Macroeconomic Persistence and Forward Information: What Else Can We Learn from Survey Expectations?

Nathan Goldstein
Bar-Ilan University

Yuiry Gorodnichenko
UC Berkley and NBER

December 2019

- Preliminary Draft -

Abstract: This study proposes two novel ideas about the expectations formation process which are found to be strongly related. Following forecasting practices, expectations could be decomposed to two components: A standard prediction based on the underlying process and an adjustment of that prediction to account for the availability of forward information. Based on the standard component, our first idea is that the underlying process could be recovered from forecast data. We apply this idea to the estimation of inflation persistence, based on the cross-sectional variation in multi-horizon inflation forecasts provided in surveys. Our method enables to track inflation persistence quarter-by-quarter, by utilizing recent responses of survey participants. Applying our measure to the Survey of Professional Forecasters, we find a gradual decline in inflation persistence over the last decades from the maximum degree of 1 to about 0.5. Our findings offer a compromise in the debate about whether inflation persistence in the US dropped dramatically or didn't change at all. However, we present several empirical patterns, related to our persistence measure, which are not consistent with prominent models of expectations formation and therefore introduce the forward information component. In contrast to recent models of expectations, we suggest that forecast disagreement and predictable errors could arise due to an advantage rather than a deficiency in expectations. This advantage is due to informative noisy signals about the future. Forecasters combine model-based predictions and multi-horizon forward signals. We show this framework could account for all the puzzling patterns found in the SPF data, and present direct evidence for the presence of forward information. We further decompose expectations to the standard and forward components and show how the forward information component could be used to quantify news shocks, with significant impact on macroeconomic fluctuations.

Keywords: Expectations formation, forward information, inflation persistence, news shocks.

JEL codes: E3;E4;E5.

1. Introduction

Since the Rational Expectations Revolution, the expectations formation process has been in an ever-growing debate. This issue has become especially important recently, when new macroeconomic policies, such as forward guidance, are designed to affect the economy directly through the expectations channel. During the last years, amounting evidence for deviations from rational expectations, based on survey forecasts, has led to different proposals to model expectations. An important element in the growing variety of models is the introduction of some deficiency in the expectations formation process, which could account for deficiencies and great dispersion in forecast data. One leading approach proposes deficiencies in information acquisition, such as in the sticky information (Mankiw and Reis, 2002), noisy information (Woodford, 2002) and rational inattention (Sims, 2003, Mackowiack and Weiderhot, 2009) models (an earlier model of imperfect information is in Lucas, 1972). Alternative approaches stress behavioral deficiencies (e.g. asymmetric loss-aversion in Capistran and Timmermann, 2010, natural expectations in Fuster et al., 2010, diagnostic expectations in Bordalo et al., 2018, behavioral inattention in Gabaix, 2019). The approach of learning suggests a deficiency in modelling the economy (Evans and Honkapohja, 2012).

In this paper, we suggest a new approach which deviates from current expectations models, by ruling out those deficiencies. Our idea is that people have different beliefs about the future, not because they disagree about the *past*, but just because they disagree about the *future*. They receive *forward* information at various horizons about the fundamental, but naturally, this information is imperfect and dispersed. Thus, information varies in two dimensions: across people and across horizons in the future. Despite this forward-looking nature of expectations, deficiencies in forecasts as observed in the data could still be addressed. For example, predictable forecast errors could arise due to gradual improvement in information about the *future* over time, and not because of information limitations with respect to realized events.

The approach we propose is motivated by the way which is often used to describe forecasting in practice. For example, in a special survey conducted among participants in the well-studied Survey of Professional Forecasters, ran by the Federal Reserve Bank of Philadelphia, a decisive majority (80%) has described their forecasting method as a combination of "model" and "subjective adjustment" (see a summary by Stark, 2013)[1]. Current approaches for modelling expectations concentrate only on the first component: People apply their underlying "model" for the fundamental, conditional on the deficiency by which their expectations formation is characterized. Our approach, instead, considers the two components together: People apply their underlying "model" for the fundamental, but then adjust the model prediction to account for new information they obtain about the future. As an example, a forecaster could predict inflation based on a model she uses, but then would adjust her prediction due to a new forward guidance announcement, representing a

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

piece of information about future inflation, which is beyond the model. Surprisingly, although this component of forward information is constructive by reducing the forecast error, it is also responsible for dispersion in forecasts and predictable errors. We use our approach to decompose survey forecasts to the two components of modelling the fundamental and the forward information adjustment. We then quantify the role of forward information.

The notion of forward information is also driven by new survey evidence we provide, which seems to be at odds with current approaches of modelling expectations, and point to the presence of this form of information in the reported forecasts. Interestingly, this evidence is derived indirectly due to additional idea we propose, which is actually related to the first, more standard component of forecasting– the modelling of the fundamental.

The idea from which we start our analysis is that this model, applied by the forecasters could be recovered from the forecasts they deliver. Moreover, while the underlying process is usually estimated with limited data of macroeconomic time-series, expectations data would provide an additional dimension of data, due to the *cross-sectional* dispersion of survey forecasts. Thus, we could estimate the underlying process quarter-by-quarter with cross-sections of forecasts and track changes in parameters over time, which are typical and quite frequent in many macroeconomic relationships. Of course, there is always the question of how much accurate is the model applied by the forecasters. However, if the final forecasts produced by this model can beat predictions from well-estimated econometric models, like inflation forecasts from the SPF survey (Ang et al., 2007), then the model recovered from these survey forecasts could be reliable as well.

We apply this idea to the estimation of a key macroeconomic parameter known as inflation persistence. We focus on this single parameter due to the simplicity of estimation on one hand, and the important debate about the great decline in inflation persistence over the years, on the other hand. The degree of inflation persistence should reflect the extent to which current shock to inflation would affect future inflation. A simple standard measure can be provided by an auto-regressive coefficient in an AR(1) (Fuhrer, 2011). Similar to former studies, using US inflation data, we obtain in a rolling estimation a decline from the maximum degree of 1 to the zero level in the decades following the Volcker disinflation. Our proposed method is to estimate inflation persistence quarter-by-quarter using cross-sections of survey forecasts of inflation. Specifically, we mimic the AR(1) process by estimating a regression of a quarterly inflation forecast on the forecast made for the previous quarter, by the same forecaster in the same survey. Interestingly, when applying this estimation to CPI inflation forecasts from the SPF survey, we find a sizable but more moderate decline in persistence from 1 to about 0.5, since the beginning of the 1980s.

Strikingly, when using this new measure of persistence to predict inflation out-of-sample over the period, we get more accurate predictions comparing to regular AR(1) forecasts. This illustrate the potential in using richness of cross-sectional data provided by survey forecasts. We further interpret our findings as a compromise in the debate about inflation persistence, found in the recent literature. Some studies rule out a decline in inflation persistence, suggesting that the decline in auto-regressive coefficients is driven by changes in the volatilities of the two components of inflation: the long run trend component which always follows a unit root process and the transitory component which always has no persistence. We show that our expectations-based measure of persistence is not sensitive to these variations in the volatilities and that even when considering this inflation decomposition in our forecast-based specification, we still document a sizable persistence for the cycle component of inflation. This finding has important implications for monetary policy, as it gives more room for a better control over inflation by a further reduction of its persistence, even after the success in decreasing trend inflation since the Volcker disinflation.

Our approach of recovering the inflation process from expectations data is in line with a variety of recent models of expectations, which assume that forecasters follow the underlying process when forming their (deficient) expectations. At the same time, there is a need to understand what determines deviations from this process, which are reflected by a sizable error component in the cross-sectional regressions we estimate to fit the AR(1) to the relation between forecasts for consecutive quarters. We show that including more "lags" of inflation forecasts for previous quarters or forecasts for other macro-variables, as in a VAR process, could not account for this lack of fit. Moreover, we find an interesting variation in the persistence estimates across different forecasting horizons. Our baseline results were based on fitting the AR(1) to quarterly forecasts from the SPF with horizon around a year ahead. When estimating similar regressions using forecasts with shorter horizon, like running the forecast for the next quarter on the forecast for the current quarter, we find much lower coefficient estimates, even in the 1980s, when persistence should be high. In fact, there is a pattern of convergence in the coefficient estimates when moving to the longer horizons, suggesting that the estimate of persistence is biased downward at shorter horizons, but the bias diminishes for longer horizons. A similar pattern is documented for the fit of the cross-sectional regressions, where the R-squared statistic is increasing in the forecast horizon and seems to converge. There is also an interesting co-movement over time of the R-squared and the estimated persistence. We demonstrate that these empirical patterns cannot be explained by recent models of expectations, suggested in the literature.

Motivated by these empirical patterns, we introduce our new approach of forward information. Intuitively, when forecasters adjust their model-based predictions to account for new signals about the future, it would lead to a bias in the model relationship that is estimated by forecast data. However, for longer horizons, for which forward signals are no longer informative, the forecasts would only be based on the model relation without a biasing adjustment. This explains the convergence pattern across horizons in the estimates of inflation persistence, and further indicate on the future horizon for which SPF forecasters are applying useful forward information, which is until about a year ahead.

We present a formal model of expectations formation with forward information, in which agents receive multiple idiosyncratic signals referring to several horizons in the future. They exploit these informative signals and optimally filter the noises when forming forecasts for multiple periods ahead. We show how the optimal forecast could be decomposed to the standard prediction component, based on the underlying process and the deviation from the state process, or the adjustment, due to the exploitation of forward signals and the smoothing process involved in that. We then use simulations to demonstrate that this framework of expectations formation can generate all the empirical patterns we found in the SPF data. The theoretical mechanism behind these patterns is further explained using a tractable example with few forward signals.

Another advantage of our framework is that it nests the standard noisy information framework, by allowing a noisy signal about the *realized* fundamental, in addition to the forward signals. Our generalized framework differs from the standard model not only by introducing the notion of forward information, but also by recognizing that information could vary not only across agents, but also across horizons in future. Thus, our framework could account for the realistic mechanism of improvement in the signals over time, when getting closer to the target period. In addition, the generalized framework allows us to directly test the restricted form implied by the standard noisy information model against the unrestricted form which includes forward information. We propose a simple specification to perform this test and estimate weight placed on forward information. Our results, using the same inflation forecasts from the SPF survey, do not only point out to a significant role of forward information, but also rules out any significant amount of noisy information in the sense of the standard model. These findings, coupled with our initial indirect evidence, support the view that SPF forecasts are not driven by dispersed information about the future, rather than about the past.

We then examine SPF forecasts for additional key macroeconomic variables. We apply our method for estimating underlying persistence from forecast data, and document empirical patterns which are similar to those obtained for inflation forecasts, indicating on a broader use of forward information by professional forecasters. The measured persistence is much more stable for the other variables relative to inflation. We also find that persistence in core inflation is considerably higher comparing to regular inflation and remained high in recent years as well. We also apply inflation forecast data from the European Central Bank SPF. Although this survey provides only annual forecasts and does not include quarterly forecasts with shorter horizons, we are able to bring some evidence for forward information utilization in this survey as well, by comparing rolling-year forecasts with fixed horizon to calendar-year forecasts, for which the horizon changes over the year. The estimated inflation persistence in recent years is sizable, between 0.4 to 0.5, and close to our estimates for the US.

The study ends up with a demonstration of another important application of the forward information concept. We propose to quantify the forward information component in survey forecasts and to use it as a measure of "news shock". A prominent strand of literature emphasizes the role of news about future shocks in explaining current business cycle fluctuations through the expectations channel (see a recent review by Beaudry and Portier, 2014). However, quantifying news shocks is not a simple task. News are usually identified by estimating a structural model or a VAR system. Our notion of forward information is closely related to the idea of news shocks. Moreover, it offers a straightforward way to extract news directly from expectations data. We consider again the SPF inflation forecasts and quantify inflation news as the difference between the reported forecast and the prediction based on the underlying process, where we use the AR(1) process which was quantified from the same forecast data in the previous sections. We then show that our measure of news have strong and highly significant impact on inflation in a simple VAR. For comparison, we also identify inflation news with the structural VAR method of Barsky and Sims (2011). By contrast, the impulse response of inflation to news obtained with this method is not significant, therefore demonstrating the potential of our approach for quantifying meaningful measures of news directly from expectations data in a systematic way.

Our study is related to several strands of literature. It is strongly related to the vast and growing literature about the expectations formation process, including all the variety of recent models mentioned above, from which we deviate (see a review in Coibion, Gorodnichenko and Kamdar, 2018). Survey forecasts are widely used by many studies in order to test and quantify these models (e.g. Andrade et al., 2013, Coibion and Gorodnichenko, 2012, 2015, Coibion, Gorodnichenko and Kumar, 2018, Fuhrer, 2018, Goldstein, 2019). Our forward information approach calls for a reconsideration of previous evidence from surveys. For example, the significant relationship between ex-post forecast errors and ex-ante forecast revisions in Coibion and Gorodnichenko (2015) could be driven by dispersed forward information, rather than models of information rigidities, which focus on dispersed information about the past.

Another related literature studies the inflation process and its persistence. The decline in inflation persistence over the last decades and its determinants are debated in several papers including Cogley and Sargent (2001), Pivetta and Reis (2007), Benati (2008), Cogley and Sbordone (2008)), Gorodnichenko (2010), Cogley, Primiceri and Sargent (2010) and a review by Fuhrer (2011). In other papers inflation persistence is estimated for the aim of modelling and forecasting inflation, such as in Stock and Watson (2007, 2016), Faust and Wright (2013), Chan, Clark and Koop (2018). Jain (2019) and Ryngeart (2017) estimate inflation persistence using survey forecasts which is close to the approach proposed in this paper. However, our approach emphasizes the advantage of using *cross-sectional* data to obtain quarterly estimates of persistence and to track changes over time. They also didn't document the important variation across horizons and the other empirical patterns shown in this paper, which motivates our new approach of forward information.

Our paper is also related to the recent literature which studies the role of news in the business cycle, including Beaudry and Portier (2006, 2014), Jaimovich and Rebelo (2009), Barski and Sims (2011, 2012), Schmitt Grohé and Uribe(2012), Angeletos and La'O (2013), Blanchard et al. (2013), Benhabib et al. (2015), Chahrour and Durado (2018). News are modeled as advanced information about the future fundamental, which is close to our notion of forward information. Here we enrich the idea of news by a more detailed information structure highlighting heterogeneity across both agents and future horizons, and applying it directly to the expectations formation process. Our findings could be used to quantify forward information systematically using forecast data and to study in more detail its important role in macroeconomic dynamics, as demonstrated with the inflation news we extract from the SPF inflation forecasts.

The remainder of the paper is organized as follows. Section 2 presents the expectation-based measure of inflation persistence and describes the results for the US inflation persistence over the last decades. Section 3 discusses the results and their contribution to the ongoing debate in the literature. Section 4 presents the puzzling empirical patterns obtained with our method and discusses the relation to prominent models of expectations formation. In Section 5, we present our model of forward information, resolve the puzzling evidence, and provide further evidence from a direct estimation of the model. Section 6 examines persistence in other key macroeconomic variables, based on the forecasts from the US SPF. Section 7 brings evidence from the ECB SPF. Section 8 demonstrates how the forward information framework could be used to quantify news shocks and estimate their impact on macroeconomic fluctuations. Section 10 concludes.

2. Expectations-based measure of inflation persistence

The simplest form of measuring inflation persistence is to estimate coefficients on lagged inflation from an auto-regression (Fuhrer, 2010). For instance, in an AR(1) specification, the coefficient on lagged inflation represents the extent to which shocks to inflation will have an effect on future inflation. Two extreme cases are when the AR coefficient is 0 or 1. In the former case, there is no persistence in inflation, since current inflation is only influenced by the current shock, while the effect of past shocks dies out completely. In the latter case, which is the unit root case, the effect of previous shocks is permanent so that inflation is fully persistent. More generally, the higher is the AR coefficient, the higher would be the degree of persistence in inflation.

Standard techniques such as a rolling window estimation could be used to assess changes in inflation persistence over time. The middle pair of figures in figure 1 describes the results from such an exercise. We use quarterly US data of CPI inflation and estimate AR(1) and AR(4) specifications, with a rolling window of 40 quarters. For the AR(1) specification, the left-hand figure shows coefficient estimates over time based on inflation data from the last 40 quarters (black line, with 95% confidence interval in the shaded area). For the AR(4) specification, the right-hand figure shows the estimated sum of auto-regressive coefficients as an approximate measure of persistence. The thick red lines are non-parametric smoothers which illustrates the long-run trend. For both specifications there is a similar declining trend in inflation persistence since the beginning of the 1980s. In fact, the estimated persistence was close to 1 at the beginning of the period and fell down all the way until hitting the zero level in the last decade. This represents a radical transformation of the inflation process from almost a unit root to a state where inflation shocks have no effect on future inflation.

This major change in the inflation process has been recognized and discussed in the recent literature. Some have argued, though, that the decline in AR coefficients should not be associated with a decline in persistence (Pivetta and Reis, 2007, Stock and Watson, 2007). We will return to this debate in the next section. At this stage, we would introduce our proposed expectations-based measure for inflation persistence as a *cross-sectional* analogue to the time-series AR specification. The dramatic changes in AR coefficient estimates, as described in figure 1, also emphasize the limitation of using time-series data to estimate a frequently changing parameter, as in our case. Rolling-window estimation or other techniques of detecting regime changes could overcome this problem only to a limited extent and are not able to provide up-to-date estimates of the parameter. In the rolling-window exercise, for example, each point estimate is based only on 40 observations, while this would result in an estimate of current persistence which is based on a history of 10 years (40 quarters). Considering this limitation, which is typical to time-series data, our idea is to use forecasts from surveys, thereby offering an additional cross-sectional dimension of data, due to forecast dispersion. This could be helpful in trying to track changes in persistence faster and with greater precision.

More specifically, consider the CPI inflation forecasts provided in the Survey of Professional Forecasters ran by the Federal Reserve Bank of Philadelphia. This survey provides in each quarter inflation forecasts for multiple horizons. Let $F_t^i x_{t+h}$ denotes the forecast of a variable x (e.g. inflation) with a horizon of h steps ahead, made by an individual i. Suppose that the forecasters deliver their forecasts based on an AR(4) specification, as we have used before. We could recover the AR coefficients which they utilize in their forecasting procedure by estimating the following specification:

$$F_t^i x_{t+h} = c + \rho_1 F_t^i x_{t+h-1} + \rho_2 F_t^i x_{t+h-2} + \rho_3 F_t^i x_{t+h-3} + \rho_4 F_t^i x_{t+h-4} + u_t \qquad (1)$$

The specification relates the h-step ahead forecast of an individual to the forecasts provided in the same survey for each of the four preceding quarters. It therefore defines the persistence in expectations of survey participants. However, this persistence should also reflect the persistence in the AR inflation process, as it is perceived by survey forecasters.

In principle, the regression could be estimated using a cross-section of participants in the survey made in quarter t, and the estimated some of coefficients $\sum \rho$ would be a measure of inflation persistence, perceived by survey participants at time t. In the SPF, the number of participants is around 40 on average. This might not be enough to obtain a satisfying precision. However, including few additional cross-sections from previous surveys could be enough to get a precise measure which is still based on up-to-date data. Thus, our proposed measure is able to assess inflation persistence at time t, using very recent data, instead of relying on relatively old data, as in a standard time-series AR.

We apply specification (1) to the SPF forecasts with $h = 3$, namely, we run the three quarters ahead forecast on the forecasts for the four previous quarters (this includes the backcast of inflation in quarter $t - 1$, which is also provided in the SPF). The regression was estimated quarter-by-quarter using the last eight cross-sections from the survey, that is, survey data from the last eight quarters. This provides us with more than 300 observations from the last two years in each estimation, as opposed to the only 40 observations from the last ten years, which have been used in the rolling-window estimation of the time-series AR. We also estimate an "AR(1)" version of specification (1), which includes only the forecast $F_t^i x_{t+h-1}$ on the right-had-side.

The top pair of figures in figure 1 displays the results. The black lines show the estimated persistence in each quarter, according to our measure, with 95% confidence interval in the shaded area. For the "AR(1)" case (left-hand figure), the persistence measure is just the coefficient on $F_t^i x_{t+h-1}$, while for AR(4) it is the summation $\sum \rho$ in specification (1). Long run variation is illustrated by the blue thick lines, which are based on a non-parametric smoother (using Epanechnikov kernel). Our measure shows again, as in the regular time-series AR estimation, a clear declining trend in inflation persistence since the 1980s. Thus, SPF forecasters seem to recognize the important change in the inflation process over the last decades.

However, the reduction in persistence perceived by survey participants is notably more moderate. The difference is well-illustrated by the bottom pair of figures in figure 1, which puts together the trend of persistence according to our expectation-based measure (blue lines from the top pair of figures) and the standard AR measure (red line from middle figures). While the standard measure shows a decline to zero degree of persistence, the expectation-based measure makes only half of the way, varying around 0.5 in recent years. Moreover, our measure of persistence is consistently higher than the standard measure over the whole period. This clear difference between our proposed measure of persistence and the standard measure would be a key to understand why a measure of perceived persistence based on a survey could describe better the transition to the inflation process, as will be explained in the next section.[2]

3. Expectation-based or actual-based measure?

Our proposed measure of persistence is based on persistence perceived by forecasters. Why would a measure of perceived persistence be better than estimating persistence with real inflation data? As a first step to understand the advantages in perceived persistence, it should be noticed that because persistence is an unobserved parameter, any estimate of persistence could eventually be viewed as some form of perceived persistence. Thus, the question should actually be whether persistence perceived by the SPF forecasters could be a better estimate than persistence perceived by the econometrician. We would highlight two main advantages in using perceived persistence from a survey. The first advantage is related to the issue of limited time-series data, which was our initial motivation to use survey data. The second advantage relies on the particular debate in the literature about what happened to inflation persistence.

3.1. The data advantage

[2] Our finding is also in contrast to the conjecture in Gabaix (2019), that due to behavioral inattention agents will tend to average persistence across different processes and thus should *underestimate* the persistence of inflation if this variable is relatively persistent.

As emphasized above, time series data of actual inflation is quite limited. It could still be sufficient for estimating the persistence in inflation, if persistence is steady over time. Unfortunately, this is not the case as evident in figure 1. This provides a room for cross-sectional data from a survey to evaluate more precisely the current state of inflation persistence. The SPF in particular, has been proven to be a leading predictor of inflation, which usually beats most available econometric models (Ang et al., 2007). If SPF forecasters are doing relatively well in predicting inflation itself, it should not be surprising that their assessment of the current state of inflation persistence is adjusting more quickly to regime changes, relative to time-series techniques such as rolling window estimation.

