signal, running from 0 to H periods ahead. The state-space model could simply be written as:

State:

$$x_t = \rho x_{t-1} + \omega_t, \tag{16}$$

where $\omega_t \sim iid\ N(0, \sigma^2_\omega)$.

---

[8] Recent studies have documented forecast error predictability even at the individual level, mostly with the opposite sign. Thus, forecasts underreact to information at the aggregate level but overreact at the individual level. Motivated by this evidence, they advocated a hybrid approach that combines informational frictions and behavioral frictions (Broer and Kohlhas, 2019; Bordalo et al., 2020; Angeletos et al., 2020). Such a combination is still required when information frictions are due to forward information. However, note that according to equation (14) the impulse response function of forecast errors to macroeconomic shocks may flip signs. This property speaks to the evidence on IRF of forecast errors, studied in Angeletos et al. (2020), which is explained there by over-extrapolation. Forward information may also call for a reinterpretation of individual-level results as well. For example, do forecasters overreact to all types of information in the same way? Or do they tend to do so only with respect to past information rather than forward information or vice versa?

Measurement:

$$y^i_{t,t+h} = x_{t+h} + v^i_{t,t+h}, \qquad (17)$$

where $h = 0, \ldots, H$ and $v^i_{t,t+h} \sim iid\ N(0, \sigma^2_h)$ is an idiosyncratic noise. Thus, agents have some forward noisy information referring to future values of the fundamental up until H periods ahead. There are no useful signals from $H + 1$ period and onward. We could view this as if the sequence of σ^2_h goes to infinity with h and σ^2_{H+1} is so large that a forward signal for horizon $H + 1$ is too noisy to be practically useful. Although it is natural to assume that σ^2_h increases in h (which is consistent with the data as we describe below), we do not restrict our framework to such monotonicity. For instance, some forward guidance relating to the path of inflation a year from now may represent an improved forward signal with at a horizon of a year ahead.

It is also interesting to consider a case where the current fundamental is observed perfectly, by imposing $\sigma^2_0 = 0$. In this case, all heterogeneity in expectations is driven by signals referring to the future and not because of imprecise current/past data, as it is assumed by the standard noisy-information framework. As before, the standard model is nested in our general framework by imposing σ^2_h to be infinite for any h, except for $\sigma^2_0 > 0$. Under these restrictions, there is no valuable forward information and realizations are not perfectly observed. Section 4 will use these restrictions to test our framework against the standard model.

The augmented state-space model, corresponding to equations (16) and (17), takes the following form:

State:

$$\boldsymbol{x}_t \equiv \begin{bmatrix} x_{t+H} \\ x_{t+H-1} \\ \vdots \\ x_t \end{bmatrix} = \begin{bmatrix} \rho & 0 & \cdots & 0 \\ 1 & 0 & 0 & 0 \\ \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & 1 & 0 \end{bmatrix} \boldsymbol{x}_{t-1} + S'\omega_{t+H} = \mathrm{P}\boldsymbol{x}_{t-1} + S'\omega_{t+H}, \qquad (18)$$

where $S = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$, so that the variance-covariance matrix of $S'\omega_t$ is $\Sigma_\omega = S'S\sigma^2_\omega$.

Measurement:

$$\boldsymbol{y}^i_t \equiv \begin{bmatrix} y^i_{t,t+H} \\ y^i_{t,t+H-1} \\ \vdots \\ y^i_{t,t} \end{bmatrix} = \begin{bmatrix} x_{t+H} \\ x_{t+H-1} \\ \vdots \\ x_t \end{bmatrix} + \begin{bmatrix} v^i_{t,t+H} \\ v^i_{t,t+H-1} \\ \vdots \\ v^i_{t,t} \end{bmatrix} = \boldsymbol{x}_t + \boldsymbol{v}^i_t, \qquad (19)$$

where the variance-covariance matrix of $\boldsymbol{v}^i_t$ is $\Sigma_v = diag\{\boldsymbol{\sigma}^2_v\}$ with $\boldsymbol{\sigma}^{2'}_v = \begin{bmatrix} \sigma^2_H & \sigma^2_{H-1} & \cdots & \sigma^2_0 \end{bmatrix}$, i.e., the idiosyncratic noise in the forward signals is uncorrelated across horizons. We also assume that the noise is

uncorrelated across agents and that shocks to the fundamentals and noise in the forward signal are uncorrelated (i.e., $E\left(\omega_{t+h}\boldsymbol{v}_t^i\right) = \boldsymbol{0}$).[9]

Based on this representation of the state-space, the Kalman filter can be applied to derive the optimal forecast of an agent i:[10]

$$\boldsymbol{x}_{t|t}^i = \boldsymbol{x}_{t|t-1}^i + G\left(\boldsymbol{y}_t^i - \boldsymbol{x}_{t|t-1}^i\right), \tag{20}$$

where $\boldsymbol{x}_{t|t}^i$ is a vector of forecasts made at time t, with horizons running from 0 to H steps ahead, and G is the gain matrix (with dimension $H + 1$).

Importantly, the h step-ahead forecast, $x_{t+h|t}^i$, when $h > H$, should simply obey

$$x_{t+h|t} = \rho x_{t+h-1|t}, \tag{21}$$

which resembles equation (4) in the standard noisy information framework. However, we can show (Appendix A) that the forecasts for shorter horizons ($0 < h \leq H$) in vector $\boldsymbol{x}_{t|t}^i$ has a different dynamic:

$$\begin{aligned} x_{t+h|t}^i = {}& \rho x_{t+h-1|t}^i + \left(x_{t+h|t-1}^i - \rho x_{t+h-1|t-1}^i\right) \\ & + \left(\mathbf{G}_j - \rho\mathbf{G}_{j+1}\right)\left(\boldsymbol{x}_t - \boldsymbol{x}_{t|t-1}^i\right) + \left(\mathbf{G}_j - \rho\mathbf{G}_{j+1}\right)\boldsymbol{v}_t^i, \end{aligned} \tag{22}$$

where $\mathbf{G}_j$ denotes the vector of elements in row $j = H - h + 1$ of matrix G. Equation (22) corresponds to equation (11) in our simple example above. As stressed earlier, introducing forward information that varies across horizons creates a variation across horizons in the optimal weights which did not exist in the standard noisy information framework. In equation (11), the variation in the weights was captured by $(w_k - W_k)$. Here it is captured by the term $\left(\mathbf{G}_j - \rho\mathbf{G}_{j+1}\right)$, where $\mathbf{G}_j$ and $\mathbf{G}_{j+1}$ summarize the optimal weights applied in the forecasts for consecutive horizons ($\mathbf{G}_{j+1}$ is adjusted to the subsequent horizon h by the loading ρ).

Also similar to equation (11), the weight differential is multiplied by the signals. In the more general framework, the signals are divided into three components:

1. The lag component $\left(x_{t+h|t-1}^i - \rho x_{t+h-1|t-1}^i\right)$, capturing signals from the past. This recursive form of the Kalman filter was absent in our above example, since the example above assumed that eventually the fundamental was perfectly observed.

2. The ex-post "errors" of lagged forecasts $\left(\boldsymbol{x}_t - \boldsymbol{x}_{t|t-1}^i\right)$, representing the additional information in the new forward signals.

---

[9] Although we rule out these types of correlations to simplify the analysis, our framework still imposes a structural correlation of the signals across horizons, due to the correlation between future fundamentals (i.e., x_{t+h} is correlated with x_{t+h+s}). Alternatively, the signals can refer to the future shocks ω_{t+h}s and therefore be uncorrelated. However, we will demonstrate that these types of forward signals could not account for the empirical patterns in the previous section.

[10] See Appendix A. The analysis there also covers the case with common noise in the signals.

3. The noise in the new signals (ν_t^i).

Equation (22) is key for understanding the empirical patterns documented in the previous section. It defines the missing error term in the regression of $x_{t+h|t}^i$ on $x_{t+h-1|t}^i$ as in specification (4). Furthermore, it shows that the error-term is correlated with the regressor $x_{t+h-1|t}^i$ since the forecast $x_{t+h-1|t}^i$ applies the same signals that are captured by the error term. Consequently, the estimated coefficient would be biased, while the sign of the bias depends on the signs of the elements in the weight differential vector $\left(\mathbf{G}_j - \rho\mathbf{G}_{j+1}\right)$.[11] When the horizon is beyond H, the weight differential would shrink so that the bias goes to zero and the R^2 converges to 1. As in the simple model above, equation (22) further allows us to extract the forward signals about future fundamentals which we cover in Section 5.

3.3. Simulation

The purpose of the simulation is to examine whether our forward information framework can qualitatively replicate the empirical patterns documented in the previous section. We perform simulations that use different degrees of persistence in the state process. Each simulation is based on 1,000 draws and includes the four following steps:

I. *Simulating the state equation*: The state process is simulated with a certain degree of persistence (ρ) for a period similar to the SPF survey (about forty years of quarterly data). The variance of the shocks to the fundamental is set to $\sigma_\omega^2 = 1$.

II. *Simulating the measurement equation*: Forward noisy signals are simulated for a group of 40 forecasters (similar to the number of participants in the SPF survey). The horizon of the signals runs from 0 to 7, and the vector of the noise variance is set to $\boldsymbol{\sigma}_\nu^{2'} = [10000 \quad 100 \quad 4 \quad 3 \quad 2 \quad 1 \quad 0.5 \quad 0.2]$. This structure assumes that noise increases in the horizon, and signals become extremely uninformative for horizons $h = 6$ and $h = 7$, where the noise variance goes to 100 and 10000, respectively (and then to infinity). To examine if the pattern of increasing noise is essential, we also conduct a second set of simulations in which all the variance parameters in $\boldsymbol{\sigma}_\nu^{2'}$ are set to 2.

III. *Computing the forecasts*: The gain matrix G is computed, based on the Kalman filter, and then used to calculate optimal forecasts, with horizons running from 0 to 7, for the 40 simulated forecasters, using equation (20).

---

[11] By contrast, if forward signals refer purely to future shocks, as mentioned in footnote 9, the error-term would not be correlated with $x_{t+h-1|t}^i$ and the estimated coefficient would not be biased.

IV. *Estimating regressions*: For a certain "quarter" in the middle of the simulated sample, we run cross-sectional regressions of $x_{t+h|t}^i$ on $x_{t+h-1|t}^i$, for each h between 1 to 7, and obtain the coefficient estimate and R^2 statistic.[12]

Figure 5 reports the estimated persistence $\hat{\rho}$ (Panel A) and R^2 statistics (Panel B), measured on the vertical axis. The value of ρ, applied in each simulation, is indicated by the first horizontal axis. The second horizontal axis indicates the horizon of each regression (i.e. $h = 3$ corresponds to a regression of $x_{t+3|t}^i$ on $x_{t+2|t}^i$). The description is therefore equivalent to Figure 1, with a further dimension of variation in ρ.

The simulations qualitatively reproduce all patterns documented in the previous section. Specifically, both the estimated persistence and R^2 demonstrate the pattern of convergence across horizons. For short horizons their values are low. As the regression is estimated for longer horizons the coefficient gets closer to ρ and the fit of the regression gets stronger. In addition, for each horizon, $\hat{\rho}$ and R^2 get higher when ρ is higher. Appendix C provides more details of the simulation results, which are in line with the SPF evidence. It also makes a comparison with another set of simulations, in which the variance of the noise in the signals is fixed across the horizons.

Besides rationalizing the empirical patterns, our model and simulations are instructive about the detection of horizon H, namely, the point in the future from which forward signals are no longer informative. The simulation results imply a convergence of the coefficient estimate to the value of ρ. On one hand, we have seen that estimating the persistence of inflation with specification (4) could lead to a biased measure, mainly downward. On the other hand, this bias would vanish if we use forecasts with long enough horizons, for which forward signals are almost uninformative. Similarly, in Figures 1 and 2, we observe signs of convergence for $\hat{\rho}$ at horizons $h = 3,4$. This convergence thus indicates that the signals at these horizons should be weak. Hence, SPF forecasters obtain useful forward information until about a year ahead on average. Furthermore, at these horizons, the estimated persistence is less likely to be biased.

4. Forward information: Direct evidence

Building on Goldstein (2021), we can have a more direct test of forward information by focusing on the deviation of an individual forecast from the consensus forecast. Note that the optimal vector of forecasts in equation (20) could be expressed as follows (using the measurement equation (19)):

---

[12] As in Figure 2, we apply a narrow window of 8 cross-sections ending at the chosen "quarter". Similar qualitative results are obtained with a single cross-section.

$$x_{t|t}^i = x_{t|t-1}^i + G(y_t^i - x_{t|t-1}^i) = x_{t|t-1}^i + G(x_t + v_t^i - x_{t|t-1}^i)$$
$$= (I - G)x_{t|t-1}^i + G(x_t + v_t^i). \tag{23}$$

Next, we take the average across individuals (hence we drop superscript i) and obtain:

$$x_{t|t} = (I - G)x_{t|t-1} + Gx_t, \tag{24}$$

where $x_{t|t}$ and $x_{t|t-1}$ are cross-sectional averages (consensus forecasts) of $x_{t|t}^i$ and $x_{t|t-1}^i$, respectively. Subtracting equation (24) from equation (23) we get

$$x_{t|t}^i - x_{t|t} = (I - G)(x_{t|t-1}^i - x_{t|t-1}) + Gv_t^i. \tag{25}$$

Equation (25) describes a simple relationship between the deviation of an individual forecast from the consensus forecast in period t and the lagged deviation from period $t - 1$. Because idiosyncratic noise v_t^i is uncorrelated with forecasts made at time $t - 1$,[13] we can use OLS to directly estimate the elements in the gain matrix, row by row, by running a regression of the deviation from the mean on lagged deviation, for each forecast horizon:

$$x_{t+h|t}^i - x_{t+h|t} = \beta_0(x_{t+H|t-1}^i - x_{t+H|t-1})$$
$$+ \beta_1(x_{t+H-1|t-1}^i - x_{t+H-1|t-1}) + \cdots \tag{26}$$
$$+ \beta_H(x_{t|t-1}^i - x_{t|t-1}) + error_t,$$

where the β coefficients are elements of row $H - h + 1$ in the matrix $(I - G)$. Note that the set of regressors in equation (26) is the same for all h.

Specification (26) can be viewed as an augmented version of the standard noisy information model where agents receive noisy signals about the current value of the fundamental. Indeed, Goldstein (2021) has proposed the following specification for the standard model:

$$x_{t+h|t}^i - x_{t+h|t} = \beta_{NOISY}(x_{t+h|t-1}^i - x_{t+h|t-1}) + error_t, \tag{27}$$

where β_{NOISY} is equal to $(1 - G_{NOISY})$ and G_{NOISY} is the Kalman gain, representing the weight placed on the single noisy signal. The standard framework imposes a restriction on the estimated matrix $(I - G)$, requiring that all the off-diagonal elements should be equal to zero. This provides a straightforward test of our model with forward information against the null of the standard noisy-information version without forward signals: We simply need to estimate specification (26) and test the significance of all the coefficients, other than the coefficient on $(x_{t+h|t-1}^i - x_{t+h|t-1})$.

---

[13] A similar specification is obtained, when adding a common noise besides the individual-specific noise. As demonstrated in appendix A, by taking the deviation of the individual forecast from the mean, the term with the common noise will be dropped out, since it appears in the same way in both forecasts.

Specification (26) is also useful for testing the null that heterogeneity in information may only be due to forward signals, while information about realized inflation is not noisy as in the standard framework. Under the null, the coefficients in the regression for $h = 0$ (i.e., backcasts of the last quarter are also provided by SPF participants) should all be zeroes. Intuitively, when all forecasters observe realized inflation in the same way, the deviation of their backcasts from the mean should not be persistent.[14] We implement this test in Table 2. We find that none of the coefficients are significant when the regression is estimated on the full sample. Although some coefficients are statistically significant when we estimate specification (26) decade by decade, the estimates are economically small and R^2 stays close to zero. These results suggest that information is noisy because of forward signals rather than signals about the current state of the fundamental.

To explore the importance of forward information further, we estimate specification (26) for longer horizons ($h > 0$) and test the restrictions implied by specification (27). As reported in Table 3, the coefficients on $\left(x_{t+h|t-1}^i - x_{t+h|t-1}\right)$, which represent diagonal elements, are strongly significant. But in each estimation, there is at least one additional coefficient that is highly significant, despite the fact that regressors tend to be correlated. In other words, we have at least one non-zero off-diagonal element in each row of the gain matrix (columns in Table 3 correspond to rows in the matrix $I - G$). This is consistent with SPF forecasters utilizing forward signals. Furthermore, information criteria suggest that including other horizons in specification (26), as opposed to (27), improves the fit considerably, which is consistent with important forward information. Although it is hard to provide economic interpretation for the estimated coefficients, we estimate specification (26) on simulated data to check if the estimates in Table 3 are plausible. We find (Appendix C) that in the more realistic case, where the variance of the noise increases in the horizon, the simulation provides patterns that are similar to Table 3.[15] For example, the dominant coefficients are those that correspond to the diagonal elements in $(I - G)$ and they tend to increase along the diagonal, namely, when the specification is estimated for longer horizons. The standard noisy information model cannot reproduce this pattern. In summary, forward information appears to be quantitatively important.

5. Forward information: Measurement and applications

---

[14] More formally, the null imposes a restriction on the last row of the gain matrix, specifically, $\mathbf{G}_{H+1} = [0 \quad 0 \quad \cdots \quad 1]$. Hence, the corresponding row in $(I - G)$, estimated by (23), should be a vector of zero coefficients. The tractable example in Section 3.1. implies such a restriction. See also the analysis in Appendix B.2.

[15] In Appendix C, we also verify that truncation of longer horizons with additional information, which can happen in practice, does not raise concerns about the available coefficient estimates.

Measurement of news about the future is usually a challenging task that requires structural restrictions or additional variables. We propose an alternative approach to recover forward information from expectations data. We first illustrate how the quantification of forward information directly follows from our framework, and then use several applications to demonstrate our method, focusing on SPF inflation forecasts.

5.1.　Quantifying forward information

We can re-write equation (22) in the following form:

$$x_{t+h|t}^i = \rho x_{t+h-1|t}^i + FI_{t+h|t}^i, \tag{28}$$

where

$$FI_{t+h|t}^i \equiv \left(x_{t+h|t-1}^i - \rho x_{t+h-1|t-1}^i\right) + (\mathbf{G}_{h+1} - \rho \mathbf{G}_h)\left(\mathbf{x}_t - \mathbf{x}_{t|t-1}^i\right) + (\mathbf{G}_{h+1} - \rho \mathbf{G}_h)\boldsymbol{v}_t^i$$

and $FI_{t+h|t}^i$ represents the adjustment to the forecast induced by forward information. The expression for $FI_{t+h|t}^i$ has three terms. The first term in the expression is forward information inherited from the previous period. The second term is new information from observing $\boldsymbol{x}_t$. The last term is the noise contained in the signal about $\boldsymbol{x}_t$. One can also use equation (11) to re-write $FI_{t+h|t}^i$ as:

$$FI_{t+h|t}^i = \sum_{k=1}^K \left(W_{k,h} - W_{k,h-1}\right) Signal_{k,t+h|t}^i, \tag{29}$$

which underscores that forward information is a function of signals about the future.

Although signals in equation (29) are not observed, one can readily recover forward information using equation (28):

$$FI_{t+h|t}^i = x_{t+h|t}^i - \rho x_{t+h-1|t}^i. \tag{30}$$

By taking the average across forecasters, we obtain an aggregate time series of forward information:

$$FI_{t+h|t} = x_{t+h|t} - \rho x_{t+h-1|t}. \tag{31}$$

Equation (31) suggests a simple method for quantifying forward information. To illustrate how it works, suppose we are interested in constructing forward information at horizon $h = 0$. To this end, we compute:

$$FI_{t|t} = x_{t|t} - (\hat{c}_{t-1} + \hat{\rho}_{t-1} x_{t-1}), \tag{32}$$

where c_t and ρ_t are time-varying intercept and slope that generalize equation (31) by applying the following steps:

(i)　Forecast data: For $x_{t|t}$, we use the average inflation forecasts for the current quarter. The backcast $x_{t-1|t}$ that refers to inflation in the previous quarter is replaced by actual lag x_{t-1}. As

we show above, the information on realized inflation is almost perfect, making x_{t-1} equivalent to the backcast.

(ii)   Persistence parameter: Because forecasters may use persistence ρ that it is different from the actual value, we need to recover ρ from forecasts. As we show above, ρ estimated by OLS varies with the horizon because OLS is biased when forward information is present. However, as we increase the forecast horizon, one may expect forward information to become less precise and thus the bias in the OLS estimate of ρ should decline. Consistent with this notion, we observe that R^2 and $\hat{\rho}$ increase with the horizon and stabilize at long horizons. Because some forecasts are missing at $h = 4$, we use $h = 3$ to estimate ρ, that is, we regress $x_{t+3|t}^i$ on $x_{t+2|t}^i$ and we essentially use the time-varying estimates in Figure 2. Note that we use $\hat{\rho}_{t-1}$ rather than $\hat{\rho}_t$ to recover $FI_{t|t}$ because the time-varying $\hat{\rho}_t$ may contain some news with respect to changes in the persistence parameter ($\hat{\rho}_t$ will be used to evaluate $FI_{t+h|t}$ for $h > 0$).

(iii)   Constant parameter: The AR(1) state equation in equation (16) (and equation (31)) did not include a constant. In practice, we include a constant that can capture time-varying trend inflation. The constant is estimated by the time-varying regressions from the previous step, and we take again a lag $\hat{c}_{t-1}$, due to news about changes in trend inflation that are embedded in $\hat{c}_t$.[16]

Using the same logic, we can recover forward information for other horizons from:

$$FI_{t+h|t} = x_{t+h|t} - \left(\hat{c}_t + \hat{\rho}_t x_{t+h-1|t}\right). \tag{33}$$

Note that this equation measures only the "marginal" forward information across horizons.[17]

Figure 6 plots the series of forward information produced by equations (32) and (33) for CPI inflation and the corresponding projections in the SPF. Panel A shows a strong co-movement of $FI_{t|t}$ and actual inflation. In particular, times of high variation in forward information corresponds to times of high variation in inflation. Other panels show series for $h = 1, \dots, 4$. The series of forward information fluctuate quite closely to each other and variation diminishes as the horizon increases.

More generally, equations (28) and (29) suggest several properties for forward information. First, the variation of forward information over time decreases in the horizon. This pattern is in line with diminishing information in forward signals for longer horizons. It is also driven by the decay in weight

---

[16] The results are quite similar though, when employing $\hat{c}_t$ and $\hat{\rho}_t$, as well as when employing the survey backcasts instead of the actual lag x_{t-1}.

[17] Although we stick here to a simplified AR(1) process, our method can be easily extended to processes of higher order by the same principles. For example, the representation of equation (32) for VAR (1) would be $FI_{t|t} = x_{t|t} - (\hat{c}_{t-1} + \hat{\boldsymbol{\rho}}_{t-1}\boldsymbol{x}_{t-1})$ where $\boldsymbol{x}_{t-1}$ is a vector of variables and $\hat{\boldsymbol{\rho}}_{t-1}$ is a vector of coefficient estimates. The coefficients will again be estimated using forecast data, by regressing $x_{t+h|t}$ on a vector of forecasts $\boldsymbol{x}_{t+h-1|t}$.

differentials $\left(W_{k,h} - W_{k,h-1}\right)$ across horizons. As illustrated in Section 3, $W_{k,h-1}$ approaches $W_{k,h}$, as the horizon increases. Second, the series for forward information are correlated across horizons because the same signals are applied at each horizon. The correlation should eventually decay due to the diminishing variation. Third, series for forward information should be serially correlated due to the overlap in forward signals over time. That is, previous forward signals which look beyond time t are still useful for the forecast made at time t. These properties are broadly supported by the summary statistics presented in Table 4.[18]

5.2.   Forward information and predictability of inflation

If useful, forward information should help to predict future inflation. To assess the quality of forward information, we regress future values of inflation on lagged inflation and forward information. We report results in Table 5. Column (2) provides estimates for the nowcast. This regression has two predictors: the lagged inflation and forward information $FI_{t|t}$. The coefficient on forward information is highly significant and R^2 rises dramatically by ≈ 0.6 compared to a regression without $FI_{t|t}$ (column (1)). Thus, forward information can explain a large share of current inflation.

Each subsequent column reports estimates from a regression of inflation at time $t + h$ on lagged inflation and all forward information components available at time t, which refer to future periods up to $t + h$. We find a significant coefficient for $FI_{t+1|t}$. For longer horizons the effect is not significant, but so is the effect of lagged inflation. Overall, these findings are in line with the evidence in Section 2. As demonstrated above, the estimates of persistence regressions in Figure 2 were biased due to forward information, which is more dominant at the shorter horizons. The results in Figure 2 illustrate that the horizons with the main biases are $h = 0,1$. Thus, the main predictability of forward information should be obtained for these horizons.[19]

Another way to gauge the importance of forward information is to examine the ratio $\sqrt{\sum_t\left(FI_{t+h|t}\right)^2}\Big/\sqrt{\sum_t\left(x_{t+h|t}\right)^2}$ where the numerator is the variation in marginal forward information (that is, information that applies only to horizon h) and the denominator is the variation in the consensus forecast $x_{t+h|t}$. A higher value of the ratio represents a larger contribution of forward information to forecasts. Table

---

[18] Another related implication of our measure of forward information, based on the theoretical analysis in Section 3, is that variation in news over time may induce variations in expectations' persistence, especially at the shorter horizons. In times of big events, with potentially high provision of forward information, this biasing effect can be particularly strong, even at longer horizons. Appendix E explores such events and reports evidence in this direction, especially following the outbreak of COVID-19.

[19] In Appendix E we present another related application in which we analyze the impulse response to our inflation forward information in VARs. We also compare our approach to the approach of news shocks in Barsky and Sims (2011).

4 reports that the variance of $FI_{t|t}$ is relatively close to the variance of actual inflation which points toward a high contribution of forward information. We find that the ratio for the nowcast is 0.356 for the full sample but the ratio has been increasing over time.[20] For longer horizons, the ratio is 0.157 for $h = 1$ and declines to 0.033 for $h = 4$. The results suggest that professional forecasters have a lot of additional information about current and near-future inflation but the quality of (marginal) forward information rapidly declines in forecast horizon.

5.3. Forward information and disagreement

As we discussed above, variation in forward information across agents can be a source of disagreement in forecasts. Although previous subsections focus on consensus forecasts, we can apply the same algorithm to the data at the forecaster level. In particular, we can estimate persistence ρ^i and intercept c^i for each forecaster i and then compute the "model" component of the forecast with $\left(\hat{c}_t^i + \hat{\rho}_t^i x_{t+h-1|t}^i\right)$ and hence express forward information as $FI_{t+h|t}^i = x_{t+h|t}^i - \left(\hat{c}_t^i + \hat{\rho}_t^i x_{t+h-1|t}^i\right)$.[21] In the next step, we compare the cross-sectional variation in forecasts $(x_{t+h|t}^i)$ with the cross-sectional variation in forward information $(FI_{t+h|t}^i)$. We report results in Table 6.

We find that forward information accounts for a large fraction of cross-sectional variation in the data. For example, nowcasts and one-step-ahead forecasts are characterized by $var\left(FI_{t+h|t}^i\right) > var\left(x_{t+h|t}^i\right)$. This pattern also underscores the negative correlation between forward information ("the error term" in equation (28)) and the model-implied forecast $(\hat{c}_t^i +\hat{\rho}_t^i x_{t+h-1|t}^i)$, which rationalizes the bias in the estimates of inflation persistence. As we increase the horizon, the cross-sectional variation in forward information declines faster than the cross-sectional variation in forecasts, which yields $var\left(FI_{t+h|t}^i\right) < var\left(x_{t+h|t}^i\right)$ for $h \geq 2$. However, even at longer horizons ("marginal") forward information accounts for more than a third of the variation in forecast disagreement and thus remains a meaningful contributor to forecast dispersion.

5.4. Forward information and the Taylor rule

---

[20] Appendix Figure G.5. presents time series for the ratio for the nowcast. The ratio fluctuates between 18% to 66% over the sample period. The contribution of forward information deteriorates following the years of high inflation which can result from a growing inattention to inflation following the Great Moderation and the decline in trend inflation. However, the ratio recovers in the recent years which can result from higher provision of informative forward signals in recent years, potentially due to forward guidance provided by the central bank.

[21] We also follow the analysis in Figure 4 and estimate time-varying persistence and intercept for each forecaster.

If forward information helps predict inflation, one may expect policymakers to act on this information. Indeed, given that monetary policy affects the economy with significant lags, policymakers should react to news about changes in future inflation rather than wait until changes in inflation materialize. Following Romer and Romer (2004) and Coibion and Gorodnichenko (2011), we apply the following specification to test this prediction:

$$r_t = c + \gamma \pi_{t|t}^{GB} + \theta_1 gap_{t|t}^{GB} + \theta_2 gr_{t|t}^{GB} + \rho_1^r r_{t-1} + \rho_2^r r_{t-2} + \varepsilon_t, \qquad (34)$$

where r_t is the federal funds rate, $\pi_{t|t}^{GB}$ is expected CPI inflation (nowcast) and $gap_{t|t}^{GB}$ and $gr_{t|t}^{GB}$ are expectations (nowcast) of the output gap and GDP growth, respectively. We use forecast data from the Fed Greenbooks (hence, the superscripts GB), which are prepared by the Fed's staff before FOMC meetings. Our focus is on the effect of $FI_{t+h|t}$, which we construct following equations (32) and (33).

Table 7 presents estimation results of specification (34) for 1983Q1-2015Q4.[22] Each column applies different measures of expected inflation. Column (1) presents results for the standard policy reaction function estimated in the literature. The estimates are broadly consistent with the results reported in earlier work. For example, the policy response is highly inertial ($\hat{\rho}_1 + \hat{\rho}_2 \approx 0.95$) and the long-run response to inflation is generally consistent with determinacy ($\hat{\gamma}/(1 - \hat{\rho}_1 - \hat{\rho}_2) \approx 1.05$). The dynamics of the fed funds rate is largely explained by macroeconomic conditions ($R^2 = 0.982$).

In column (2), we use $\pi_{t|t}^{GB} = (\hat{c}_{t-1} + \hat{\rho}_{t-1} \pi_{t-1}) + FI_{t|t}^{GB}$ (i.e., equation (32)) to separate the contributions due to the "model" component $(\hat{c}_{t-1} + \hat{\rho}_{t-1} \pi_{t-1})$ and forward information $(FI_{t|t}^{GB})$. We modify equation (34) to have different responses to these two components. We find that only the forward-information component is significant, while the "model" component of $\pi_{t|t}^{GB}$ is not. These results suggest that, like nowcasts in the SPF, the Greenbook nowcasts for inflation contain information useful for policymakers and they respond to this information strongly.

Column (3) presents results for an augmented specification where in addition to $FI_{t|t}^{GB}$ we control for forward information at longer horizons. We find that policymakers significantly respond to $FI_{t|t}^{GB}$ and $FI_{t+1|t}^{GB}$. The estimated coefficients are positive for longer horizons ($h > 1$ in column (3)) but they are imprecisely estimated. These findings are in line with our earlier results for the SPF, i.e., forward information is useful at short horizons but its ability to forecast at longer horizons is limited.

---

[22] The Greenbooks are published with a lag of five years. To apply (32) and (33) we use the time-varying estimates of $\hat{c}_t$ and $\hat{\rho}_t$ that were obtained above using cross-sections of SPF forecasts. For this reason, the sample period starts at 1983 which is also right after the Volcker disinflation. For lagged inflation, we take backcasts of the Greenbooks to align with the Greenbook's definition of quarterly inflation. There are 8 Greenbooks every year which precede the meetings of the FOMC, usually two in a quarter. In each quarter we take the forecasts from the earlier Greenbook which are provided around the same time of the SPF forecasts. The Greenbooks are available at the website of the Philadelphia Fed.

To assess the stability of this pattern, we split the sample roughly in halves (1983-1999 and 2000-2015) and report results in columns (4) and (5). We find that forward information can predict policy in both subsamples. However, forward information at longer horizons appears to be more important in the second subsample: the coefficient on $FI_{t+2|t}^{GB}$ is now also statistically significant. We conclude that forward information plays a significant role in shaping monetary policy.

6. Evidence from additional macro forecasts

To preserve space, our empirical analysis has been focused on inflation expectations from the SPF. This section briefly reports results for additional macroeconomic expectations. For a detailed description of the results, see Appendix F.

More SPF forecasts. We examine forecast data of other measures of inflation (core, PCE, and GDP deflator inflation) and additional key variables (GDP growth, unemployment and interest rates). We first estimate the persistence regressions from Section 2 across the forecast horizons. The results broadly confirm the stylized patterns documented in Section 2: both the persistence estimate and the fit of the regression tend to increase with the forecast horizon. Some variables show more sensitivity to the underlying process so that the pattern is more clearly demonstrated for higher-order AR processes. We also apply the direct approach proposed in Section 4. We again find evidence consistent with forward information. We then extract the forward information component for some variables, applying the method from Section 5. We find that this measure of forward information has significant predictive power for future movements in the corresponding variables.

Fed forecasts. In Section 5.4, we estimate a Taylor rule using projections from the Fed's Greenbooks. The availability of Greenbook forecasts for multiple horizons enables us to examine if the stylized pattern of persistence that was documented in the SPF (Section 2) also exists in forecasts of the Fed. Since there is no dispersion in Greenbook forecasts, we estimated time-series regressions. Yet, the pattern of increasing persistence across horizons is detected in Greenbook forecasts for several key variables (for comparison we also estimated similar time-series regressions with mean-level SPF forecasts). Thus, beyond the specific application of the Taylor rule, this evidence points to a more general utilization of forward information by the Fed staff.

ECB SPF inflation forecasts. Another Survey of Professional Forecasters is managed by the European Central Bank. We investigate the presence of forward information in forecasts of European CPI inflation. In the U.S. SPF we used quarterly forecasts for multiple horizons which are not available in the ECB SPF.

Instead, we took advantage of two types of annual forecasts that are provided in each quarter: calendar forecasts referring to current and next calendar years and rolling forecasts referring to a year and two years ahead. We estimate persistence in both types of forecasts and document a new type of variation in the first type. Specifically, persistence in calendar-year forecasts decreases over the calendar quarters of the year. This pattern is consistent with the presence of forward information that varies across horizons. In the calendar forecasts, the horizon changes when moving from quarter to quarter so that forward information increases over the year with a biasing effect on the estimated persistence. This pattern does not appear in the rolling-year forecasts, for which the horizon is fixed in each quarter. Interestingly, we find the same effect in annual forecasts of the U.S. SPF. These findings provide additional support for the presence of forward information in both surveys.

7. Conclusion

Although there has been an explosion of research on how expectations are formed and departures from full-information rational expectations are increasingly clear, much remains to be learned. There is also a new sense of urgency to shed more light on expectations as central bankers and other government bodies must increasingly rely on tools based on the management of expectations (e.g., forward guidance) in the hopes of appropriately changing the beliefs about future macroeconomic aggregates and hence stabilize the economy.

We propose a new approach to characterize the expectations formation process. We formalize the fact that projections are a combination of model-based prediction and add-factoring, which uses forward information that is not incorporated into the model and current/past values of macroeconomic variables. Specifically, we extend the canonical noisy-information model by introducing signals about future fundamentals (forward information) at multiple horizons. Thus, we emphasize information variation not only across agents but also across future horizons. We find that cross-sectional dispersion in SPF forecasts is driven by forward information. That is, forecasters disagree about the future while holding (almost) perfect information about the realized values. The presence of forward information can also rationalize why perceived persistence of fundamentals increases in the forecast horizon, an empirical pattern that we document in the SPF data. We propose a simple method to extract the forward information component in forecast data and demonstrate the usefulness of this direct measure for forecasting and policymaking.

We view these findings as a first step in utilizing forward information to understand how economic agents form expectations. For example, we use forward information to directly measure the impact of central bank communication on the beliefs of economic agents. We focused on professional forecasters who are some of the most informed players in the economy and who likely place a lot of weight on forward information.

Consumers and firms may be less sophisticated and more inattentive to incoming macroeconomic data. As a result, they could be better characterized by a mix of noisy information about past, current and future fundamentals. To determine what type of information is a good description of consumers and firms, one may need to run customized surveys of these agents that mimic the structure of the SPF.

References

Andrade, Philippe, and Herve Le Bihan (2013). "Inattentive Professional Forecasters," *Journal of Monetary Economics* 60(8), 967-982.

Ang, Andrew, Geert Bekaert, and Min Wei (2007). "Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?" *Journal of Monetary Economics* 54(4), 1163-1212.

Angeletos, George-Marios, Zhen Huo, and Karthik A. Sastry (2020). "Imperfect Macroeconomic Expectations: Evidence and Theory," *NBER Macroeconomics Annual 2020*, 35.

Barsky, Robert B., and Eric R. Sims (2011). "News Shocks and Business Cycles." *Journal of Monetary Economics* 58(3), 273–89.

Barsky, Robert B., and Eric R. Sims (2012). "Information, Animal Spirits, and the Meaning of Innovation in Consumer Confidence." *American Economic Review* 102 (4), 1343–77.

Beaudry, Paul, and Franck Portier (2014). "News-Driven Business Cycles: Insights and Challenges." *Journal of Economic Literature* 52(4), 993–1074.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma and Andrei Shleifer (2020). "Over-reaction in Macroeconomic Expectations," *American Economic Review* 110(9), 2748–2782.

Broer, Tobias, and Alexandre N. Kohlhas (2019). "Forecaster (Mis-)Behavior," Working Paper.

Capistrán, Carlos, and Allan Timmermann (2009). "Disagreement and Biases in Inflation Expectations." *Journal of Money, Credit and Banking* 41(2-3), 365-396.

Chahrour, Ryan, and Kyle Jurado (2018). "News or Noise? The Missing Link" *American Economic Review* 108(7), 1702-1738.

Chahrour, Ryan, and Kyle Jurado (2021). "Optimal Foresight" *Journal of Monetary Economics* 118, 245-259.

Coibion, Olivier, and Yuriy Gorodnichenko (2011). "Monetary Policy, Trend Inflation and the Great Moderation: An Alternative Interpretation." *American Economic Review* 101(1), 341-370.

Coibion, Olivier, and Yuriy Gorodnichenko (2012). "What Can Survey Forecasts Tell Us About Informational Rigidities?" *Journal of Political Economy* 120(1), 116-159.

Coibion, Olivier, and Yuriy Gorodnichenko (2015). "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *American Economic Review* 105(8), 2644-2678.

Coibion, Olivier, Yuriy Gorodnichenko, and Rupal Kamdar (2018). "The Formation of Expectations, Inflation and the Phillips Curve," *Journal of Economic Literature* 56(4), 1447-1491.

Driscoll, John C., and Aart C. Kraay (1998). "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data," *Review of Economics and Statistics* 80(4), 549-560.

Elliott, Graham, Ivana Komunjer, and Allan Timmermann (2008). "Biases in Macroeconomic Forecasts: Irrationality or Asymmetric Loss?" *Journal of European Economic Association* 6(1), 122–157.

Fair, Ray C. (1986). "Evaluating the Predictive Accuracy of Models," in *Handbook of Econometrics*, Volume III, edited by Zvi Griliches and Michael D. Intriligator, Elsevier, 1979-1995.

Fuhrer, Jeffrey. C. (2011). "Inflation persistence." In B. M. Friedman and M. Woodford (Eds.), *Handbook of Monetary Economics*, Chapter 9, Volume 3, pp. 423—486. San Diego CA: Elsevier.

Fuhrer, Jeffery, (2018). "Intrinsic Expectations Persistence: Evidence from Professional and Household Survey Expectations." Federal Reserve Bank of Boston Research Department Working Paper 18-9.

Giacomini, Raffaella, Vasiliki Skreta, and Javier Turen (2020). "Heterogeneity, Inattention, and Bayesian Updates," *American Economic Journal: Macroeconomics* 12(1), 282–309.

Goldstein, Nathan (2021). "Tracking Inattention." Manuscript.

Jain, Monica (2019). "Perceived Inflation Persistence," *Journal of Business and Economic Statistics* 37(1), 110-120.

Kohlhas, Alexandre. N. and Ansgar Walther (2021). "Asymmetric Attention," *American Economic Review*, Forthcoming.

Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers (2004). "Disagreement about Inflation Expectations," *NBER Macroeconomics Annual 2003* 18, 209–248.

Nguyen, Thuy Lan, and Wataru Miamoto (2019). "The Expectational Effects of News in Business Cycles: Evidence from Forecast Data." *Journal of Monetary Economics*, Forthcoming.

Patton, Andrew J., and Allan Timmermann (2010). "Why Do Forecasters Disagree? Lessons from the Term Structure of Cross-sectional Dispersion," *Journal of Monetary Economics* 57(7), 803–820.

Romer, Christina D., and David H. Romer (2004). "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review* 94(4), 1055-1084.

Stark, Tom (2013). "SPF Panelists Forecasting Methods: A Note on the Aggregate Results of a November 2009 Special Survey", Federal Reserve Bank of Philadelphia Research Department.

Woodford, Michael (2002). "Imperfect Common Knowledge and the Effects of Monetary Policy." In: *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, edited by Philippe Aghion, Romain Frydman, Joseph Stiglitz and Michael Woodford, Princeton University Press, 25-28.

Figure 1: Persistence Across Forecast Horizons

Panel A: 1981-2017 (full sample)


Note: p-value (equality of estimated coefficients) <0.01
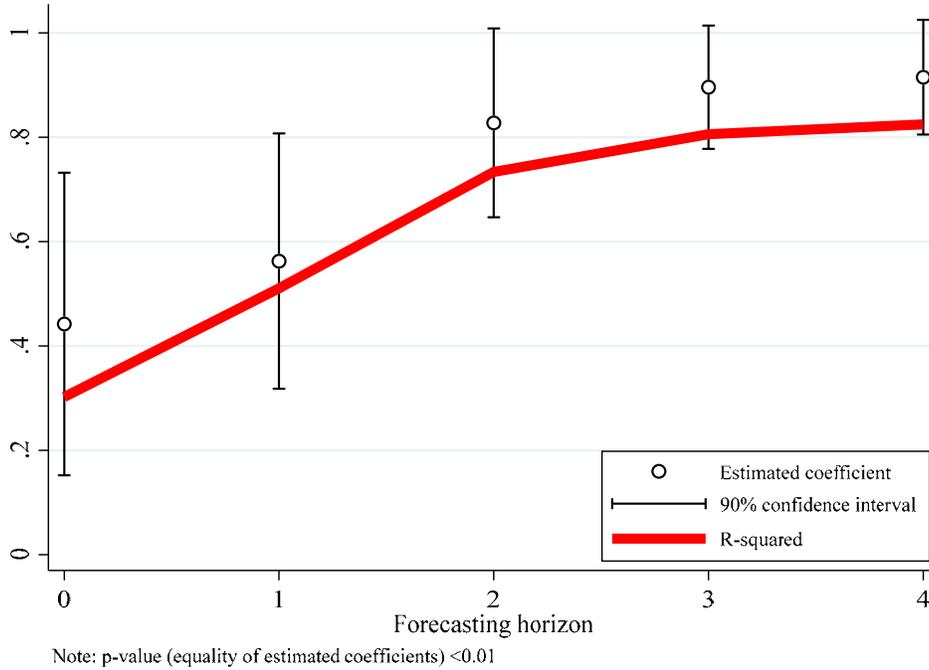
Panel B. Results by decade

Note: p-value (equality of estimated coefficients) <0.01
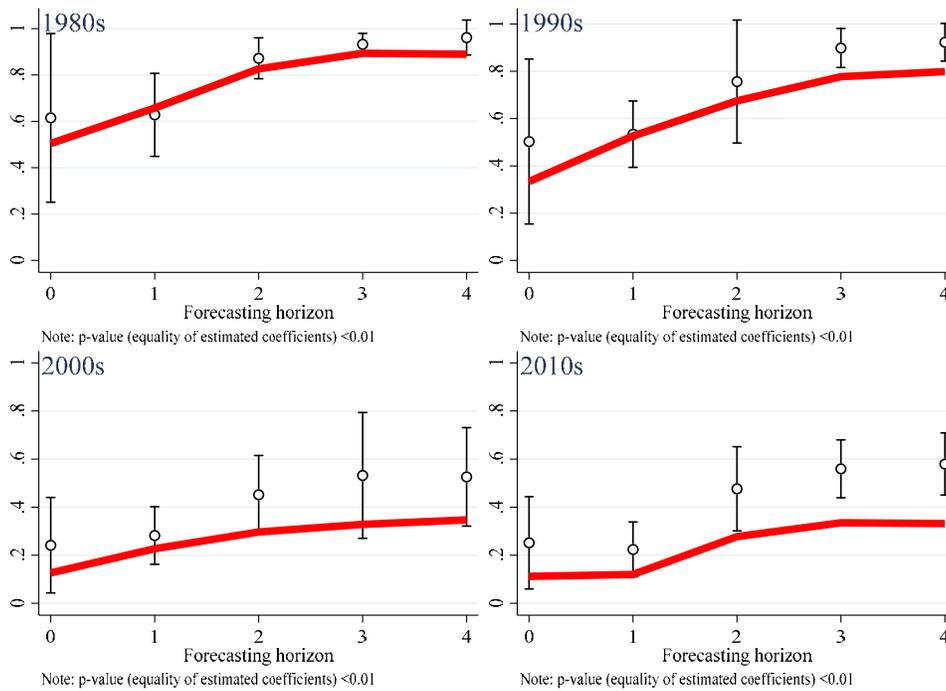
Notes: The figure plots persistence coefficients and R-squared statistics, based on estimating specification (4) for different forecast horizons, using different CPI inflation forecasts from the SPF survey. The whiskers show the confidence interval around the point estimate, based on Driscoll and Kraay (1998) standard errors. P-values refer to the test of equality of coefficients across forecasting horizons.

Figure 2: Expectation-Based Persistence by Forecasting Horizon

Panel A: Persistence Estimates

Panel B: R-Squared of Persistence Regressions
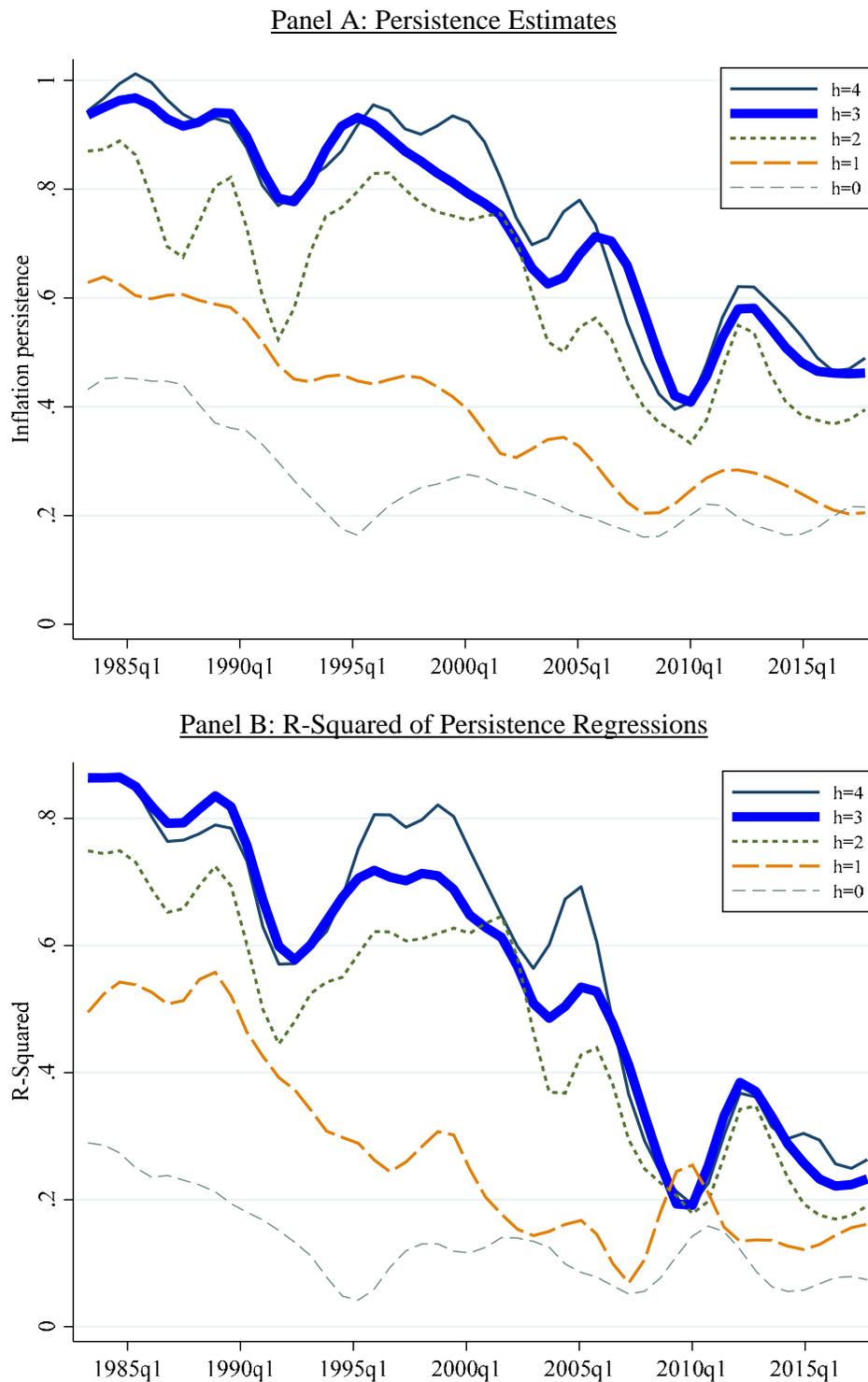
Notes: The figure plots smoothed estimates of persistence $\hat{\rho}$ (panel A) and R-squared measures (panel B) based on estimating specification (4) for different forecast horizons in the SPF survey. Each quarterly point is based on OLS estimation using the forecasts data from the last 8 quarters for a specific horizon. The smoother is a local mean which uses an Epanechnikov kernel.