Our proposed measure essentially extracts this assessment of inflation persistence from the survey, by simply utilizing the additional cross-sectional dimension of survey forecasts. The estimates of persistence in the survey are very precise, as can also be seen in figure 1: The 95% confidence interval is usually no wider than 0.2, while for the time-series rolling-window AR the confidence interval is at least twice larger, especially in recent years. In addition, an indication for how changes in the inflation process are quickly incorporated by SPF forecasters can be obtained from examining the high frequency variation in their perceived persistence (the black lines in the top pair of graphs). Interestingly, there are three notable drops in persistence around the start of the Great Recession, as well as the recessions at the beginning of the 1990s and 2000s. This could reflect the view of SPF professionals that sizable shocks to inflation during the time of recession would have only temporary effect and, as a consequence, future inflation would be less related to current inflation. This kind of information cannot be incorporated on real time by econometric techniques, such as rolling-window estimation.

Another indication on the quality of persistence measure from the survey, relative to the standard measure, is provided by the following simple exercise: We took the persistence estimates produced by the two methods for each quarter in the period, as reported in figure 1, and use them to predict inflation out-of-sample. For instance, suppose that in 2000Q1 the AR(1) coefficient estimate, based on a window of the last 40 quarters (until 1999Q4) is $\hat{\rho}_{AR}$. Using the SPF forecasts, instead, and running a regression of three on two quarter-ahead forecasts, we obtain a coefficient estimate of $\hat{\rho}_{SPF}$. Based on these two estimates of persistence, the inflation in 2000Q1 (still not known to survey participants in 2000Q1) is predicted to be $\hat{\rho}_{AR}\pi_{t-1}$ and $\hat{\rho}_{SPF}\pi_{t-1}$, respectively. In the same way, inflation is predicted for each quarter in the whole period (1983-2017). We then compare the quality of the predictions using the root mean squared error (RMSE).

Table 1 reports results from this out-of-sample prediction exercise and standardizes the RMSE of the SPF-based prediction to 1. Longer horizon predictions were also considered ($h = 0,4,8$. See in the column headers), as well as different window sizes for the $\hat{\rho}_{AR}$ persistence estimate (20, 40 and 60). Surprisingly, in all cases the RMSE of inflation predictions based on persistence measured with real inflation data is higher than the RMSE produced survey-based persistence by around 5% to 25%, where the gap is getting larger for the longer horizons. Changing to AR(4) specification (lower panel in the table), increases this gap even more.

Thus, although the true value of inflation persistence can never be verified, it seems that our expectation-based measure reflects the current state of the inflation process, better than the standard time series auto-regression. This is due to the use of more up-to-date data in a cross-sectional form, which is available from surveys. Still, there is a need to understand more deeply the reason for the deviation of the two measures from each other, especially, why the expectation-based measure is consistently higher than the standard measure during the whole period, as evident in figure 1, and why the difference was widening in recent years, with the decline in the degree of persistence by both measures.

3.2. The bias advantage

The patterns of persistence shown in figure 1 suggest that one of the measures is biased in a certain direction and this bias has grown with the decline in inflation persistence over the years. The debate in the literature about what happened to inflation persistence could shed light on this issue. Pivetta and Reis (2007) and Stock and Watson (2007) began to forcefully argue against a significant decline in inflation persistence (see Reis et al., 2018, for a recent more general discussion). In particular, Stock and Watson (2007) have suggested their influential unobserved-component modelling of inflation, according to which inflation is decomposed to trend and cyclical component. Specifically, the inflation process is described as:

$$\begin{aligned} \pi_t &= \tau_t + \eta_t \\ \tau_t &= \tau_{t-1} + \epsilon_t \end{aligned} \qquad (2)$$

where τ_t is the trend component which follows a random-walk and is shifted by the permanent shock ϵ_t. η_t is the cyclical component which is another form of a shock that affects inflation only at time t. Thus, it could be viewed as the process in (2) is characterized by two kinds of persistence: The long-run persistence in the trend component, which is at the highest degree of 1 (unit root), and the short run persistence in the cyclical component, which is virtually zero, since η_t is just a normally distributed mean zero and serially uncorrelated shock to current inflation. According to Stock and Watson (2007), this process has not changed over the years, implying no changes in persistence parameters. The decline in persistence, as measured by AR coefficients, is driven by a decline in the volatility of the permanent shock (ϵ_t) relative to the volatility of the transitory shock (η_t) during the post-Volcker era. The change in volatilities is described by augmenting (2) with stochastic volatilities for both types of shocks (named UC-SV model), and it is associated with an improved monetary policy which has gained better control over trend inflation.

This argument has an important implication for the new measure of inflation persistence we propose. The standard measure of persistence, based on the AR coefficient, should be biased downward, driven by the relative decline in the volatility of the permanent shock. By contrast, our expectation-based measure is not sensitive to variations in the volatilities of inflation shocks, since survey forecasts should not incorporate the effect of future shocks, which are not known to forecasters. If inflation follows the process in (2), it also follows that the persistence in survey forecasts should solely reflect the persistence in trend inflation and should not be biased by variations in volatilities. This reveals another advantage from using forecast data rather than inflation data to estimate the persistence of the underlying process. The growing departure of the two measures from each other, as appears in figure 1, illustrates how this advantage is more emphasized in recent years.

In fact, the simple framework we propose, for estimation of persistence with survey data, is also compatible with unobserved-component modelling of inflation, as in equation (2). For example, specification (1) is a cross-sectional version of AR(4) process, which relates h-step ahead forecasts to the forecasts for the four previous quarters. However, unlike regular AR(4), it allows long-run or trend inflation to change from period to period. Long-run inflation (which is τ_t in terms of equation (2)) could be recovered as the regression constant divided by one minus the persistence - $c/(1-\sum\rho)$. Hence, we could repeat the same procedure used in figure 1 to construct a time-varying measure of the trend component of inflation, based on cross-sections of SPF forecast data.

Figure 2 describes the estimates of $c/(1-\sum\rho)$ from running specification (1) quarter-by-quarter. The graph of estimated trend inflation is very similar to typical results from estimating the UC-SV model of Stock and Watson (2007): Trend inflation has gradually declined, following the Volcker disinflation, until converging to a steady level close to 2%. This also demonstrates how our approach could be useful not just for obtaining an up-to-date measure for inflation persistence, but also for trend inflation.

At the same time, our results provide evidence against the simple modelling of Stock and Watson (2007), which restricts the persistence of the cyclical component (η_t) to be zero. If SPF forecasters obey the process in (2), their forecast should correspond to their belief about the trend component of inflation (τ_t) and should not differ across horizons. This would imply that the degree of persistence in the forecasts should always be equal to 1, which clearly contradicts the results in figure 1.

We also have a direct access to data about the long-run beliefs of SPF participants. Since 1991, survey participants have also been asked to provide a 10-year ahead forecast for CPI inflation. Figure 3 shows the average reply over time, as well as the forecast for four quarters ahead and the actual quarterly inflation. While the two types of forecasts are much less volatile than the actual inflation, they do not seem to coincide either. Rather, the 4 quarters-ahead forecast is fluctuating to some extent around the 10-year ahead forecast. Again, this is not in line with the process defined in (2), which implies that the 10-year ahead belief, representing the trend component, would be the only forecast that should apply to any horizon.

The SPF evidence therefore suggests that short-run persistence should not be restricted to zero and should also be allowed to change over time. In fact, several recent studies have suggested to modify the unobserved component model in this direction, by looking on the inflation gap (e.g. Cogley, Primiceri and Sargent, 2010, Faust and Wright, 2013, Chan, Clark and Koop, 2018). The inflation gap is defined as the gap between inflation and its trend component, or $\pi_t-\tau_t$, using the above notations. According to the process in (2), the inflation gap should be equal to the cyclical shock η_t and thus would lack any persistence. Instead, we now allow the inflation gap to follow an AR process with some positive degree of persistence. We can also adjust our approach to directly measure the persistence in the inflation gap, based on expectations data. Consider the following specification:

$$\left(F_t^i x_{t+h} - F_t^i \tau_t\right) = c + \rho\left(F_t^i x_{t+h-1} - F_t^i \tau_t\right) + u_t \qquad (3)$$

where $F_t^i \tau_t$ is the forecast of trend inflation made by individual i at time t. Thus, we run the forecast for inflation gap t$+h$ quarters ahead on the forecast for the gap a quarter before. This corresponds to "AR(1)" version of specification (1), where the only difference is that we allow heterogenous beliefs about trend inflation (in specification (1) there is a uniform trend component captured by $c/(1-\sum\rho)$). The data on 10-year ahead inflation could be used to estimate specification (3) quarter-by-quarter, at least since the beginning of the 1990s, when these long-run forecasts are available in the SPF survey.

The estimated inflation-gap persistence is depicted in figure 4. Interestingly, the pattern over time is very similar to the one described in figure 1. Although the graph begins in the 1990s, a very high degree of persistence, above 0.8, can still be traced, followed by a gradual decline to around 0.4 in recent years. Some temporary fall around the Great Recession could also be noticed. In addition, confidence intervals are very tight, as in figure 1, indicating highly precise estimates of persistence. Thus, our evidence using the expectation-based measure of persistence suggest that even if inflation is modelled with unobserved component, where the trend follows a random walk, there is still additional sizable persistence in short-run cyclicality. This is again contradicting the restrictive modelling of Stock and Watson (2007) and offers a form of compromise in the recent debate in the literature about inflation persistence: There was a substantial decline in persistence after all, but it was only the short-run persistence which has experienced this transformation. This has some important implications for monetary policy, because it implies that the role of a central bank does not end in stabilizing trend inflation. There is also a need to control the impact of cyclical shocks, keeping them from being long-lasting.

The results in figure 4 highlight how our proposed measure, despite its simplicity, can be flexible enough to consider different ways of modelling inflation, including advanced unobserved component specifications, which are especially difficult to estimate with a limited time-series data. For example, Chan, Clark and Koop (2018) estimated a model which includes varying persistence in the inflation gap using time-series data (including average forecasts for long-run inflation). They were able to find some reduction in this short run persistence which is overall in line with the results in figure 4, but the path of persistence is considerably smoother and exhibits more moderate change. Using the cross-sectional dimension of data provided in survey forecasts is thus especially useful in assessing sizable changes in the inflation process and tracking them on time, without placing too much weight on inflation history due to data limitations.

4. Relation to expectations theories: Three puzzles

So far, we have not considered the relation of our expectation-based measure of persistence to theories of the expectations formation process. Our implied assumption in applying specification (1) was that forecasters use similar AR model to produce their forecasts and the parameters of that model could therefore be extracted from a cross-section of their multi-horizon inflation projections. This assumption could be justified based on rational expectations where the heterogeneity in the forecasts is due to some iid reporting error. In that case, all agents could follow the relation in specification (1), while the reporting error is absorbed by the regression error term.

More recent approaches for modelling expectations could nonetheless justify the application of specification (1). A prominent strand of literature highlights the role of heterogeneous information as the underlying reason for forecast heterogeneity. The heterogeneity in information is driven by forms of information frictions. Two leading examples are the sticky information model by Mankiw and Reis (2002) in which information acquisition is gradual, and the framework of noisy information, following Sims (2003) and Woodford (2002), in which information is acquired immediately but contains a noise component. Importantly, according to both frameworks, agents are assumed to form their expectations based on the same underlying process for the fundamental, conditional on their varying information. Thus, our strategy to recover the process parameters from cross-sections of survey forecasts should still hold.

Of course, the particular process used by forecasters to model the fundamental is not known, but finding the appropriate specification is an empirical task which is analogue to the regular task of an econometrician who attempts to fit the best specification to the fundamental itself. For instance, in the previous section we have estimated inflation persistence in the SPF survey with specifications based on AR(1) and AR(4). It is clear from figure 1 (top pair of figures), though, that the results are very similar, so that the simple AR(1) seems a satisfying approximation to the underlying model used by survey participants.

However, if all agents are assumed to follow the same model as assumed by the informational approach. There is no clear explanation for the presence of error term in our regression specification (1) (or its AR(1) form). The confidence intervals in figure 1, although being quite tight, still demonstrate non-negligible error component which requires more rigorous explanation. As mentioned above, imposing an additional horizon specific reporting error component could account for the lack of a perfect fit. Nevertheless, even this ad-hoc explanation could not address further puzzling evidence obtained with the SPF data. After presenting the evidence, we will discuss in more detail the difficulty of available recent theories of expectations formation to explain the findings.

4.1.Three puzzling patterns

Puzzle 1

As mentioned above estimating inflation persistence with a specification based on a simple AR(1) seems a satisfying approximation. This could improve our measure of persistence even more, because it enables to utilize additional data across different forecast horizons. Specifically, we use the inflation forecasts in the SPF survey which are available for multiple quarters ahead and estimate the specification:

$$F_t^i x_{t+h} = c + \rho F_t^i x_{t+h-1} + u_t \quad (4)$$

for forecast horizons h running from 0 to 4. Thus, we are running the 4-quarter ahead forecast on the 3-quarter ahead forecast, the 3-quarter ahead forecast on the 2-quarter ahead forecast and so on. This was not possible with the higher order specification (1), because of the need to include several "lags" in the form of forecasts made for several quarters before $t + h$. Accordingly, the results in figure 1 were obtained with $h = 3$, and there is only one additional horizon, $h = 4$, for which specification (1) could be applied.

The estimated persistence should not vary across forecast horizons since inflation persistence is dictated by the underlying process. Surprisingly, though, it is clearly not the case as evident in panel A of figure 5. The figure shows persistence estimates based on estimating specification (4) for each of the five available horizons in the SPF survey, separately. As in figure 1, we run the regression quarter-by-quarter using the last 8 cross-sections of the survey. The five lines in figure 5, are non-parametric smoothers of the coefficient point estimates corresponding to the five horizons $h = 0,1,2,3,4$. Notice that the thick blue line corresponding to $h = 3$ is the same line as in panel A of figure 1, which was also based on first order auto-regression applied to the same forecast horizon.

Overall, the lines for the longest horizons $h = 3$ and $h = 4$ show a high similarity in the estimated persistence over time, but as we move to shorter horizons persistence estimates get lower and lower and this hierarchy is consistent over the whole sample period. For example, at the beginning of the 1980s inflation persistence is very high and close to 1 when estimating specification (4) for horizons $h = 3$ and $h = 4$. However, when using the current quarter forecasts ($h = 0$, ran on previous quarter backcasts provided in the SPF), the coefficient estimate is just around 0.4 – a remarkably lower degree of persistence. Interestingly, all the lines show a decrease in estimated persistence over the years, so that the differences between horizons are diminishing but the pattern is preserved. In recent years the estimated persistence ranges between 0.2 (for $h = 0$) to around 0.5 (for $h = 3$ and $h = 4$). Moreover, it seems that the estimated persistence is converging as we move to the longer horizon. This may indicate that the estimated persistence is, for some reason, biased downward when applying forecast data with shorter horizons, but the bias diminishes when estimating persistence with longer horizon forecasts. The reason for such a form of bias is yet not understood.

Puzzle II

As noted above, unless introducing some ad-hoc horizon-specific reporting errors, estimation of specification (4) should provide a high fit, if AR(1) is a good approximation to the inflation process modeled by forecasters, as the data seems to indicate. It is therefore interesting to investigate not only the coefficient estimates from specification (4), but also the R-squared of the regressions.

This is presented in panel B of figure 5. The figure in panel B is based on the same estimation of specification (4), over time and across the different horizons, as in panel B, but instead of the coefficient estimates it provides the R-squared statistics. However, the R-squared figure is overwhelmingly similar to the coefficient figure. In particular, there is again a systematic difference in the fit across forecast horizons, where the R-squared of regressions with shorter horizon forecasts is consistently lower. There is no clear explanation for such a pattern for the R-squared.

Puzzle III

Another clear pattern from the same figure of R-squared statistic is that the fit of specification (4) decline over time for any horizon. Thus, the fit of the regression deteriorates with the decline in the estimated persistence documented above. This kind of relation is expected when estimating a regular *time-series* autoregression: Keeping all else constant, if the persistence of the fundamental is lower, the current shocks to the fundamental will play a relatively higher role in explaining volatility of that fundamental. In terms of a regression, the relative variance of the error term will be higher and the R-squared will be lower.

However, this should not be the case with specification (4), which is estimated using a cross-section or panel data of forecasts. Here there is no role for unknown shocks, so there should not be a difference whether forecasters compute their forecasts based on a high or low persistence parameter. Hence, there is no clear reason for a decline in the fit of specification (4) over time.

It could be argued, though, that while AR(1) provides a good approximation for the multiple-horizon forecasts at the beginning of the sample, SPF participants start to use richer specification overtime. For example, their model could include additional variables in a form of a VAR specification, and failing to modify specification (4) to account for that change in modelling, would result in a declining fit.

To examine this possibility, we tried alternative specifications to approximate the forecasting model of SPF participants. Figure 6 presents R-squared statistics over time from several alternatives: Higher order AR (panel B), AR for the inflation gap (panel C), and VAR (panel D). For example, the VAR specification is based augmenting (4) with additional forecasts of inflation, interest and unemployment rates, which were made for the four quarters preceding to $t + h$ (h is set to 3 in order to have enough "lags"). Yet, it could easily be noticed that the figures are all quite similar and resemble the results based on AR(1) specification, as presented in panel A. The remarkable decline in the fit over time is consequently our third puzzling evidence.

4.2. Relation to expectations theories

Rational expectations with a reporting error

As discussed above, using specification such as (4) to estimate the persistence in the underlying process could be justified based on standard rational expectations, with additional reporting error term. The idiosyncratic reporting error would account for cross-sectional heterogeneity in the forecasts and for the presence of an error term in the regression specification. However, it could not resolve the puzzles introduced above. In particular, the estimate of persistence should not vary across forecast horizons, in contrast to our first puzzle. The variation in the R-squared of our estimated specification, across horizons and over time could not be explained as well by a simple reporting error.

In the last years, several alternative mechanisms for explaining the expectations formation process have been proposed in the literature, which deviate from the standard full-information rational expectations hypothesis. The new theories of expectations where motivated by amounting empirical evidence from surveys of forecasts, such as the SPF, against rational expectations. It seems though that the empirical puzzles presented here poses a challenge even towards those recent approaches. We will demonstrate this difficulty by briefly considering some prominent recent approaches of expectations modelling.

Learning

The approach of learning, reviewed by Evans and Honkapohja (2012), generally assumes that agents follow the practice of econometricians and estimate the underlying process of the fundamental, while updating the data from period to period. Their expectations are then derived from the estimated process. Our strategy of recovering the persistence of the process from expectations data would be straightforward according to this approach. If agents estimate an AR process and use it to construct their predictions, we could use specification such as (4) to track their estimated persistence. Nevertheless, the coefficient estimate should not vary across forecast horizons, in contrast to our evidence in puzzle I, since people are using the same estimated AR to form their forecasts for any horizon. The other puzzles remain unexplained as well.

Behavioral approach

Another line of research introduces some behavioral characteristics to explain deviations from rationality. The asymmetric loss function suggested by Elliott, Komunjer and Timmermann (2008) and Capistran and Timmemann (2009) is an example for this approach. In their model, people are allowed to prefer positive over negative prediction errors or vice versa and this would result in biasing the rational expectation in a certain direction depending on the parameter of asymmetric preference, which is assumed to be heterogenous across agents.

We show in appendix A, that with such expectations, the coefficient estimate in specification (4) would be a biased estimate of the underlying inflation persistence. At the same time, we also show that the bias is positive and does not vary or diminish when applying specification (4) for different horizons. This is at odds with the pattern observed in figure 5 (panel A), which seems as a downward bias in the estimated persistence for shorter forecast horizons that tends to decay when moving to longer horizons. Similarly, the behavioral approaches in Fuster et al. (2010) and Gabaix (2019), which assume that agents follow some simplified misspecified model cannot address he difference in estimated persistence across horizons. More broadly, if a deviation from rationality is behavioral it is expected to persist to longer horizon thus affecting our measure of persistence in a similar way across horizons, unlike the convergence pattern documented in figure 5.

Informational approach

As mentioned above, a prominent strand of literature relaxes the assumption of full information in the standard rational expectations hypothesis. By introducing heterogeneity in information across agents as in the sticky and noisy information models, the dross-sectional dispersion of forecasts can be clearly explained ,as well as other departures from rational expectations observed in survey data. Coibion and Gorodnichenko (2012, 2015) found strong empirical support for this approach in the SPF survey in particular, by documenting how survey forecasts adjust gradually to the flow of information.

However, the evidence provided here seems challenging for this approach as well. Our methodology of recovering the underlying persistence from survey forecasts assumes that forecasters follow the same model, which is a regular assumption according to models of information rigidity, in which cross-sectional variation is only due to heterogeneity in information while all agents know the underlying process. Thus, conditional on their varying information all agents predict in the same way. The estimated coefficient and R-squared in specification (4) should not vary across horizons, which is against the evidence documented in figure 5.[3]

Disagreement about persistence

An important assumption according to all the above approaches of expectations formation is that all agents agree about the underlying model for the fundamental. If this model is AR(1), we could estimate the persistence in the model from the persistence in expectations as in specification (4). It could be argued, however, that there is a fundamental disagreement about the correct model. In fact, disagreement about the model could be the main source for forecast dispersion, instead of information heterogeneity as in the dominant models of information rigidities described above. An example for this approach is the model of Patton and Timmermann (2010), where agents have different views about the long-run mean of the fundamental. It is straightforward to extend this idea to our parameter of interest and suggest that agents have different beliefs about the underlying persistence.

A simple way to formalize this notion is to assume that each agent applies her own ρ_i in the AR(1) model, where ρ_i is iid normally distributed across agents around the true persistence ρ. Hence, the forecasts of agents would follow:

$$F_t^i x_{t+h} = \rho_i F_t^i x_{t+h-1} = \rho F_t^i x_{t+h-1} + (\rho_i - \rho) F_t^i x_{t+h-1} \quad (5)$$

[3] The same would apply to the model of diagnostic expectations in Bordalo et al. (2018), which introduces a form of behavioral overreaction to the noisy information framework.

Accordingly, if we regress $F_t^i x_{t+h}$ on $F_t^i x_{t+h-1}$ as in specification (4), we would have an error term corresponding to $(\rho_i - \rho) F_t^i x_{t+h-1}$, which is the disagreement about the value of persistence interacted with the forecast on the left-hand-side. Thus, unlike the informational approach, the source for the estimation error is now clear. Variation in the error-variance over time as in puzzle III, could be explained by variation in disagreement about persistence. Nevertheless, it can easily be verified that $E_i\left[(\rho_i - \rho) F_t^i x_t\right] = E_i\left[(\rho_i - \rho) F_t^i {x_t}^2\right] = 0$, so that the OLS estimate of the coefficient on $F_t^i x_{t+h-1}$ is a consistent estimate of ρ. This is in contrast to the differences in the coefficient estimates and the fit of the regression across forecast horizons, as documented in puzzles I and II.