Notes: The figure plots R-squared measures from four different specifications which were estimated for each quarter of the sample period, using SPF inflation forecasts. In panel A, inflation expectations follow an AR(1) process as in specification (4). In panel B inflation expectations follow an AR(4) process, by including forecasts for additional quarters as regressors. In panel C expectations of the inflation gap follow an AR(1) process. The forecasted inflation gap is computed as the difference between forecasts of inflation and 10-year inflation (sample starts on 1990Q2 because of 10-year forecasts). In panel D inflation expectations follow a VAR(4) process which augments the specification in panel B with four "lags" of unemployment and 3-month interest rate forecasts. All specifications are estimated for $h = 3$. Each point on the black lines is based on OLS estimation using the forecasts data from the last 8 quarters. The brown line is a local mean smoother which uses an Epanechnikov kernel.

Figure 4: Disagreement about Persistence

### Panel A: Disagreement about Persistence
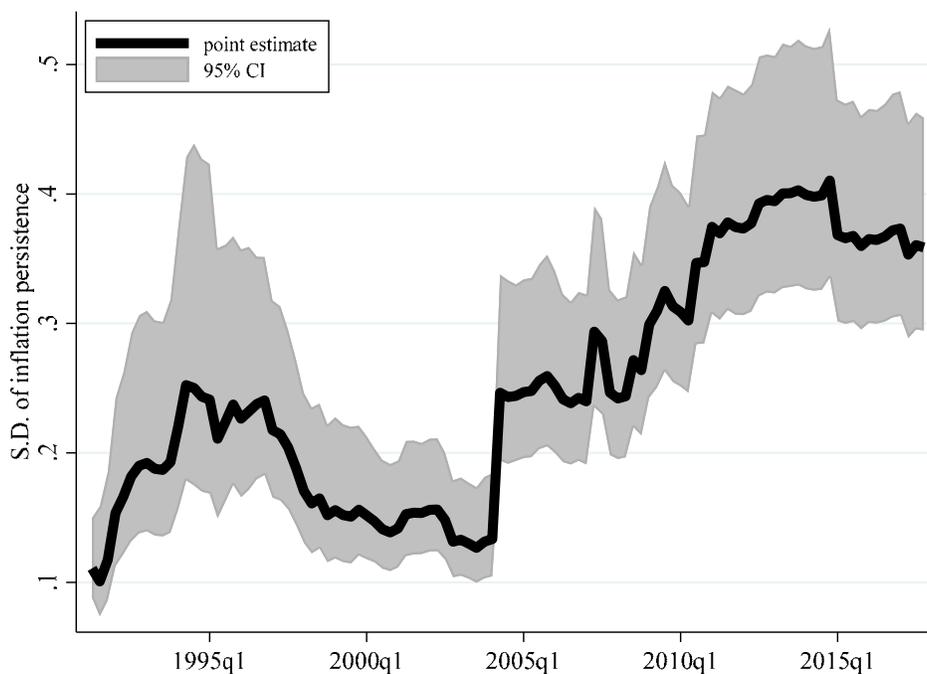
Panel B: Persistence Estimates
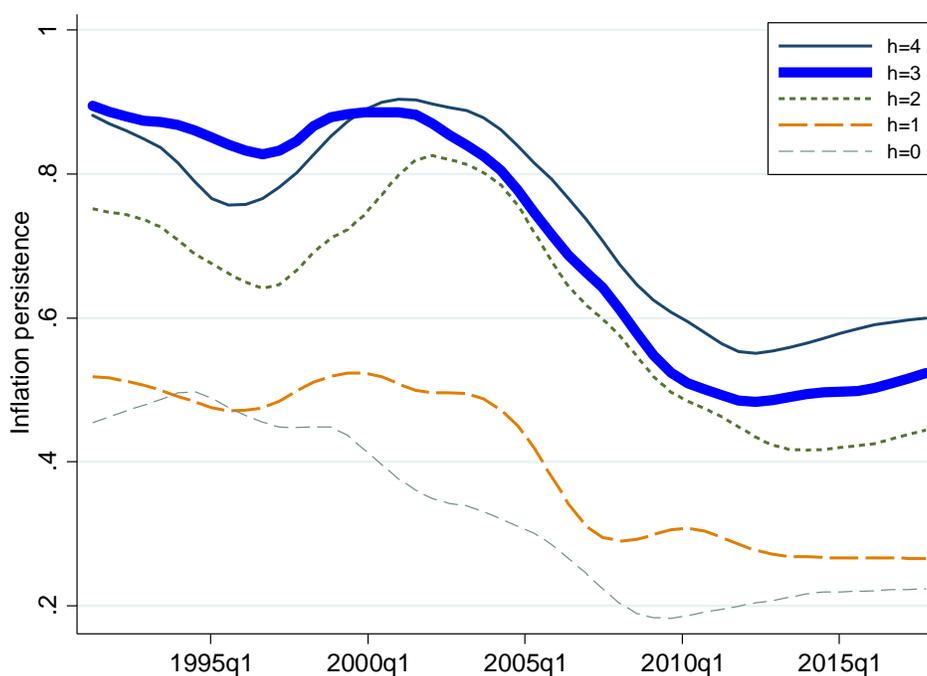
Notes: The figure plots smoothed estimates of persistence $\hat{\rho}$ (Panel B) and cross-sectional standard deviations of persistence (Panel A) based on estimating specification (6) for individuals in the SPF survey. The specification is estimated for each forecaster who deliver at least 20 observations in a rolling window of 40 quarters. For each time window, the mean across forecasters of persistence estimates for a certain horizon is displayed in Panel B. The smoother is a local mean which uses an Epanechnikov kernel. The standard deviation of persistence estimates (at $h = 4$) across forecasters is displayed in Panel A.

Figure 5: Simulation Results

Panel A: Persistence Estimates
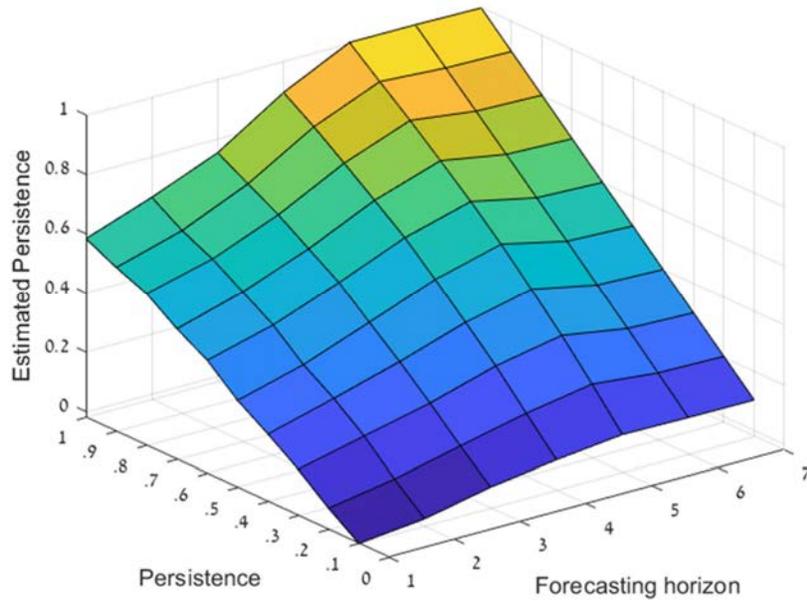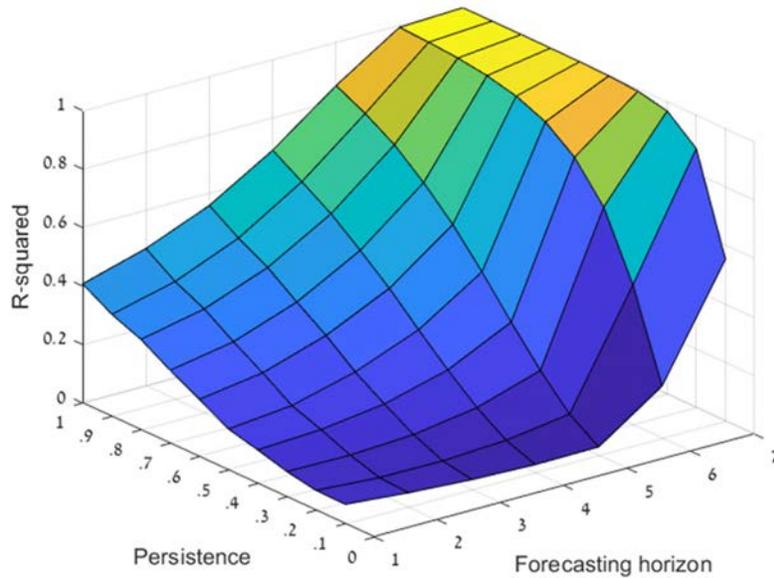
Panel B: R-Squared of Persistence Regressions

Notes: The two panels in the figure show estimation results of specification (4) applied to a simulated data of forecasts, according to the model presented in section 3. Estimated persistence (panel A) and R-squared (panel B) are averaged across 1000 draws of the simulation. Each simulation applies a different value of persistence in the fundamental. The variance of the shock in the state process is standardized to one. Regressions were estimated for seven forecast horizons out of eight horizons for which noisy signals are available (two consecutive horizons in each regression). The noise-to-signal ratio at horizons $h = 0, \dots, 7$ is set to 0.2, 0.5, 1 ,2 ,3 ,4, 100, 10000, respectively.

Figure 6: Inflation Forward Information

Panel A: h=0

Panel B: h=1

Panel C: h=2

Panel D: h=3

Panel E: h=4

Forward information

Inflation, right axis
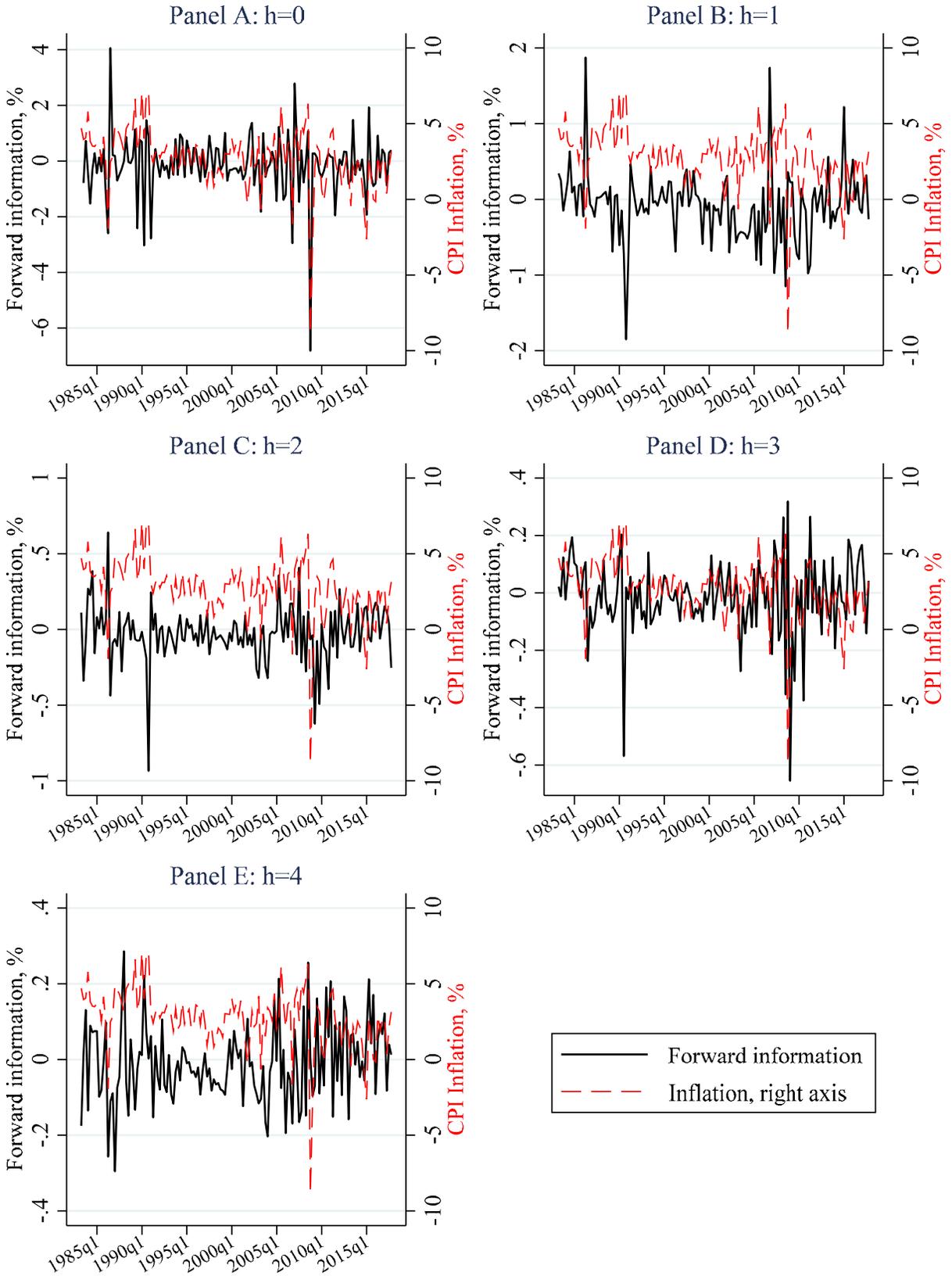
Notes: The figure plots time series of forward information for inflation. Forward information is computed according to equations (32) and (33).

TABLE 1
Summary of regression properties in the example of section 5.2

	Forecast Horizon			Persistence		
	$h = 1$	$h = 2$	$h = 3$	$\rho = 0$	$0 < \rho < 1$	$\rho = 1$
Estimated persistence	0	Positive but Biased. Bias decreases in σ_2^2 if $\sigma_2^2 > \sigma_1^2 > \sigma_\omega^2$.	ρ	0	Increases in ρ if $\sigma_2^2 > \sigma_1^2 > \sigma_\omega^2$.	Positive but Biased
R-squared	0	Between 0 and 1. Increases in σ_2^2 if $\sigma_2^2 > \sigma_1^2 > \sigma_\omega^2$.	1	0	Increases in ρ	Between 0 and 1

Notes: The table summarizes the theoretical predictions by the model analyzed in section 3.3. Model predictions refer to the estimated persistence $\hat{\rho}$ and R-squared of a regression of the forecast $x_{t+h|t}^i$ on $x_{t+h-1|t}^i$.

TABLE 2
Regressions of the deviation from the mean backcast

Dependent variable: $x_{t\|t}^{i} - x_{t\|t}$ (backcasts)	Full Sample	1980s	1990s	2000s	2010s
$x_{t\|t-1}^{i} - x_{t\|t-1}$	0.018	0.118	0.017**	-0.000	0.002**
	(0.014)	(0.082)	(0.008)	(0.000)	(0.001)
$x_{t+1\|t-1}^{i} - x_{t+1\|t-1}$	0.009	0.014	0.032**	0.000	0.002
	(0.016)	(0.084)	(0.015)	(0.000)	(0.001)
$x_{t+2\|t-1}^{i} - x_{t+2\|t-1}$	-0.029	-0.202**	-0.034***	0.001	0.000
	(0.018)	(0.085)	(0.013)	(0.001)	(0.003)
$x_{t+3\|t-1}^{i} - x_{t+3\|t-1}$	0.009	0.078	0.013	-0.000	0.002
	(0.018)	(0.123)	(0.015)	(0.001)	(0.002)
$x_{t+4\|t-1}^{i} - x_{t+4\|t-1}$	0.001	-0.007	-0.009	-0.000	0.002**
	(0.014)	(0.067)	(0.009)	(0.001)	(0.001)
Constant	0.004	0.027	0.002	0.000	0.001
	(0.005)	(0.037)	(0.002)	(0.000)	(0.002)
Obs.	3,849	559	1,068	1,272	950
R^2	0.004	0.023	0.023	0.003	0.002

Notes: The table reports coefficient estimates from regressions of the individual deviation from the mean backcast, based on specification (26) with $h = 0$. Each column reports results for a specified sample period. Each panel refers to a different horizon. Driscoll-Kraay standard errors are in parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 levels.

TABLE 3
Regressions of the deviation from the mean forecast

Dependent variable:	$x^i_{t+1|t} - x_{t+1|t}$	$x^i_{t+2|t} - x_{t+2|t}$	$x^i_{t+3|t} - x_{t+4|t}$	$x^i_{t+4|t} - x_{t+4|t}$
$x^i_{t|t-1} - x_{t|t-1}$	-0.013	-0.013	-0.063***	-0.062***
	(0.025)	(0.021)	(0.012)	(0.020)
$x^i_{t+1|t-1} - x_{t+1|t-1}$	0.220***	0.003	0.021	0.032
	(0.052)	(0.040)	(0.044)	(0.034)
$x^i_{t+2|t-1} - x_{t+2|t-1}$	0.130***	0.458***	-0.095**	-0.056*
	(0.050)	(0.057)	(0.046)	(0.032)
$x^i_{t+3|t-1} - x_{t+3|t-1}$	-0.126**	-0.120**	0.486***	0.103*
	(0.061)	(0.057)	(0.069)	(0.059)
$x^i_{t+4|t-1} - x_{t+4|t-1}$	0.071	0.056	0.037	0.362***
	(0.066)	(0.043)	(0.038)	(0.044)
Constant	-0.008	-0.001	0.005	0.007
	(0.009)	(0.006)	(0.008)	(0.008)
Obs.	3,854	3,856	3,855	3,853
R^2	0.053	0.146	0.213	0.178
BIC	10,515	8,565	7,434	7,323
BIC for specification (27)	10,822	8,763	7,635	7,484

Notes: The table reports coefficient estimates from regressions of the individual deviation from the mean forecast, based on specification (26). Each column reports results for a specified dependent variable, using the whole sample period. BIC reports the value of Bayes Information criterion for each specification as well as the BIC corresponding to specification (27). Driscoll-Kraay standard errors are in parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 levels.

TABLE 4

Summary statistics of series for actual inflation and forward information about inflation

Variable:	Mean	Standard deviation	Serial correlation	Correlation between horizons
$FI_{t|t}$	-0.150	1.122	-0.275	
$FI_{t+1|t}$	-0.078	0.454	0.204	-0.258
$FI_{t+2|t}$	-0.029	0.182	-0.073	0.324
$FI_{t+3|t}$	-0.017	0.134	-0.209	0.069
$FI_{t+4|t}$	-0.010	0.104	0.011	-0.094
Actual inflation	2.704	1.992	0.350	

Notes: The table reports summary statistics for the time series of (marginal) forward information computed according to equations (32) and (33). The correlation between horizons in the right column shows the correlation between $FI_{t+h|t}$ and $FI_{t+h-1|t}$. The sample period is 1983Q3-2007Q4. ***, **, * denote significance at 0.01, 0.05, and 0.10 levels.

TABLE 5

Inflation predictability and forward information 1983-2017

Dependent variable:	π_t	π_t	π_{t+1}	π_{t+2}	π_{t+3}	π_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
$FI_{t|t}$		1.554***	0.670***	-0.013	0.411	0.405**
		(0.146)	(0.140)	(0.191)	(0.257)	(0.197)
$FI_{t+1|t}$			1.275***	0.405	0.415	0.853
			(0.397)	(0.437)	(0.380)	(0.605)
$FI_{t+2|t}$				0.097	0.373	0.512
				(0.936)	(0.584)	(1.112)
$FI_{t+3|t}$					-0.134	0.191
					(1.035)	(1.814)
$FI_{t+4|t}$						-0.782
						(1.033)
π_{t-1}	0.349***	0.758***	0.368**	0.186	0.211	0.345**
	(0.083)	(0.095)	(0.164)	(0.193)	(0.175)	(0.165)
Constant	1.757***	0.877***	1.898***	2.210***	2.192***	1.856***
	(0.221)	(0.295)	(0.417)	(0.432)	(0.405)	(0.405)
R^2	0.122	0.721	0.109	0.035	0.038	0.082

Notes: The table reports coefficient estimates for regressions of current and future inflation on realized inflation and forward information. $x_t \equiv \pi_t$ is CPI inflation. The forward information variables are computed according to equations (32) and (33). Newey-West standard errors are in parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 levels.

TABLE 6

Cross-sectional variation in forecasts and forward information

	$var(x_{t+h|t}^i)$	$var(\hat{c}_t^i + \hat{\rho}_t^i x_{t+h-1|t}^i)$	$var(FI_{t+h|t}^i)$	$\sqrt{\dfrac{var(FI_{t+h|t}^i)}{var(x_{t+h|t}^i)}}$
$h = 0$	0.889	0.398	1.143	1.134
$h = 1$	0.610	0.533	0.700	1.071
$h = 2$	0.498	0.336	0.324	0.807
$h = 3$	0.481	0.308	0.279	0.762
$h = 4$	0.459	0.304	0.178	0.623

Notes: The table reports cross-sectional variation in inflation forecasts and their two components: model $\left(\hat{c}_t^i + \hat{\rho}_t^i x_{t+h-1|t}^i\right)$ and forward information $(FI_{t+h|t}^i)$. The components are computed according to equations (32) and (33) at the individual level in each quarter. Individual-level perceived parameters of inflation are estimated as in Figure 4 for each quarter, using rolling windows of 40 quarters. Within a window, the cross-section includes forecasters with at least 15 forecasts. The entries in the table report the mean of cross-sectional variation over the sample period.

TABLE 7
Taylor Rule and Forward Information 1983-2015

Dependent variable: r_t	Full sample			1983-1999	2000-2015
	(1)	(2)	(3)	(4)	(5)
$\pi_{t\|t}^{GB}$	0.051**				
	(0.021)				
$(\hat{c}_{t-1} + \hat{\rho}_{t-1}\pi_{t-1})$		0.031	0.124	0.310***	-0.021
		(0.053)	(0.083)	(0.085)	(0.078)
$FI_{t\|t}^{GB}$		0.058***	0.151***	0.286***	0.088***
		(0.018)	(0.039)	(0.079)	(0.028)
$FI_{t+1\|t}^{GB}$			0.176***	0.266**	0.096**
			(0.059)	(0.107)	(0.044)
$FI_{t+2\|t}^{GB}$			0.069	0.112	0.208**
			(0.087)	(0.136)	(0.089)
$FI_{t+3\|t}^{GB}$			0.134	0.165	0.276
			(0.191)	(0.196)	(0.440)
$FI_{t+4\|t}^{GB}$			0.205	0.223	-0.515*
			(0.276)	(0.425)	(0.302)
$gap_{t\|t}^{GB}$	0.025*	0.024	0.027	0.070***	0.026
	(0.014)	(0.015)	(0.017)	(0.021)	(0.023)
$gr_{t\|t}^{GB}$	0.149***	0.149***	0.143***	0.216***	0.135***
	(0.039)	(0.039)	(0.038)	(0.026)	(0.046)
r_{t-1}	1.134***	1.151***	1.089***	0.889***	1.268***
	(0.099)	(0.111)	(0.128)	(0.155)	(0.093)
r_{t-2}	-0.184**	-0.196**	-0.185*	-0.028	-0.323***
	(0.089)	(0.097)	(0.103)	(0.116)	(0.092)
R^2	0.982	0.982	0.984	0.962	0.980

Notes: The table reports coefficient estimates for regressions based on specification (34). The quarterly forecasts are taken from the Fed Greenbooks. The "model" component of inflation forecasts, $(\hat{c}_{t-1} + \hat{\rho}_{t-1}\pi_{t-1})$, uses time-varying persistence estimates from cross-sections of SPF forecasts, as explained in the text. The forward information variables are computed according to equations (32) and (33). Newey-West standard errors are in parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 levels.

Online Appendix

Appendix A: Forward information with a common noise

This appendix introduces a common signal to the forward information framework, presented in section 3, and shows how it affects the main results derived in sections 3 and 4.

Consider the following state-space representation:

State:

$$\boldsymbol{x}_t \equiv \begin{bmatrix} x_{t+H} \\ x_{t+H-1} \\ \vdots \\ x_t \end{bmatrix} = \begin{bmatrix} \rho & 0 & \cdots & 0 \\ 1 & 0 & 0 & 0 \\ \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & 1 & 0 \end{bmatrix} \boldsymbol{x}_{t-1} + S'\omega_{t+T} = \mathrm{P}\boldsymbol{x}_{t-1} + S'\omega_{t+H} \tag{A.1}$$

where $\omega_t \sim iid\ N(0,\sigma_\omega^2)$ and $S = [1 \quad 0 \quad \cdots \quad 0]$, so that the variance-covariance matrix of $S'\omega_{t+H}$ would be $\Sigma_\omega = S'S\sigma_\omega^2$.

Measurement:

$$\boldsymbol{y}_t^i \equiv \begin{bmatrix} y_{t,t+H}^i \\ y_{t,t+H-1}^i \\ \vdots \\ y_{t,t}^i \end{bmatrix} = \begin{bmatrix} x_{t+H} \\ x_{t+H-1} \\ \vdots \\ x_t \end{bmatrix} + \begin{bmatrix} v_{t,t+H}^i \\ v_{t,t+H-1}^i \\ \vdots \\ v_{t,t}^i \end{bmatrix} + \begin{bmatrix} e_{t+H} \\ e_{t+H-1} \\ \vdots \\ e_t \end{bmatrix} = \boldsymbol{x}_t + \boldsymbol{v}_t^i + \boldsymbol{e}_t \tag{A.2}$$

where $\boldsymbol{v}_t^i$ is an iid normally distributed idiosyncratic noise and $\Sigma_v = \boldsymbol{I}_H\boldsymbol{\sigma}_v^2$ with $\boldsymbol{\sigma}_v^{2'} = [\sigma_H^2 \quad \sigma_{H-1}^2 \quad \cdots \quad \sigma_0^2]$. We also introduce the common iid noise $\boldsymbol{e}_t$ with a variance-covariance matrix $\Sigma_e = \boldsymbol{I}_H\boldsymbol{\sigma}_e^2$ where the variance may also vary across horizons. All types of shocks are uncorrelated contemporaneously and at all leads and lags.

To solve for the optimal forecast by the Kalman filter, we use the following Riccati equation:

$$\Psi = \mathrm{P}\{\Psi - \Psi(\Psi + \Sigma_v + \Sigma_e)^{-1}\Psi\}\mathrm{P}' + \Sigma_\omega \tag{A.3}$$

where Ψ is the variance of the one-step ahead forecast error. The Kalman gain matrix, denoted by G (with dimension $H + 1$), would then be obtained by:

$$\mathrm{G} = \Psi(\Psi + \Sigma_v + \Sigma_e)^{-1} \tag{A.4}$$

Thus, the optimal forecast of an agent i would be:

$$\boldsymbol{x}_{t|t}^i = \boldsymbol{x}_{t|t-1}^i + \mathrm{G}(\boldsymbol{y}_t^i - \boldsymbol{x}_{t|t-1}^i) \tag{A.5}$$

where $\boldsymbol{x}_{t|t}^i$ is a vector of forecasts made at time t, with horizons running from 0 to H steps ahead. The forecast takes the same form as the forecast in equation (20) in the main text, except that the gain matrix would be different and signals in $\boldsymbol{y}_t^i$ also contain common noises.

From (A.5), the h step-ahead forecast $x_{t+h|t}^i$, which is an element in $\boldsymbol{x}_{t|t}^i$, can be written as:

$$\begin{aligned} x_{t+h|t}^i = x_{t+h|t-1}^i &+ \mathrm{G}_{j,1}(y_{t,t+H}^i - x_{t+H|t-1}^i) + \mathrm{G}_{j,2}(y_{t,t+H-1}^i - x_{t+H-1|t-1}^i) + \cdots \\ &+ \mathrm{G}_{j,H+1}(y_{t,t}^i - x_{t|t-1}^i) \end{aligned} \tag{A.6}$$

where the coefficients $\mathrm{G}_{j,1}$, $\mathrm{G}_{j,2}$, …, $\mathrm{G}_{j,H+1}$ are elements of row $j = H - h + 1$ in the gain matrix G.

As in the case without a common noise, when forecasting more than H steps ahead ($h > H$), the forecast would simply be:

$$x_{t+h|t}^i = \rho x_{t+h-1|t}^i \tag{A.7}$$

which resembles equation (4) in the standard noisy information framework. We now examine the relationship between $x_{t+h|t}^i$ and $x_{t+h-1|t}^i$ for shorter horizons ($0 < h \leq H$). From (A.5) the forecast $x_{t+h-1|t}$ would be

$$
\begin{aligned}
x_{t+h-1|t}^i = x_{t+h-1|t-1}^i &+ G_{j+1,1}\left(y_{t,t+H}^i - x_{t+H|t-1}^i\right) \\
&+ G_{j+1,2}\left(y_{t,t+H-1}^i - x_{t+H-1|t-1}^i\right) + \cdots \\
&+ G_{J+1,H+1}\left(y_{t,t}^i - x_{t|t-1}^i\right)
\end{aligned}
\tag{A.8}
$$

where the coefficients $G_{j+1,1}$, $G_{j+1,2}$, ..., $G_{j+1,H+1}$ are the elements of row $j+1$ in the gain matrix G. Multiplying (A.8) by ρ and subtracting it from (A.6), we obtain:

$$
\begin{aligned}
x_{t+h|t}^i &- \rho x_{t+h-1|t}^i \\
&= x_{t+h|t-1}^i - \rho x_{t+h-1|t-1}^i + \left(G_{j,1} - \rho G_{j+1,1}\right)\left(y_{t,t+H}^i - x_{t+H|t-1}^i\right) \\
&+ \left(G_{J,2} - \rho G_{j+1,2}\right)\left(y_{t,t+H-1}^i - x_{t+H-1|t-1}^i\right) + \cdots \\
&+ \left(G_{j,H+1} - \rho G_{j+1,H+1}\right)\left(y_{t,t}^i - x_{t|t-1}^i\right)
\end{aligned}
\tag{A.9}
$$

Comparing to (A.7), the term $x_{t+h|t}^i - \rho x_{t+h-1|t}^i$ does not equal to zero. The RHS of (A.9) expresses the deviation from the relation $x_{t+h|t} = \rho x_{t+h-1|t}$, for each horizon in $0 < h \leq H$.

To get a compact version of (A.9) that corresponds to the key result in equation (22) in the main text, denote the vector that is equal to row j in matrix G by $\mathbf{G}_j$. Moving $\rho x_{t+h-1|t}$ to the RHS and using substitutes from (A.2.) we obtain:

$$
\begin{aligned}
x_{t+h|t}^i = \rho x_{t+h-1|t}^i &+ \left(x_{t+h|t-1}^i - \rho x_{t+h-1|t-1}^i\right) + \left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\left(x_t - x_{t|t-1}^i\right) \\
&+ \left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)v_t^i + \left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)e_t
\end{aligned}
\tag{A.10}
$$

Thus, the forecast $x_{t+h|t}^i$ is a composition of a prediction based on the underlying process ($\rho x_{t+h-1|t}^i$), and "subjective adjustment" which account for forward information. In the standard noisy information framework, as in section 2, the second component would be absent, so that the optimal forecast would obey $x_{t+h|t}^i = \rho x_{t+h-1|t}^i$. The need for an adjustment component in the presence of forward signals is driven by the variation across horizons in the optimal weights placed on the signals. This variation is captured by the term $\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)$, where $\mathbf{G}_j$ and $\mathbf{G}_{j+1}$ are consecutive rows in the gain matrix.

Hence, equation (A.10) has the same interpretation as equation (22). The only difference is the presence of an additional term $\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)e_t$ in the adjustment component, which represents the effect of the common noise. The empirical implications for the estimation of the parameter of persistence by specification (4) would also be similar: the adjustment component introduces an error to the regression. The error would bias the coefficient estimate from ρ because it is correlated with the RHS forecast $x_{t+h-1|t}^i$. The common noise introduces another source of correlation, but only in the time dimension. In a cross-sectional estimation of specification (4), common noise can be captured by including a constant in the regression. In a panel estimation, the time correlation between e_t and $x_{t+h-1|t}^i$ can also be captured by time fixed-effects. Thus, the only remaining source of bias would be the idiosyncratic noise v_t^i, leading to the same empirical pattern that is addressed in the main text. Specifically, when estimating persistence by specification (4) for different horizons, the coefficient estimate would converge to ρ and the fit of the regression will get close to perfect, for longer horizons, due to the decay in forward information at farther horizons.

Next, we turn to specification (26) from Section 4 which is used to directly estimate the gain matrix in the forward information framework and to test this framework against the standard setup of noisy information. Specification (26) is derived by subtracting the mean forecast from the individual forecast. With common noise, the individual forecast from (A.5) could be written as (using a substitute from (A.2)):

$$
x_{t|t}^i = (I - G)x_{t|t-1}^i + G\left(x_t + v_t^i + e_t\right)
\tag{A.11}
$$

Taking the average across individuals (denoted by dropping superscript i) we obtain:

$$x_{t|t} = (I - G)x_{t|t-1} + G(x_t + e_t) \qquad (A.12)$$

Notice that only v_t^i is vanished in the mean forecast while the common noise e_t remains. Thus, when subtracting the mean from the individual forecast, the common noise disappears, obtaining:

$$x_{t|t}^i - x_{t|t} = (I - G)(x_{t|t-1}^i - x_{t|t-1}) + Gv_t^i \qquad (A.13)$$

which is similar to (25) in the main text. It follows that even in the presence of a common noise, specification (26) could be applied in the same way to estimate $(I - G)$ and to preform tests. The error component in the regression is only due to the individual-specific noise.

In a mean-level estimation such as a regression of average forecast errors on average forecast revisions, which is discussed in section 3.1, the results would be affected by the presence of a common noise. Coibion and Gorodnichenko (2015) who proposed this mean level specification to estimate the Kalman gain in the standard noisy information setup, show that common noise introduces estimation bias. Similarly, common noise will affect the estimation properties under forward information. This highlights an additional advantage in using specification (26): since we take the deviation of individual forecasts from the mean, the specification is not sensitive to the presence of common noise. This advantage is also highlighted by Goldstein (2021) in the context of quantifying information frictions under the standard noisy framework, as in equation (27).

Appendix B: Forward information model: Derivations

B.1. Deriving the optimal forecasts

In section 3.2., we consider the optimal forecasting for following state-space representation:

State:

$$\boldsymbol{x}_t \equiv \begin{bmatrix} x_{t+2} \\ x_{t+1} \\ x_t \end{bmatrix} = \begin{bmatrix} \rho & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \boldsymbol{x}_{t-1} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \omega_{t+2} = \mathrm{P}\boldsymbol{x}_{t-1} + S'\omega_{t+2} \qquad \text{(B.1)}$$

where $\omega_t \sim iid\ N(0, \sigma_\omega^2)$.

Measurement:

$$\boldsymbol{y}_t^i \equiv \begin{bmatrix} y_{t,t+2}^i \\ y_{t,t+1}^i \\ y_{t,t}^i \end{bmatrix} = \begin{bmatrix} x_{t+2} \\ x_{t+1} \\ x_t \end{bmatrix} + \begin{bmatrix} v_{t,t+2}^i \\ v_{t,t+1}^i \\ 0 \end{bmatrix} = \boldsymbol{x}_t + \boldsymbol{v}_t^i \qquad \text{(B.2)}$$

where $v_{t,t+1}^i \sim iid\ N(0, \sigma_1^2)$ and $v_{t,t+2}^i \sim iid\ N(0, \sigma_2^2)$.

Because x_t is perfectly observed, the forecast for time t would simply be $x_{t|t} = x_t$. The one step-ahead forecast $x_{t+1|t}$ should apply four useful signals:

1. ρx_t, with ex-post forecast error equal to ω_{t+1}.
2. $y_{t-1,t+1}^i$ (forward signal from the previous period), with ex-post forecast error equal to $-v_{t-1,t+1}^i$.
3. $y_{t,t+1}^i$, with ex-post forecast error equal to $-v_{t,t+1}^i$.
4. $\rho^{-1} y_{t,t+2}^i$, with ex-post forecast error equal to $-\rho^{-1}(\omega_{t+2} + v_{t,t+2}^i)$.

All other noisy signals from previous periods are no longer useful after x_t is perfectly observed.

Accordingly, the one step-ahead forecast could be represented as a weighted sum of the four signals:

$$x_{t+1|t}^i = W_1 \rho x_t + W_2 y_{t-1,t+1}^i + W_3 y_{t,t+1}^i + W_4 \rho^{-1} y_{t,t+2}^i \qquad \text{(B.3)}$$

where $\sum_{k=1}^4 W_k = 1$. The forecaster should choose the weights which minimize the following expected squared error:

$E_t\{x_{t+1} - x_{t+1|t}^i\}^2 = E_t\left\{W_1\omega_{t+1} + W_2(-v_{t-1,t+1}^i) + W_3(-v_{t,t+1}^i) + W_4\left(-\rho^{-1}(\omega_{t+2} + v_{t,t+2}^i)\right)\right\}^2 = (W_1)^2\sigma_\omega^2 + (W_2)^2\sigma_2^2 + (W_3)^2\sigma_1^2 + (W_4)^2\rho^{-2}(\sigma_\omega^2 + \sigma_2^2)$

Thus, the optimization problem can be written as:

$$min_{W_k}\{(W_1)^2\sigma_\omega^2 + (W_2)^2\sigma_2^2 + (W_3)^2\sigma_1^2 + (1 - W_1 - W_2 - W_3)^2\rho^{-2}(\sigma_\omega^2 + \sigma_2^2)\}$$

After setting $\sigma_\omega^2 = 1$ for normalization, the FOCs are:

$2W_1 = 2(1 - W_1 - W_2 - W_3)\rho^{-2}(1 + \sigma_2^2)$
$2W_2\sigma_2^2 = 2(1 - W_1 - W_2 - W_3)\rho^{-2}(1 + \sigma_2^2)$
$2W_3\sigma_1^2 = 2(1 - W_1 - W_2 - W_3)\rho^{-2}(1 + \sigma_2^2)$

The solution to this system would obtain the following optimal weights:

$$W_1 = \frac{\sigma_2^2\sigma_1^2(1 + \sigma_2^2)}{m}$$

$$W_2 = \frac{\sigma_1^2(1 + \sigma_2^2)}{m}$$

$$W_3 = \frac{\sigma_2^2(1 + \sigma_2^2)}{m} \qquad \text{(B.4)}$$

$$W_4 = \frac{\rho^2\sigma_2^2\sigma_1^2}{m},$$

where $m = \sigma_2^2\sigma_1^2(1 + \sigma_2^2) + \sigma_1^2(1 + \sigma_2^2) + \sigma_2^2(1 + \sigma_2^2) + \rho^2\sigma_2^2\sigma_1^2$.

Next, we follow the same steps to derive the two steps-ahead optimal forecasts $x_{t+2|t}^i$. The same four signals used in the one-step-ahead forecast apply for $x_{t+2|t}^i$, after multiplying by ρ to adjust to the new horizon. Specifically, the four available predictions for two steps ahead are:

1. $\rho^2 x_t$, with ex-post forecast error equal to $(\rho\omega_{t+1} + \omega_{t+2})$.

2. $\rho y_{t-1,t+1}^i$, with ex-post forecast error equal to $\left(-\rho v_{t-1,t+1}^i + \omega_{t+2}\right)$.
3. $\rho y_{t,t+1}^i$, with ex-post forecast error equal to $\left(-\rho v_{t,t+1}^i + \omega_{t+2}\right)$.
4. $y_{t,t+2}^i$, with ex-post forecast error equal to $-v_{t,t+2}^i$.

Hence, we write the optimal forecast as

$$x_{t+2|t}^i = w_1 \rho^2 x_t + w_2 \rho y_{t-1,t+1}^i + w_3 \rho y_{t,t+1}^i + w_4 y_{t,t+2}^i \qquad \text{(B.5)}$$

The expected squared error could therefore be expressed as follows:

$$E_t\{x_{t+2} - x_{t+2|t}^i\}^2 = E_t\{w_1(\rho\omega_{t+1} + \omega_{t+2}) + w_2(-\rho v_{t-1,t+1}^i + \omega_{t+2}) + w_3(-\rho v_{t,t+1}^i + \omega_{t+2}) +$$

$$w_4(-v_{t,t+2}^i)\}^2 = (w_1)^2\rho^2 + (w_1 + w_2 + w_3)^2 + (w_2)^2\rho^2\sigma_2^2 + (w_3)^2\rho^2\sigma_1^2 + (w_4)^2\sigma_2^2$$

and the optimization problem is consequently:

$$min_{w_k}\{(w_1)^2\rho^2 + (w_1 + w_2 + w_3)^2 + (w_2)^2\rho^2\sigma_2^2 + (w_3)^2\rho^2\sigma_1^2 + (w_4)^2\sigma_2^2\}$$

The FOCs are:

$2w_1\rho^2 = -2(w_1 + w_2 + w_3) + 2(1 - w_1 - w_2 - w_3)\sigma_2^2$
$2w_2\rho^2\sigma_2^2 = -2(w_1 + w_2 + w_3) + 2(1 - w_1 - w_2 - w_3)\sigma_2^2$
$2w_3\rho^2\sigma_1^2 = -2(w_1 + w_2 + w_3) + 2(1 - w_1 - w_2 - w_3)\sigma_2^2$

and the solution for the optimal weights is:

$$w_1 = \frac{\sigma_2^2\sigma_2^2\sigma_1^2}{m}$$

$$w_2 = \frac{\sigma_2^2\sigma_1^2}{m}$$

$$\qquad \text{(B.6)}$$

$$w_3 = \frac{\sigma_2^2\sigma_2^2}{m}$$

$$w_4 = \frac{(1 + \rho^2)\sigma_2^2\sigma_1^2 + \sigma_1^2 + \sigma_2^2}{m}$$

When moving further to forecast three steps ahead, there is no further change in the optimal weights, and thus the optimal forecast obeys $x_{t+3|t}^i = \rho x_{t+2|t}^i$. The optimal weights do not change for $h \geq 3$, since there are no forward signals referring to these horizons. To see that, we repeat the same steps as above to derive $x_{t+3|t}$. Based on the same four signals, the forecast $x_{t+3|t}$ uses the following predictions:

1. $\rho^3 x_t$, with ex-post forecast error equal to $(\rho^2\omega_{t+1} + \rho\omega_{t+2} + \omega_{t+3})$.
2. $\rho^2 y_{t-1,t+1}^i$, with ex-post forecast error equal to $\left(-\rho^2 v_{t-1,t+1}^i + \rho\omega_{t+2} + \omega_{t+3}\right)$.
3. $\rho^2 y_{t,t+1}^i$, with ex-post forecast error equal to $\left(-\rho^2 v_{t,t+1}^i + \rho\omega_{t+2} + \omega_{t+3}\right)$.
4. $\rho y_{t,t+2}^i$, with ex-post forecast error equal to $\left(-\rho v_{t,t+2}^i + \omega_{t+3}\right)$.

The optimal forecast follows:

$$x_{t+3|t}^i = \widetilde{w}_1\rho^3 x_t + \widetilde{w}_2\rho^2 y_{t-1,t+1}^i + \widetilde{w}_3\rho^2 y_{t,t+1}^i + \widetilde{w}_4\rho y_{t,t+2}^i \qquad \text{(B.7)}$$

Note that in this case the expected squared error is expressed as:

$$E_t\{x_{t+3} - x_{t+3|t}^i\}^2 = E_t\{\rho(x_{t+2} - x_{t+2|t}) + \omega_{t+3}\}^2 = \rho^2 E_t\{x_{t+2} - x_{t+2|t}\}^2 + 1$$

So, minimizing $E_t\{x_{t+3} - x_{t+3|t}^i\}^2$ is equivalent to the minimizing $E_t\{x_{t+2} - x_{t+2|t}^i\}^2$ and produces the same optimal weights, w_k, derived above. The same reasoning holds for any horizon beyond the range of forward signals. Thus, the relationship $x_{t+h|t}^i = \rho x_{t+h-1|t}^i$ applies for any $h \geq 3$.

B.2. Kalman filter representation

In this section, we represent the optimal forecast in the Kalman filter framework. The filter is given by:

$$x_{t|t}^i = x_{t|t-1}^i + G\left(y_t^i - x_{t|t-1}^i\right) \qquad \text{(B.8)}$$

where the gain matrix G must be specified. Expanding the matrix notation, we have:

$$\begin{bmatrix} x^i_{t+2|t} \\ x^i_{t+1|t} \\ x^i_{t|t} \end{bmatrix} = \begin{bmatrix} x^i_{t+2|t-1} \\ x^i_{t+1|t-1} \\ x^i_{t|t-1} \end{bmatrix} + \begin{bmatrix} G_{1,1} & G_{1,2} & G_{1,3} \\ G_{2,1} & G_{2,2} & G_{2,3} \\ G_{3,1} & G_{3,2} & G_{3,3} \end{bmatrix} \times \begin{bmatrix} y^i_{t,t+2} - x^i_{t+2|t-1} \\ y^i_{t,t+1} - x^i_{t+1|t-1} \\ y^i_{t,t} - x^i_{t|t-1} \end{bmatrix} \qquad \text{(B.9)}$$

Each element in G corresponds to a weight placed on one of the current signals, in each of the three forecasts. Hence, we could use the optimal weights derived above to guess the elements in the gain matrix, using the fact that that the filter algorithm is also based on the minimization of the squared error.

Specifically, the last row in the matrix should include 0, 0 and 1 in order to obtain $x^i_{t|t} = y^i_{t,t} = x_t$, which is due to the perfect signal of the realized fundamental at period t. The second row corresponds to the weights in the forecast $x_{t+1|t}$. According to (B.3), the elements of this row should be $G_{2,1} = W_4\rho^{-1}$, $G_{2,2} = W_3$ and $G_{2,3} = W_1\rho$. Similarly, the elements in the first row of the gain matrix should correspond to the optimal weights in (B.5). Thus, we get that $G_{1,1} = w_4$, $G_{1,2} = w_3\rho$ and $G_{1,3} = w_1\rho^2$ and the gain matrix is therefore:

$$G = \begin{bmatrix} w_4 & w_3\rho & w_1\rho^2 \\ W_4\rho^{-1} & W_3 & W_1\rho \\ 0 & 0 & 1 \end{bmatrix}$$

To validate this result, we derive the variance-covariance matrix Ψ of the one-step-ahead forecast error, by using $G = \Psi(\Psi + \Sigma_v)^{-1}$, and then verify that Ψ solves the Riccati equation.

It should be noted that in the Kalman filter representation of the optimal forecast there is no explicit reference for the signal $y^i_{t-1,t+1}$, as in (B.3) and (B.5). However, this signal is implicit in the lagged forecasts in (B.9). For instance, the one-step-ahead forecast in the Kalman filter representation of (B.9) follows:

$$x^i_{t+1|t} = x^i_{t+1|t-1} + W_1\rho\left(y^i_{t,t} - x^i_{t|t-1}\right) + W_3\left(y^i_{t,t+1} - x^i_{t+1|t-1}\right) + W_4\rho^{-1}\left(y^i_{t,t+2} - x^i_{t+2|t-1}\right) \qquad \text{(B.10)}$$

The lagged forecasts on the right-hand-side follow

$$x^i_{t|t-1} = x^i_{t|t-2} + W_1\rho\left(y^i_{t-1,t-1} - x^i_{t-1|t-2}\right) + W_3\left(y^i_{t-1,t} - x^i_{t|t-2}\right) + W_4\rho^{-1}\left(y^i_{t-1,t+1} - x^i_{t+1|t-2}\right)$$

$$x^i_{t+1|t-1} = x^i_{t+1|t-2} + w_1\rho^2\left(y^i_{t-1,t-1} - x^i_{t-1|t-2}\right) + w_3\rho\left(y^i_{t-1,t} - x^i_{t|t-2}\right) + w_4\left(y^i_{t-1,t+1} - x^i_{t+1|t-2}\right)$$

$$x^i_{t+2|t-1} = \rho x^i_{t+1|t-1}$$

Plugging into (B.10) and rearranging terms, we obtain:

$$x^i_{t+1|t} = W_1\rho y^i_{t,t} + [(1 - W_3 - W_4)w_4 - W_1W_4]y^i_{t-1,t+1} + W_3 y^i_{t,t+1} + W_4\rho^{-1}y^i_{t,t+2} +$$
$$other\ signals \qquad \text{(B.11)}$$

where *other signals* represent all the signals in the lagged forecasts, which are not the signal $y^i_{t-1,t+1}$. The signal $y^i_{t-1,t+1}$ is the only lagged signal which is still informative in period t. All other lagged signals, which refer to period t and before, are not informative after x_t is perfectly observed.

By using the expressions for the optimal weights from (B.4) and (B.6), it can be verified that the term *other signals* in (B.11) is equal to zero. Furthermore, the coefficient on the lagged signal $y^i_{t-1,t+1}$ in (B.11), which is given by $[(1 - W_3 - W_4)w_4 - W_1W_4]$, is just equal to W_2, which is the same weight placed on this signal in (B.3). Thus, the optimal forecast in (B.11), based on the Kalman filter, is the same optimal forecast derived in (B.3) by directly optimizing the weights to minimize the squared forecast error. In a similar way, it can be shown that the optimal forecast $x^i_{t+2|t}$ according to the Kalman filter is the same forecast derived in (B.5). Thus, our guessed solution for the gain matrix is verified again.

B.3. Patterns of regression properties across forecast horizons

The simulation results, as presented in Figure 5, show that in a cross-sectional regression of $x^i_{t+h|t}$ on $x^i_{t+h-1|t}$ with low h (short horizon), the coefficient estimate and R-squared would be low, and as h increases they would converge to the values of ρ and 1, respectively. In our tractable example, because there are only three signals (one perfect signal of realized fundamental and two forward signals), this pattern is demonstrated in a compact form by increasing the horizon from $h = 1$ to $h = 3$.

$h = 1$: Suppose that we run a cross-sectional regression of $x^i_{t+1|t}$ on $x^i_{t|t}$. Based on (13), the coefficient on $x^i_{t|t}$ is ρ. However, the OLS coefficient estimate would be zero. This is easily verified, when recalling that $x^i_{t|t} = x_t$, due to the

perfect signal $y_{t,t}^i$ in (B.2). so that there is no cross-sectional variation in $x_{t|t}^i$. It also follows that all the cross-sectional variation in $x_{t+1|t}^i$ is determined by the "error term" specified in (13), and the R^2 should therefore be 0.