In appendix B, we suggest a more complicated framework, in which disagreement about persistence is combined with noisy information. Agents get noisy signals not only about the value of the fundamental, but also about the persistence parameter. When filtering the noises with Kalman filter and due to the optimal gain dependency on the persistence, we show that in this setting, ρ_i is correlated with $F_t^i x_t$. As a result, the OLS coefficient would now be a biased estimate of the true persistence ρ. However, unlike in figure 5, the bias is not necessarily negative and should not decay for longer horizons. It therefore seems that assuming disagreement about the model would not resolve all the above puzzles.

5. A model of forward noisy information

The key insight from the previous section is that the expectations formation process should include two important ingredients in order to address our empirical findings. The first ingredient is some additional factor which forecasters apply beyond the underlying model. In fact, this kind of ingredient is familiar in the practice of forecasting as "add-factoring", which represents some adjustment to the model prediction, based on some additional information used by forecasters. The add-factoring could account for the regression error in specification (4). If it is further correlated with the forecasts across agents, it would bias our estimate of model persistence. The second important ingredient is some variation of the first one across forecast horizons, which could account for the variations observed in figure 5.

Our model would include these two ingredients in a noisy information framework, which will be different from the standard framework in two important aspects. Unlike in the standard noisy information model where there is only information about the past, in the current framework there will signals about the future fundamental, which represent the add-factoring ingredient. Furthermore, in the standard framework information varies only across agents whereas in our model information is heterogenous in two dimensions: across agents and across horizons. People obtain various signals of forward information for multiple horizons and, as a consequence, their add-factoring would differ across horizons.

We will first present a general framework and use simulations to show how the puzzling patterns from the previous section could be obtained with the new model. We would then use a tractable example to derive some analytical results.

5.1.General Framework

Consider a state-space representation of a fundamental x_t, which follows an AR(1) process. At time t, Agents receive multiple signals denoted by $y^i_{t,t+h}$, which refer to a time $t+h$ in the future, where h is the horizon of the signal, running from 0 to T periods ahead. The state-space model could simply be written as:

State:

$$x_t = \rho x_{t-1} + \omega_t \tag{5}$$

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

Measurement:

$$y^i_{t,t+h} = x_{t+h} + \nu^i_{t,t+h} \tag{6}$$

where $h = 0, \dots, T$ and $\nu^i_{t,t+h} \sim iid\ N(0, \sigma^2_h)$ is an idiosyncratic noise. Thus, agents have some forward noisy information referring to future values of the fundamentals. There are no useful signals from $T+1$ periods ahead onward. We could view this as if, eventually, the forward signal would not contain any valuable information, so that σ^2_{T+1} is infinite and therefore the sequence of σ^2_h goes to infinity with h. Naturally, we would expect the variance of the signals to follow a monotonic ordering, $\sigma^2_0 < \sigma^2_1 < \cdots < \sigma^2_T$, which implies a deterioration in the precision of the signals, as the horizon increases. However, the structure could be less restrictive. For instance, some forward guidance relating to the path of inflation a year from now, may represent some improved forward signal for a year horizon. The important restriction is therefore that eventually, for a long enough horizon the signals become uninformative.

It is also interesting to consider a case where the fundamental is finally observed after it is realized, by imposing $\sigma^2_0 = 0$. In that case, x_t is perfectly observed at time t , so that all the heterogeneity in expectations is derived from the signals about the future and not because of imprecise data, as it is assumed by the standard noisy information framework. On the other hand, the standard model is nested in our 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 realization are not perfectly observed. In the next section, we will use these restrictions to test the new framework against the standard one.

In order to calculate the optimal forecast that would filter the sequence of forward signals, we could augment the state-space representation, so as to consider a vector of fundamentals x_{t+h}s, with h running from 0 to T. We could then apply a Kalman filter to compute the forecast for all the elements in this vector jointly. The augmented state-space model would be the following:

State:

$$\boldsymbol{x}_t \equiv \begin{bmatrix} x_t \\ x_{t+1} \\ \vdots \\ x_{t+T} \end{bmatrix} = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \ddots & \ddots & 1 \\ 0 & 0 & \cdots & \rho \end{bmatrix} \boldsymbol{x}_{t-1} + S'\omega_{t+T} = \mathrm{P}\boldsymbol{x}_{t-1} + S'\omega_{t+T} \qquad (7)$$

where $S = [0 \quad \cdots \quad 0 \quad 1]$, so that the variance-covariance matrix of $S'\omega_t$ would be $\Sigma_\omega = S'S\sigma_\omega^2$.

Measurement:

$$\boldsymbol{y}_t^{\boldsymbol{i}} \equiv \begin{bmatrix} y_{t,t}^i \\ y_{t,t+1}^i \\ \vdots \\ y_{t,t+T}^i \end{bmatrix} = \begin{bmatrix} x_t \\ x_{t+1} \\ \vdots \\ x_{t+T} \end{bmatrix} + \begin{bmatrix} v_{t,t}^i \\ v_{t,t+1}^i \\ \vdots \\ v_{t,t+T}^i \end{bmatrix} = \boldsymbol{x}_t + \boldsymbol{v}_t^{\boldsymbol{i}} \qquad (8)$$

where the variance-covariance matrix of $\boldsymbol{v}_t^{\boldsymbol{i}}$ would be $\Sigma_v = \boldsymbol{I}_T\boldsymbol{\sigma}_v^2$ with $\boldsymbol{\sigma}_v^{2\prime} = [\sigma_0^2 \quad \sigma_1^2 \quad \cdots \quad \sigma_T^2]$. That is, the idiosyncratic noise in the forward signals is uncorrelated across horizons. It is also assumed that the noise is not correlated across agents. Finally, it is assumed that shocks to the fundamentals and noise in the forward signal are independent, that is $\boldsymbol{E}\left(S'\omega_t\boldsymbol{v}_t^{\boldsymbol{i}\prime}\right) = \boldsymbol{0}$.

Now we could apply the Ricatti equation to solve for the variance of the (one-step ahead) forecast error, denoted by Ψ:

$$\Psi = \mathrm{P}\{\Psi - \Psi(\Psi + \Sigma_v)^{-1}\Psi\}\mathrm{P}' + \Sigma_\omega \qquad (9)$$

The Kalman gain matrix, denoted by G (with dimension $T + 1$) , would then be obtained by

$$\mathrm{G} = \Psi(\Psi + \Sigma_\omega)^{-1} \qquad (10)$$

Thus, the optimal forecast would be

$$\boldsymbol{x}_{t|t} = \boldsymbol{x}_{t|t-1} + \mathrm{G}\left(\boldsymbol{y}_t^{\boldsymbol{i}} - \boldsymbol{x}_{t|t-1}\right) \qquad (11)$$

It follows that the forecast for h steps ahead could be written as

$$\begin{gathered} x_{t+h|t} = x_{t+h|t-1} + \mathrm{G}_{h+1,1}\left(y_{t,t}^i - x_{t|t-1}\right) + \mathrm{G}_{h+1,2}\left(y_{t,t+1}^i - x_{t+1|t-1}\right) \\ + \cdots + \mathrm{G}_{h+1,T+1}\left(y_{t,t+T}^i - x_{t+T|t-1}\right) \end{gathered} \qquad (12)$$

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

Importantly, when forecasting more than T steps ahead ($h > T$), the forecast would simply be

$$x_{t+h|t} = \rho x_{t+h-1|t} \quad (13)$$

We would now examine the relationship between $x_{t+h|t}$ and $x_{t+h-1|t}$ for shorter horizons ($0 < h \leq T$). From (12) the forecast $x_{t+h-1|t}$ would be

$$\begin{aligned} x_{t+h-1|t} = x_{t+h-1|t-1} + \mathrm{G}_{h,1}\left(y^i_{t,t} - x_{t|t-1}\right) + \mathrm{G}_{h,2}\left(y^i_{t,t+1} - x_{t+1|t-1}\right) \\ + \cdots + \mathrm{G}_{h,T+1}\left(y^i_{t,t+T} - x_{t+T|t-1}\right) \end{aligned} \quad (14)$$

where the coefficients $\mathrm{G}_{h,1}$, $\mathrm{G}_{h,2}$, …, $\mathrm{G}_{h,T+1}$ are the elements of row h in the gain matrix G. Multiplying (14) by ρ and subtracting it from (12), we obtain

$$\begin{aligned} x_{t+h|t} - \rho x_{t+h-1|t} \\ = x_{t+h|t-1} - \rho x_{t+h-1|t-1} \\ + \left(\mathrm{G}_{h+1,1} - \rho \mathrm{G}_{h,1}\right)\left(y^i_{t,t} - x_{t|t-1}\right) \\ + \left(\mathrm{G}_{h+1,2} - \rho \mathrm{G}_{h,2}\right)\left(y^i_{t,t+1} - x_{t+1|t-1}\right) + \cdots \\ + \left(\mathrm{G}_{h+1,T+1} - \rho \mathrm{G}_{h,T+1}\right)\left(y^i_{t,t+T} - x_{t+T|t-1}\right) \end{aligned} \quad (15)$$

Comparing to (13), the term $x_{t+h|t} - \rho x_{t+h-1|t}$ does not equal to zero, but rather is determined by the differences between the horizons in the weight placed on the forward information. Those differences are captured by subtracting ρ times the elements in row h from corresponding elements in row $h+1$ of the gain matrix.

Denote the vector which equals to row h in matrix G by $\mathbf{G}_h$. Moving $\rho x_{t+h-1|t}$ to the RHS and using substitutes from (8) we get

$$\begin{aligned} x_{t+h|t} = \rho x_{t+h-1|t} + \left(x_{t+h|t-1} - \rho x_{t+h-1|t-1}\right) \\ + \left(\mathbf{G}_{h+1} - \rho \mathbf{G}_h\right)\left(\boldsymbol{x}_t - \boldsymbol{x}_{t|t-1}\right) + \left(\mathbf{G}_{h+1} - \rho \mathbf{G}_h\right)\boldsymbol{\nu}^{\boldsymbol{i}}_t \end{aligned} \quad (16)$$

Suppose that, similar to specification (4), we run a cross-sectional OLS regression of $x_{t+h|t}$ on $x_{t+h-1|t}$, using forecasts made at time t for two consecutive horizons, h and $h-1$. If we take $h > T$, it is clear from (13) that a perfect fit should be obtained, and the coefficient estimate would be an unbiased estimator of the persistence in the state equation, ρ.

However, when the regression is estimated for shorter horizons, it would contain a disturbance term as characterized in specification (16). The disturbance includes three components:

1. The lagged disturbance from period $t-1$.
2. Ex-post errors of lagged forecasts.
3. The noises in current forward signals ($\boldsymbol{\nu}^{\boldsymbol{i}}_t$).

Consequently, the fit of the regression would be imperfect, that is, R-squared statistic will be lower than one, when the regression is estimated for $0 < h \leq T$. Furthermore, the OLS estimate of the coefficient on $x_{t+h-1|t}$

would be a biased estimator of ρ. This is because $\boldsymbol{v}_t^i$ in the disturbance term is correlated with $x_{t+h-1|t}$. The sign of the bias would be determined by the signs of the elements in the vector $(\mathbf{G}_{h+1} - \rho\mathbf{G}_h)$. This vector represents the change in the optimal weights placed on the signals, when moving the forecast horizon ahead from h to $h+1$. Using a tractable example, we will demonstrate later that this change in the weights is crucial for understanding the empirical patterns documented in the previous section. In particular, we will show how it varies with the forecast horizon and the degree of persistence, thereby determining the size of the bias in the coefficient estimate and the fit of the regression.

5.2. Simulation

The forward noisy information model, unlike the standard version of noisy information, is able to account for the source of the regression error term in specification (4). It further implies that the coefficient estimate could differ when using forecasts at different horizons. For a long enough horizon, though, the coefficient estimate would converge to ρ, while the variance of the regression error component decays to zero. These results are generally in line with the puzzling findings documented in the previous section. We further use some simulations in order to examine in more detail if our model could produce empirical patterns, which are similar to those reported above. Specifically, we are interested in tracking a convergence pattern from *below*, for both the coefficient estimate and the R-squared, as in figure 5, and a positive relation between them, as in figure 6.

We report the results of three simulations. Each simulation uses a different degree of persistence in the state process $\rho = 0.2, 0.5, 0.8$. Each simulation is based on 1000 draws and includes the four following steps:

I. Simulating the state process: The state process is simulated with a certain degree of persistence for a period similar to the SPF survey (from the beginning of the 1980s fro CPI inflation forecasts). 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 run from 0 to 7, and the vector of the noise variance is set to $\boldsymbol{\sigma}_v^{2\prime} = [0.2 \quad 0.5 \quad 1 \quad 2 \;\; 3 \quad 4 \quad 100 \quad 10000]$. 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.
III. Computing the forecasts: The Kalman gain matrix is computed based on equation (9) and (10) and then used to calculate optimal forecasts, with horizons running from 0 to 7, for the 40 simulated forecasters, using equation (11).
IV. Running regressions: For a certain "quarter" in the middle of the simulated sample we run cross-sectional regressions as in specification (4) and obtain the coefficient estimate and R-squared statistic.

Figure 7 reports average simulation results of the coefficient estimates (panel A) and R-squared statistics (panel B). The horizontal axis measures the inverse of the signal to noise ratio which equals to the noise variance since

σ_{ω}^{2} is set to one. This corresponds to the different forecast horizons, for which we ran the regressions. The three lines in the figure correspond to the three simulations, which take different values of persistence for the state-process.

Strikingly, the simulations reproduce all the three puzzling patterns, documented in the previous section. First, we observe in panel A that the coefficient estimate is a way below the true persistence for the short horizon. For example, when the true persistence is 0.5 the coefficient estimate is below 0.2. However, as we estimate the regression with longer horizon forecasts (higher noise to signal ration) 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 a little bit hump-shaped path. This illustrate the possibility of upward bias as well, although the bias in this direction is quite small in the simulation.

Second, there is a similar pattern of convergence for the R-squared statistic. Interestingly, the fit is very low when using shorter horizon forecasts, starting from below 0.1, but finally increases towards a perfect fit. Furthermore, the convergence could be quite slow, especially with low degree of persistence. 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 could explain the low R-squared values documented in the previous section in recent years even for the longer horizons in the SPF (figure 5, panel B).

Third, the simulation results in figure 7 also demonstrate a clear dependence of the regression properties on the underlying degree of persistence. Importantly, not only 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. This pattern is similar to the third puzzle in the previous section, where the fit of the regression deteriorated over time along with the decline in inflation persistence.

Finally, besides resolving the puzzling empirical patterns, our model and simulations are instructive about the proper way of applying our proposed measure of inflation persistence. 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. In panel A of figure 5, we observe a convergence of the coefficient estimate for horizons $h = 3{,}4$, for which the lines are very similar. Our model suggests that this indicates the point in future for which signals are very week and, as a consequence, persistence is estimated precisely enough. This validates our initial results of measuring inflation persistence with four and three quarters ahead SPF forecasts, as reported in figure 1. More generally, the conclusion from our model is that a measure of inflation persistence based on expectations data would be reliable if applied cautiously.

In order to gain further insight on the empirical patterns driven by our model and their relation to the information structure of the model, we present a tractable example, in which we can demonstrate several results analytically.

5.3. A case of two forward signals

Consider a case with a perfect signal about realization (current date value of fundamental) and two forward signals referring to two subsequent periods, $t+1$ and $t+2$. Accordingly, the state-space representation would be

State:

$$\boldsymbol{x}_t \equiv \begin{bmatrix} x_t \\ x_{t+1} \\ x_{t+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & \rho \end{bmatrix} \boldsymbol{x}_{t-1} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \omega_{t+2} = \mathrm{P}\boldsymbol{x}_{t-1} + S'\omega_{t+2} \tag{17}$$

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

Measurement:

$$\boldsymbol{y}_t^{\boldsymbol{i}} \equiv \begin{bmatrix} y_{t,t}^i \\ y_{t,t+1}^i \\ y_{t,t+2}^i \end{bmatrix} = \begin{bmatrix} x_t \\ x_{t+1} \\ x_{t+2} \end{bmatrix} + \begin{bmatrix} 0 \\ \nu_{t,t+1}^i \\ \nu_{t,t+2}^i \end{bmatrix} = \boldsymbol{x}_t + \boldsymbol{\nu}_t^{\boldsymbol{i}} \tag{18}$$

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

Notice that the perfect signal is $y_{t,t}^i$, which does not contain noise, so that x_t is perfectly known at time t. Hence, there is no noisy information in this case, in the sense of the standard noisy information framework. Only the forward signals $y_{t,t+1}^i$ and $y_{t,t+2}^i$ are imperfect. It follows that the forecast for the current period (which could be considered a backcast, since x_t has been realized) is simply

$$x_{t|t} = x_t \tag{19}$$

This would also simplify the analytical derivation of the optimal weights in the forecasts for the future periods. Instead of solving the Ricatti equation, we can equivalently consider the optimal forecast as a weighted sum of all the relevant signals, where the weights should minimize the squared forecast error[4]. For a forecast made at time t, and due to the assumption that realized values are perfectly known, there are only four signals to take into account: The perfect signal about the realized x_t , the forward signal $y_{t-1,t+1}^i$ which was received in the previous period but refers to a date after the realization, and the two new forward signals received now, $y_{t,t+1}^i$ and $y_{t,t+2}^i$. Any other signal received in the past should not be considered, since it refers to a realization which

[4] All the results remain the same, when the optimal forecast is derived by the regular algorithm of the (steady-state) Kalman filter applied to the state-space representation in (17) and (18). In appendix C.2., we show the direct mapping from the optimal weights derived here to the elements in the gain matrix **G.**

has been already perfectly observed. Hence, after adjusting the signals to the forecast horizon, the one-step ahead forecast will take the form:

$$x_{t+1|t} = W_1 \rho x_t + W_2 y^i_{t-1,t+1} + W_3 y^i_{t,t+1} + W_4 \rho^{-1} y^i_{t,t+2} \qquad (20)$$

where W_1, W_2, W_3 and W_4 are the optimal weights placed on each informative signal, which obeys $W_1 + W_2 + W_3 + W_4 = 1$. The optimal weights should minimize the expected squared forecast error, that is, $E_t\left(x_{t+1} - x_{t+1|t}\right)^2$. Setting $\sigma^2_\omega = 1$, without loss of generality, we derive, in appendix C.1., that the optimal weights are given by

$$\begin{aligned} W_1 &= \frac{\sigma_2^2 \sigma_1^2 (1+\sigma_2^2)}{m} \\ W_2 &= \frac{\sigma_1^2 (1+\sigma_2^2)}{m} \\ W_3 &= \frac{\sigma_2^2 (1+\sigma_2^2)}{m} \\ W_4 &= \frac{\rho^2 \sigma_2^2 \sigma_1^2}{m} \end{aligned} \qquad (21)$$

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

In a similar way, the forecast for two steps ahead will be a weighted sum of the same four signals, which will be moved one more period to the future. Specifically:

$$x_{t+2|t} = w_1 \rho^2 x_t + w_2 \rho y^i_{t-1,t+1} + w_3 \rho y^i_{t,t+1} + w_4 y^i_{t,t+2} \qquad (22)$$

Comparing to the one-step ahead forecast in (19), each of the four informative signals is further multiplied by ρ, because of the change in the horizon. If the corresponding optimal weights in (19) and (22) 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 be $x_{t+2|t} = \rho x_{t+1|t}$, which follows the AR(1) process of the fundamental. This is the regular result with the standard noisy information model (and other models of expectations, as discussed above). However, this is not the case when introducing the forward signals. As shown in appendix C.1., the optimal weights that minimizes the two-steps ahead squared forecast error, are given by

$$\begin{aligned} w_1 &= \frac{\sigma_2^2 \sigma_2^2 \sigma_1^2}{m} \\ w_2 &= \frac{\sigma_2^2 \sigma_1^2}{m} \\ w_3 &= \frac{\sigma_2^2 \sigma_2^2}{m} \\ w_4 &= \frac{(1+\rho^2)\sigma_2^2 \sigma_1^2 + \sigma_1^2 + \sigma_2^2}{m} \end{aligned} \qquad (23)$$

It is easy to see that the weights on the first three signals, w_1, w_2 and w_3 are lower than the corresponding weights used in the one-step ahead forecast - W_1, W_2 and W_3, respectively. In contrast, w_4 is greater than W_4, so that the decline in the weight placed on the first three signals by the two steps ahead forecast turns to an excessive weight placed on the fourth signal. Intuitively, because the fourth signal ($y^i_{t,t+2}$) refers directly to x_{t+2}, it is given an extra weight in the two steps ahead forecast. This variation in the optimal weights would shift the relation between the two consecutive forecasts from being simply $x_{t+2|t} = \rho x_{t+1|t}$. Following (20) and (22) the relation could be expressed as

$$x_{t+2|t} = \rho x_{t+1|t} + \sum_{k=1}^{4} (w_k - W_k)\, Signal_k \qquad (24)$$

where $Signal_k$ corresponds to each of the four signals used in (22) ($Signal_1 = \rho^2 x_t$, $Signal_2 = \rho y^i_{t-1,t+1}$, etc.). Equation (24) is a specific case of equation (16) derived above in the more general framework. It provides two important insights: First, the change in the optimal weights across forecasting horizons is responsible for the error term in a cross-sectional regression of $x_{t+2|t}$ on $x_{t+1|t}$. In the general equation (16), the difference in the optimal weights across horizons is expressed by $(\mathbf{G}_{h+1} - \rho\mathbf{G}_h)$. The form of the error term can explain the empirical patterns we have previously obtained. For example, the signals in the error term are correlated with regressor $x_{t+1|t}$, thus, explaining the bias of the OLS coefficient estimate. Second, equation (24) could also be interpreted as a decomposition of the forecast $x_{t+2|t}$ to two components. The first component, $\rho x_{t+1|t}$, 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. Thus, our model of the expectations formation provides a formal description of a common forecasting practice, which includes some subjective adjustment component (e.g. Stark, 2013). Furthermore, in section 9, we use this interpretation to quantify the second forecast component as a news shock.

When increasing the horizon of the forecast by a further step, however, the optimal weights would not change and remain w_k, as we show in appendix C.1.. This is because the forward signals do not refer to future periods beyond $t+2$, so that $x_{t+2|t}$ is sufficient for an optimal forecast for $t+3$. Accordingly, we obtain:

$$x_{t+3|t} = w_1\rho^3 x_t + w_2\rho^3 y^i_{t-1,t+1} + w_3\rho^2 y^i_{t,t+1} + w_4\rho y^i_{t,t+2} = \rho x_{t+2|t} \qquad (25)$$

More generally, the relation between two consecutive forecasts would obey the simple form of $x_{t+h|t} = \rho x_{t+h-1|t}$ for any $h \geq 3$.

Finally, we could use (19) and (20) to express the relation between the forecasts with the shortest horizons, $x_{t+1|t}$ and $x_{t|t}$:

$$\begin{aligned} x_{t+1|t} = \rho x_{t|t} + (W_1 - 1)\rho x_t + W_2\rho y^i_{t-1,t+1} + W_3 y^i_{t,t+1} \\ + W_4\rho^{-1} y^i_{t,t+2} \end{aligned} \qquad (26)$$

The interpretation is similar to (24), where the simple relation based on the state process is modified due to variations in the optimal weights placed on the signals, when the forecast horizon changes from $h = 0$ to $h = 1$. The optimal weights in $x_{t+1|t}$ are W_k, while for $x_{t|t}$ the whole weight is placed on the first signal, which is the perfectly known realization of x_t. Equation (26) express those differences in the weights, multiplied by the four signals.

In sum, equations (24), (25) and (26) summarize, for our tractable case, the varying relation between consecutive forecasts $x_{t+h|t}$ and $x_{t+h-1|t}$, for each forecast horizon. This sheds light on the empirical patterns obtained with our cross-sectional regression of $x_{t+h|t}$ on $x_{t+h-1|t}$. More specifically, 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 $x_{t+h-1|t}$. As a consequence, the OLS coefficient estimate could be biased away from the underlying persistence and the fit of the regression could be poor. However, as we move to the longer horizons and estimate the regression, the error term of the regression would shrink and the relation between $x_{t+h|t}$ and $x_{t+h-1|t}$ would get closer to $x_{t+h|t} = \rho x_{t+h-1|t}$, therefore, implying both a reduction in the bias of the coefficient estimate and an increase in the regression R-squared. This convergence patterns are obtained due to the tendency of the variation in the optimal weights across horizons to diminish, when increasing the horizon.

In Appendix C we examine more specifically how the properties of the regression changes across forecast horizons and across different degrees of persistence in the state process. A summary of the results is presented in table 2, which highlights the empirical patterns by considering the extreme cases, that is, shortest versus longest horizon ($h \geq 3$, as explained above), as well as, zero persistence versus a unit root ($\rho = 1$). More generally, our key insight is that the empirical patterns are driven by the gap $(w_k - W_k)$ between the optimal weights, which vary by both the forecast horizon and the underlying persistence.

5.4. Predictability of forecast errors

A key finding in the literature which examines survey forecasts, is that in contrast to rational expectations, forecast errors are predictable by information available to forecasters. Thus, one of the main tasks of any model of expectations formation, which deviates from rational expectation, is to explain this predictability. The models which have been proposed so far in the literature do so by introducing some form of deficiency in the expectations formation. A notable example is the study by Coibion and Gorodnichenko (2015), which presents pervasive evidence for predictability of ex-post forecast errors by forecast revisions at the aggregate level. They show that this outcome is expected according to the prominent models of sticky and noisy information because of the gradual processing of information at the aggregate level, and estimate the parameters of information rigidity from this relation between forecast errors and revision.

Here we use our tractable example, in which information about the past is perfect, to demonstrate that this predictability could arise from gradual processing of *forward* information, even in the absence of any information rigidity in the sense of a standard noisy information model. Consider the one step-ahead forecast error, which is obtained using (20):

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

Taking the average across agents, would drop all the idiosyncratic terms. Hence, we get:

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

Where the upper bar in $\bar{x}_{t+1|t}$ denotes the cross-sectional average.

The forecast $\bar{x}_{t+1|t}$ revises the forecast $\bar{x}_{t+1|t-1}$, which is the two step-ahead forecast from the last period. Using (20) and (22), and averaging across agents, we show in appendix C.5. that the revision to the average forecast could be expressed as:

$$\bar{x}_{t+1|t} - \bar{x}_{t+1|t-1} = \rho\omega_t + (w_1 + w_2 + w_3 - W_1)\omega_{t+1} + W_4\rho^{-1}\omega_{t+2}$$

Notice that both the forecast error and forecast revision contains the shocks ω_{t+1} and ω_{t+2}, which produces a correlation between them. Notice that when assuming rational expectations, the forecast error would only contain the one step ahead shock ω_{t+1} and the revision would only contain the current shock ω_t, which are not correlated. According to the standard noisy information model, the error and revision are correlated since they both contain shocks up to time t, due to the gradual processing of imperfect information about the past. In our model, by contrast, the correlation between forecast errors and revisions is due to future shocks, as a result of gradual processing of forward information. In appendix C.5., we derive the explicit expression for the OLS coefficient in a regression of the forecast error on forecast revision, and show that it should be positive, in line with the estimates in Coibion and Gorodnichenko (2015), using several surveys of forecasts, including the SPF in particular. Similarly, it can be shown that forecast errors are also predictable by lagged forecast errors, which is also a typical finding in survey forecasts (e.g. Coibion and Gorodnichenko, 2012).

5.5.Forward information in SPF forecasts

So far, we have provided indirect evidence, which support the presence of forward information in the SPF inflation forecasts. The interpretation of the empirical patterns described in the previous sections, following our model, is that forward signals lead to a downward bias in our proposed measure for inflation persistence at the shorter horizon. However, Forward information seems to decay around 4-quarters-ahead horizon in the SPF and therefore, persistence is well-approximated when estimated for this horizon.

Given the more specific information structure, as described by the model, it would be desirable to asses more directly the role played by forward information. In particular, it would be helpful to distinguish between forward information to rigidity in information with respect to previous data, in the sense of standard noisy information model.

We suggest a simple strategy for this purpose, following a similar approach proposed by Goldstein (2019) to estimate the standard noisy information framework (as well as the model of sticky information). Our focus is on estimating the weight placed on forward information, which is summarized in the gain matrix G in equation (11), describing the optimal vector of forecasts in our model.

The optimal forecast in (11) could be rewritten as follows (using the measurement equation (8)):

$$\begin{aligned}\boldsymbol{x}_{t|t} &= \boldsymbol{x}_{t|t-1} + \mathrm{G}\left(\boldsymbol{y}_t^{\boldsymbol{i}} - \boldsymbol{x}_{t|t-1}\right) = \boldsymbol{x}_{t|t-1} + \mathrm{G}\left(\boldsymbol{x}_t + \boldsymbol{v}_t^{\boldsymbol{i}} - \boldsymbol{x}_{t|t-1}\right) \\ &= (I - G)\boldsymbol{x}_{t|t-1} + G\left(\boldsymbol{x}_t + \boldsymbol{v}_t^{\boldsymbol{i}}\right)\end{aligned} \tag{27}$$

Next, we take the average forecast across individuals, which will be denoted by an upper bar, and obtain:

$$\bar{\boldsymbol{x}}_{t|t} = (I - G)\bar{\boldsymbol{x}}_{t|t-1} + G\boldsymbol{x}_t \tag{28}$$

where $\bar{\boldsymbol{x}}_{t|t}$ and $\bar{\boldsymbol{x}}_{t|t-1}$ are the average of $\boldsymbol{x}_{t|t}$ and $\boldsymbol{x}_{t|t-1}$, respectively. Notice that by taking the average over the cross-section of forecasters, the idiosyncratic noise in the signals is vanished.

Subtracting equation (28) from (27) we get

$$\boldsymbol{x}_{t|t} - \bar{\boldsymbol{x}}_{t|t} = (I - G)\left(\boldsymbol{x}_{t|t-1} - \bar{\boldsymbol{x}}_{t|t-1}\right) + G\boldsymbol{v}_t^{\boldsymbol{i}} \tag{29}$$

Equation (29) describes a simple relationship between the deviation of individual forecast from the mean in period t and lagged deviation from period $t-1$. Thus, we can 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. The regressors for each horizon will be the same and include lagged deviations from the mean forecast running over all relevant horizons, as summarized in the vector $\left(\boldsymbol{x}_{t|t-1} - \bar{\boldsymbol{x}}_{t|t-1}\right)$. Specifically, the regressions will take the following form:

$$\begin{aligned}x_{t+h|t} - \bar{x}_{t+h|t} &= \beta_0\left(x_{t|t-1} - \bar{x}_{t|t-1}\right) + \beta_1\left(x_{t+1|t-1} - \bar{x}_{t+1|t-1}\right) + \cdots \\ &+ \beta_T\left(x_{t+T|t-1} - \bar{x}_{t+T|t-1}\right) + error_{t+h|t}\end{aligned} \tag{30}$$

where the β coefficients are elements of row $h+1$ in the matrix $(I-G)$.

Fortunately, this simple specification provides unbiased estimates of the elements in $(I-G)$, since, following (29), the error term in the regression is a corresponding element in the vector $G\boldsymbol{v}_t^{\boldsymbol{i}}$, where the noise components in current signals are uncorrelated with *lagged* forecasts.

The specification can be viewed as an augmented version of the standard noisy information model. For the standard model, Goldstein (2019) has proposed the following specification:

$$x_{t+h|t} - \bar{x}_{t+h|t} = \beta_{NOISY}\left(\bar{x}_{t+h|t-1} - x_{t+h|t-1}\right) + error_{t+h|t} \qquad (31)$$

In the standard framework there is only one noisy signal y_t^i, which refers to the last realization. As a result, the deviation from the mean forecast would be related only to the lagged deviation which refers to $t+h$. All the other lagged deviations in (30) should therefore be omitted. As shown in Goldstein (2019), in this simple framework β_s should equal $(1 - G_{NOISY})$, where G_{NOISY} is the gain parameter representing the weight placed on the single noisy signal.

This provides a straightforward way to test our model of forward information against the null of standard noisy information version: We simply need to estimate specification (30) and test the significance of all the coefficients other than the coefficient on $\left(\bar{x}_{t+h|t-1} - x_{t+h|t-1}\right)$. Furthermore, specification (30) offers a second type of test which would examine the null of ruling out noisy information, that is, the possibility that heterogeneity in information may only be due to forward signals, while information about realized inflation is perfectly utilized by SPF forecasters. As illustrated in our tractable example above, this null would impose a restriction on the gain matrix. Specifically, if information about realized x_t is perfectly available, the first row in the gain matrix, would take the form of $\mathbf{G}_1 = [1 \quad 0 \quad \cdots \quad 0]$, since forecasters would place all the weight on the perfect signal $y_{t,t}^i = x_t$, when forming their forecast $x_{t|t}$ for time t (which is actually a backast as far as x_t is realized). In terms of specification (30), which estimates the rows in the matrix $(I - G)$ as in (29), we would expect all the elements in the first row to be zeroes under the null. Luckily, the SPF survey also provide backasts of realized inflation in the previous quarter. Hence, we can estimate specification (30) with the data of the backasts and test the significance of the coefficients (this corresponds to the case of $h = 0$, using our notations for the model).

Table 3 reports estimates of specification (30), using SPF survey data of CPI inflation forecasts with multiple horizons, which was utilized in the previous sections. Each panel reports estimation results for a different h, which corresponds to a different row in the matrix $(I - G)$, as explained. In principal we should have $T + 1$ regressors representing the number of available signals. Due to the availability of the data, we include all lagged forecasts referring to the quarters from t to $t + 4$. Recall, though, that our previous results indicate that forward signals die out around a year ahead, so these number of regressors seems sufficient. Finally, in each panel we also break the sample to the different decades in an attempt to address the change in inflation persistence over time, as documented in the previous sections, which should affect the weights in the gain matrix.

The results in table 3 support the rejection of both null hypotheses. First, when focusing on panel A, which reports results for backasts $h = 0$, we strikingly find that all the coefficients are not significant when the regression is estimated for the whole sample. In the estimation for sub-periods there are few significant cases but even in those cases the coefficients are very close to zero in magnitude. R-squared statistics are all very low as well. This evidence suggests that noisy information in the sense suggested by the standard model is negligible.

In contrast, the findings in the other panels of the table, support the presence of forward information. Overall, the coefficients on $(\bar{x}_{t+h|t-1} - x_{t+h|t-1})$, which represents diagonal elements, are strongly significant, but in each estimation there is at least one additional coefficient which is highly significant. As explained above, this is in line with the utilization of forward signals by SPF forecasters reflected in non-zero elements in the gain matrix.

In fact, a concern could be raised that the off-diagonal elements are not properly represented and distinguished in the regression estimates due to the high correlation between the regressors, which are all deviations from the mean forecast at different horizons. For this reason, we also provide in table 4 a comparison between specifications (31) and (30) using the BIC statistic. For all forecast horizons, specification (30), which augments (31) with forward information is preferred over specification (31). Thus, in line with our results from the previous sections, the direct estimation of our model provides strong evidence that expectations in the SPF are driven heterogenous information with respect to future inflation, rather than rigidities in information about realized inflation.

6. Persistence in other macroeconomic variables

Up to now, our empirical analysis has been focused on forecasts of inflation. Our interest in applying our expectation-based measure of persistence to inflation was motivated by the important debate in the literature about the changes in this key parameter over the years. Beyond shedding light on this debate though, the empirical patterns we found, opened the window to a new understanding of the expectations formation process. According to the model of forward information laid down the previous section, our measure of inflation persistence is biased downward at shorter forecast horizons, because forecasters apply some information about future inflation. The measure of persistence converges around a horizon of a year ahead, indicating that forward information is no longer valuable beyond this horizon, and consequently the measure of persistence become valid.

It is interesting to examine whether this pattern is common in forecasts of other macroeconomic variables as well. In particular, if we run our simple cross-sectional "AR(1)" specification, using consecutive forecasts of other macroeconomic variables, available in the SPF, would we obtain a similar variation in the coefficient estimates and R-squared across forecast horizons? This would point to a role of forward information in forecasting other key macroeconomic variables. Figure 8 presents the results from applying such an exercise to SPF forecasts of unemployment rate, interest rate and real GDP growth. The results are presented in the same form as in figure 5. We ran regressions quarter-by-quarter, using the cross-sections of forecasts from the last eight quarters. The figure describes for each variable the coefficient estimate (left side) and R-squared statistic (right side) from the regressions, estimated over time and across the different horizons available in the survey ($h = 0,1,2,3,4$). Figure 9 presents results for forecasts of additional measures of inflation: GDP deflator, PCE inflation and core inflation. Sample period varies depending on the availability of forecast data for each variable: For unemployment, GDP growth and GDP inflation sample starts at the beginning of the 1970s, for interest rate sample starts at the beginning of the 1980s and for PCE and core inflation it only begins in 2009 (eight quarters after the first available forecasts because of the small window size).

Interestingly, for most variables, both the coefficient-estimate and R-squared show patterns of variation across horizons and convergence in a similar way to our baseline results in figure 5. For unemployment and interest rate forecasts (panels A and B in figure 8), the estimated persistence, which should be well estimated in the convergence region, is quite steady around the level of 1 over the years. Thus, SPF participants consistently associate a random walk process with movements of unemployment and interest rate. Interestingly, comparing to inflation, the convergence of estimates across horizons for the unemployment and interest rate forecasts is quite fast, since the graphs for $h = 1,2,3$ are quite similar, while only for $h = 0$ the coefficient-estimates and R-squared are noticeably lower. This suggests that utilization of forward information with respect to these variables is lower compared to inflation. This could be the outcome of a higher degree of attention to inflation by the forecasters or a greater availability of valuable forward signals about inflation, due to central bank communication measures 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 us unemployment, or a high frequency variable such as the Treasury-Bill rate.

For the forecasts of PCE and core inflation (Panels B and C in figure 9), there is more variation across horizons and the convergence is around $h = 3, 4$, implying 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 which 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 quite similar to the persistence in CPI inflation during these years, as described in figure 5. For the core inflation persistence is estimated around 0.8. This higher degree of persistence is expected since the core measure of inflation excludes the components with the highest short-run volatility from the CPI inflation, and serves as an approximation to trend inflation. This approximation is however limited, as the measured persistence is still lower than 1.

For two variables, there is no clear pattern across horizons for the coefficient estimate and the R-squared of the regression: The GDP growth (panel C in figure 8) and GDP inflation (panel A in figure 9). 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 a simple 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 GDP and GDP deflator forecasts refer to the levels of GDP and GDP price index. The original forecasts are then transformed to the growth rate of GDP and GDP inflation rate, and the estimation is applied to these transformed forecasts in order to measure the degree of persistence. 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, the strongest downward bias in our measure of persistence, is expected to be obtained when estimating the coefficient with $h = 0$. However, the regression cannot be estimated for the horizon $h = 0$, with the rate-of-change forecasts because we need to regress forecasts for the current quarter on backasts, and data of backasts is only available in levels and not in terms of the rate of change.

The second issue could be addressed by including more "lags" in the cross-sectional regression as in specification (1), which imitate an AR(4) process by regressing a forecast on forecasts for the four preceding quarters. Recall that in our baseline findings, using the CPI inflation forecasts, there was no difference between AR(1) to AR(4), as presented in figure1. Figure 10 presents a similar comparison using the forecasts of GDP inflation and GDP growth and estimating the regression coefficient based on AR(1) and AR(4) (left and right panels, respectively). Here we observe a noticeable difference between the specifications, especially for GDP inflation (top pair of figures). The estimated persistence based on the sum of coefficients when using AR(4) specification, is consistently higher over time, relative to estimates based on AR(1). Furthermore, extending the specification smooths much of the volatility over time in the estimates based on AR(1). This indicates that AR(1) does not provide a satisfying approximation for these variables. If AR(4) is a better fit to the model used by SPF forecasters in these cases, then a comparison across horizons cannot be done properly, since we don't have enough multiple horizons of forecasting in the survey. Thus, it could well be case that forward information is applied by the professional forecasters to predict the GDP and GDP inflation.

Turning to the estimates of persistence themselves, according to the AR(4) specification, the persistence in GDP growth is quite steady over the years around the level of 0.5. Interestingly, exceptional drops are documented in times of recessions at the beginning of the 1980s, the beginning of the 1990s, and more weekly during the great recession. The drop in persistence at the beginning of the 1980s was particularly strong, heating the zero level. This demonstrates how high frequency regime changes are embedded in the modelling of professional forecasters, reflecting their views about how long would be the effect of adverse shocks, hitting the economy during recessions, in a way which is not applicable by econometric modelling, especially in real time. This illustrates again the benefit from using our expectation-based measure of persistence.

The course of persistence over time in GDP inflation is of particular interest, as it is comparable to our baseline findings for CPI inflation. Nevertheless, the path of inflation persistence looks quite different in figure 10 comparing to figure 1. For CPI inflation we have documented a gradual decline in the degree of persistence from 1 at the beginning of the 1980s to about 0.5 in recent years. For the GDP inflation, though, such a clear trend is not documented. According to the AR(4) specification persistence was not as high in the 1980s. It finally reached the degree of 1 in the 1990s. A drop in inflation persistence only occurred around 2000. Another sharp decrease appears around the great recession, but after then persistence returns to high levels between 0.7 to 1 in recent years.

Figure 11 sheds some more light on the distinguished properties of GDP inflation. Panel A describes again the expectations-based measure of persistence in the blue line and compares it to an ordinary measure based on time-series AR(4) estimation (with rolling window of 40 quarters). Recall that for CPI inflation the decline in the sum of AR coefficients since the 1980s was dramatic, from 1 to zero. However, this is not the case when fitting time-series AR to the time series of GDP inflation, where the estimated sum of coefficients is above 0.6 even in the 2000s, and only drops in the last quarters of the sample. In fact, the expectations-based measure and the standard AR measure of persistence in GDP inflation, as displayed in figure 11, were quite close to each other. This is also a departure from the findings in figure 1 for CPI inflation where consistent difference between the two measures over the sample period, which was explained by the shifts in the volatilities of short-run vs. long-run shocks to inflation. It seems therefore the transitions in the process of GDP inflation during the last decades didn't follow a certain trend.

We also examine the R-squared of the cross-sectional regressions used to estimate persistence from the forecast data. For CPI inflation the R-squared declined with the decline in persistence. However, for GDP inflation, as presented in panel B of figure 11, the R-squared is relatively low over the whole period, even in time where estimated persistence was relatively high. This could also be associated with the more complex dynamics of GDP inflation, as indicated in figure 10, as opposed to the simple AR(1) approximation for CPI inflation, which provided a clear measure of persistence.

Finally, in panel C of figure 11, we examine more directly the link between inflation measures by computing the simple correlation between CPI and GDP inflation over time, with a rolling window of 40 quarters. Strikingly, there is a consistent deterioration in the correlation of the two series since the 1970s. The two series which were highly correlated at the beginning, but this correlation dropped dramatically from about 0.8 to only 0.2 in recent years. The dashed line shows how the correlation between CPI and PCE inflation evolved during that time. Here the results are very different, demonstrating a high and steady correlation over the years, between 0.8 to 1. The similarity of CPI and PCE inflation confirms our previous findings, when measuring persistence with PCE inflation forecasts in figure 10 and the similarity to the results for CPI inflation forecasts. It should also be mentioned that PCE is also based on GDP inflation, but only concentrates on the consumption component in the GDP. This suggests that weakening link between GDP inflation to CPI (and PCE) measures should be associated with the prices of the other GDP components, namely, investment exports and imports. These interesting findings raises issues which are beyond the focus of this paper, but certainly deserves a deeper examination by the literature, For our purpose, this preliminary findings demonstrate again why persistence measures for GDP inflation are not similar to our findings for CPI inflation.

To summarize, it seems overall that our model of forward information and the properties of our measure of persistence applies not just for inflation but to other macroeconomic variables. Our findings also demonstrate how patterns could be changed when the underlying process is more complex.

7. Inflation persistence in the European union

In this section, we apply our methodology to estimate inflation persistence in the European Union, using forecast data from the equivalent European Survey of Professional Forecasts, which is a quarterly survey managed by the European Central Bank since 1999Q1. Inflation forecasts refer to Harmonized Index of Consumer Prices (HICP), which summarizes consumer price inflation across all countries in the European Union. Unlike the US SPF, though, there are no quarterly forecasts in the ECB SPF. Instead, all the forecasts refer to annual changes. Specifically, participants provide rolling-year forecasts for the inflation during a year from now, and the year afterwards. Additionally, they provide forecasts for calendar years, mainly the current and the next one (third calendar year forecasts are included only in some of the quarters). There are also long-run forecasts for the fifth or sixth calendar year.