In the simulation presented in Figure 5, this limiting case was avoided by introducing some low degree of noise in the signal $y_{t,t}^i$ ($\sigma_0^2 = 0.2$), thereby, allowing some variation in $x_{t|t}^i$. Consequently, the simulated coefficient estimates and R^2 are positive but still very low.

$\boldsymbol{h = 3}$: At the other extreme, if we regress $x_{t+3|t}^i$ on $x_{t+2|t}^i$, we should get a perfect fit where the OLS coefficient estimate is equal to ρ. As shown in Section B.1., the optimal weights in the two forecasts $x_{t+3|t}^i$ and $x_{t+2|t}^i$ are the same, so $x_{t+3|t}^i$ is exactly equal to $\rho x_{t+2|t}^i$. This applies to any $h \geq 3$. More generally, this second limiting case would apply when h is sufficiently high so that there is no informative signal referring to that horizon. Also, recall that according to our simulation results, even if the noise in the forward signal is very high (but not infinite), the coefficient estimate would be very close to ρ, while the fit of the regression could be considerably far from a perfect fit.

$\boldsymbol{h = 2}$: This is the intermediate case, which is described by equation (11) in the main text. The coefficient on $x_{t+1|t}^i$ should be ρ, but the OLS estimate would be biased since $x_{t+1|t}^i$ is correlated with $Signal_{k,t+2}^i$ in the error term. The R^2 would be between 0 and 1.

It is also interesting to examine how regression properties for $h = 2$ vary when increasing the noise of the two steps ahead signal, that is, when increasing σ_2^2. This is, in a sense, a way to imitate our above simulation in which the noise in forward signals increases when moving to longer horizons. As shown in section 3.1 in the main text, by combining (B.3) and (B.5), the relation between the optimal forecasts with consecutive horizons $x_{t+2|t}$ and $x_{t+1|t}^i$ can be expressed as:

$$x_{t+2|t}^i = \rho x_{t+1|t}^i + \sum_{k=1}^{4}(w_k - W_k)\,Signal_{k,t+2}^i \qquad \text{(B.12)}$$

Thus, a deviation from the simple state relation of $x_{t+2|t}^i = \rho x_{t+1|t}^i$ is due to changes in the optimal weights $(w_k - W_k)$ across forecast horizons.

It is useful to begin with the special case where σ_2^2 goes to infinity so that the signal $y_{t,t+2}^i$ become meaningless. As a consequence, forecasts at time t only apply two signals: the perfect signal about realized x_t and the forward signal $y_{t,t+1}^i$. The one-step-ahead forecast is then:

$$x_{t+1|t}^i = W_1^{lim}\rho x_t + W_3^{lim} y_{t,t+1}^i \qquad \text{(B.13)}$$

where $W_1^{lim} + W_3^{lim} = 1$. The expected squared forecast error is:

$$E_t\{x_{t+1} - x_{t+1|t}^i\}^2 = E_t\{W_1^{lim}\omega_{t+1} + W_3^{lim}(-v_{t,t+1}^i)\}^2 = (W_1^{lim})^2 + (W_3^{lim})^2\sigma_1^2$$

so that optimal weights which minimize the expected squared error are simply given by $W_1^{lim} = \sigma_1^2(1 + \sigma_1^2)^{-1}$ and $W_3^{lim} = (1 + \sigma_1^2)^{-1}$.

Similarly, the two-step-ahead forecast is also a weighted average of two signals:

$$x_{t+2|t}^i = w_1^{lim}\rho^2 x_t + w_3^{lim}\rho y_{t,t+1}^i, \qquad \text{(B.14)}$$

where $w_1^{lim} + w_3^{lim} = 1$. The expected squared forecast error is:

$$E_t\{x_{t+2} - x_{t+2|t}^i\}^2 = E_t\{w_1^{lim}(\rho\omega_{t+1} + \omega_{t+2}) + w_3^{lim}(-\rho v_{t,t+1}^i + \omega_{t+2})\}^2 = (w_1^{lim})^2\rho^2 + (w_3^{lim})^2\rho^2\sigma_1^2 +$$
$$(w_1^{lim} + w_3^{lim})^2 = (w_1^{lim})^2\rho^2 + (w_3^{lim})^2\rho^2\sigma_1^2 + 1$$

Hence, the optimal weights which minimize the squared error should again be $w_1^{lim} = \sigma_1^2(1 + \sigma_1^2)^{-1}$ and $w_3^{lim} = (1 + \sigma_1^2)^{-1}$. This corresponds to the previous result that going beyond informative horizons, there will be no variation in the optimal weights across consecutive horizons and we obtain the simple relation of $x_{t+2|t}^i = \rho x_{t+1|t}^i$.

More generally, we now show that, under the plausible assumption $\sigma_2^2 > \sigma_1^2 > \sigma_\omega^2=1$, the gap between corresponding optimal weights W_k and w_k is getting closer to zero when σ_2^2 increases, leading to the empirical patterns observed for the coefficient estimate and R-squared across horizons. From (B.4) and (B.6), it is easy to see that $w_k < W_k$ for $k = 1,2,3$, while $w_4 > W_4$. Thus, we need to show that $(w_k - W_k)$ increases in σ_2^2 for $k = 1,2,3$, while $(w_4 - W_4)$ decreases in k:

<u>$k = 1$</u>:

From (B.4) and (B.6) the difference between w_1 and W_1 is:

$$w_1 - W_1 = \frac{\sigma_2^2\sigma_2^2\sigma_1^2 - \sigma_2^2\sigma_1^2(1 + \sigma_2^2)}{m} = \frac{-\sigma_2^2\sigma_1^2}{m}$$

where the expression for m, as specified above, is:

$$m = \sigma_2^2\sigma_1^2(1 + \sigma_2^2) + \sigma_1^2(1 + \sigma_2^2) + \sigma_2^2(1 + \sigma_2^2) + \rho^2\sigma_2^2\sigma_1^2 = (2 + \rho^2)\sigma_1^2\sigma_2^2 + (\sigma_2^2)^2 + (\sigma_2^2)^2\sigma_1^2 + \sigma_2^2 + \sigma_1^2$$

Taking the derivative with respect to σ_2^2 we obtain

$$\frac{\partial(w_1 - W_1)}{\partial\sigma_2^2} = \frac{-\sigma_1^2 m + \sigma_2^2\sigma_1^2[(2 + \rho^2)\sigma_1^2 + 2\sigma_2^2 + 2\sigma_2^2\sigma_1^2 + 1]}{m^2}$$

Plugging m into the derivative and rearranging we get:

$$\frac{\partial(w_1 - W_1)}{\partial\sigma_2^2} = \frac{\sigma_1^2(\sigma_2^2)^2 + (\sigma_1^2)^2(\sigma_2^2)^2 - (\sigma_1^2)^2}{m^2} = \frac{\sigma_1^2[(\sigma_2^2)^2(1 + \sigma_1^2) - \sigma_1^2]}{m^2} > 0$$

under the assumption of $\sigma_2^2 > \sigma_1^2 > \sigma_\omega^2 = 1$.

<u>$k = 2$</u>:

From (B.4) and (B.6) the difference between w_2 and W_2 is:

$$w_2 - W_2 = \frac{\sigma_2^2\sigma_1^2 - \sigma_1^2(1 + \sigma_2^2)}{m} = \frac{-\sigma_1^2}{m}$$

Taking the derivative with respect to σ_2^2 we obtain:

$$\frac{\partial(w_2 - W_2)}{\partial\sigma_2^2} = \frac{\sigma_1^2[(2 + \rho^2)\sigma_1^2 + 2\sigma_2^2 + 2\sigma_2^2\sigma_1^2 + 1]}{m^2} > 0$$

<u>$k = 3$</u>:

From (B.4) and (B.6) the difference between w_3 and W_3 is:

$$w_3 - W_3 = \frac{\sigma_2^2\sigma_2^2 - \sigma_2^2(1 + \sigma_2^2)}{m} = \frac{-\sigma_2^2}{m}$$

Taking the derivative with respect to σ_2^2 we obtain:

$$\frac{\partial(w_3 - W_3)}{\partial\sigma_2^2} = \frac{-m + \sigma_2^2[(2 + \rho^2)\sigma_1^2 + 2\sigma_2^2 + 2\sigma_2^2\sigma_1^2 + 1]}{m^2} = \frac{1}{\sigma_1^2} \cdot \frac{\partial(w_1 - W_1)}{\partial\sigma_2^2} > 0$$

under the assumption of $\sigma_2^2 > \sigma_1^2 > \sigma_\omega^2 = 1$, as demonstrated for the case of $k = 1$.

<u>$k = 4$</u>:

Recall that $\sum_{k=1}^{4} W_k = \sum_{k=1}^{4} w_k = 1$. Accordingly, we obtain:

$$\frac{\partial(w_4 - W_4)}{\partial\sigma_2^2} = \frac{\partial[-(w_1 - W_1) - (w_2 - W_2) - (w_3 - W_3)]}{\partial\sigma_2^2} = -\frac{\partial(w_1 - W_1)}{\partial\sigma_2^2} - \frac{\partial(w_2 - W_2)}{\partial\sigma_2^2} - \frac{\partial(w_3 - W_3)}{\partial\sigma_2^2} < 0$$

under the assumption $\sigma_2^2 > \sigma_1^2 > \sigma_\omega^2 = 1$, by using our above results for $k = 1,2,3$.

At the limit, all the weight-differentials converge to zero when σ_2^2 goes to infinity, which again demonstrates why the estimated coefficient and the R-squared should converge to ρ and 1, respectively, when the horizon becomes uninformative. Finally, we note that the sign of the bias in the coefficient estimate can change from negative to positive before the estimate eventually convergences to ρ. This possibility can be noticed in the simulation results presented in Figure 5. Here it is demonstrated using our tractable case, by increasing the noise σ_2^2. Using equations (B.3) – (B.6), we express the OLS coefficient estimate (for $h = 2$) as follows:

$$\beta_{OLS} = \frac{Cov(x_{t+2|t}, x_{t+1|t})}{Var(x_{t+1|t})} = \frac{W_2 w_2 \rho \sigma_2^2 + W_3 w_3 \rho \sigma_1^2 + W_4 w_4 \rho^{-1} \sigma_2^2}{W_2{}^2 \sigma_2^2 + W_3{}^2 \sigma_1^2 + W_4{}^2 \rho^{-2} \sigma_2^2}$$

$$= \rho \frac{W_2 w_2 \sigma_2^2 + W_3 w_3 \sigma_1^2 + W_4 w_4 \rho^{-2} \sigma_2^2}{W_2{}^2 \sigma_2^2 + W_3{}^2 \sigma_1^2 + W_4{}^2 \rho^{-2} \sigma_2^2}$$

$$= \rho + \frac{W_2(w_2 - W_2)\rho \sigma_2^2 + W_3(w_3 - W_3)\rho \sigma_1^2 + W_4(w_4 - W_4)\rho^{-1} \sigma_2^2}{W_2{}^2 \sigma_2^2 + W_3{}^2 \sigma_1^2 + W_4{}^2 \rho^{-2} \sigma_2^2}$$

$$= \rho + \frac{-\sigma_1^2 \sigma_1^2 (1 + \sigma_2^2)\rho \sigma_2^2 - \sigma_2^2 \sigma_2^2 (1 + \sigma_2^2)\rho \sigma_1^2 + (\sigma_2^2 \sigma_1^2 + \sigma_1^2 + \sigma_2^2)\rho^2 \sigma_2^2 \sigma_1^2 \rho^{-1} \sigma_2^2}{\sigma_1^2 (1 + \sigma_2^2)\sigma_1^2 (1 + \sigma_2^2)\sigma_2^2 + \sigma_2^2 (1 + \sigma_2^2)\sigma_2^2 (1 + \sigma_2^2)\sigma_1^2 + \rho^2 \sigma_2^2 \sigma_1^2 \sigma_2^2 \sigma_1^2 \sigma_2^2} \quad \text{(B.15)}$$

$$= \rho + \frac{-\sigma_1^2 (1 + \sigma_2^2)\rho - \sigma_2^2 (1 + \sigma_2^2)\rho + (\sigma_2^2 \sigma_1^2 + \sigma_1^2 + \sigma_2^2)\rho \sigma_2^2}{\sigma_1^2 (1 + \sigma_2^2)(1 + \sigma_2^2) + \sigma_2^2 (1 + \sigma_2^2)(1 + \sigma_2^2) + \rho^2 \sigma_2^2 \sigma_1^2 \sigma_2^2} =$$

$$= \rho + \frac{\rho(\sigma_2^2 \sigma_2^2 \sigma_1^2 - \sigma_1^2 - \sigma_2^2)}{\sigma_1^2 (1 + \sigma_2^2)(1 + \sigma_2^2) + \sigma_2^2 (1 + \sigma_2^2)(1 + \sigma_2^2) + \rho^2 \sigma_2^2 \sigma_1^2 \sigma_2^2} =$$

$$= \rho + \frac{\rho(\sigma_2^2 \sigma_2^2 \sigma_1^2 - \sigma_1^2 - \sigma_2^2)}{(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2 + \rho^2 \sigma_2^2 \sigma_1^2 \sigma_2^2}$$

Thus, we obtain an expression for the bias which consists of competing negative and positive terms:

$$\beta_{OLS} - \rho = \frac{-\rho(\sigma_1^2 + \sigma_2^2)}{(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2 + \rho^2 \sigma_2^2 \sigma_1^2 \sigma_2^2} + \frac{\rho \sigma_2^2 \sigma_2^2 \sigma_1^2}{(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2 + \rho^2 \sigma_2^2 \sigma_1^2 \sigma_2^2}$$

Specifically, for low values of σ_2^2, the negative bias dominates, whereas the positive bias dominates for high values of σ_2^2. Nevertheless, it is easy to see that both biases eventually converge to zero when σ_2^2 goes to infinity.

B.4. Patterns of regression properties when changing persistence in the state

Another empirical pattern in the regression of $x_{t+h|t}^i$ on $x_{t+h-1|t}^i$, documented in section 2 and demonstrated by the simulation results in Figure 5, is the co-movement of the coefficient and R-squared. Apparently, this pattern is also associated with the differences in the optimal weights across forecasting horizons, as we illustrate with our tractable version of our model (see the right side of Table 1). We first look at the two extreme cases of $\rho = 0$ and $\rho = 1$.

$\boldsymbol{\rho = 0}$: the lack of persistence in the state process implies that the forecasts are uncorrelated across horizons. That is, a forecast $x_{t+h|t}^i$ relies only on signals referring to $t + h$, and is uncorrelated with $x_{t+h-1|t}^i$ (nor with $x_{t+h+1|t}^i$). In terms of our system of equations, the optimal forecasts would be:

$$x_{t|t}^i = x_t \quad \text{(B.16)}$$

$$x_{t+1|t}^i = W_2 y_{t-1,t+1}^i + W_3 y_{t,t+1}^i$$

$$x_{t+2|t}^i = w_4 y_{t,t+2}^i$$

$$x_{t+3|t}^i = 0$$

where $W_2 = 1 - W_3 = \sigma_1^2(\sigma_1^2 + \sigma_2^2)^{-1}$ and $w_4 = 1$. The forecasts are uncorrelated with each other since the fundamental is uncorrelated across different periods and the noise in forward signals is uncorrelated across horizons. It follows that in a cross-sectional regression of $x_{t+h|t}^i$ on $x_{t+h-1|t}^i$, the coefficient-estimate and R-squared should be zero. In terms of equation (B.12), for example, all cross-sectional variation in $x_{t+2|t}^i$ is driven by the regression error term, which collapses to $w_4 y_{t,t+2}^i$ (according to (B.16)).

$\boldsymbol{\rho = 1}$: When the fundamental follows a random walk, the general results obtained in Section B.3. would hold. This is not a limiting case where the coefficient and R-squared should converge to 1 ($\rho = 1$). Rather, the regression properties

10

would depend on the forecast horizon in the same way described above. Specifically, for the short horizon $h = 1$ the coefficient-estimate and R-squared should still be zero, while for the long horizon $h = 3$ the fit becomes perfect. The interesting case for investigating the response to change in persistence is the middle horizon, $h = 2$. As implied by equation (B.12), the coefficient on $x_{t+1|t}^i$ should be $\rho = 1$, but the OLS estimate would be biased, due to the correlation of $x_{t+1|t}^i$ with the error term, and the R-squared would be between zero and one. Nevertheless, the difference in regression properties between zero persistence and random walk still demonstrates an increase in the coefficient estimate and R-squared from zero to positive values, following a rise in persistence.

More generally, we now show how the relation between $x_{t+2|t}^i$ and $x_{t+1|t}^i$, as specified in (B.12) would vary when changing the persistence of the state process. First, notice that the difference between optimal weights placed on the same signals in $x_{t+2|t}^i$ and $x_{t+1|t}^i$, which is $(w_k - W_k)$, tends to diminish when the persistence ρ increases. From our previous results, the weight differential can be expressed as weights can be expressed as $w_k - W_k = a_k m^{-1}$. For $k = 1,2,3$, a_k(numerator) is negative and only m (denominator) is a function of ρ. Hence, it is sufficient to show that the derivative of m^{-1} with respect to ρ is negative:

$$\frac{\partial m^{-1}}{\partial \rho} = \frac{\partial [(2 + \rho^2)\sigma_1^2\sigma_2^2 + (\sigma_2^2)^2 + (\sigma_2^2)^2\sigma_1^2 + \sigma_2^2 + \sigma_1^2]^{-1}}{\partial \rho} = -\frac{2\rho\sigma_1^2\sigma_2^2}{m^2} < 0$$

Thus, the negative difference $w_k - W_k$, $k = 1,2,3$, diminishes as persistence increases. For $k = 4$ the difference $w_4 - W_4$ is positive. Following the same above argument, this gap should decrease in ρ because the weights w_k and W_k sum to one. Overall, this implies that the correlation between $x_{t+2|t}^i$ and $x_{t+1|t}^i$ will tend to be higher with more persistence since the error term in (B.12) would diminish due to the diminishing weight differential.

Second, we show more formally that the (biased) OLS coefficient estimate and fit of the regression are increasing in ρ, even in the presence of a non-diminishing bias. Based on (B.16), we begin with:

$$\beta_{OLS} - \rho = \frac{-\rho(\sigma_1^2 + \sigma_2^2)}{(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2 + \rho^2\sigma_2^2\sigma_1^2\sigma_2^2} + \frac{\rho\sigma_2^2\sigma_2^2\sigma_1^2}{(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2 + \rho^2\sigma_2^2\sigma_1^2\sigma_2^2} > \frac{-\rho(\sigma_1^2 + \sigma_2^2)}{(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2 + \rho^2\sigma_2^2\sigma_1^2\sigma_2^2} > -\rho$$

Thus, the downward bias is no greater than $-\rho$, so that β_{OLS} should always be positive. Taking the derivative with respect to ρ, we get:

$$\frac{\partial \beta_{OLS}}{\partial \rho} = 1 + \frac{(\sigma_2^2\sigma_2^2\sigma_1^2 - \sigma_1^2 - \sigma_2^2)[(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2 + \rho^2\sigma_2^2\sigma_1^2\sigma_2^2]}{[(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2 + \rho^2\sigma_2^2\sigma_1^2\sigma_2^2]^2} - \frac{2\rho^2\sigma_2^2\sigma_1^2\sigma_2^2(\sigma_2^2\sigma_2^2\sigma_1^2 - \sigma_1^2 - \sigma_2^2)}{[(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2 + \rho^2\sigma_2^2\sigma_1^2\sigma_2^2]^2}$$

$$= 1 + \frac{(\sigma_2^2\sigma_2^2\sigma_1^2 - \sigma_1^2 - \sigma_2^2)}{(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2 + \rho^2\sigma_2^2\sigma_1^2\sigma_2^2} - \frac{2\rho^2\sigma_2^2\sigma_1^2\sigma_2^2(\sigma_2^2\sigma_2^2\sigma_1^2 - \sigma_1^2 - \sigma_2^2)}{[(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2 + \rho^2\sigma_2^2\sigma_1^2\sigma_2^2]^2}$$

$$= 1 + (\beta_{OLS} - \rho)\rho^{-1} - (\beta_{OLS} - \rho)\frac{2\rho\sigma_2^2\sigma_1^2\sigma_2^2}{(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2 + \rho^2\sigma_2^2\sigma_1^2\sigma_2^2}$$

$$= \beta_{OLS}\rho^{-1} - (\beta_{OLS} - \rho)\frac{2\rho\sigma_2^2\sigma_1^2\sigma_2^2}{(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2 + \rho^2\sigma_2^2\sigma_1^2\sigma_2^2}$$

$$= \beta_{OLS}\left[\rho^{-1} - \frac{2\rho\sigma_2^2\sigma_1^2\sigma_2^2}{(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2 + \rho^2\sigma_2^2\sigma_1^2\sigma_2^2}\right] + \rho\frac{2\rho\sigma_2^2\sigma_1^2\sigma_2^2}{(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2 + \rho^2\sigma_2^2\sigma_1^2\sigma_2^2} > 0$$

It is easy to see that the term in brackets is positive under the assumption $\sigma_2^2 > \sigma_1^2$. Since β_{OLS} is also positive as shown above, the derivative is positive. Thus, we can conclude that the coefficient-estimate β_{OLS} increases in ρ.

The R-squared follows a similar pattern. The R-squared between $x_{t+2|t}^i$ and $x_{t+1|t}^i$ is given by:

$$R^2 = \frac{[Cov(x_{t+2|t}, x_{t+1|t})]^2}{Var(x_{t+1|t})Var(x_{t+2|t})} = \frac{(W_2 w_2 \rho\sigma_2^2 + W_3 w_3 \rho\sigma_1^2 + W_4 w_4 \rho^{-1}\sigma_2^2)^2}{(W_2^2\sigma_2^2 + W_3^2\sigma_1^2 + W_4^2\rho^{-2}\sigma_2^2)(w_2^2\rho^2\sigma_2^2 + w_3^2\rho^2\sigma_1^2 + w_4^2\sigma_2^2)}$$

After plugging in the optimal weights from (B.4) and (B.6) and rearranging, we can express the R-squared in the following way:

$$R^2 = \frac{B^2}{B^2 + \Delta},$$

Where:

$$B = m^2 Cov(x_{t+2|t}, x_{t+1|t}) = \sigma_1^2(\sigma_2^2)^2[\rho^3\sigma_1^2\sigma_2^2 + \rho(1 + \sigma_2^2)(\sigma_1^2 + \sigma_2^2) + \rho(\sigma_2^2\sigma_1^2 + \sigma_1^2 + \sigma_2^2)]$$

$$\Delta = \sigma_1^2(\sigma_2^2)^2(\sigma_1^2 + \sigma_2^2)[\rho^4(\sigma_1^2)^2(\sigma_2^2)^2 + 2\rho^2\sigma_1^2\sigma_2^2(1 + \sigma_2^2)(\sigma_2^2\sigma_1^2 + \sigma_1^2 + \sigma_2^2) + (1 + \sigma_2^2)^2(\sigma_2^2\sigma_1^2 + \sigma_1^2 + \sigma_2^2)^2]$$

11

Taking the derivative with respect to ρ yields:

$$\frac{\partial R^2}{\partial \rho} = \frac{2B\frac{\partial B}{\partial \rho}(B^2 + \Delta) - B^2\left(2B\frac{\partial B}{\partial \rho} + \frac{\partial \Delta}{\partial \rho}\right)}{(B^2 + \Delta)^2} = \frac{B\left(2\Delta\frac{\partial B}{\partial \rho} - B\frac{\partial \Delta}{\partial \rho}\right)}{(B^2 + \Delta)^2} = K\left(2\Delta\frac{\partial B}{\partial \rho} - B\frac{\partial \Delta}{\partial \rho}\right)$$

Where $K = B/(B^2 + \Delta)^2 > 0$. The derivative, $\partial B/\partial \rho$, can be expressed as:

$$\frac{\partial B}{\partial \rho} = B\rho^{-1} + 2\rho^2(\sigma_1^2)^2(\sigma_2^2)^3$$

Plugging this into the derivative of R^2, we obtain:

$$\frac{\partial R^2}{\partial \rho} = K\left\{B\left(2\Delta\rho^{-1} - \frac{\partial \Delta}{\partial \rho}\right) + 4\Delta\rho^2(\sigma_1^2)^2(\sigma_2^2)^3\right\}$$

$$= K\{B[-2\rho^3(\sigma_1^2)^3(\sigma_2^2)^4(\sigma_1^2 + \sigma_2^2) + 2\rho^{-1}\sigma_1^2(\sigma_2^2)^2(\sigma_1^2 + \sigma_2^2)(1 + \sigma_2^2)^2(\sigma_2^2\sigma_1^2 + \sigma_1^2 + \sigma_2^2)^2]$$
$$+ 4\Delta\rho^2(\sigma_1^2)^2(\sigma_2^2)^3\}$$
$$= K\{B\sigma_1^2(\sigma_2^2)^2(\sigma_1^2 + \sigma_2^2)[-2\rho^3(\sigma_1^2)^2(\sigma_2^2)^2 + 2\rho^{-1}(1 + \sigma_2^2)^2(\sigma_2^2\sigma_1^2 + \sigma_1^2 + \sigma_2^2)^2]$$
$$+ 4\Delta\rho^2(\sigma_1^2)^2(\sigma_2^2)^3\}$$

The term in the squared brackets is positive, and consequently the whole derivative. It follows that the R-squared increases in ρ. Hence, the R-squared and the coefficient estimate should demonstrate a pattern of co-movement in response to a change in ρ.