Accordingly, we estimate inflation persistence quarter by quarter, using a cross-sectional specification based on AR(1), as in specification (4), but using annual forecasts. Denote the rolling forecasts for one and two years ahead by $F_t^i x_{1Y}$ and $F_t^i x_{2Y}$, respectively. Our regression is the following:

$$F_t^i x_{2Y} = c + \rho F_t^i x_{1Y} + u_t \tag{32}$$

There are no further horizons for which this form of specification could be estimated. However, considering our previous empirical and theoretical results, forward information does not play a role beyond a horizon of one year ahead among US professional forecasters. If this applies to professionals in the European survey as well, then (32) could provide a satisfying approximation for inflation persistence, without a concern for a bias due to forward signals, which was evident in the US SPF for shorter horizons.

On the other hand, due to the same considerations, the calendar year forecasts in the ECB SPF would not be so suitable for our purpose, because they are expected to produce a biased estimate of persistence. When forecasting for calendar years, target years would not change across the surveys conducted in Q1 to Q4 of the same year. As a result, signals are getting more informative when moving from Q1 to Q4. not just about the current year but also about the next one. This may bias the coefficient estimate due to a correlation between the right had side forecast in the regression and the error term, along similar lines demonstrated by our model in section 5.

Figure 12 describes the results from estimating specification (32) with the rolling-year forecasts. Similar to our previous procedures, we estimate (32) for each quarter using the last eight cross-sections of inflation forecasts. Panel A show the coefficient estimates, by the black line, with 95% confidence interval in the shaded area. The blue line describes the low-frequency variation with a non-parametric smoother and the dashed line show the results from the US SPF, as presented in figure 1 (for $h = 3$). Although the sample period only includes the 2000s, inflation persistence is still sizable, as in the US, fluctuating mainly between 0.4 to 0.6, and estimated very precisely. The lowest degree, around 0.2, is obtained during the Great Recession but recovers in the following years. The inflation persistence in the US has started from a higher degree (around 0.8), but became quite similar to the persistence in the European inflation, during the last years. Recall, however, the difference between quarterly and annual inflation forecasts reported in the two surveys. The R-squared from the regressions is described in panel B in the figure, and demonstrates again a co-movement with the estimated persistence – a pattern which was explained by our model of forward information.

Finally, like in figure 4, we also provide a measure trend inflation in the European Union, which is estimated from regression (32) as the constant divided by one minus the persistence coefficient. Interestingly, the estimates are very steady since 2001, a little bit below the 2%. This also matches the views about long-run inflation in the form of five year ahead forecasts provided in the survey, and reconfirms the usefulness of our method for providing a reliable measure of trend inflation as well. The dashed line shows equivalent estimates from the US SPF, as described in figure 4. As expected, US trend inflation is consistently higher, where the gap is above 0.5% at the beginning of the 2000s, but diminishes over the years.

We return again to the calendar-year forecasts in order to illustrate the role of forward information. As explained above, the information structure with respect to calendar quarters should change when moving from the Q1 survey to the Q4 survey. A simple way to handle that is to estimate the regression separately for each calendar quarter during the survey period. Denote the forecasts for inflation in the current and the next calendar years by $F_t^i x_{1C}$ and $F_t^i x_{2C}$, respectively. We could then estimate the following specification:

$$F_t^i x_{2C} = c + \rho F_t^i x_{1C} + u_t \qquad (33)$$

where we only use the forecasts provided in Q1 of each year, then, similarly, the forecasts provided in Q2, and so on. The results from this exercise are reported in table 5 (panel B). For comparison we repeat the same exercise with specification (32), using the rolling-year forecasts (panel A). Without forward information, for instance, according to the standard noisy information model, coefficient estimates should not change across calendar quarters or between specifications. In contrast, if forward information plays a role, the estimated coefficient in (33) may change when moving across calendar quarter, since the forecast horizon is getting shorter, but should not change with rolling-year forecasts as in (32), in which the forecast horizon, and thus, the information structure, is kept fixed across quarters. Strikingly, the results in table 5 confirm the pattern predicted by forward information: For the calendar-year forecasts there is a sizable reduction in the coefficient estimate when moving from Q1 to the other quarters. When (33) is estimated for only Q1 forecasts over the sample period (1999Q1-2017Q4) the coefficient estimate 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.

However, with the rolling-year forecasts coefficient estimates are higher and vary little across quarters, between 0.453 and 0.502.

Another interesting piece of evidence is provided from the US SPF data, which includes besides the quarterly forecasts, annual forecasts for calendar years, enabling the estimation of (33) using forecasts for US inflation during the same period (1999Q1-2017Q4). The results, presented in panel C in table 5, show even more clear pattern of reduction in the coefficient estimate, when moving from utilizing Q1 forecasts to Q2 forecasts, from Q2 to Q3 and from Q3 to Q4. This is a clear illustration of how forward information is accumulated when the horizon is getting shorter, leading to a growing downward bias in estimated persistence, which resembles the pattern in figure 2, when using quarterly forecasts. Notice also that when using Q1 forecasts the coefficient estimate for the US SPF is the same as the estimate for the ECB SPF – 0.429, implying the same degree of persistence for recent years, when considering a far enough horizon to rule out the forward information bias. Interestingly, the constant is a little bit higher for the US, in line with the difference in trend inflation, as observed in figure 12. Overall, these findings provide additional support for the presence of forward information in both surveys of professionals, which is limited to the approximate horizon of one year ahead.

8. Forward information as news shocks

An important strand of literature emphasizes the role of "news shocks" in driving business cycle fluctuations. The idea of news shocks is that shocks which will happen in the future may cause movements of the fundamental in the current period through the expectational channel, that is, due to information available today about the future shocks. Surprisingly, although this idea is very similar to our notion of forward information, the literature does not refer directly to expectations data for studying the nature of news. Instead, news shocks are usually identified in structural models or VARs, which are estimated only with data on fundamentals[5]. By contrast, here we provided direct evidence for the presence of forward information or news in reported expectations. Moreover, our approach could offer a systematic method for quantifying the news component in expectations data. While the literature on news shocks focuses mainly on news about technology innovations, which are considered as important in the context of business cycle models, the approach of forward information could be applied generally to expectations of any variable, in order to examine how future movements in the variable are predictable by the news component. Here we demonstrate briefly how our approach could be used to quantify news shocks, focusing mainly on the SPF inflation forecasts.

Recall the forecast decompositions obtained in equations (16) or (24) above. The standard component of the forecast is based on the underlying AR(1) process ($\rho x_{t+h-1|t}$ in (16)). The rest is an adjustment of this standard

[5] There are few exceptions. Barsky and Sims (2012) make an ad-hoc use of certain consumer confidence questions from the Michigan survey and show that confidence innovations convey new shocks which explain a sizable share of future activity. Another approach is taken by Miyamoto and Nguyen (2019) who employ a standard DSGE framework to estimate the role of news shocks, but incorporate additional data on expectations in the estimation of the model.

prediction due to available forward information. This adjustment component could therefore be treated as a news shock. More specifically, applying this idea to the current quarter average forecast in the SPF survey, news could be quantified as:

$$news_{t|t} = \bar{x}_{t|t} - (\hat{c}_{t-1} + \hat{\rho}_{t-1} x_{t-1}) \quad (34)$$

where $\bar{x}_{t|t}$ is the average forecast for the current quarter and the term in brackets is the prediction based on AR(1) process (which includes a constant). According to our previous results, the AR(1) provides a good approximation for SPF inflation forecasts. The cross-sectional method proposed in section 2 could be applied to estimate the persistence and constant parameter quarter-by-quarter, using SPF inflation forecasts. We use the estimates from specification (4) with $h = 3$ (Persistence estimates are the same estimates reported in panel A of figure 1). Importantly, we rely on our above conclusion that $h = 3$ is a long enough horizon to provide unbiased estimates of persistence. The estimates $\hat{c}_{t-1}$ and $\hat{\rho}_{t-1}$ in (34) are incorporated in lags, because the current values may include some additional news component which induces forecasters to change their perceived process[6]. Finally, we also rely on our finding in table (3) which rules out noisy information with respect to realized inflation. As a consequence, we could apply actual lagged inflation in the AR(1) component (x_{t-1}), instead of a backcast of lagged inflation. Overall, the estimation of news shocks with (34), demonstrates how the methods and results from the previous sections should be incorporated together in order to properly quantify news according to our forward information approach.

As mentioned above, in the news literature, identified news shocks should not only affect expectations, but rather predict movements in the fundamental itself. Thus, after quantifying inflation news shocks from forecast data, we examine their effect on future inflation. Table 6 provides estimates from a simple univariate autoregressions of inflation on lagged inflation and our measured news. The coefficient on news is highly significant in all estimations. The R-squared rises dramatically by more than 0.6, when adding the news variable to the auto regression (comparing columns (1) with (2) and (3) with (4)). Thus, news can explain a large share of future inflation variation. The coefficient estimates are above one, which may represent the inherent under-response of forecasters to forward information, which is imperfect by nature. Notice also the sizable increase of the coefficients on lagged inflation when adding the news variable. This could shed more light on our above finding of the consistent gap between inflation persistence based on expectations data and the persistence based on standard AR coefficients.

Next, we extend the analysis to a VAR framework. We estimate a VAR with four macroeconomic variables and our series of inflation news shocks. The four variables are inflation rate, unemployment rate, 3-month and 10-year treasury bill rates. The estimated impulse responses of the four variables to inflation news are presented in panel A of figure 13. Again, we obtain a strong response of inflation to its news. The initial response is close to

[6] The results are quite similar though, even when employing $\hat{c}_t$ and $\hat{\rho}_t$.

the estimates in table 6 and then dies out quite quickly. The response of the other variables to inflation news is quite weak and insignificant.

In order to compare our results with the literature about news shocks, we also apply the influential method suggested in Barsky and Sims (2011) for identifying news shocks. Their approach provides a useful benchmark, since it is applied within a VAR analysis. Focusing on technology shocks, they propose to identify a TFP news shock as a shock orthogonal to the current TFP innovation, which best explains future TFP movements. This general framework could be applied to identify news in other fundamentals. Hence, we apply their method to identify inflation news and estimate impulse responses to those news in a VAR framework, including the same four fundamentals we have used before (inflation, unemployment, short and long run interest rates). The estimated impulse responses are reported in panel B of figure 13. In contrast to our previous findings (panel A), the news shocks identified by the Barsky-Sims method do not induce any significant response even by inflation. Based on a forecast error variance decomposition, table 7 provides the share of inflation volatility at various horizons explained by the news shocks according to each measure. 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 news quantified by our forward information approach can explain more than a half of the same variation. This illustrates how our simple expectations-based approach could be useful in systematically quantifying news shocks which are important for explaining future movements in the fundamental.

We conclude this section with two additional results. First, we briefly examine how our approach can identify meaningful news shocks in other variables. We employ SPF forecasts of unemployment, interest rate and GDP growth and estimate a series of news shocks for each variable using (34). As for inflation, it involves the estimation of quarter-by-quarter persistence for each variable with the proper horizon, based on our results in figure 8 (constant term is also estimated). Table 8 reports the effect of those news series in a simple univariate AR. The effect of news shocks is highly significant for all three variables. This may also indicate on a long-run effect of news shocks produced by our method, which influences highly persistent variables, such as the unemployment and interest rate[7].

Second, we also consider the multi-horizon information structure suggested by our forward information framework. According to this structure, which is not explicitly modeled in the literature about news shocks, some additional news might be incorporated in the forecasts for longer horizons. Although the forecaster should smooth forward signals across forecasting horizons, according to our model in section 5, the news component in a forecast for $t + h$ could still have some additional predictive ability for this specific horizon, which could be identified and estimated. Equation (34) extracts news only from current quarter forecasts. A simple generalization to longer horizons could be as follows:

[7] We also considered the unit root possibility, by taking the differenced series of unemployment and interest rates and running them on the news series. The coefficient on news was again highly significant in these specifications.

$$news_{t+h|t} = \bar{x}_{t+h|t} - \left(\sum_{k=0}^{k=h-1} \hat{\rho}_{t-1}{}^{k} \hat{c}_{t-1} + \hat{\rho}_{t-1}{}^{h+1} x_{t-1} \right) \tag{35}$$

Thus, the standard prediction component based on the underlying AR(1) process is the same as in (34) except that the AR(1) prediction is moved forward to period $t + h$. The news component is then the deviation of the forecast for this period from the AR(1) prediction. Another approach could track more closely equation (16) and isolate only the news beyond those already embedded in forecasts for shorter horizons. Specifically:

$$news_{t+h|t} = \bar{x}_{t+h|t} + \left(\hat{c}_t - \hat{\rho}_t \bar{x}_{t+h-1|t} \right) \tag{36}$$

Thus, the realization of x_{t-1} is replaced by the mean forecast for period $t + h - 1$. The AR(1) parameters are those reflected by the forecasts of time t. Those parameters were also already embedded in forecast for period $t + h - 1$.

We compute news shocks for inflation using the additional horizons available in the SPF beyond the current quarter. We then examine their predictive ability for inflation variation at time $t + h$. The results based on a univariate framework are reported in table 9. Each column reports the results from a regression of inflation at time $t + h$ on lagged inflation and all the news components available at time t which refer to $t + h$ or before. This could be viewed as an estimation of the impulse responses to news at the various horizons. Significant coefficient is documented for $h = 1$, especially with the second approach. For longer horizons the effect is not significant, but so is the effect of lagged inflation. Results for other variables, which are not reported here show significance of quantified news even for longer horizons in some cases. In sum, these findings suggest that the multi-horizon structure of information should also be taken into account, when analyzing the role of news shocks. Our approach provides a straightforward way to quantify news at any forecast horizon, for this purpose.

9. Conclusion

This paper proposes a new approach to characterize expectations formation as a combination of model-based prediction and subjective adjustment based on available forward information. We rule out the need to introduce deficiencies in expectations in order to obtain forecast heterogeneity and error predictability. Instead they could be driven by signals about the future, which must be imperfect by their nature. We bring new evidence from the SPF survey forecasts, supporting the presence of forward information. We provide a simple method for quantifying the forward information component, which could also be useful for studying the role of news shocks in business cycle dynamics.

Our approach has important implications for recent macroeconomic policies using Central bank communication, as it offers an explicit framework to analyze the effect on expectations in the form of forward signals. It also sheds light on the New Keynesian Phillips Curve which still seems to perform better when it is based on nominal rigidities rather than rigidities in expectations.

Our new measure of inflation persistence provides the possibility of tracking important changes in this key parameter over-time. Our findings suggest that the importance of inflation persistence in recent years was under-estimated, demanding more efforts from monetary policy makers in order to further improve control over inflation.

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, and Jennifer La'O (2013). "Sentiments." *Econometrica* 81 (2), 739–79.

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 (2006). "Stock Prices, News, and Economic Fluctuations." *American Economic Review,* 96 (4), 1293–1307.

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

Benati, Luca (2008). "Investigating Inflation Persistence Across Monetary Regimes." *Quarterly Journal of Economics*, 123 (3), 1005-1060.

Benhabib, Jess, Pengfei Wang, and Yi Wen (2015). "Sentiments and Aggregate Demand Fluctuations." *Econometrica,* 83 (2), 549–85.

Blanchard, Olivier J., Jean-Paul L'Huillier, and Guido Lorenzoni (2013). "News, Noise, and Fluctuations: An Empirical Exploration." *American Economic Review,* 103 (7), 3045–70.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2018). "Diagnostic expectations and Credit Cycles." *Journal of Finance*, 73 (1), 199-227.

Bordalo, Pedro, Nicola Gennaioli, Yuaren Ma and Andrei Shleifer (2018). "Over-reaction in Macroeconomic Expectations." NBER working paper No. 24932.

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.

Chan, Joshua C.C., Todd E. Clark, and Gary Koop (2018). "A New Model of Inflation, Trend Inflation, and Long-Run Inflation Expectations." *Journal of Money, Credit and Banking*, 50 (1), 5-53.

Cogley, T. and Thomas J. Sargent (2001). "Evolving post-World War II U.S. inflation dynamics." In B. S. Bernanke and K. Rogoff (Eds.), *NBER Macroeconomics Annual 2001*, pp. 331—338. Boston, MA: National Bureau of Economic Research, Inc.

Cogley, Timothy. W., Georgio Primiceri, and Thomas J. Sargent (2010). "Inflation-Gap Persistence in the U.S." *American Economic Journal: Macroeconomics,* 2 (1), 43—69.

Cogley, Timothy W., and Argia M. Sbordone (2008). "Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve." *The American Economic Review* 98 (5), 2101-2126.

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, 1447-1491.

Coibion, Olivier, Yuriy Gorodnichenko, and Saten Kumar (2018). "How do Firms Form Their Expectations? New Survey Evidence" *American Economic Review*, 108 (9), 2671-2713.

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.

Evans, George W., and Seppo Honkapohja (2012). *Learning and Expectations in Macroeconomics*. Princeton University Press.

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 North-Holland Publishing Co, 1979-1995

Faust, Jon, and Jonathan H., Wright (2013). "Forecasting Inflation," in *Handbook of Economic Forecasting*, Volume II, edited by Graham Elliot and Allan Timmermann, Elsevier North-Holland Publishing Co, 2-56

Fuhrer, J. 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 Papers, No. 18-9.

Fuster, Andreas, David Laibson, and Brock Mendel (2010). "Natural Expectations and Macroeconomic Fluctuations." *Journal of Economic Perspectives*, 24 (4), 67-84.

Gabaix, Xavier (2019). “Behavioral Inattention.” In: *Handbook of Behavioral Economics: Applications and Foundations 1*, edited by B. Douglas Bernheim, Stefano DellaVigna, David Laibson, Elsevier: North Holland, Vol. 2, 261-343.

Goldstein, Nathan (2019). "Tracking Inattention: A New Expectation-Based Measure." Manuscript.

Gorodnichenko, Yuriy (2008). "Endogenous Information, Menu Costs and Inflation Persistence." NBER Working Paper 14184.

Jaimovich, Nir, and Sergio Rebelo (2009). “Can News about the Future Drive the Business Cycle?” *American Economic Review*, 99 (4), 1097–1118.

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

Lucas, Robert E. (1972). “Expectations and the Neutrality of Money.” *Journal of Economic Theory*, 4 (2), 103–24.

Maćkowiak, Bartosz, and Mirko Wiederholt (2009). “Optimal Sticky Prices under Rational Inattention.” *The American Economic Review*, 99 (3), 769-803.

Mankiw, N. Gregory, and Ricardo Reis (2002). “Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve.” *Quarterly Journal of Economics*, 117 (4), 1295-1328.

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

Pivetta, Frederic. and Ricardo Reis (2007). "The Persistence of Inflation in the United States." *Journal of Economic Dynamics and Control*, 31(4), 1326—1358.

Ryngaert, Jane (2017). "What Do (and Don’t) Forecasters Know About U.S. Inflation?" Manuscript.

Schmitt-Grohé, Stephanie, and Martín Uribe (2012). “What’s News in Business Cycles.” *Econometrica*, 80 (6), 2733–64.

Sims, Christopher A. (2003). “Implications of Rational Inattention.” *Journal of Monetary Economics*, 50 (3), 665-690.

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.

Stock, James H., and Mark W. Watson (2007). "Why has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking*, 39, 3—33.

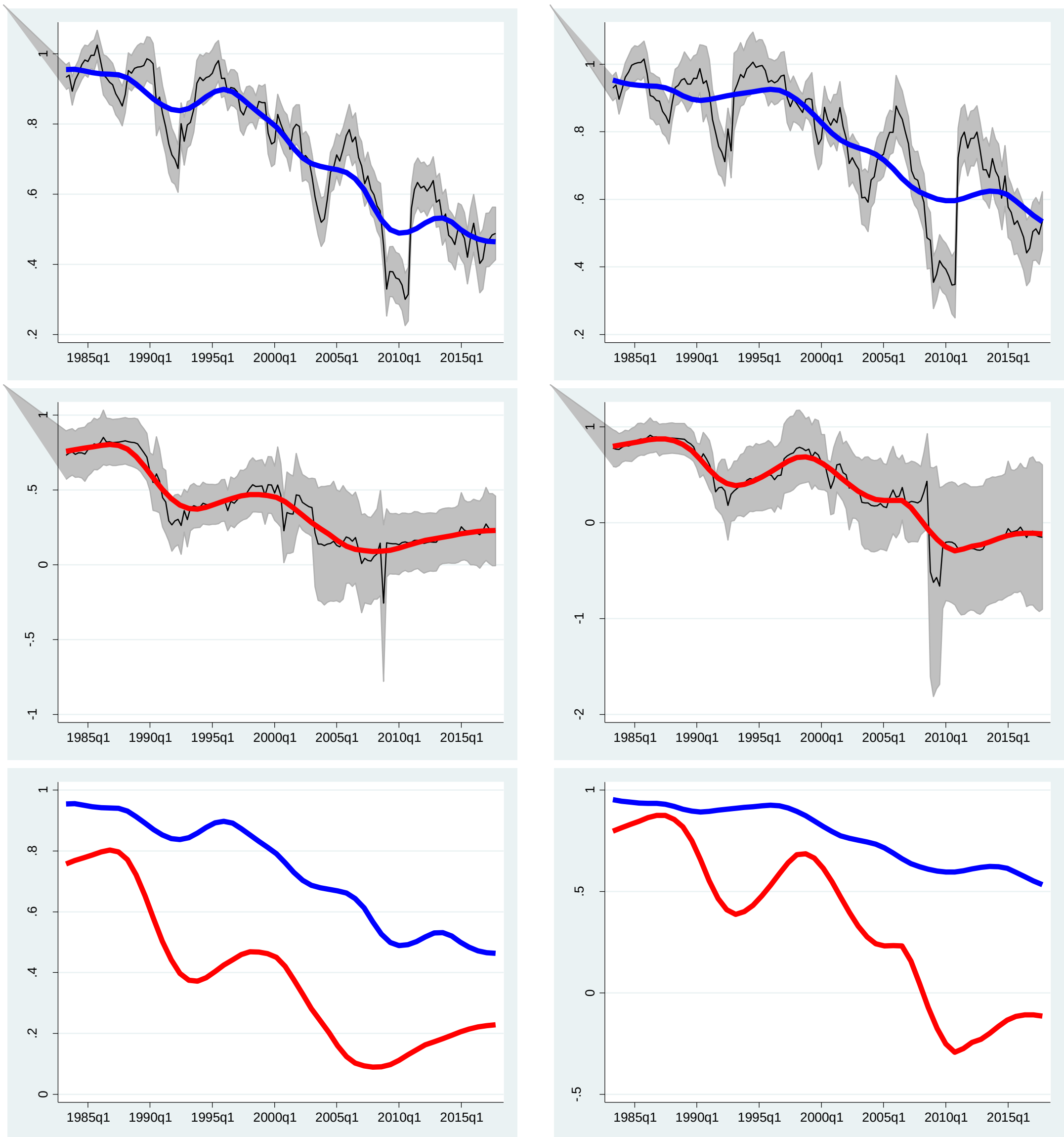
Stock, James H., and Mark W. Watson (2016). "Core Inflation and Trend Inflation." *Review of Economics and Statistics*, 98 (4), 770—784.

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: Estimates of Inlation Persistence over Time

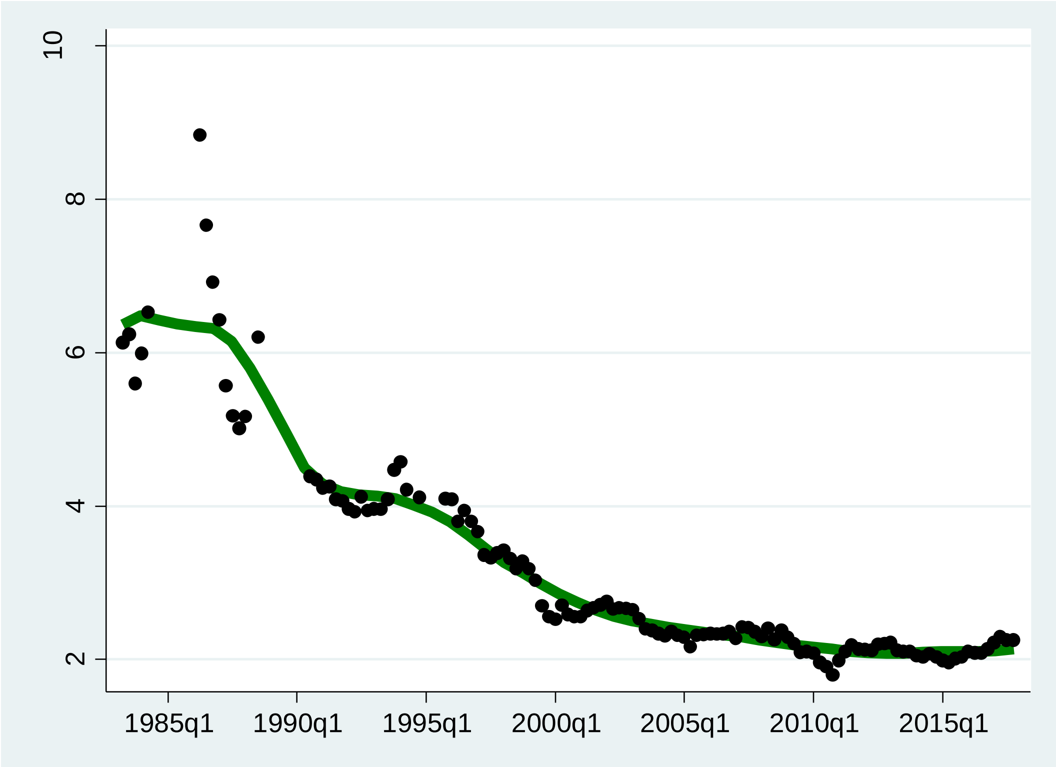
Panel A: "AR(1)" Coefficient

Panel B: "AR(4)" Sum of Coefficients



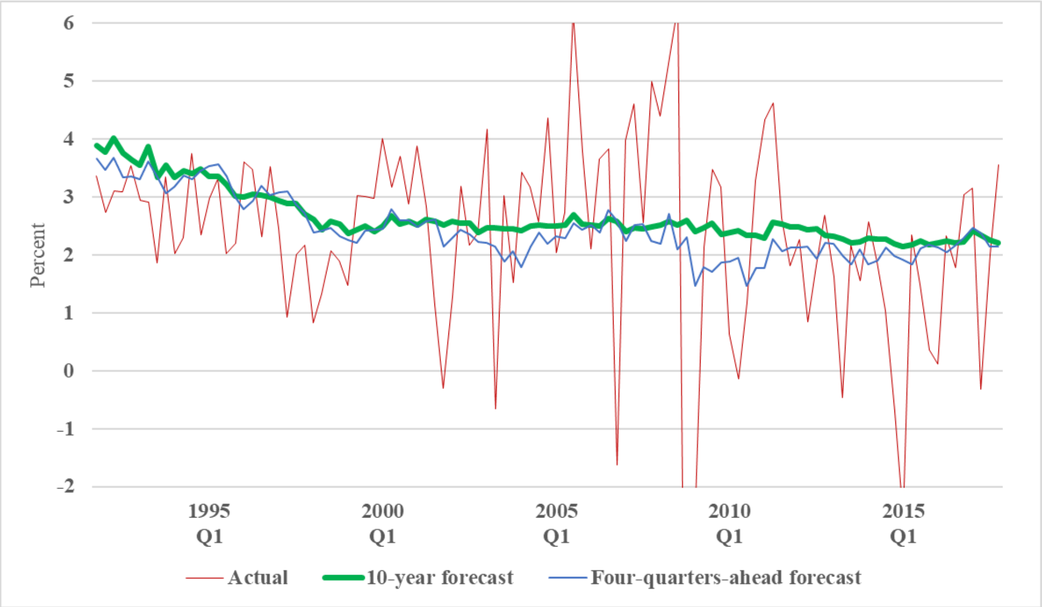
Notes: The figure plots estimates of inflation persistence over time, based on AR(1) and AR(4) specifications (panels A and B on the left and right columns, respectively). First row presents expectation-based measures based on specifications (1), for a forecast horizon $h = 3$. Each point on the black lines is an OLS estimate using the forecasts data from the last 8 quarters with the shaded area showing the 95% confidence interval. The blue lines are local mean smoothers which uses Epanechnikov kernel. Second row presents persistence measures from actual inflation data, based on a rolling-window estimation of auto-regressions, with a window size of the last 40 quarters. Confidence interval in the shaded area is based on Newey-West standard errors. The red lines are local mean smoothers which uses Epanechnikov kernel. The third row compares the inflation persistence estimates from the first and second rows.

Figure 2: Expectations-Based Estimates of Trend Inflation



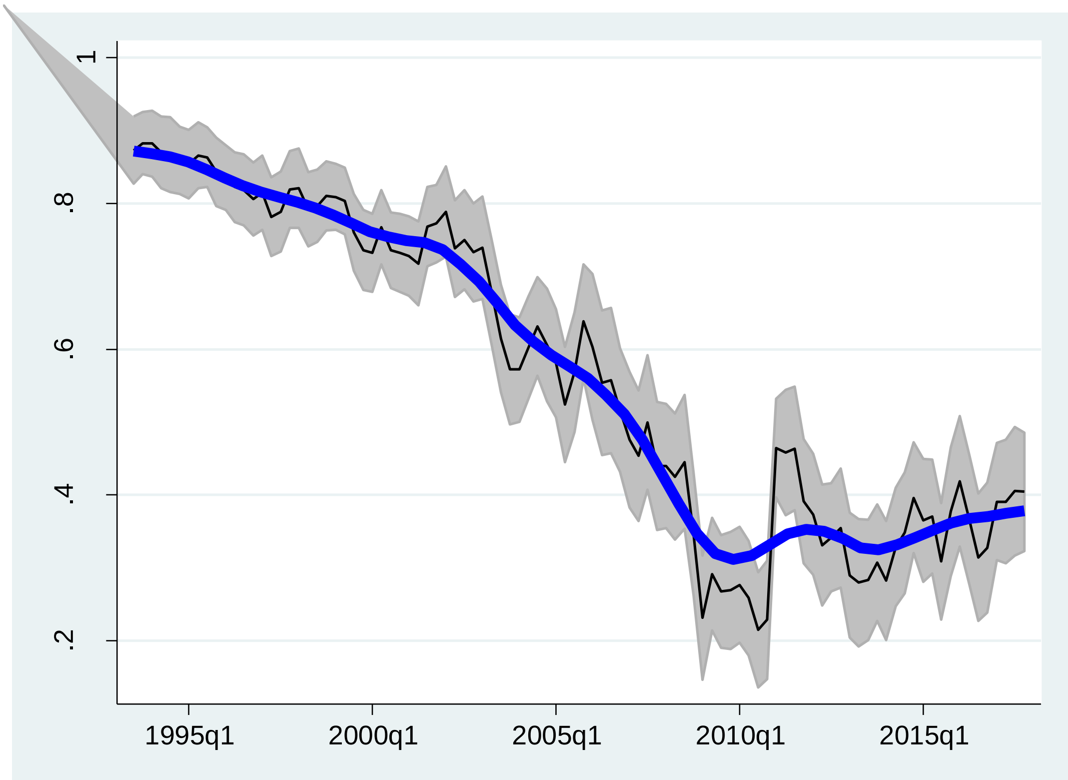
Notes: The figure plots estimates of trend inflation over time, based on specification (1), for a forecast horizon $h = 3$. Trend inflation is estimated as the regression constant divided by one minus the persistence coefficient. Estimates which are not significantly different from one are excluded. Each black point is based on OLS estimation using the forecasts data from the last 8 quarters. The green line is a local mean smoother of the black points which uses Epanechnikov kernel.

Figure 3: Forecasts of Inflation Gap



Notes: The figure plots actual CPI inflation data (annualized quarterly rates), and average forecast of inflation for 10-year and four-quarters ahead from the Survey of Proffesional Forecasts.

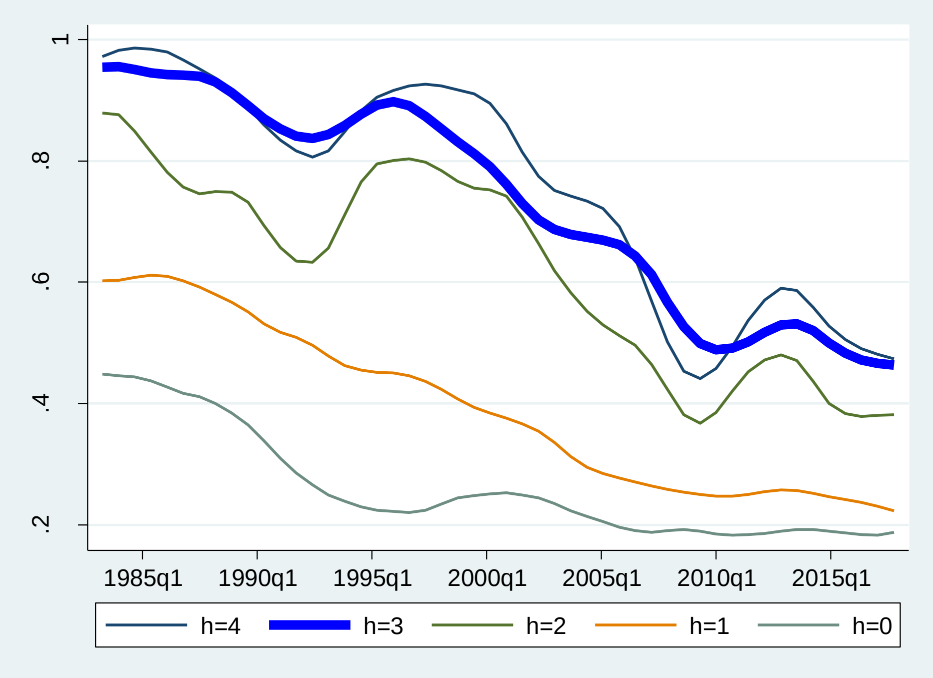
Figure 4: Estimates of Inlation-Gap Persistence over Time



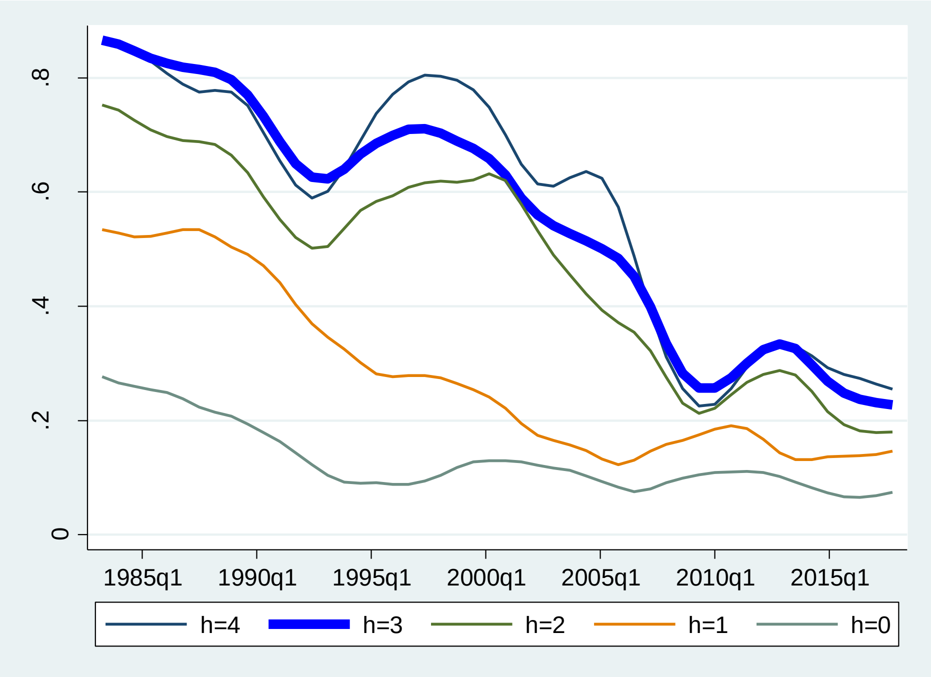
Notes: The figure plots estimates of inflation-gap persistence over time, based on specification (3) for a forecast horizon $h = 3$. The inflation-gap is measured as the difference between the forecast for h (or $h - 1$) quarters ahead and the 10-year forecast. Each point on the black lines is an OLS estimate using the forecasts data from the last 8 quarters with the shaded area showing the 95% confidence interval. The blue line is a local mean smoother which uses Epanechnikov kernel.

Figure 5: Expectation-Based Persistence by Forecasting Horizon

Panel A: Persistence Estimates



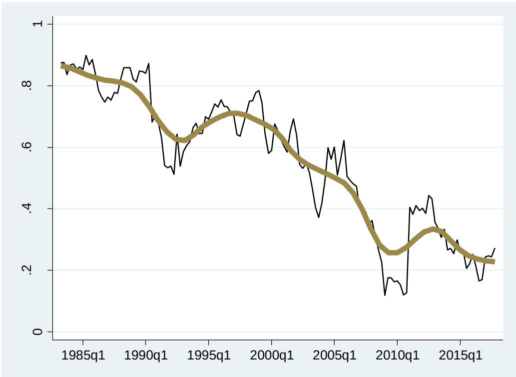
Panel B: R-Squared of Persistence Regressions



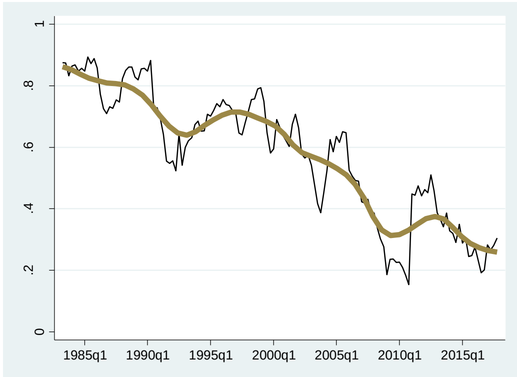
Notes: The figure plots smoothed persistence coefficients (panel A) and R-squared measures (panel B) based on estimating specification (4) for different forecast horizon in the SPF survey. Each point on the lines is based on OLS estimation using the forecasts data from the last 8 quarters. The smoother is a local mean which uses Epanechnikov kernel.

Figure 6: The Declining Trend of R-Squared Measures

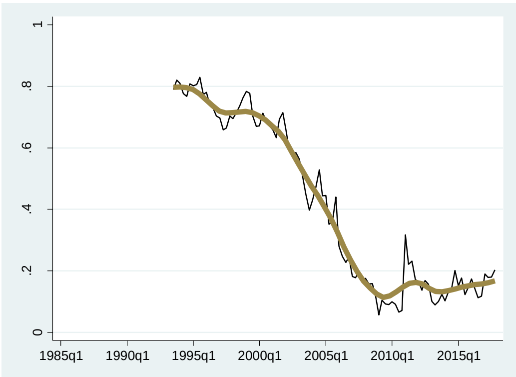
Panel A: "AR(1)"



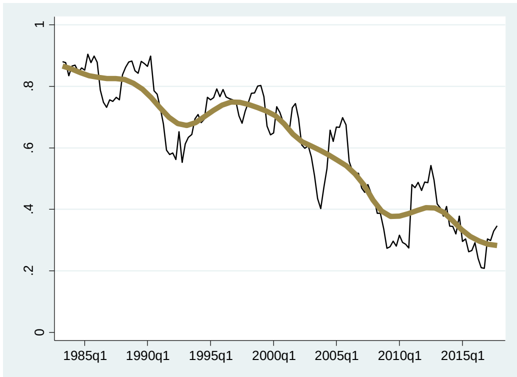
Panel B: "AR(4)"



Panel C: "AR(1)" with Inflation-Gap



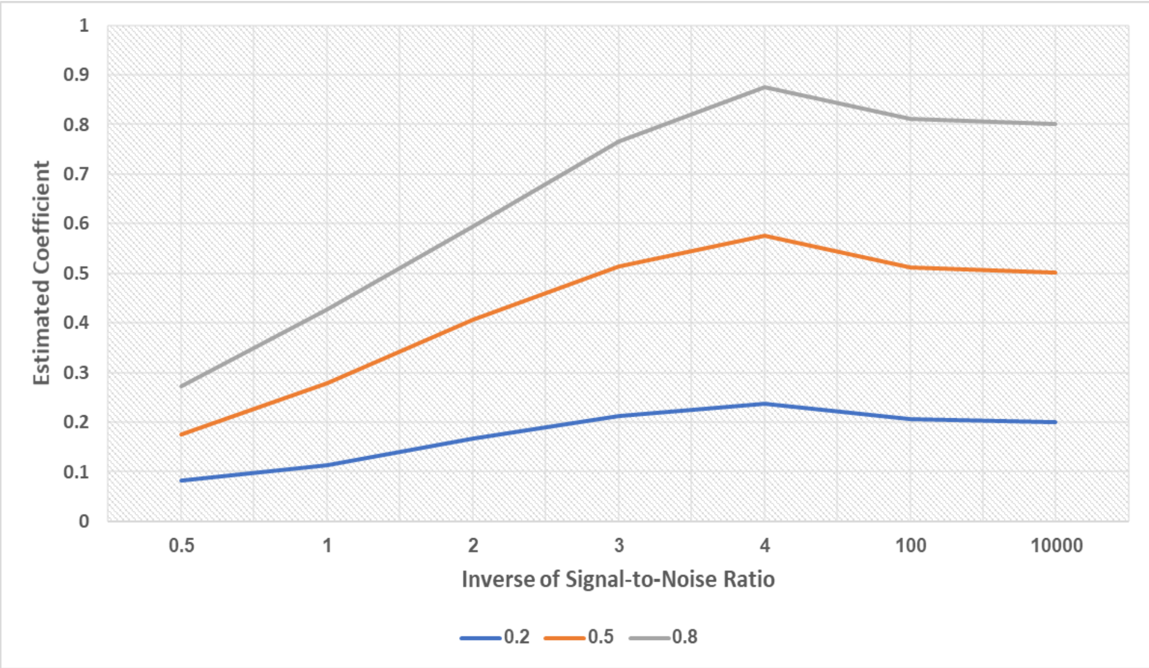
Panel D: "VAR(4)"



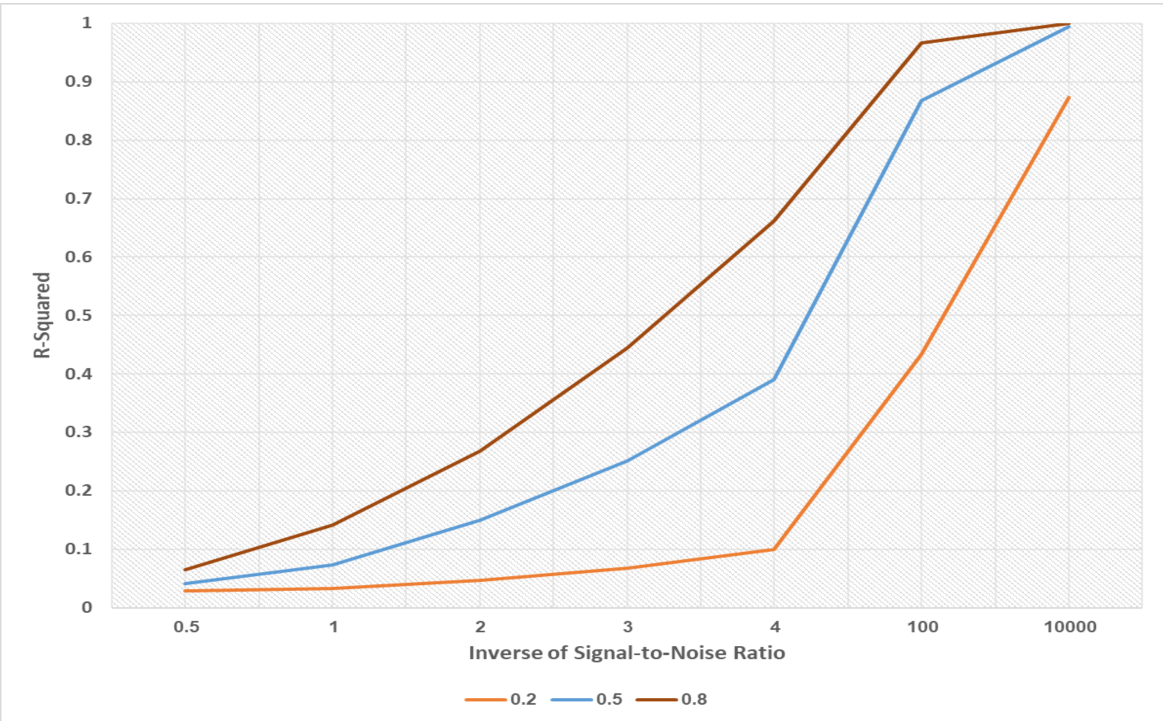
Notes: The figure plots R-squared measures from four different specifications which were estimated for each quarter of the sample period. In panel A inflation expectations follow an AR(1) process as in specification (4). In panel B inflation expectations follow an AR(4) process as in specification (1). In panel C expectations of the inflation-gap follow an AR(1) process as in specification (3) (sample starts on 1990Q2 because of 10-year forecasts). In panel D inflation expectations follow a VAR(4) process which augments specification (1) 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 Epanechnikov kernel.