B.5. Predictability of forecast errors

This section shows that forecast errors are predictable by forecast revisions at the aggregate level due to forward signals. Consider the one-step-ahead forecast error. Using (B.3), we obtain:

$$x_{t+1} - x_{t+1|t}^i = W_1\omega_{t+1} + W_2(-v_{t-1,t+2}^i) + W_3(-v_{t,t+1}^i) + W_4\left(-\rho^{-1}(\omega_{t+2} + v_{t,t+2}^i)\right)$$

By taking the average across agents, all terms with idiosyncratic noise drop out. Hence, we get:

$$x_{t+1} - x_{t+1|t} = W_1\omega_{t+1} + W_4(-\rho^{-1}\omega_{t+2})$$

where $x_{t+1|t}$ (without superscript i) denotes the cross-sectional average.

The forecast $x_{t+1|t}$ revises the forecast $x_{t+1|t-1}$, which is the two-step-ahead forecast from the last period. Using (B.3) and (B.5), and averaging across agents, the revision to the average forecast could be expressed as (exploiting the property that the optimal weights amount to 1):

$$x_{t+1|t} - x_{t+1|t-1} = (W_1\rho x_t + W_2 x_{t+1} + W_3 x_{t+1} + W_4\rho^{-1}x_{t+2}) - (w_1\rho^2 x_{t-1} + w_2\rho x_t + w_3\rho x_t + w_4 x_{t+1}) =$$
$$W_1(-\omega_{t+1}) + W_4\rho^{-1}\omega_{t+2} + w_1(\rho\omega_t + \omega_{t+1}) + (w_2 + w_3)\omega_{t+1} = w_1\rho\omega_t + (w_1 + w_2 + w_3 - W_1)\omega_{t+1} + W_4\rho^{-1}\omega_{t+2}$$

Consider a regression of forecast error with the forecast revision as the explanatory variable. Then, the expected OLS coefficient estimate would be (setting $\sigma_\omega^2 = 1$ and assuming $E[\omega_t\omega_{t-1}]$) :

$$\beta_{CG} = \frac{Cov(x_{t+1} - x_{t+1|t}, x_{t+1|t} - x_{t+1|t-1})}{Var(x_{t+1|t} - x_{t+1|t-1})} = \frac{E\left[(x_{t+1} - x_{t+1|t})(x_{t+1|t} - x_{t+1|t-1})\right]}{E\left[(x_{t+1|t} - x_{t+1|t-1})^2\right]}$$

$$= \frac{E\left[(W_1\omega_{t+1} + W_4(-\rho^{-1}\omega_{t+2}))(w_1\rho\omega_t + (w_1 + w_2 + w_3 - W_1)\omega_{t+1} + W_4\rho^{-1}\omega_{t+2})\right]}{E\left[(w_1\rho\omega_t + (w_1 + w_2 + w_3 - W_1)\omega_{t+1} + W_4\rho^{-1}\omega_{t+2})^2\right]}$$

$$= \frac{W_1(w_1 + w_2 + w_3) - (W_1)^2 - (W_4)^2\rho^{-2}}{(w_1)^2\rho^2 + (w_1 + w_2 + w_3 - W_1)^2 + (W_4)^2\rho^{-2}}$$

We can further verify that the expected coefficient is positive by showing that the numerator is positive. Using the expressions for the optimal weights from (B.4) and (B.6), we obtain (m is defined in (B.4)):

$$W_1(w_1 + w_2 + w_3) - (W_1)^2 - (W_4)^2\rho^{-2} = \frac{\sigma_2^2\sigma_1^2(1 + \sigma_2^2)\sigma_2^2\sigma_2^2 - \rho^2\sigma_2^2\sigma_1^2\sigma_2^2\sigma_1^2}{m^2} =$$

$$= \frac{\sigma_2^2\sigma_1^2\sigma_2^2\left((1 + \sigma_2^2)\sigma_2^2 - \rho^2\sigma_1^2\right)}{m^2}$$

The expression in brackets $\left((1 + \sigma_2^2)\sigma_2^2 - \rho^2\sigma_1^2\right)$ is positive due to $\sigma_2^2 > \sigma_1^2 > 1$ and $\rho \leq 1$. Consequently, the whole numerator is positive, making the coefficient on the forecast revision positive as well ($\beta_{CG} > 0$).

Appendix C: Simulations

This appendix provides more details on the simulation results reported in Section 3.3. Figure 5 replicates the persistence regression of $x_{t+h|t}^i$ on $x_{t+h-1|t}^i$ that was estimated in Section 2, using SPF inflation forecasts. The results in Figure 5 are in line with the patterns documented in the SPF data along several dimensions.

First, we observe in Panel A that the coefficient estimate is substantially lower than the true persistence for the short horizons. For example, when the true persistence is 0.5 the coefficient estimate is below 0.2 in the regression applied to $h = 1$. However, as we estimate the regression using forecasts with longer horizons (higher noise to signal ratios), the coefficient estimate gets closer to the persistence value and finally converges to it. It is important to note that the convergence is not monotonic, but rather takes somewhat hump-shaped path. This result illustrates the possibility of upward bias as well, although the bias in this direction is quite small in the simulation (the positive bias is also possible in our above tractable example. See Appendix B.3.). The hump shape is also consistent with the evidence in Figure 2, in which the lines corresponding to the two longest horizons are flipping over time.

Second, there is a similar pattern of convergence for the R-squared statistic (Panel B). Interestingly, the fit is very poor in regressions with forecasts at short horizons, representing the higher role of forward information in the error term. The values of the R-squared start below 0.1 for each persistence value, but eventually increase towards a perfect fit. Furthermore, the convergence could be quite slow, especially in cases where the degree of persistence is low. For instance, with $\rho = 0.2$, the R-squared is below 0,5, even when estimating the regression for $h = 6$, for which the signal is effectively uninformative. This property is consistent with the low R-squared values documented in recent years, even for the longer horizons in the SPF (Figure 2, Panel B).

Third, the simulation results in Figure 5 also demonstrate a dependence of the regression properties on the underlying degree of persistence. Importantly, not only does the coefficient-estimate increases in the degree of persistence (Panel A), but so does the R-squared statistic (Panel B). This pattern is consistent across the different horizons (horizontal axis), for which we estimate the regressions. The co-movement of the regression estimates and persistence resembles the pattern documented in Section 2, where the coefficient and the fit of the regression deteriorate over time along with the decline in inflation persistence.

Appendix Figure G.2 describes results from another set of simulations, in which the variance of the noise is constant instead of monotonically increasing in the horizon. Overall, it seems that the patterns of the regression properties are similar to those in Figure 5. However, notice that the rate of convergence is different. Due to the concavity of the lines, the convergence rate is slow for the short horizons and becomes faster for the longer horizons. This type of convergence seems less consistent with the evidence in Figure 2. The next simulation provides further evidence that supports the more realistic pattern of increasing noisiness in forward signals.

As mentioned in Section 4, we further simulate the estimation of specification (26) on simulated forecasts as in Section 3.3, with ρ set to 0.5. Appendix Table G.1. resembles the structure of Table 3, where each column is a simulated estimation for a certain horizon. The coefficient estimates are in bold, while the regular numbers report the true parameters from the matrix $(I - G)$. Although $H = 7$ in the simulation, the regressions were estimated as if there were only 5 available horizons in the data, thereby, checking the sensitivity of the results to this practical limitation. The results confirm that such truncation is not a concern. However, the estimates in Appendix Table G.1. are still very close to the true values. Thus, the estimation of specification (26) seems very reliable despite a truncation. In addition, as in our results for the SPF the diagonal coefficients are very dominant. Interestingly, the simulation provides a good approximation even to the off-diagonal elements. In practice, as evident in Table 3 these estimates could still suffer from imprecision due to the small values, and some multicollinearity due to measurement error.

Another useful point is demonstrated by the results in Appendix Table G.2., which are based on the simulation as in Appendix Figure G.2, where the variance of the noise stays constant across horizons. Compared to the previous simulation with increasing noise, there is a notable difference in the pattern of the diagonal elements across horizons. when noise increases in the horizon (Appendix Table G.1.), the diagonal estimates also increase in the horizon, while for fixed noise there is a decline. Our results in Table 3 for the SPF are in line with the first pattern. Thus, this result is another form of support for a general pattern of deterioration in forward information as the horizon gets longer.

This result also sheds light on the estimation of the restricted specification in (27). Because this version focuses on the diagonal elements of $(I - G)$ in the forward information framework, the coefficients should tend to increase when the specification is estimated for longer horizons. This pattern was documented in the SPF data by Goldstein (2021). Thus, according to our simulation, this variation provides another form of evidence against the standard noisy information framework, in which the coefficient in (27) should not vary across horizons.

Appendix D: Estimating Persistence with Asymmetric Loss Function

According to Elliot et al. (2008) and Capistrán and Timmerman (2009), biased forecast errors observed in surveys could result from asymmetric preference of positive over negative forecast errors or vise-versa. Different asymmetrical tendencies across individuals would explain forecast disagreement. More formally, following Capistrán and Timmermann (2009), this asymmetry is modeled by a LINEX loss-function over forecast errors:

$$L\big(FE_t^i x_{t+h};\ \theta_i\big) = \big[exp\big(\theta_i FE_t^i x_{t+h}\big) - \theta_i(h)FE_t^i x_{t+h} - 1\big]/\theta_i^2$$

where $FE_t^i x_{t+h} \equiv x_{t+h} - x_{t+h|t}^i$ is the forecast error of forecaster i and θ_i is the asymmetry parameter. A positive value of θ_i corresponds to positive error loss-aversion, while a negative θ_i corresponds to the opposite. As θ_i shrinks to zero the function converges to the regular (symmetric) squared error loss-function.

The optimal individual forecast which minimizes the specified loss-function is:

$$x_{t+h|t}^i = E_t x_{t+h} + \frac{1}{2}\theta_i \sigma_{t+h|t}^2 \qquad\qquad (D.1)$$

where the variable x is assumed to be normally distributed with conditional mean, represented by the rational expectation term $E_t x_{t+h}$, and with conditional variance $\sigma_{t+h|t}^2$. Thus, the individual forecast is biased relative to the rational expectation by a term that depends on the asymmetric tendency parameter θ_i, and the variance of the x.

Similarly, the forecast for $h - 1$ steps ahead is:

$$x_{t+h-1|t}^i = E_t x_{t+h-1} + \frac{1}{2}\theta_i \sigma_{t+h-1|t}^2 \qquad\qquad (D.2)$$

Suppose that the fundamental follows an AR(1) process $x_t = \rho x_{t-1} + \omega_t$, where $\omega_t \sim iid\ N\big(0, \sigma_{t|t-1}^2\big)$. The implied rational expectation would therefore be:

$$E_t x_{t+h} = \rho^h x_t = \rho E_t x_{t+h-1} = \rho x_{t+h-1|t}^i - \frac{1}{2}\rho\theta_i \sigma_{t+h-1|t}^2$$

where we use (D.2) to substitute for $E_t x_{t+h-1}$. We then substitute this expression in (D.1) and rearrange to obtain

$$x_{t+h|t}^i = \rho x_{t+h-1|t}^i + \frac{1}{2}\theta_i\big(\sigma_{t+h|t}^2 - \rho\sigma_{t+h-1|t}^2\big) \qquad\qquad (D.3)$$

The last term on the right-hand side would correspond to the error term in a cross-sectional regression of the forecast $x_{t+h|t}^i$ on the forecast $x_{t+h-1|t}^i$. The mean of the error term would be zero only if θ_i has a zero mean (symmetry). In the general case, where the mean of θ_i is θ, which is different from zero, the regression would include a constant term and the error term would be $\frac{1}{2}(\theta_i - \theta)\big(\sigma_{t+h|t}^2 - \rho\sigma_{t+h-1|t}^2\big)$.

In any case, it is clear from (D.2) that θ_i is positively correlated with $x_{t+h-1|t}^i$, so that the OLS estimate of the coefficient on $x_{t+h-1|t}^i$ would be a biased estimate of the persistence parameter ρ. In fact, the OLS coefficient in a cross-sectional regression is unrelated to ρ, because heterogeneity in forecasts is driven only by the asymmetric bias component, while the persistence component, $\rho^h x_t$, is the same across agents. More formally:

$$\beta_{OLS} = \frac{Cov(x_{t+h|t}, x_{t+h-1|t})}{Var(x_{t+h-1|t})} = \frac{Cov(\rho^h x_t + \frac{1}{2}\theta_i \sigma_{t+h|t}^2, \rho^{h-1}x_t + \frac{1}{2}\theta_i \sigma_{t+h-1|t}^2)}{Var(\rho^{h-1}x_t + \frac{1}{2}\theta_i \sigma_{t+h-1|t}^2)} = \frac{Cov(\frac{1}{2}\theta_i \sigma_{t+h|t}^2, \frac{1}{2}\theta_i \sigma_{t+h-1|t}^2)}{Var(\frac{1}{2}\theta_i \sigma_{t+h-1|t}^2)}$$

$$= \frac{\sigma_{t+h|t}^2}{\sigma_{t+h-1|t}^2}$$

Thus, if the conditional variance of the fundamental is time-independent, the OLS coefficient in a cross-sectional regression of $x_{t+h|t}^i$ on $x_{t+h-1|t}^i$ equals 1. Capistrán and Timmermann (2009) have assumed that the conditional variance follows a GARCH(1,1) process:

$$\sigma_{t+1|t}^2 = \alpha_0 + \alpha_1 \omega_t^2 + \beta_1 \sigma_{t|t-1}^2$$

Thus, the resulting OLS coefficient $(\alpha_0/\sigma_{t+h-1|t}^2 + \alpha_1 + \beta_1)$ depends on the conditional variance. Still, it does not depend on ρ, nor does it follow a particular pattern of variation across forecast horizons.

Appendix E: More Applications of Forward Information

In Section 6, we proposed a simple measure of forward information, based on the forward information component in forecast data. This appendix complements Section 6 with two additional applications of our measure, which are described briefly in Sections 6.2.1. and 6.2.3.

E.1. Big news

As demonstrated in the main text, our method for extracting forward information from forecasts and the forecast persistence regressions from Section 2 are closely related. The forward information component shifts the persistence estimates downward at the shorter horizons. Thus, it is expected that times of big news will induce big shifts in forecast persistence. A recent prominent example of a major event that should deliver a high amount of forward information is the outbreak of COVID-19. In terms of our framework, forecasters would interpret the outbreak of the pandemic as a series of multi-horizon shocks that will hit the economy in the following quarters. According to our analysis, the adjustment of forecasts to such significant news may lead to a decline in the persistence of their forecasts at the various horizons.

The results in Panel A of Appendix Figure G.3 support this conjecture. The figure shows estimates of forecast persistence before and after the outbreak of COVID-19, by applying specification (4) to the SPF waves of 2020Q1 and 2020Q2. The forecasts in these waves were collected in February and May 2020, respectively, which is before and after the outbreak of the pandemic outside China. Interestingly, there is a substantial decline in the estimates of persistence in the survey wave of 2020Q2 relative to 2020Q1, for all four major macro variables presented in the figure (inflation, GDP growth, interest rate and unemployment). Moreover, the decline is observed not only in short-horizon estimates ($h = 1$, left graph), but also in estimates for the longer horizon ($h = 3$, left graph). This finding suggests that the scope of forward information embedded in the outbreak of COVID-19 is long enough to have a pervasive biasing effect even on the year-ahead estimates. Still, in line with the decaying effect of forward information across horizons, as demonstrated above, the decline in the estimates at the longer horizon is more moderate. Specifically, as shown in the figure, estimated persistence has declined by around 0.5 (!) in 2020Q2 relative to 2020Q1 for $h = 1$, and by around a half of this size for $h = 3$.

As a reference to the evidence from COVID-19, Appendix Figure G.3 presents the results from a similar exercise applied to two other major events: the financial collapse on the eve of the Great Recession and the 9/11 terror attack (Panels B and C, respectively). For the collapse of the big US financial firms in September 2008, we compare the SPF waves of 2008Q3 and 2008Q4. For the terror attack of September 2001, we compare the SPF waves of 2001Q3 and 2001Q4. In contrast to the COVID-19 results, the pattern of decline in persistence is much weaker, following these events, and it is not consistent across variables and horizons. Hence, the sharp pattern of decline in persistence following COVID-19, which was obtained despite a small sample of forecasters in single waves, points to an exceptional amount of forward information brought by the burst of this unprecedented crisis.

E.2. Comparison with a methodology of news shocks

This application provides a comparison of the forward-information measure of news with news shocks identified by a familiar method in the literature. For this purpose, we extend the analysis in Section 5.2 to a VAR framework. We estimate the effect of the series of forward information about inflation, $FI_{t|t}$, in a VAR with four macroeconomic variables, including the inflation rate, the unemployment rate, 3-month and 10-year treasury bill rates. The estimated impulse responses of the four variables to $FI_{t|t}$ are presented in panel A of Appendix Figure G.4. Again, we obtain a strong response of inflation to its forward information, despite taking into account the effect of other variables. The initial response is close to the estimates in Table 6 (column (2)) and then dies out quite quickly. The response of the other variables to forward information about inflation is quite weak and insignificant.

We then identify inflation news shocks, using the influential method suggested in Barsky and Sims (2011). The idea of their approach is to identify a news shock as a shock orthogonal to the current innovation in the fundamental, which best explains future movements in that fundamental. In line with the literature on news shocks, they apply their method to identify technology shocks. Here, we implement their method to identify inflation news shocks, in a VAR framework that includes the same variables used above (inflation, unemployment, short and long-run interest rates). The estimated impulse responses to the identified inflation news shocks are reported in Panel B of Appendix Figure G.4. In contrast to our previous findings (Panel A), the news shocks identified by the Barsky-Sims method do not induce any significant response to it, even by inflation.

Based on a forecast error variance decomposition, Appendix Table G.4 reports, for each method, the share of inflation volatility at various horizons which is explained by the news variable. The difference between the two methods is again very clear: While according to the Barsky-Sims method news shocks account for no more than 10% of the variation in inflation, the component quantified by the forward information approach can explain more than a half of the same variation. This comparison illustrates how the straightforward application of the notion of forward information to forecast data can be useful for quantifying predictive components that are important for explaining future movements in the fundamental beyond the standard VAR.

Appendix F: Forward information in forecasts of additional macro variables

This appendix describes in more detail the evidence that was summarized in Section 6 in the main text.

F.1. SPF forecasts

The empirical evidence through the paper was based on inflation expectations, on which there is a particular focus by the literature, due to their main role in macroeconomic analysis. We examine how forward information is incorporated in expectations more generally, by first investigating SPF forecasts of additional key macroeconomic variables, as well as forecasts of other measures of inflation.

We first examine if the patterns of persistence documented in Section 2, characterize forecasts of other macroeconomic variables as well. Thus, we regress the forecast $x^i_{t+h|t}$ on the forecast $x^i_{t+h-1|t}$, using multi-horizon forecasts of other macroeconomic variables, available in the SPF. A pattern of increase in the coefficient estimates and R-squared across forecast horizons, as in Figure 2, would point to a utilization of forward information in expectations of other variables. Appendix Figure G.6 presents the results for SPF forecasts of unemployment rate, interest rate and real GDP growth. The results are presented in the same form as in Figure 2. Regressions are estimated quarter-by-quarter, using the cross-sections of forecasts from the last eight quarters. For each variable, the figure describes the coefficient estimate (left side) and R-squared statistic (right side), estimated over time and across the different horizons available in the survey ($h = 0,1,2,3,4$. For GDP growth forecasts $h = 0$ is not available due to the conversion of level forecasts from the survey to growth forecasts). Appendix Figure G.7 presents results for forecasts of additional measures of inflation: GDP deflator, PCE inflation and core inflation (due to a similar conversion of prices-level forecasts to inflation forecasts, $h = 0$ estimates are also not available for the GDP deflator). The sample period varies depending on the availability of forecast data for each variable: For unemployment, GDP growth and GDP inflation, the sample starts at the beginning of the 1970s. For the interest rate, the sample starts at the beginning of the 1980s and for PCE and core inflation, it begins in 2009 (eight quarters after the first available forecasts due to the small window size).

For most variables, both the coefficient-estimate and R-squared demonstrate patterns of variation across horizons and convergence that are similar to the baseline results in Figure 2. For unemployment and interest rate forecasts (panels A and B in Appendix Figure G.6), the estimated persistence in the "convergence" region (longer horizons), is quite steady over the years, around the level of 1. Thus, SPF participants consistently associate a random walk process with movements in unemployment and interest rate. Notably, compared to inflation, the convergence of estimates across horizons for unemployment and interest rate forecasts is quite fast: the graphs for $h = 2,3,4$ are quite similar, whereas only for $h = 0$ the coefficient-estimates and R-squared are noticeably lower. The pattern of faster convergence suggests that utilization of forward information with respect to these variables might be lower, compared to inflation. This result may reflect a higher degree of attention to inflation by the forecasters or a greater availability of valuable forward signals about inflation, due to measures of central bank communication, especially in recent years. The convergence pattern could also be affected by the noise in realized data. Recall, that according to our above evidence, survey participants are well-informed about realized inflation. This, however, could be different for variables with more data releases such as unemployment, or a high-frequency variable such as the Treasury-Bill rate. This point will be examined later using the methodology proposed in Section 4.

For the forecasts of PCE and core inflation (Panels B and C in Appendix Figure G.7), there is more variation across horizons and the convergence is around $h = 3, 4$, implying the availability of forward signals up to a year ahead, similar to the baseline findings for CPI inflation forecasts. The coefficient estimates for the longer horizons provides a measure of persistence that is quite steady over the years, since the data is available for these variables only for the last decade, after the great decline in inflation persistence. Also notice that the persistence in PCE inflation is around 0.6, which is close to the persistence in CPI inflation during these years, as described in Figure 2. For the core inflation, persistence is estimated at around 0.8. This higher degree of persistence in core inflation is in line with the notion that this measure approximates trend inflation, by excluding CPI components with high-frequency volatility. However, the measured persistence is still lower than 1.

For two variables, there is no clear pattern across horizons in the persistence regressions: The GDP growth (Panel C in Appendix Figure G.6) and the inflation by GDP deflator (Panel A in Appendix Figure G.7). In order to understand the reason for this exception, we should recall that the pattern of convergence across horizons depends on two important factors: First, the structure of information and secondly, the underlying process, specifically, how well this process is approximated by AR(1). Considering the first issue, it should be noticed that unlike other variables in the SPF which are forecasted in terms of the rate of change, the original forecasts of GDP and GDP deflator are reported in levels. The original forecasts are then transformed to GDP growth and inflation rates, and the estimation is applied to the rate-of-change forecasts, as it is applied to forecasts of other variables. However, it is not clear if forward signals were applied by SPF forecasters to predict the level or the rate of change and this may obscure the pattern across horizons. For instance, suppose that the process in growth rates follows AR(1):

$$x_t = \rho x_{t-1} + \omega_t$$

where $x_t = \Delta z_t$. Thus, the process in levels follows AR(2):

16

$$\Delta z_t = \rho \Delta z_{t-1} + \omega_t \Leftrightarrow z_t = (1+\rho)z_{t-1} - \rho z_{t-2} + \omega_t$$

Forecasters receive multiple forward signals on the levels z_{t+h} ($h = 0,1,...,H$) and apply the corresponding state-space representation in terms of z_t. Accordingly, they form Kalman filter level-forecasts $z_{t+h|t}^i$. The appropriate persistence regression for this case would therefore be:

$$z_{t+h|t}^i = c + \rho_1 z_{t+h-1|t}^i + \rho_2 z_{t+h-2|t}^i + error_t$$

This specification also relates to the second issue mentioned above. We may not only need to estimate the specification using forecasts in levels, but also to expand the dynamics beyond AR(1) by including more "lags." Appendix Figure G.8 shows the results of applying this specification to the original level forecasts of GDP and GDP deflator. It describes the persistence estimate as the sum of the coefficients ($\rho_1 + \rho_2$) and the R-squared for the various horizons, where h starts at 2 due to the additional "lag". As opposed to the results in Appendix Figure G.6 (Panel C) and Appendix Figure G.7 (Panel A), the results from the modified specification are much closer to a pattern of "layers," as was documented for the other variables. The big swings in the 1980s may indicate that more "lags" are still required. Indeed, the estimation of AR(4) persistence (adding the forecasts $z_{t+h-3|t}^i$ and $z_{t+h-4|t}^i$) would diminish these swings, but the pattern of layers cannot be demonstrated with high-order dynamics. It should also be noted that some measurement issues are also possible since explicit forecasts of real GDP has been provided only since the beginning of the 1980s (see documentation at the SPF website). In sum, the results from the persistence regressions broadly confirm the patterns documented in Section 2. Forward information thus plays a more general role in expectations of macroeconomic variables. Our findings also demonstrate how patterns may change when the underlying process is more complex.

We also apply the direct method proposed in section 4, for testing and estimating the role of forward information, to SPF forecasts of the above macro variables. First, we test the presence of noise in realization, according to the standard noisy information framework, by applying specification (26) to deviations of individual backcasts from the mean ($h = 0$, the test is not performed for GDP growth and GDP inflation because of the absence of backcasts for these variables, as noted above). The estimation results are reported in Appendix Table G.5. Recall that perfect information about realizations would imply zero coefficients, as obtained above for CPI inflation forecasts (Table 2). The same result is also obtained using forecasts of PCE and core measures of inflation. For the unemployment and interest rate, some significance is documented for several coefficients, but the size of coefficient estimates is still very close to zero, as reported in Appendix Table G.5. Thus, information about realizations is close to perfect, even for these variables. At the same time, the limited amount of noise found in the backcasts of unemployment and interest rate is in line with the faster pattern of convergence across horizons in the persistence estimates, as documented above for the same variables.

Second, we estimate specification (26) using the forecasts of each variable for the various available horizons and then compare it with the restricted version in specification (27), which rules out the availability of forward signals (zero coefficients on non-diagonal elements). Appendix Table G.6. presents the BIC statistics from the two specifications, for each variable and each forecasting horizon ($h = 1,2,3,4$). Similar to the findings in section 4, the unrestricted specification (26) outperforms the restricted version. Hence, there is direct evidence for the presence of forward information, even in forecasts of GDP growth and GDP inflation, for which the indirect approach based on persistence patterns, provided less conclusive evidence.

Finally. we briefly examine how our approach for quantifying forward information in Section 5 can identify forward information with predictive power in forecasts of additional variables. We employ SPF forecasts of unemployment, interest rate and GDP growth and estimate series of $FI_{t|t}$ for each variable using (32). As for inflation, it involves the estimation of quarter-by-quarter persistence for each variable with the proper horizon, based on the steps described above (constant term is also estimated). Appendix Table G.7. reports highly significant effects of those forward information series in a simple univariate AR. The significant effect obtained for highly persistent variables, such as the unemployment and interest rate also indicate the long-lasting impact of forward information on movements in macroeconomic fundamentals (We also considered the unit root possibility, by taking the differenced series of unemployment and interest rates and running them on the forward information series. The coefficient on forward information was again highly significant in these specifications.).

F.2. Fed Forecasts

The Greenbook forecasts of the Fed staff are available for multiple horizons, thus allowing us to examine if the patterns of persistence across horizons that were documented in the SPF also exists in forecasts of the Fed. Accordingly, we regress $x_{t+h|t}$ on $x_{t+h-1|t}$ for $h = 0,...,4$ as in Section 2, with the difference that only time-series regressions can be applied (no cross-section). The results are reported in Appendix Table G.3. for CPI inflation and additional variables with a comparison to the SPF (Panels A and B). Interestingly the same pattern of an increase in the coefficient over h from Section 2 is documented in the Greenbook forecasts (and similarly in mean-level forecasts of the SPF). The

robustness of this pattern in Greenbook forecasts across several variables points to utilization of forward information by the Fed staff when preparing their periodic projections for the U.S. economy.

We also note that in the Greenbooks, unlike in the SPF, the forecasts of both GDP and GDP deflator are originally in growth rates. Indeed, the results reported in Appendix Table G.3. demonstrate an increase in persistence across horizons for both variables, when applying time-series regressions to the Greenbook forecasts (Panel B). Panel A further shows that a similar pattern arises in SPF mean-level forecasts of GDP growth based on a time-series regression, but not in the GDP inflation forecasts. These findings shed more light on issues that were raised above regarding the SPF micro-level evidence on these variables.

### F.3. ECB SPF inflation forecasts

We also investigate inflation forecasts in the European SPF, managed by the European Central Bank since 1999Q1. The forecasts also refer to consumer inflation (the Harmonized Index of Consumer Prices across European countries), and like the US SPF, the survey is quarterly. Unlike the US SPF, though, there are no quarterly forecasts in the ECB SPF. Instead, all forecasts refer to annual changes. Specifically, participants provide rolling-year forecasts for the inflation a year from now, and the year afterwards. Additionally, they provide forecasts for calendar years, mainly for the current year and one year ahead. Thus, we cannot estimate persistence regressions across quarterly horizons as in Section 2. Rather, we can obtain only two persistence estimates: the persistence in the rolling-year forecasts and the persistence in calendar-year forecasts. Notice, however, that the forward information model implies an important distinction between the two estimations. If forward information is utilized by forecasters, the calendar-year persistence should vary across calendar quarters, while the rolling-year persistence should not. The reason is that the target years of calendar forecasts are fixed when advancing from Q1 to Q4 of the year. As a result, calendar forecasts apply more forward information from survey to survey, during the calendar year, thus, leading to more biased estimates of persistence, according to our above analysis.

We examine this conjecture by estimating persistence regressions, for each calendar-quarter separately. Appendix Table G.8. reports the estimates for the two types of annual forecasts in Panels A and B (rolling and calendar forecasts, respectively). The results support the presence of a forward information effect. For the calendar-year forecasts, there is a sizable reduction in persistence estimates when moving from Q1 to the other quarters. The persistence for Q1 surveys is 0.429, while the estimates using Q2, Q3 or Q4 surveys are between 0.355 and 0.397. The null of coefficient equality across calendar quarters is also strongly rejected. By contrast, with the rolling-year forecasts, coefficient estimates are higher and vary little across quarters, between 0.453 and 0.502.

We can also check the variation across calendar quarters in the US SPF, since besides the quarterly forecasts it further provides forecasts for the calendar year, like the ECB SPF. Panel C in Appendix Table G.8. reports the results, using CPI inflation forecasts for the same period as for the European survey (1991Q1-2007Q4). Strikingly, we find even more clear pattern of reduction in the persistence estimates, across the calendar quarters. Notice also that when using Q1 forecasts, which should provide the least-biased estimation (longer horizons), the coefficient estimate for the US SPF is the same as the estimate for the ECB SPF – 0.429, implying a similar degree of inflation persistence in the two areas during the recent years, when considering a far enough horizon to rule out the forward information bias. Finally, we note that the regression constant is a little bit higher for the US, in line with a small difference in trend inflation. Overall, these findings provide additional support for the presence of forward information in both surveys of professionals.

Appendix G: Additional figures and tables

Figure G.1: Persistence Patterns with Various AR Orders

Panel A: AR(2)
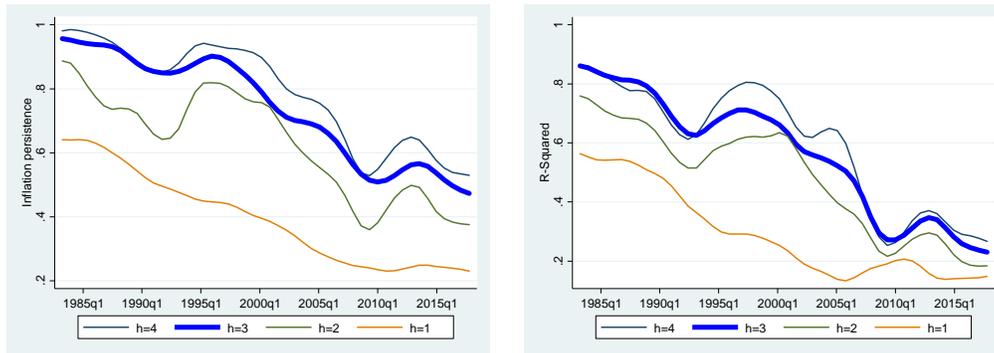
Panel B: AR(3)
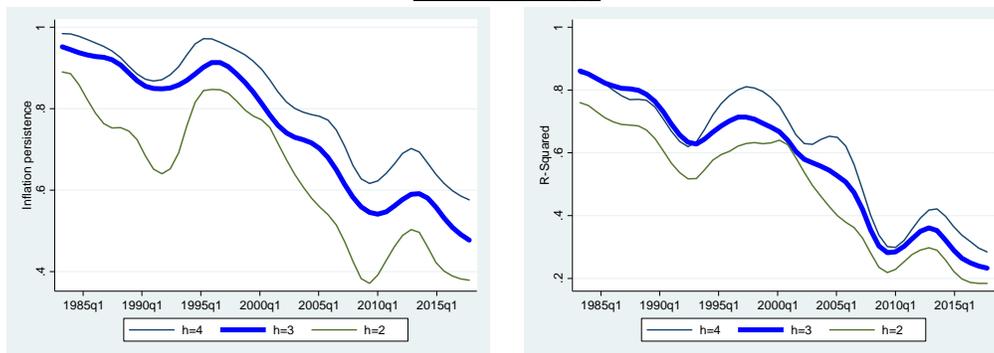
Panel C: AR(2)
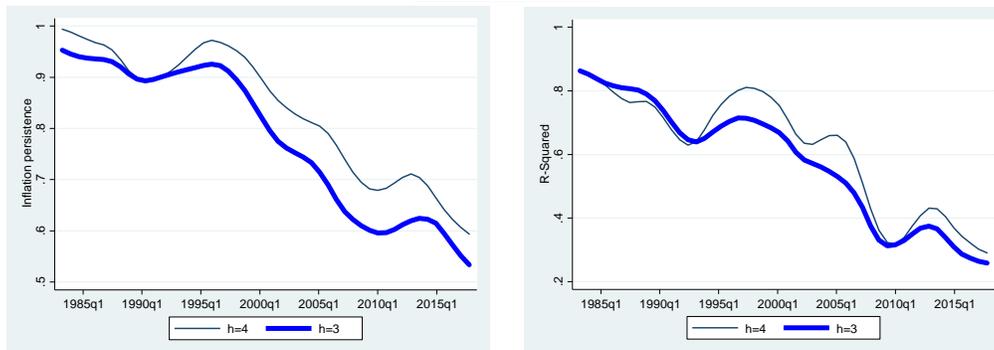
Notes: The figure plots smoothed estimates of persistence $\hat{\rho}$ and R-squared measures, based on estimating augmented version of specification (4) for different forecast horizons in the SPF survey. The augmented version includes additional forecasts for previous quarters according to the order of the autoregression, as specified in each panel. Each quarterly point is based on OLS estimation using the forecast data from the last 8 quarters for a specific horizon. The smoother is a local mean which uses an Epanechnikov kernel.

Figure G.2: Simulation Results – Fixed Noise Across Horizons

Panel A: Persistence Estimates
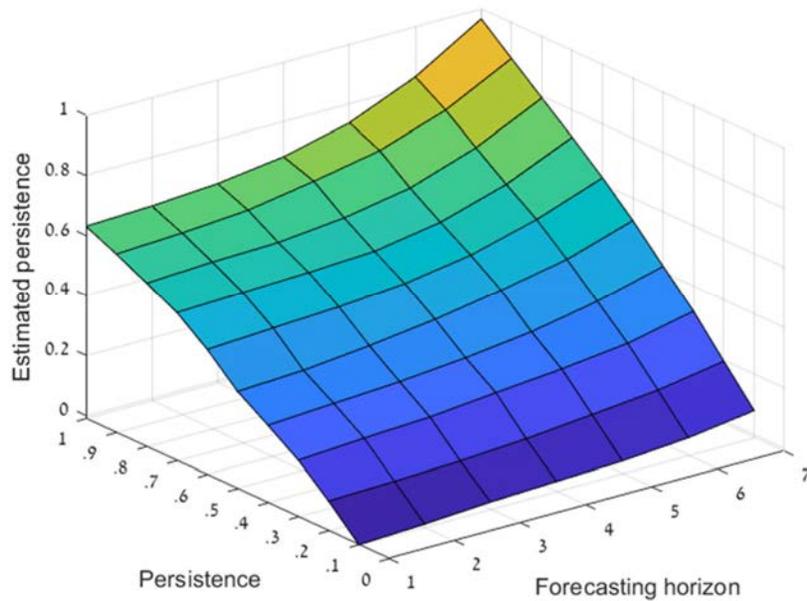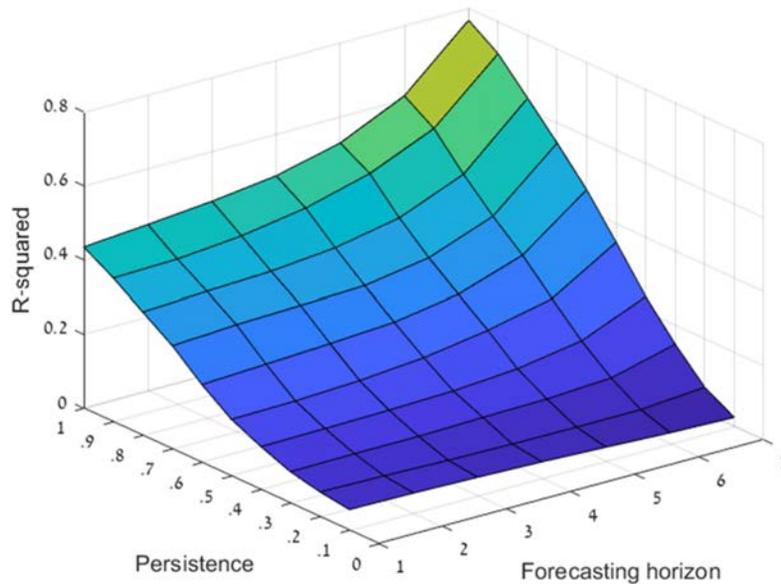
Panel B: R-Squared of Persistence Regressions

Notes: The two panels in the figure show estimation results of specification (4), applied to a simulated data of forecasts, according to the model presented in section 3. Coefficient estimate (panel A) and R-squared (panel B) are averaged across 1000 draws of the simulation. Each simulation applies a different value of persistence in the state process. The variance of the shock in the state process is standardized to one. Regressions were estimated for seven forecast horizons out of eight horizons for which noisy signals are available (two consecutive horizons in each regression). The noise-to-signal ratio across all the horizons is set to 2.

Figure G.3: Forward Information about Major Events

Panel A: COVID-19
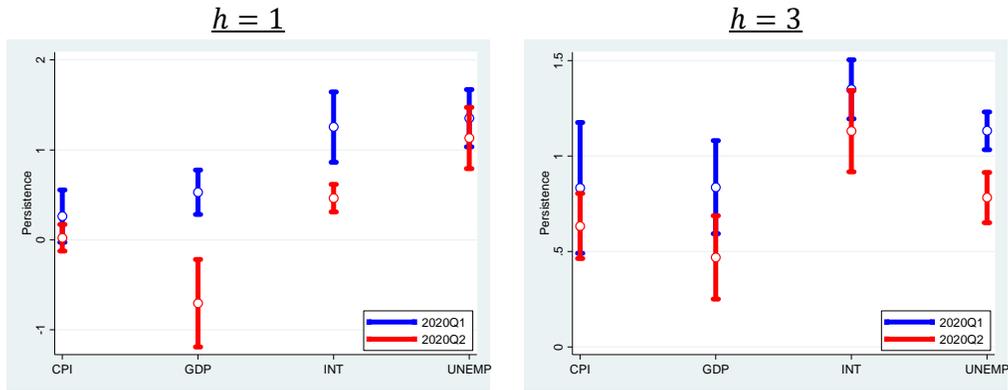
Panel B: 2008 Financial Crisis
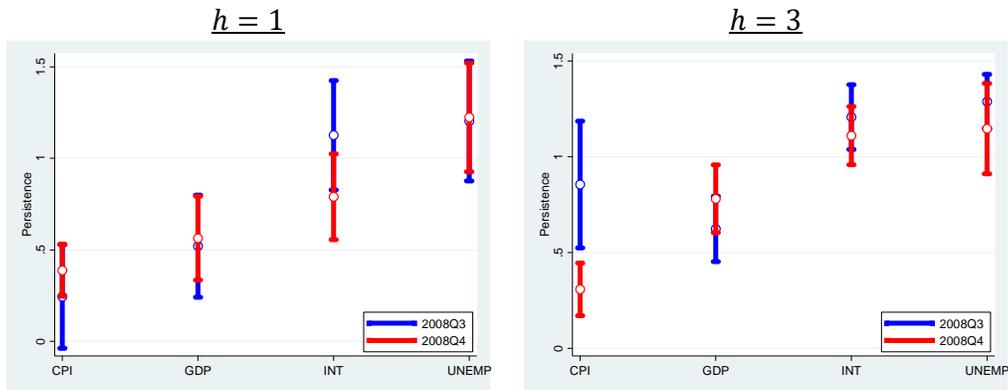
Panel C: 9/11 Attack
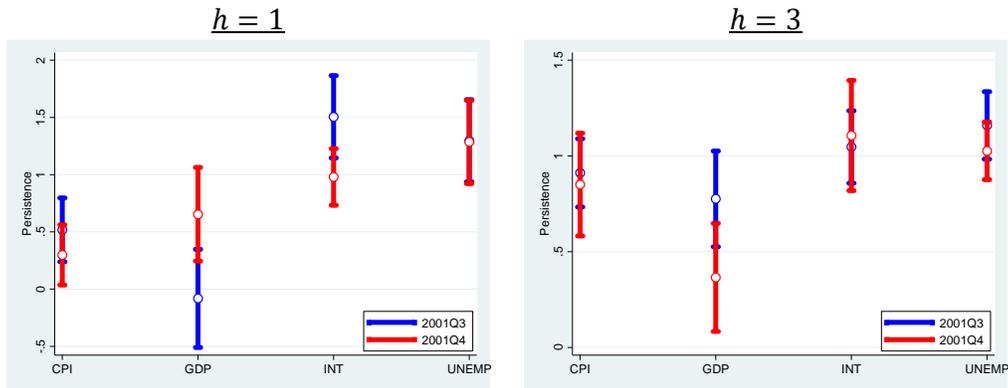
Notes: The figure plots estimates of persistence coefficients for specified quarters around three major events, based on specification (4). The estimation is applied to cross-sections of SPF forecasts of four macroeconomic variables in the specified quarters. Each circle presents a point estimate and whiskers show the 95% confidence interval. Blue and red estimates refer, respectively, to the surveys before and after the event became known. CPI = CPI inflation, GDP = real GDP growth, INT = 3-month Treasury bill interest rate, UNEMP = unemployment rate.

Figure G.4: Impulse Responses to Inflation News

Panel A: Forward Information Method
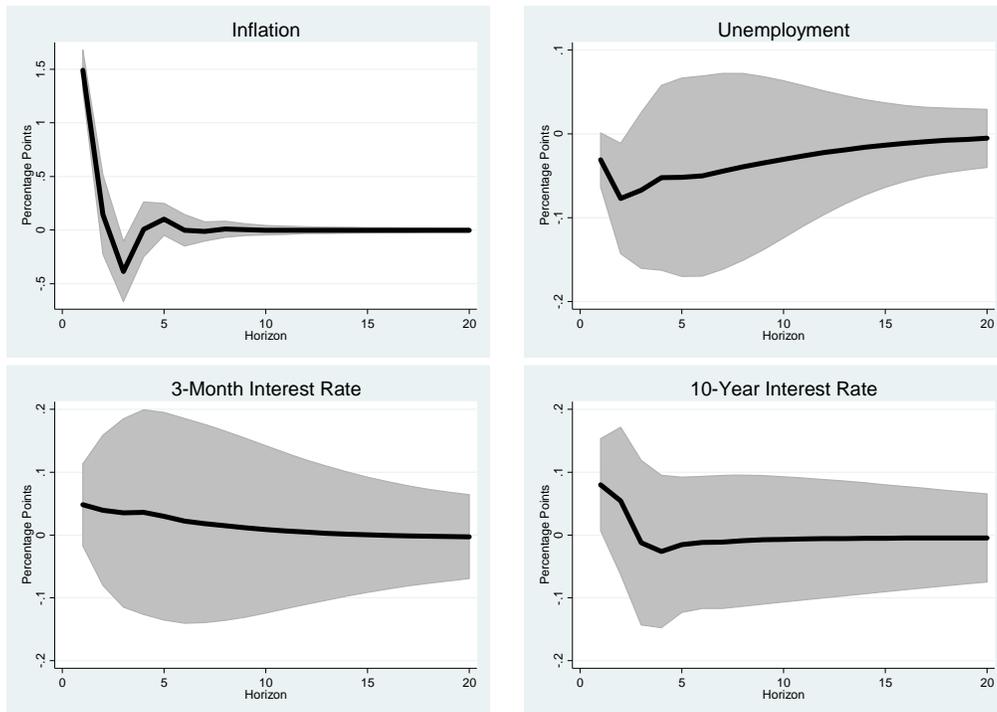
Panel B: Barsky-Sims News-Shock Method
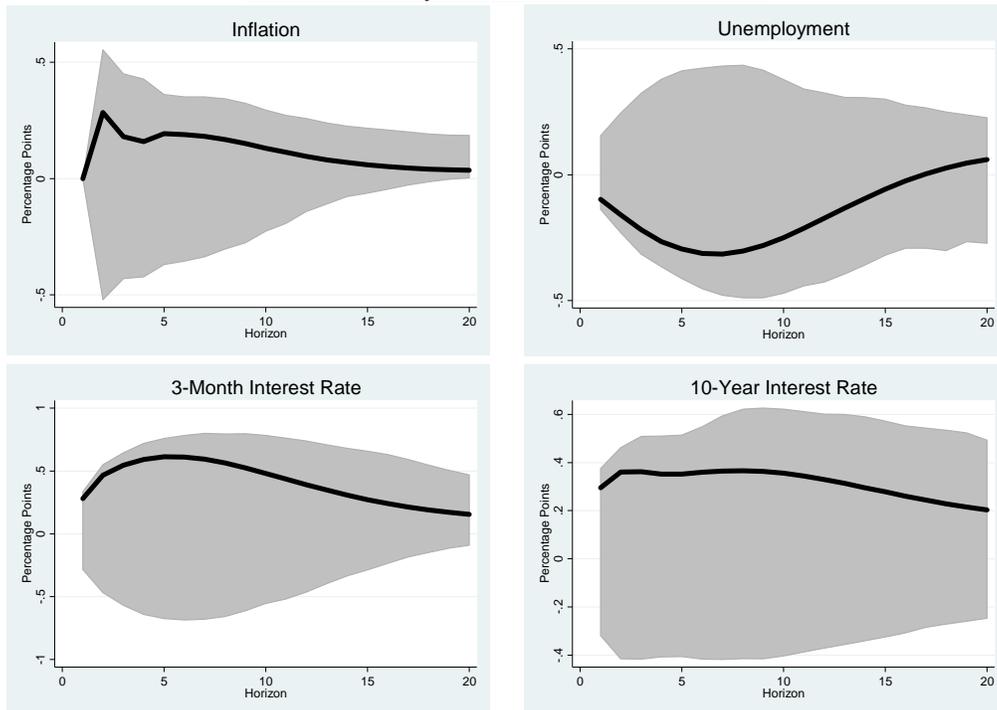
Notes: The figure plots impulse responses to inflation news in a VAR system estimated on the 1983Q1-2017Q4 sample. Inflation news is obtained by the forward information method in panel A, or by the Barsky-Sims news-shock method in panel B. The solid lines are the estimated impulse responses and the shaded areas show the 95% confidence interval based on bootstrap replications.

Figure G.5: Contribution of Forward Information over Time

Notes: The figure plots the ratio between forward information and forecast variation, using $\sqrt{\Sigma_t (FI_{t|t})^2} \Big/ \sqrt{\Sigma_t (x_{t|t})^2}$. The component of forward information $FI_{t|t}$ is evaluated by equation (32). The ratio is computed with a rolling window of 40 obs.