Figure 7: Simulation Results

Panel A: Persistence Estimates



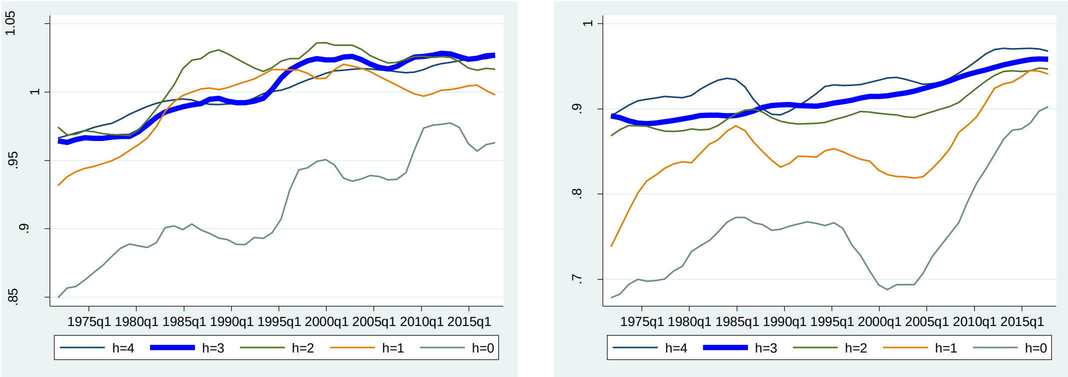
Panel B: R-Squared of Persistence Regressions



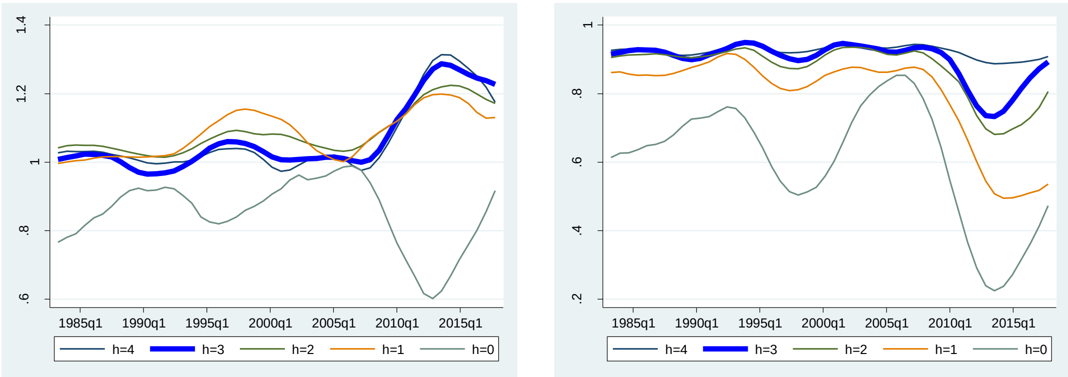
Notes: The two panels in the figure show estimation results of specification (4), applied to a simulated data of forecasts, according to the model presented in section 5. Coefficient estimate (panel A) and R-squared (panel B) are averaged across 1000 draws of the simulation. Each line corresponds to a different degree of the state persistence in the simulated data. 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 inverse of the signal-to-noise ratio (which equals the variance of the noise) is indicated by the horizontal axis. The ratio for $h = 3$ is 0.2.

Figure 8: Persistence Patterns of Additional Macroeconomic Variables

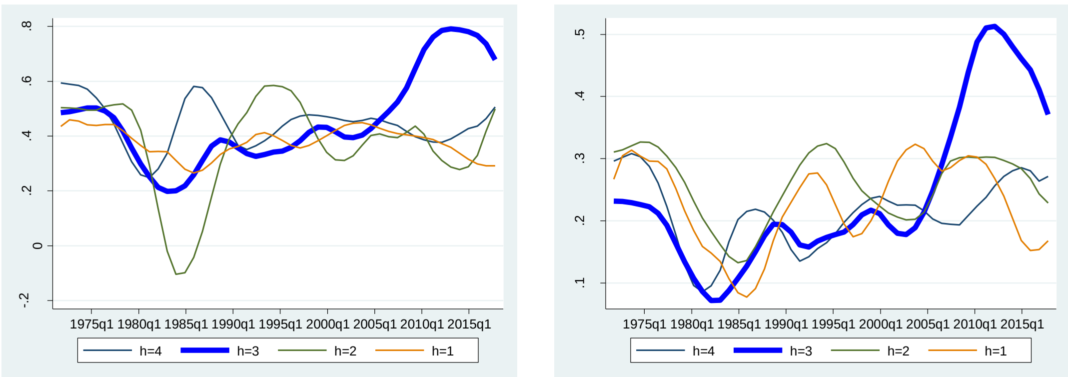
Panel A: Unemployment Rate



Panel B: Interest Rate



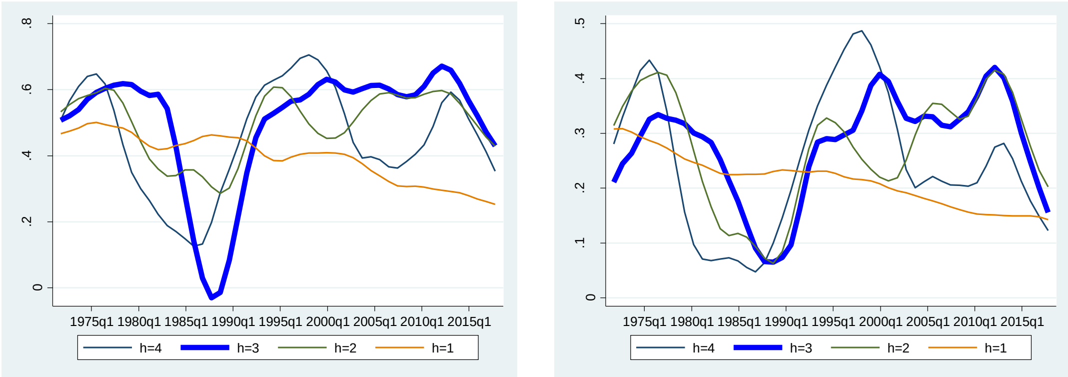
Panel C: Real GDP Growth



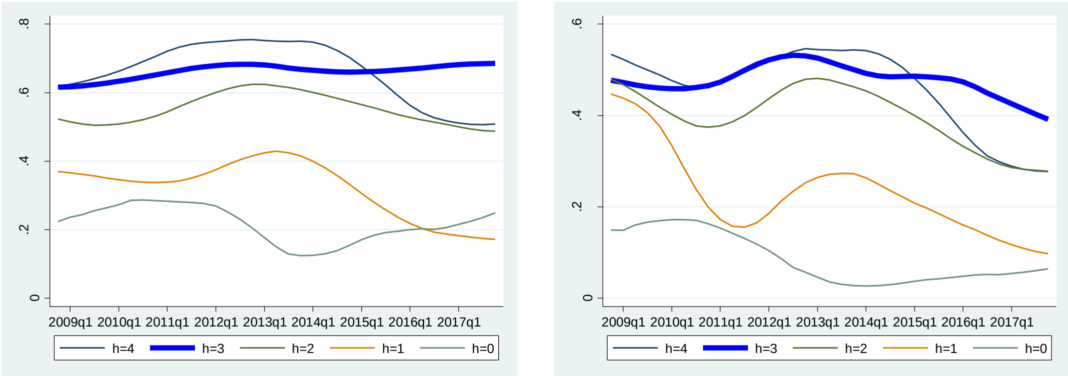
Notes: The figure plots smoothed persistence coefficients and R-squared measures, based on estimating specification (4) for different forecast horizons, and using different macroeconomic variables from the SPF survey as specified in each panel. Each point on the lines is based on OLS estimation using the forecasts data from the last 8 quarters. The smoother is a local mean which uses Epanechnikov kernel.

Figure 9: Persistence Patterns of Additional Inflation Variables

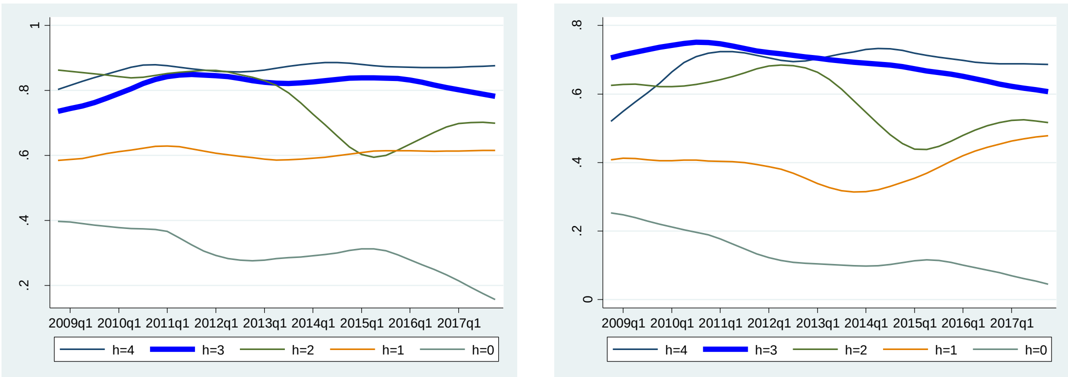
Panel A: GDP Deflator



Panel B: PCE Inflation



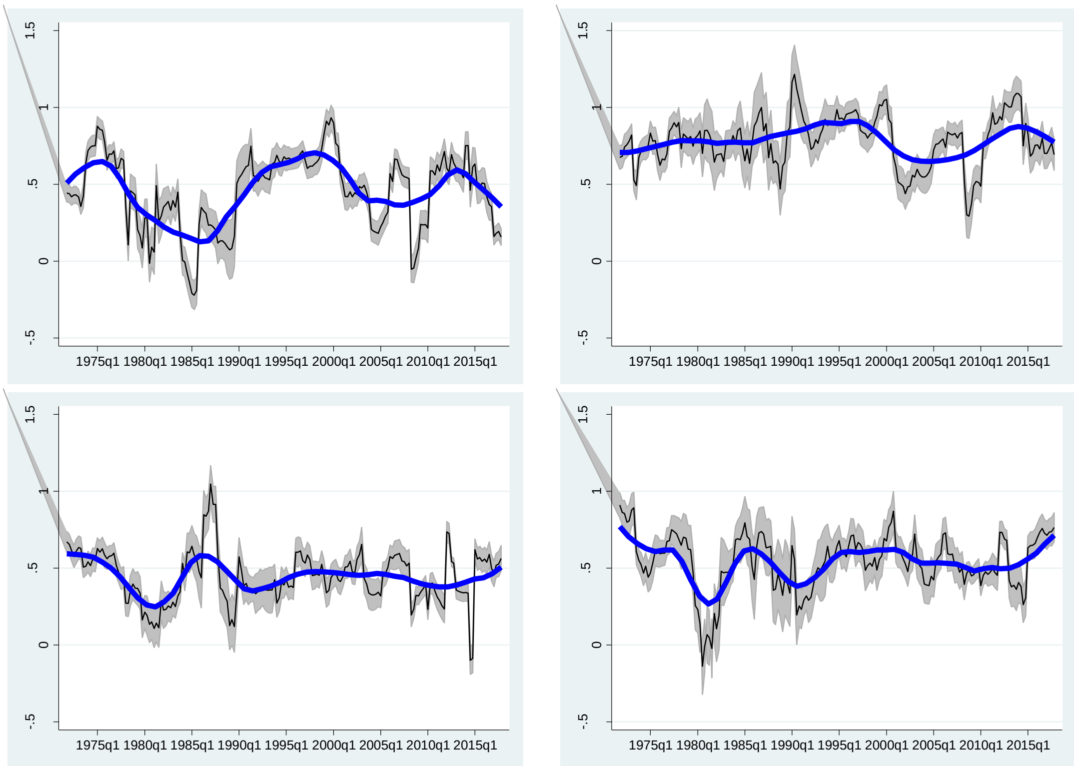
Panel C: Core Inflation



Notes: The figure plots smoothed persistence coefficients and R-squared measures, based on estimating specification (4) for different forecast horizons, and using different inflation variables from the SPF survey as specified in each panel. Each point on the lines is based on OLS estimation using the forecasts data from the last 8 quarters. The smoother is a local mean which uses Epanechnikov kernel.

Figure 10: Sensitivity of Persistence Estimates to Specification

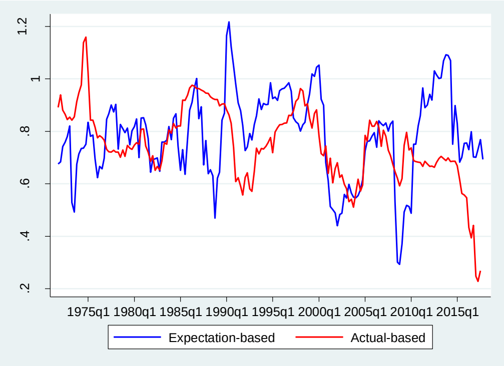
Panel A: "AR(1)" Coefficient | Panel B: "AR(4)" Sum of Coefficients



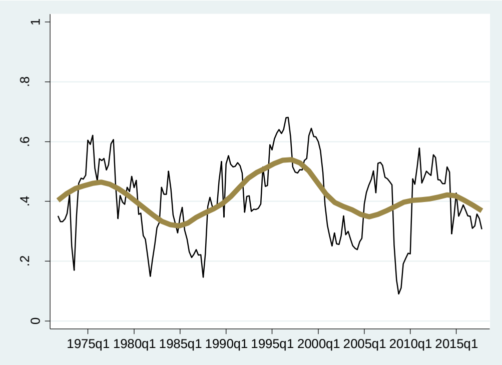
Notes: The figure plots estimates of GDP deflator (first row) and GDP growth (second row) persistence over time, based on AR(1) and AR(4) specifications (panels A and B on the left and right columns, respectively). Persistence measures are based on specifications (4) and (1) for a forecast horizon $h = 4$. Each point on the black lines is an OLS estimate using the forecasts data from the last 8 quarters with the shaded area showing the 95% confidence interval. The blue lines are local mean smoothers which uses Epanechnikov kernel.

Figure 11: Persistence of GDP Deflator

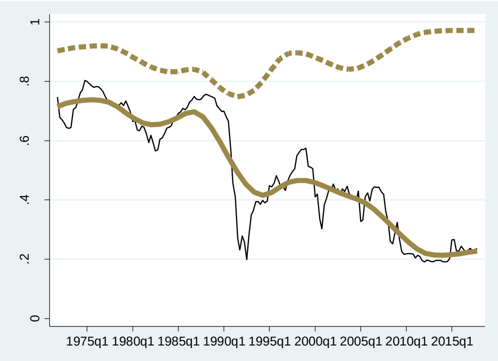
Panel A: GDP Deflator Persistence



Panel B: R-squared



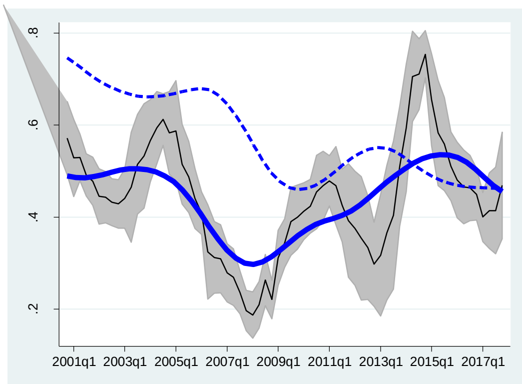
Panel C: CPI and GDP Inflation



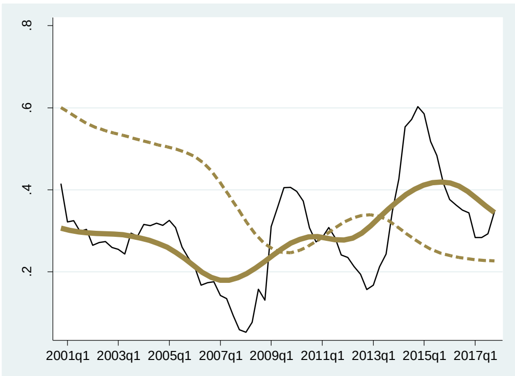
Notes: The figure plots estimates which are related to the persistence of the GDP-deflator inflation. Panel A plots two measures of persistence. The red line shows coefficient estimates from a rolling AR(4) regression using actual inflation data of GDP-deflator (window size is 40 quarters). The blue line presents expectation-based measures of persistence, using SPF forecasts of the GDP deflator. Each point is the sum of the coefficients in specification (XX), estimated by an OLS for a forecast horizon $h = 4$, using the forecasts data from the last 8 quarters. Panel B plots, by a black line, the R-squared measues from those expectations regressions. The brown line is a local mean smoother which uses Epanechnikov kernel. Panel C plots , by a black line, a rolling-window coefficient-of-correlation between the GDP-deflator and CPI inflation, using actual inflation data with a window size of 40 quarters. The brown line is a local mean smoother which uses Epanechnikov kernel. The dashed brown line is a smoothed rolling-window coefficient-of-correlation between PCE and CPI inflation.

Figure 12: Inflation Persistence in the European Union

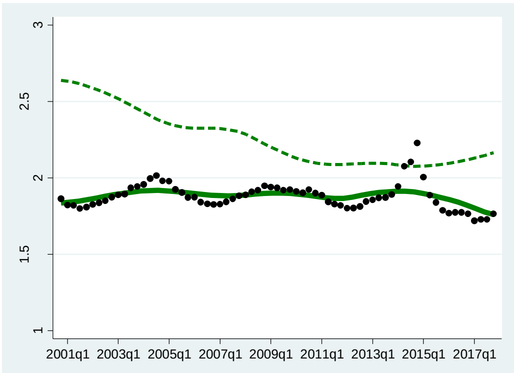
Panel A: Inflation Persistence



Panel B: R-squared



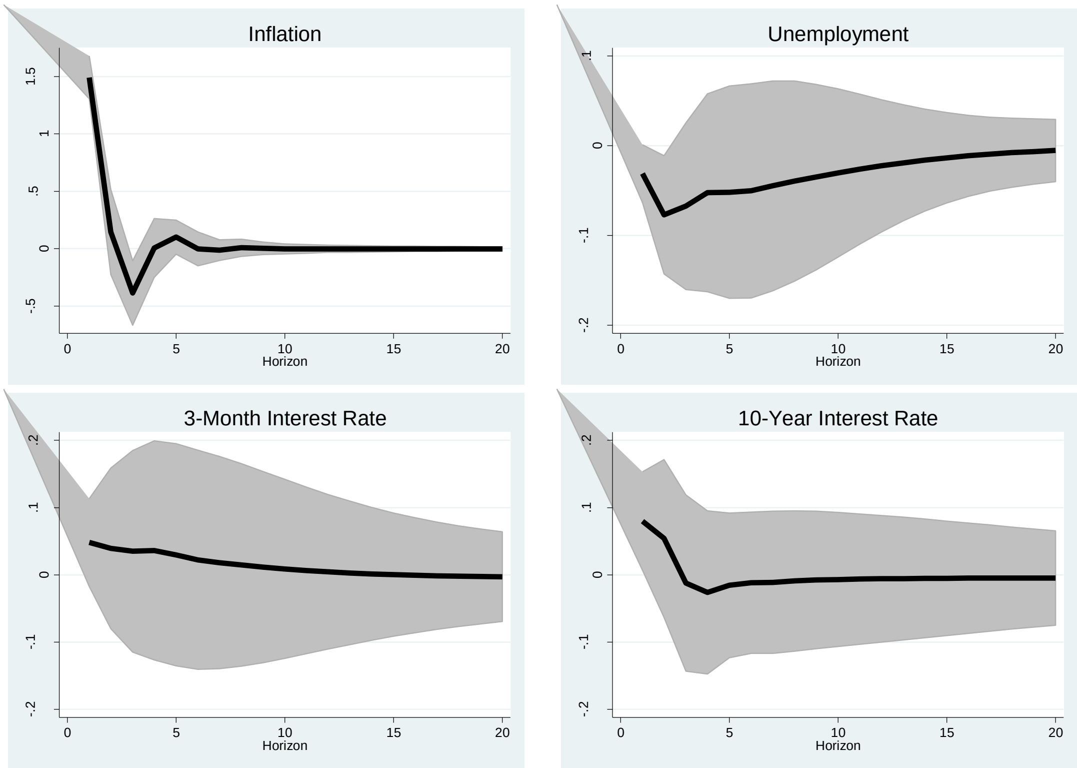
Panel C: Trend Inflation



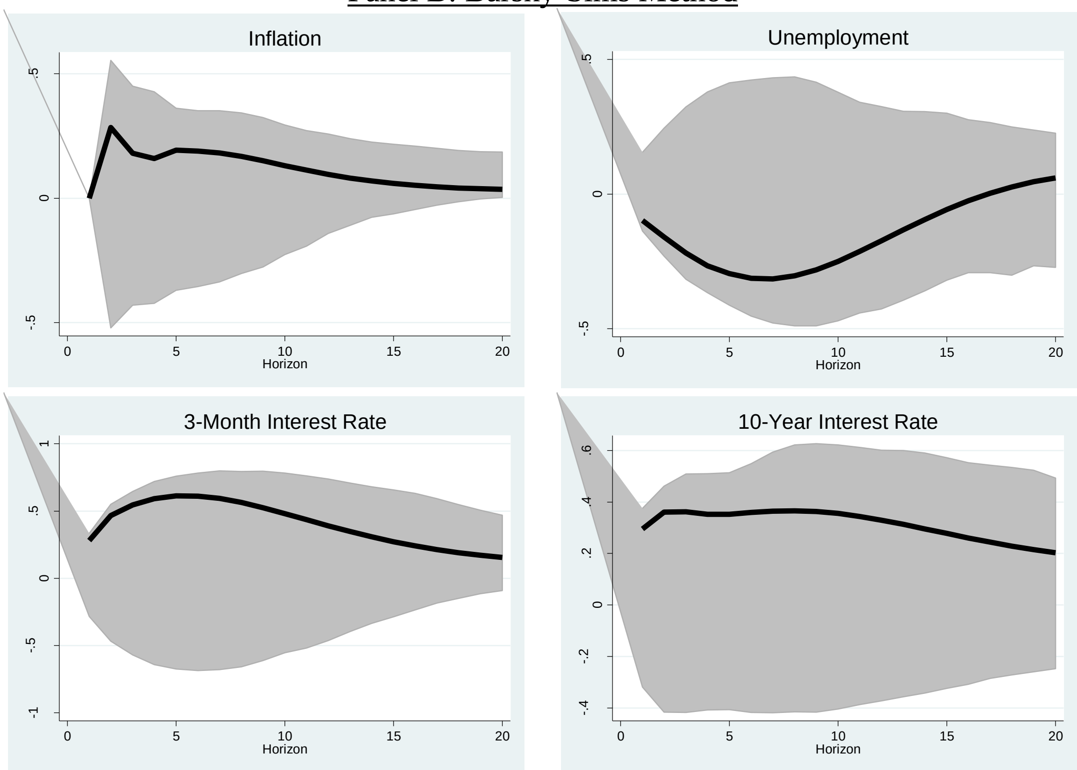
Notes: The figure plots estimates of inflation persistence (panel A), R-squared measure (panel B) and trend inflation (panel C), based on specification (34) and using rolling annual inflation forecasts for the subsequent two years from the European Central Bank SPF. Specification (34) is estimated by OLS for each quarter, using data from the last 8 quarters. Panel A plots the estimated persistence coefficients by the black line, with the shaded area showing the 95% confidence interval, and a local mean smoother which uses Epanechnikov kernel by the solid blue line. Panel B plots the R-squared of the regressions estimated for each quarter (black line) , and a local mean smoother which uses Epanechnikov kernel (solid brown line). Panel C plots trend inflation estimated as the regression constant divided by one minus the persistence coefficient (black dots), and a local mean smoother which uses Epanechnikov kernel (solid green line). The dashed lines in each panel replicates corresponding results from the US SPF, which are presented in figures 1.A. (top row), 6A and 2. Respectively.