Figure G.6: Persistence Patterns of Additional Macroeconomic Variables

Panel A: Unemployment Rate
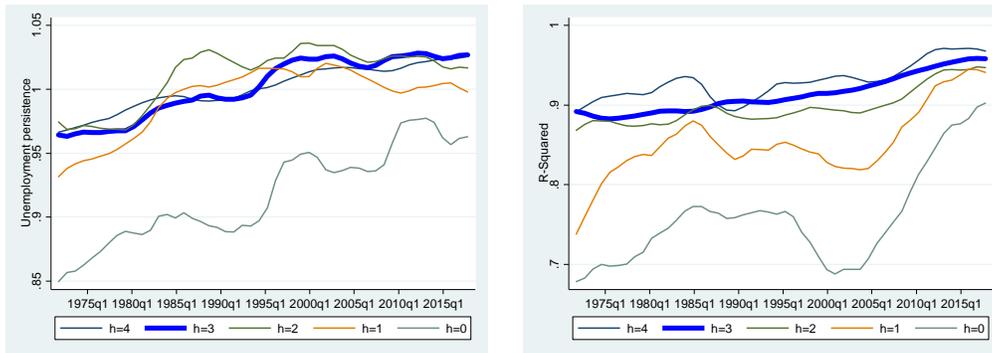
Panel B: Interest Rate
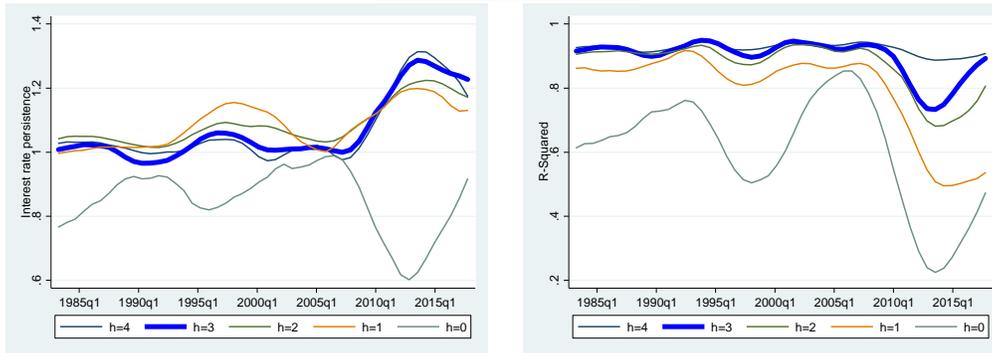
Panel C: Real GDP Growth
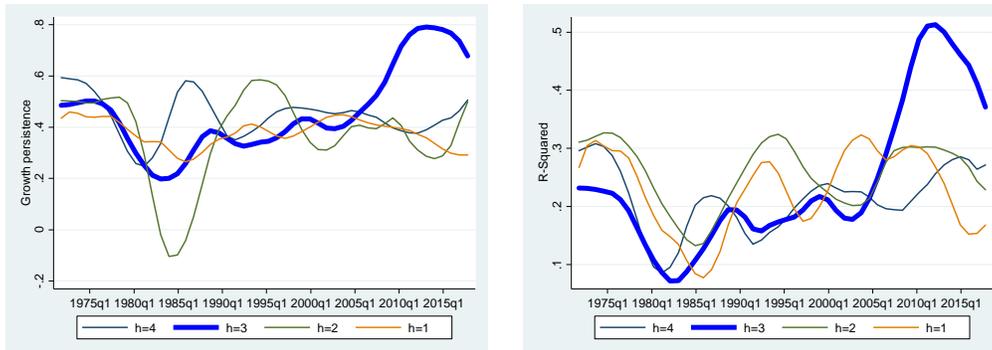
Notes: The figure plots smoothed estimates of persistence $\hat{\rho}$ and R-squared measures based on estimating specification (4) for different forecast horizons, and using different macroeconomic variables from the SPF survey, as specified in each panel. Each quarterly point is based on OLS estimation using the forecast data from the last 8 quarters for a specific horizon. The smoother is a local mean which uses an Epanechnikov kernel.

Figure G.7: Persistence Patterns of Additional Inflation Variables

Panel A: GDP Deflator
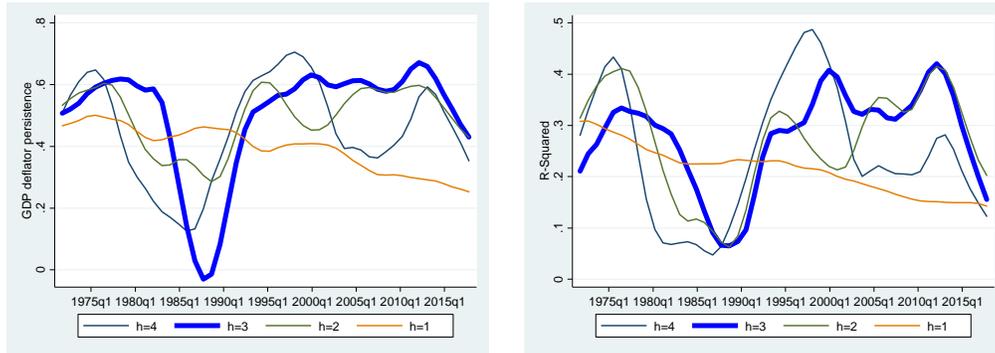
Panel B: PCE Inflation
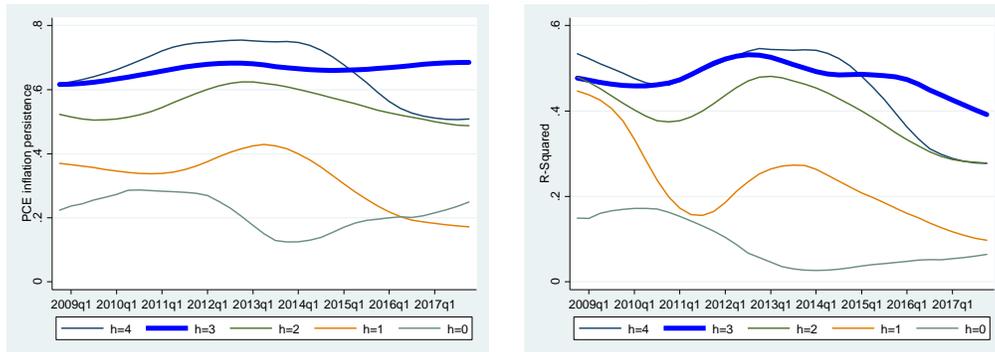
Panel C: Core Inflation
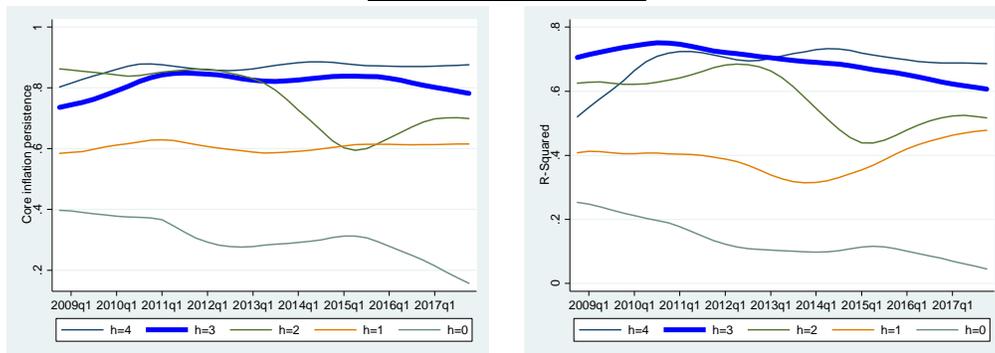
Notes: The figure plots smoothed estimates of persistence $\hat{\rho}$ and R-squared measures based on estimating specification (4) for different forecast horizons, and using forecasts for different inflation measures from the SPF survey, as specified in each panel. Each point on the lines is based on OLS estimation using the forecast data from the last 8 quarters. The smoother is a local mean which uses an Epanechnikov kernel.

Figure G.8: Real GDP and GDP Deflator – Levels and Higher Order

Panel A: Real GDP
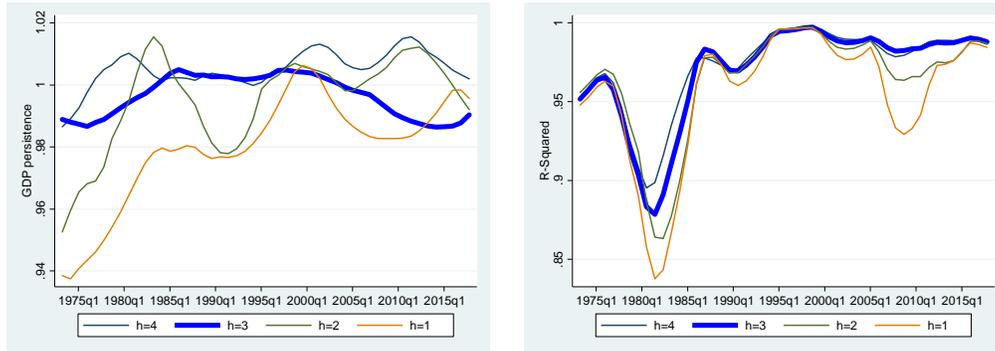
Panel B: GDP Deflator
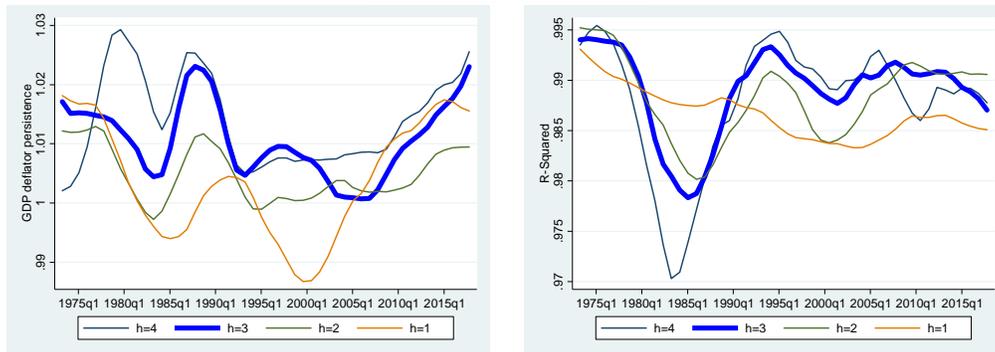
Notes: The figure plots smoothed estimates of persistence $\hat{\rho}$ and R-squared measures for various forecast horizons, using level forecasts of GDP and GDP deflator from the SPF survey. The h-step ahead forecast is regressed on $h-1$ and $h-2$-step ahead forecasts, as in AR(2), and the persistence is estimated as the sum of coefficients. Each point on the lines is based on OLS estimation using the forecast data from the last 8 quarters. The smoother is a local mean which uses an Epanechnikov kernel.

TABLE G.1.

Regressions of the deviation from the mean forecast: Simulation – Fixed noise across horizons

Dependent variable:	$x^i_{t\|t} - x_{t\|t}$	$x^i_{t+1\|t} - x_{t+1\|t}$	$x^i_{t+2\|t} - x_{t+2\|t}$	$x^i_{t+3\|t} - x_{t+4\|t}$	$x^i_{t+4\|t} - x_{t+4\|t}$	$x^i_{t+5\|t} - x_{t+5\|t}$	$x^i_{t+6\|t} - x_{t+6\|t}$	$x^i_{t+7\|t} - x_{t+7\|t}$
$x^i_{t\|t-1} - x_{t\|t-1}$	0.903	-0.010	-0.001	0.000	0.000	0.000	0.000	0.000
	0.903	**-0.011**	**-0.001**	**-0.001**	**0.000**			
$x^i_{t+1\|t-1} - x_{t+1\|t-1}$	-0.010	0.892	-0.013	-0.002	0.000	0.000	0.000	0.000
	-0.010	**0.892**	**-0.013**	**-0.002**	**-0.001**			
$x^i_{t+2\|t-1} - x_{t+2\|t-1}$	-0.001	-0.013	0.879	-0.016	-0.003	0.000	0.000	0.000
	-0.001	**-0.013**	**0.879**	**-0.017**	**-0.003**			
$x^i_{t+3\|t-1} - x_{t+3\|t-1}$	0.000	-0.002	-0.016	0.862	-0.022	-0.004	-0.001	0.000
	0.000	**-0.002**	**-0.016**	**0.862**	**-0.025**			
$x^i_{t+4\|t-1} - x_{t+4\|t-1}$	0.000	0.000	-0.003	-0.022	0.838	-0.031	-0.007	-0.002
	0.000	**-0.001**	**-0.005**	**-0.022**	**0.824**			
$x^i_{t+5\|t-1} - x_{t+5\|t-1}$	0.000	0.000	0.000	-0.004	-0.031	0.804	-0.047	-0.016
$x^i_{t+6\|t-1} - x_{t+6\|t-1}$	0.000	0.000	0.000	-0.001	-0.007	-0.047	0.749	-0.084
$x^i_{t+7\|t-1} - x_{t+7\|t-1}$	0.000	0.000	0.000	0.000	-0.002	-0.016	-0.084	0.639

Notes: The table reports true parameters and simulated coefficient estimates from regressions of the individual deviation from the mean forecast, based on specification (26). Each column reports in bold the simulated regression results for a specified dependent variable, with truncation at horizon $h = 4$. The true parameters in each column correspond to separate rows in the matrix $(I - G)$ (right column corresponds to the first row and so on). Forecasts are simulated with $\rho = 0.5$ and noise-to-signal ratio of 2 for all signals.

TABLE G.2.

Regressions of the deviation from the mean forecast: Simulation

Dependent variable:	$x^i_{t\|t} - x_{t\|t}$	$x^i_{t+1\|t} - x_{t+1\|t}$	$x^i_{t+2\|t} - x_{t+2\|t}$	$x^i_{t+3\|t} - x_{t+4\|t}$	$x^i_{t+4\|t} - x_{t+4\|t}$	$x^i_{t+5\|t} - x_{t+5\|t}$	$x^i_{t+6\|t} - x_{t+6\|t}$	$x^i_{t+7\|t} - x_{t+7\|t}$
$x^i_{t\|t-1} - x_{t\|t-1}$	0.513	-0.046	-0.007	-0.002	0.000	0.000	0.000	0.000
	0.512	**-0.045**	**-0.008**	**-0.002**	**-0.001**			
$x^i_{t+1\|t-1} - x_{t+1\|t-1}$	-0.018	0.618	-0.059	-0.013	-0.004	-0.002	-0.001	0.000
	-0.020	**0.618**	**-0.062**	**-0.016**	**-0.006**			
$x^i_{t+2\|t-1} - x_{t+2\|t-1}$	-0.001	-0.030	0.686	-0.072	-0.022	-0.009	-0.004	-0.002
	0.000	**-0.028**	**0.689**	**-0.070**	**-0.020**			
$x^i_{t+3\|t-1} - x_{t+3\|t-1}$	0.000	-0.003	-0.036	0.764	-0.072	-0.028	-0.014	-0.007
	-0.001	**-0.003**	**-0.038**	**0.764**	**-0.071**			
$x^i_{t+4\|t-1} - x_{t+4\|t-1}$	0.000	-0.001	-0.007	-0.048	0.783	-0.086	-0.043	-0.021
	0.001	**-0.002**	**-0.009**	**-0.055**	**0.751**			
$x^i_{t+5\|t-1} - x_{t+5\|t-1}$	0.000	0.000	-0.002	-0.014	-0.064	0.776	-0.111	-0.055
$x^i_{t+6\|t-1} - x_{t+6\|t-1}$	0.000	0.000	0.000	0.000	-0.001	-0.004	0.988	-0.006
$x^i_{t+6\|t-1} - x_{t+6\|t-1}$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000

Notes: The table reports true parameters and simulated coefficient estimates from regressions of the individual deviation from the mean forecast, based on specification (26). Each column reports in bold the simulated regression results for a specified dependent variable, with truncation at horizon $h = 4$. The true parameters in each column correspond to separate rows in the matrix $(I - G)$ (right column corresponds to the first row and so on). Forecasts are simulated according to the details in Section 3, with $\rho = 0.5$.

TABLE G.3.

Persistence across forecast horizons in time-series regression: SPF consensus and Greenbook

Forecast horizon:	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
Panel A: SPF consensus					
CPI Inflation	0.395***	0.606***	0.942***	1.002***	0.999***
	(0.060)	(0.081)	(0.034)	(0.019)	(0.016)
GDP Inflation		0.962***	1.003***	1.009***	0.957***
		(0.180)	(0.031)	(0.026)	(0.027)
PCE Inflation	0.299***	0.299***	0.509***	0.572***	0.678***
	(0.052)	(0.043)	(0.055)	(0.055)	(0.052)
Core CPI Inflation	0.454***	0.785***	0.868***	0.895***	0.866***
	(0.046)	(0.056)	(0.014)	(0.017)	(0.045)
Interest Rate	0.986***	0.989***	0.989***	0.982***	0.978***
	(0.013)	(0.006)	(0.007)	(0.008)	(0.008)
Unemployment Rate	0.971***	0.974***	0.970***	0.965***	0.960***
	(0.021)	(0.014)	(0.010)	(0.006)	(0.004)
GDP Growth		0598***	0.571***	0.639***	0.790***
		(0.030)	(0.042)	(0.056)	(0.069)
Industrial Production		0520***	0.628***	0.693***	0.777***
		(0.032)	(0.072)	(0.046)	(0.078)
Panel B: Greenbook					
CPI Inflation	0.426***	0.366***	0.870***	0.995***	1.019***
	(0.061)	(0.079)	(0.055)	(0.025)	(0.014)
GDP Inflation	0.555***	0.637***	0.898***	0.956***	0.975***
	(0.068)	(0.077)	(0.042)	(0.037)	(0.034)
PCE Inflation	0.361***	0.183***	0.323***	0.842***	0.944***
	(0.082)	(0.060)	(0.127)	(0.063)	(0.064)
Core CPI Inflation	0.588***	0.727***	0.983***	0.967***	0.953***
	(0.071)	(0.092)	(0.035)	(0.023)	(0.022)
Unemployment Rate	0.982***	0.971***	0.968***	0.957***	0.949***
	(0.026)	(0.012)	(0.007)	(0.007)	(0.006)
GDP Growth	0.528***	0524***	0.598***	0.717***	0.808***
	(0.071)	(0.052)	(0.071)	(0.096)	(0.112)
Industrial Production	0.553***	0346***	0.491***	0.701***	0.739***
	(0.056)	(0.036)	(0.077)	(0.130)	(0.105)

Notes: The table reports coefficient estimates from regressions of mean forecasts $x_{t+h|t}$ on $x_{t+h-1|t}$. Each entry reports an estimate from a different regression. The variable x and forecast horizon h that were applied in each regression are specified in the headers of the table's rows. Panels A and B use forecasts from SPF and Greenbook, respectively. The sample period is 1983Q1-2015Q4, except for PCE and core CPI inflation, for which the sample begins in 2007Q1. Newey-West standard errors are in parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 levels.

TABLE G.4.

Share of inflation volatility explained by news

Forecast Horizon	Forward information method	Barsky-Sims method
1	54.3	2.5
4	53.5	5.2
8	51.8	8.4
12	51.2	9.6
20	50.6	10.0

Notes: The table reports the share of inflation volatility explained by news from a variance decomposition applied to a VAR system estimated on the 1983Q1-2017Q4 sample. Inflation news is obtained by the forward information method, or by the Barsky-Sims news-shock method, as indicated in the columns' headers.

TABLE G.5.

Regressions of the deviation from the mean backcast

	Unemployment	Interest Rate	PCE Inflation	Core Inflation
Dependent variable: $x^i_{t|t} - x_{t|t}$ (backcasts)				
$x^i_{t|t-1} - x_{t|t-1}$	0.027***	0.018***	-0.000	0.001
	(0.006)	(0.003)	(0.000)	(0.001)
$x^i_{t+1|t-1} - x_{t+1|t-1}$	-0.012*	-0.013***	-0.000	-0.001
	(0.007)	(0.003)	(0.000)	(0.001)
$x^i_{t+2|t-1} - x_{t+2|t-1}$	0.011*	0.012***	0.000	0.000
	(0.007)	(0.003)	(0.000)	(0.001)
$x^i_{t+3|t-1} - x_{t+3|t-1}$	-0.009	-0.010***	0.000	-0.001
	(0.007)	(0.004)	(0.000)	(0.001)
$x^i_{t+4|t-1} - x_{t+4|t-1}$	0.003	-0.001	-0.000	0.001
	(0.004)	(0.002)	(0.000)	(0.001)
Constant	0.000	-0.000	-0.000	-0.000
	(0.001)	(0.000)	(0.000)	(0.000)
Obs.	5,476	3,825	1,288	1,394
R^2	0.006	0.019	0.001	0.005

Notes: The table reports coefficient estimates from regressions of the individual deviation from the mean backcast, based on specification (26) with $h = 0$. Each column reports results, using SPF forecasts for a specified macroeconomic variable. Driscoll-Kraay standard errors are in parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 levels.

TABLE G.6.

Forward noisy information vs. the standard model: SPF forecasts of additional variables

Dependent variable	Forward information	Standard framework	Forward information	Standard framework
	Panel A: Unemployment Rate		**Panel D: GDP Deflator**	
$x^i_{t+1\|t} - x_{t+1\|t}$	-6962	-7247	16299	17522
$x^i_{t+2\|t} - x_{t+2\|t}$	-3385	-3343	15099	16437
$x^i_{t+3\|t} - x_{t+3\|t}$	-1085	-835	15626	16796
$x^i_{t+4\|t} - x_{t+4\|t}$	73	117	16489	16739
	Panel B: Interest Rate		**Panel E: PCE Inflation**	
$x^i_{t+1\|t} - x_{t+1\|t}$	-1568	-1694	3129	3129
$x^i_{t+2\|t} - x_{t+2\|t}$	1079	1109	2267	2279
$x^i_{t+3\|t} - x_{t+3\|t}$	2277	2313	1557	1577
$x^i_{t+4\|t} - x_{t+4\|t}$	3164	3155	1453	1439
	Panel C: Real GDP Growth		**Panel F: Core Inflation**	
$x^i_{t+1\|t} - x_{t+1\|t}$	19336	20871	1247	1273
$x^i_{t+2\|t} - x_{t+2\|t}$	18891	20473	1058	1074
$x^i_{t+3\|t} - x_{t+3\|t}$	19047	20544	1043	1095
$x^i_{t+4\|t} - x_{t+4\|t}$	20541	20654	818	827

Notes: The table reports BIC statistics associated with specifications (26) and (27), for the forward information and the standard noisy information framework, respectively. Each panel presents results for a specified macroeconomic variable, using the SPF forecasts. The specifications were estimated for several forecast horizons as indicated in the first column ($h = 0,1,2,3$).

TABLE G.7.

Regressions of forward information: Additional variables

Dependent Variable (x_t):	Unemployment		Interest Rate		Real GDP Growth	
	(1)	(2)	(3)	(4)	(5)	(6)
$FI_{t|t}$		1.110***		1.144***		1.146***
		(0.084)		(0.064)		(0.133)
x_{t-1}	1.626***	1.095***	1.507***	0.972***	0.313***	0.259***
	(0.090)	(0.087)	(0.094)	(0.067)	(0.074)	(0.057)
x_{t-2}	-0.641***	-0.109	-0.479***	0.084	0.108	0.062
	(0.183)	(0.105)	(0.145)	(0.094)	(0.088)	(0.076)
x_{t-3}	-0.008	0.013	-0.033	-0.165***	0.027	0.011
	(0.168)	(0.082)	(0.141)	(0.046)	(0.078)	(0.069)
x_{t-4}	-0.017	-0.044	-0.019	0.092**	0.008	0.007
	(0.065)	(0.046)	(0.090)	(0.038)	(0.090)	(0.070)
Constant	0.256***	0.186***	0.066**	0.116***	1.546***	2.402***
	(0.078)	(0.053)	(0.033)	(0.032)	(0.379)	(0.412)
R^2	0.975	0.989	0.985	0.996	0.142	0.442

Notes: The table reports coefficient estimates from various regressions. The variable x_t is indicated in the columns' headers. The forward information variable is computed according to (32). The sample period is 1972Q1-2017Q4 for unemployment and GDP growth, and 1983Q3-2017Q4 for the interest rate. Newey-West standard errors are in parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 levels.

TABLE G.8.

Persistence in ECB and US annual forecasts across calendar quarters (1999Q1-2017Q4)

	Whole Sample	Q1	Q2	Q3	Q4
Panel A: $x_{2Y|t}^i = c + \rho x_{1Y|t}^i + u_t$ (ECB SPF)					
Constant	0.956***	0.940***	1.000***	0.927***	0.959***
	(0.047)	(0.060)	(0.067)	(0.037)	(0.086)
$F_t^i x_{1Y}$	0.483***	0.488***	0.457***	0.502***	0.484***
	(0.027)	(0.033)	(0.040)	(0.023)	(0.049)
Obs.	3378	946	836	752	844
R^2	0.439	0.438	0.406	0.507	0.405
Panel B: $x_{2C|t}^i = c + \rho x_{1C|t}^i + u_t$ (ECB SPF)					
Constant	1.031***	0.984***	1.087***	0.990***	1.013***
	(0.039)	(0.068)	(0.029)	(0.058)	(0.063)
$F_t^i x_{1C}$	0.379***	0.429***	0.355***	0.397***	0.370***
	(0.023)	(0.035)	(0.026)	(0.040)	(0.034)
Obs.	4214	1067	1057	991	1099
R^2	0.560	0.492	0.509	0.631	0.600
Panel C: $x_{2C|t}^i = c + \rho x_{1C|t}^i + u_t$ (US SPF)					
Constant	1.617***	1.391***	1.521***	1.673***	1.689***
	(0.091)	(0.179)	(0.114)	(0.103)	(0.122)
$F_t^i x_{1C}$	0.269***	0.429***	0.338***	0.236***	0.193***
	(0.041)	(0.082)	(0.043)	(0.041)	(0.042)
Obs.	2753	667	686	656	744
R^2	0.183	0.361	0.280	0.161	0.091

Notes: The table reports coefficient estimates for the specified equations at the top of each panel. $x_{1Y|t}^i$ and $x_{2Y|t}^i$ are rolling-year forecasts for one and two years ahead, respectively. $x_{1C|t}^i$ and $x_{2C|t}^i$ are calendar-year forecasts for the current and the next calendar years. Regressions were estimated for the period of 1999Q1-2017Q4 and separately for each of the calendar quarters in this period as indicated by columns' headers. Standard errors of Driscoll and Kraay (1998) are in parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 levels.