Figure 13: Impulse Responses to Inflation News Shocks

Panel A: Forward Information Method



Panel B: Barsky-Sims Method



Notes: The figure plots impulse responses to inflation news shocks in a VAR system estimated on the 1983Q1-2017Q4 sample. Inflation news shocks are obtained by the forward information method in panel A, or by the Barsky-Sims 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.

TABLE 1

Out-of-Sample Performance of Inflation Predictions, Produced by Various Measures of Inflation Persistence

Specification	Current Quarter	4 Quarters Ahead	8 Quarters Ahead
Forecast, AR(1)	1.000	1.000	1.000
Actual, AR(1), Window: 40 obs.	1.052	1.160	1.255
Actual, AR(1), Window: 20 obs.	1.233	1.075	1.148
Actual, AR(1), Window: 60 obs.	1.046	1.062	1.148
Forecast, AR(4)	1.003	1.020	1.011
Actual, AR(4), Window: 40 obs.	1.099	1.165	1.278
Actual, AR(4), Window: 20 obs.	1.583	1.460	1.872
Actual, AR(4), Window: 60 obs.	1.074	1.184	1.346

Notes: The table reports out-of-sample predictive performance from various specifications. All specifications are estimated for each quarter in the sample period (1983Q2 to 2017Q4). Based on the auto-regressive coefficients produced by each specification, inflation predictions are calculated for three different horizons, as indicated at the top of the table columns. The statistics reported in the table are mean squared prediction errors from each specification divided by the mean squared error from the first specification at the same horizons. Specifications use either actual CPI inflation or SPF forecasts. When using actual data specification follows AR(1) or AR(4) process and estimated by a rolling-window procedure, using different window sizes as indicated in the first column of the table. When using forecast data estimation follows specification (1) in the main text to mimic AR(1) and AR(4) processes, respectively. The forecast specifications are estimated for each quarter using the forecasts data from the last 8 quarters.

TABLE 2

Summary of regression properties in the example of section 5.2

	Forecast Horizon			Persistence	
	$h = 1$	$h = 2$	$h = 3$	$\rho = 0$	$\rho = 1$
Coefficient	0	Positive but Biased	ρ	0	Positive but Biased
R-squared	0	Between 0 and 1	1	0	Between 0 and 1

Notes: The table summarizes the theoretical predictions by the simple case of the model analyzed in section 5.2. Model predictions refer to the coefficient and R-squared of a regression of the forecast $x_{t+h|t}$ on $x_{t+h-1|t}$.

TABLE 3

Regressions of the deviation from the mean forecast

	Whole Sample	1980s	1990s	2000s	2010s
Panel A: Dependent variable: $x_{t\|t} - \bar{x}_{t\|t}$ (backcasts)					
$x_{t\|t-1} - \bar{x}_{t\|t-1}$	0.023	0.128	0.019**	0.000	0.001**
	(0.015)	(0.088)	(0.008)	(0)	(0.001)
$x_{t+1\|t-1} - \bar{x}_{t+1\|t-1}$	0.006	-0.007	0.03**	0.000	0.001
	(0.016)	(0.085)	(0.015)	(0)	(0.001)
$x_{t+2\|t-1} - \bar{x}_{t+2\|t-1}$	-0.016	-0.115	-0.028**	0.001	0.000
	(0.014)	(0.08)	(0.011)	(0.001)	(0.002)
$x_{t+3\|t-1} - \bar{x}_{t+3\|t-1}$	0.015	0.114	0.015	-0.001**	0.000
	(0.013)	(0.079)	(0.013)	(0)	(0.001)
$x_{t+4\|t-1} - \bar{x}_{t+4\|t-1}$	-0.005	-0.053	-0.010	0.000	0.001
	(0.016)	(0.043)	(0.01)	(0.001)	(0.001)
Constant	0.004	0.025	0.002	0.000	0.001
	(0.006)	(0.045)	(0.002)	(0)	(0.002)
Obs.	3849	559	1068	1272	950
R^2	0.004	0.019	0.025	0.004	0.002

Notes: The table reports coefficient estimates from regressions of the individual deviation from the mean forecast. 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 (continued)

Regressions of the deviation from the mean forecast

	Whole Sample	1980s	1990s	2000s	2010s
Panel B: Dependent variable: $x_{t+1\|t} - \bar{x}_{t+1\|t}$					
$x_{t\|t-1} - \bar{x}_{t\|t-1}$	0.037	0.005	0.024	0.006	0.125**
	(0.026)	(0.05)	(0.032)	(0.026)	(0.048)
$x_{t+1\|t-1} - \bar{x}_{t+1\|t-1}$	0.258***	0.091	0.331***	0.268***	0.376***
	(0.056)	(0.072)	(0.055)	(0.09)	(0.057)
$x_{t+2\|t-1} - \bar{x}_{t+2\|t-1}$	0.124**	0.17**	-0.010	0.118	0.146***
	(0.05)	(0.068)	(0.051)	(0.09)	(0.043)
$x_{t+3\|t-1} - \bar{x}_{t+3\|t-1}$	-0.092**	-0.043	0.1**	-0.136**	0.010
	(0.053)	(0.125)	(0.049)	(0.075)	(0.052)
$x_{t+4\|t-1} - \bar{x}_{t+4\|t-1}$	0.066	-0.063	-0.002	0.279***	-0.061
	(0.069)	(0.101)	(0.037)	(0.071)	(0.098)
Constant	-0.008	0.008	0.003	-0.015	-0.009
	(0.009)	(0.042)	(0.006)	(0.017)	(0.008)
Obs.	3854	566	1068	1271	949
R^2	0.091	0.026	0.142	0.121	0.168

Notes: The table reports coefficient estimates from regressions of the individual deviation from the mean forecast. 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 (continued)

Regressions of the deviation from the mean forecast

	Whole Sample	1980s	1990s	2000s	2010s
Panel C: Dependent variable: $x_{t+2\|t} - \bar{x}_{t+2\|t}$					
$x_{t\|t-1} - \bar{x}_{t\|t-1}$	0.005	0.025	-0.027	0.010	0.002
	(0.022)	(0.044)	(0.034)	(0.033)	(0.038)
$x_{t+1\|t-1} - \bar{x}_{t+1\|t-1}$	0.086**	-0.057	0.147***	0.134***	0.132***
	(0.038)	(0.093)	(0.038)	(0.046)	(0.041)
$x_{t+2\|t-1} - \bar{x}_{t+2\|t-1}$	0.468***	0.377***	0.273***	0.504***	0.527***
	(0.053)	(0.069)	(0.048)	(0.081)	(0.036)
$x_{t+3\|t-1} - \bar{x}_{t+3\|t-1}$	-0.077	-0.24**	0.136***	-0.073	0.092
	(0.055)	(0.111)	(0.043)	(0.051)	(0.082)
$x_{t+4\|t-1} - \bar{x}_{t+4\|t-1}$	0.069	0.171**	0.034	0.17**	-0.015
	(0.05)	(0.094)	(0.046)	(0.093)	(0.072)
Constant	-0.001	0.009	0.007	0.000	-0.010
	(0.006)	(0.035)	(0.005)	(0.009)	(0.008)
Obs.	3856	569	1068	1270	949
R^2	0.232	0.089	0.287	0.312	0.368

Notes: The table reports coefficient estimates from regressions of the individual deviation from the mean forecast. 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 (continued)

Regressions of the deviation from the mean forecast

	Whole Sample	1980s	1990s	2000s	2010s
Panel D: Dependent variable: $x_{t+3\|t} - \bar{x}_{t+3\|t}$					
$x_{t\|t-1} - \bar{x}_{t\|t-1}$	-0.057***	-0.07***	-0.033	-0.053***	-0.039**
	(0.012)	(0.026)	(0.028)	(0.013)	(0.019)
$x_{t+1\|t-1} - \bar{x}_{t+1\|t-1}$	0.063	-0.081	0.121***	0.144***	0.034
	(0.043)	(0.106)	(0.028)	(0.031)	(0.052)
$x_{t+2\|t-1} - \bar{x}_{t+2\|t-1}$	-0.051	-0.051	0.12**	-0.098	0.032
	(0.053)	(0.055)	(0.064)	(0.082)	(0.062)
$x_{t+3\|t-1} - \bar{x}_{t+3\|t-1}$	0.514***	0.307***	0.312***	0.609***	0.592***
	(0.074)	(0.078)	(0.05)	(0.09)	(0.046)
$x_{t+4\|t-1} - \bar{x}_{t+4\|t-1}$	0.067**	0.138**	0.144***	0.118**	0.101***
	(0.036)	(0.068)	(0.035)	(0.056)	(0.027)
Constant	0.004	0.027	0.013**	-0.004	0.000
	(0.007)	(0.033)	(0.007)	(0.011)	(0.008)
Obs.	3855	568	1067	1271	949
R^2	0.306	0.123	0.415	0.433	0.442

Notes: The table reports coefficient estimates from regressions of the individual deviation from the mean forecast. 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 4

Forward noisy information vs. the standard model

Dependent variable	Specification (30): Forward noisy information	Specification (31): Standard noisy information
$x_{t+1\|t} - \bar{x}_{t+1\|t}$	10515	10822
$x_{t+2\|t} - \bar{x}_{t+2\|t}$	8565	8763
$x_{t+3\|t} - \bar{x}_{t+3\|t}$	7434	7635
$x_{t+4\|t} - \bar{x}_{t+4\|t}$	7323	7484

Notes: The table reports BIC statistics associated with specifications (30) and (31). The specifications were estimated for several forecast horizons as indicated in the first column ($h = 0,1,2,3$).

TABLE 5

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

	Whole Sample	Q1	Q2	Q3	Q4
<u>Panel A</u>: $F_t^i x_{2Y} = c + \rho F_t^i x_{1Y} + u_t$ (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
<u>Panel B</u>: $F_t^i x_{2C} = c + \rho F_t^i x_{1C} + u_t$ (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
<u>Panel C</u>: $F_t^i x_{2C} = c + \rho F_t^i x_{1C} + u_t$ (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. Standard errors of Driscoll and Kraay (1998) are in parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 levels

TABLE 6

Regressions of inflation and news shocks

Dependent Variable (x_t):	(1)	(2)	(3)	(4)	(5)	(6)
					83-99	00-17
$news_t$		1.547***		1.546***	1.348***	1.764***
		(0.147)		(0.154)	(0.096)	(0.138)
x_{t-1}	0.351***	0.761***	0.369***	0.762***	1.054***	0.657***
	(0.085)	(0.067)	(0.069)	(0.063)	(0.085)	(0.091)
x_{t-2}			-0.073	-0.029		
			(0.075)	(0.057)		
x_{t-3}			0.158	0.018		
			(0.102)	(0.049)		
x_{t-4}			-0.050	0.033		
			(0.073)	(0.048)		
Constant	1.744***	0.869***	1.603***	0.803***	-0.109	1.135***
	(0.225)	(0.302)	(0.439)	(0.281)	(0.274)	(0.268)
R^2	0.124	0.719	0.144	0.722	0.784	0.759

Notes: The table reports coefficient estimates from various regressions. x_t is inflation. The news variable is computed according to (34). Newey-West standard errors are in parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 levels

TABLE 7

Share of inflation volatility explained by news shocks

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 shocks from a variance decomposition applied to a VAR system estimated on the 1983Q1-2017Q4 sample. Inflation news shocks are obtained by the forward information method, or by the Barsky-Sims method, as indicated in the columns' headers.

TABLE 8

Regressions of news shocks: Additional variables

Dependent Variable (x_t):	Unemployment		Interest Rate		Real GDP Growth	
	(1)	(2)	(3)	(4)	(5)	(6)
$news_t$		1.108***		1.144***		1.146***
		(0.087)		(0.066)		(0.135)
x_{t-1}	1.626***	1.096***	1.507***	0.972***	0.313***	0.259***
	(0.091)	(0.088)	(0.096)	(0.068)	(0.076)	(0.068)
x_{t-2}	-0.641***	-0.110	-0.479***	0.084	0.108	0.062
	(0.185)	(0.106)	(0.147)	(0.096)	(0.089)	(0.077)
x_{t-3}	-0.008	0.013	-0.033	0.164***	0.027	0.011
	(0.170)	(0.084)	(0.144)	(0.046)	(0.079)	(0.071)
x_{t-4}	-0.017	-0.044	-0.019	0.090**	0.008	0.007
	(0.066)	(0.047)	(0.092)	(0.039)	(0.092)	(0.071)
Constant	0.256***	0.185***	0.066*	0.116***	1.546***	2.402***
	(0.079)	(0.053)	(0.034)	(0.032)	(0.384)	(0.419)
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 news variable is computed according to (34). The sample period is 1972Q1-2017Q4 for unemployment and GDP growth, and 1983Q3-2017-Q4 for the interest rate. Newey-West standard errors are in parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 levels

TABLE 9
Inflation and news shocks across different horizons

Dependent variable:	x_t	x_{t+1}	x_{t+2}	x_{t+3}	x_{t+4}
Panel A: First approach					
$news_{t\|t}$	1.547***	0.048	-0.328	-0.045	-0.148
	(0.147)	(0.245)	(0.216)	(0.286)	(0.309)
$news_{t+1\|t}$		0.764**	1.240	0.625	0.192
		(0.378)	(1.041)	(0.702)	(1.256)
$news_{t+2\|t}$			-1.111	0.812	1.108
			(0.871)	(0.902)	(0.861)
$news_{t+3\|t}$				-0.911	1.058
				(0.577)	(1.574)
$news_{t+4\|t}$					-1.801
					(1.222)
x_{t-1}	0.761***	0.334**	0.192	0.268*	0.355**
	(0.067)	(0.168)	(0.194)	(0.156)	(0.159)
Constant	0.869***	1.917***	2.158***	2.037***	1.805***
	(0.302)	(0.419)	(0.422)	(0.368)	(0.380)
R^2	0.719	0.080	0.044	0.048	0.093
Panel B: Second approach					
$news_{t\|t}$	1.547***	0.676***	-0.005	0.0402	0.389*
	(0.147)	(0.143)	(0.196)	(0.261)	(0.199)
$news_{t+1\|t}$		1.258***	0.412	0.429	0.851
		(0.400)	(0.449)	(0.388)	(0.620)
$news_{t+2\|t}$			0.019	0.405	0.529
			(0.963)	(0.606)	(1.149)
$news_{t+3\|t}$				-0.306	0.170
				(1.098)	(1.896)
$news_{t+4\|t}$					-0.892
					(1.074)
x_{t-1}	0.761***	0.376**	0.192	0.209	0.346**
	(0.067)	(0.173)	(0.200)	(0.176)	(0.171)
Constant	0.869***	1.868***	2.189***	2.189***	1.841***
	(0.302)	(0.448)	(0.452)	(0.408)	(0.422)
R^2	0.719	0.112	0.035	0.037	0.083

Notes: The table reports coefficient estimates from various regressions. x_t is inflation. The news variables are computed according to (35) and (36), in panels A and B, respectively. Newey-West standard errors are in parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 levels

Appendix A: 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 individual would explain forecast disagreement. More formally, following Capistrán and Timmermann (2009), this asymmetry is modeled by a LINEX loss-function over forecast errors:

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

where $FE_t^i x_{t+h} \equiv x_{t+h} - F_t^i x_{t+h}$ is the forecast error of forecaster i and θ_i is the asymmetry parameter. Positive 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

$$F_t^i x_{t+h} = E_t x_{t+h} + \frac{1}{2}\theta_i \sigma^2_{t+h|t} \qquad \text{(A.1)}$$

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 $\sigma^2_{t+h|t}$. 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.

In a similar way the forecast for $h-1$ steps ahead is

$$F_t^i x_{t+h-1} = E_t x_{t+h-1} + \frac{1}{2}\theta_i \sigma^2_{t+h-1|t} \qquad \text{(A.2)}$$

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

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

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

$$F_t^i x_{t+h} = \rho F_t^i x_{t+h-1} + \frac{1}{2}\theta_i\left(\sigma^2_{t+h|t} - \rho\sigma^2_{t+h-1|t}\right) \qquad \text{(A.3)}$$

The last term on the right-hand-side would correspond to the error term in a cross-sectional regression of the forecast $F_t^i x_{t+h}$ on the forecast $F_t^i x_{t+h-1}$. The mean of the error term would be zero only if θ_i is distributed around 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 $\frac{1}{2}(\theta_i - \theta)\left(\sigma^2_{t+h|t} - \rho\sigma^2_{t+h-1|t}\right)$.

In any case, it is clear from (A.2) that θ_i is positively correlated with $F_t^i x_{t+h-1}$, so that the OLS estimate of the coefficient on $F_t^i x_{t+h-1}$would be a biased estimate of the persistence parameter ρ. The sign of the bias depends on the sign of $\left(\sigma^2_{t+h|t} - \rho\sigma^2_{t+h-1|t}\right)$ which has no cross-sectional variation. For example, if the variance of the shock to the fundamental is fixed this term will be positive and the coefficient estimate will be an upward-biased estimate of ρ.

Capistrán and Timmermann (2009) have assumed that the conditional variance follows a GARCH(1,1) process:

$$\sigma^2_{t+1|t} = \alpha_0 + \alpha_1 \omega^2_t + \beta_1 \sigma^2_{t|t-1}$$

For this case we get

$$\sigma^2_{t+h|t} - \rho\sigma^2_{t+h-1|t} = \alpha_0 + \alpha_1 \sigma^2_{t+h-1|t} + \beta_1 \sigma^2_{t+h-1|t} - \rho\sigma^2_{t+h-1|t} = \alpha_0 + (\alpha_1 + \beta_1 - \rho)\sigma^2_{t+h-1|t}$$

which could also be negative, and thus result in a negative bias. However, whether the bias is positive or negative, there is no tendency of the bias to shrink when increasing the forecast horizon h, which contradicts the evidence in section 4.

Appendix B: Estimating Persistence with Disagreement about Persistence

Consider the following state-space representation:

State:

$$x_t = \rho_t x_{t-1} + \omega_t \tag{B.1}$$

$$\rho_t = \rho_{t-1} + \epsilon_t$$

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

Measurement:

$$y_t^i = x_t + \nu_t^i \tag{B.2}$$

$$\rho_t^i = \rho_t + \varepsilon_t^i$$

where $\nu_t^i \sim iid\ N(0, \sigma_\nu^2)$ and $\varepsilon_t^i \sim iid\ N(0, \sigma_\varepsilon^2)$.

Thus, the underlying persistence in x_t is not a fixed parameter. Instead, it follows a stochastic process, where it can be shifted permanently by the shock ϵ_t. Forecasters receive noisy signal about both the fundamental and its persistence.

The forecaster first applies a Kalman Filter to evaluate the persistence:

$$F_t^i \rho_t = (1 - G_\rho) F_{t-1}^i \rho_t + G_\rho \rho_t^i \tag{B.3}$$

Where $F_t^i \rho_t$ is an individual's forecast of ρ_t, made at time t, and G_ρ is the Kalman gain. Using the state and measurement equations, this can be rewritten as:

$$F_t^i \rho_t = \rho_t + \left[(1 - G_\rho)\left(F_{t-1}^i \rho_t - \rho_t\right) + G_\rho \varepsilon_t^i\right] \tag{B.4}$$

where the term in squared brackets is an individual's zero-mean deviation from the "mean" persistence ρ_t. Denote this term by κ_t^i .

Next, given the evaluated persistence, the forecaster applies a Kalman Filter to predict the fundamental:

$$F_t^i x_t = (1 - G_x^i) F_{t-1}^i x_t + G_x^i y_t^i \tag{B.5}$$

$$F_t^i x_{t+1} = F_t^i \rho_{t+1} F_t^i x_t \tag{B.6}$$

Where $F_t^i x_t$ is an individual's forecast of x_t, and G_x^i is the Kalman gain that weighs the signal about the fundamental, given the signal-to-noise ratio and the evaluation of the fundamental's persistence made in the first stage (Hence, the superscript i).

Combining (B.4) and (B.6) and noting that $F_t^i \rho_{t+1} = F_t^i \rho_t$, would result in the following regression equation:

$$F_t^i x_{t+1} = \rho_t F_t^i x_t + u_t^i \tag{B.7}$$

where $u_t^i = \kappa_t^i F_t^i x_t$. Thus, when running the regression cross-sectionally, using the data of forecasts reported at time t, the properties of the error term will depend on the cross-sectional correlation between the individual forecast of the fundamental and his evaluation of the underlying persistence. In particular, by substituting (B.2) and (B.5):

$$E_i\left[\kappa_t^i F_t^i x_t\right] = E_i\left[\kappa_t^i\left(\left(1 - G_x^i\right)F_{t-1}^i x_t + G_x^i y_t^i\right)\right] = E_i\left[\kappa_t^i G_x^i\left(x_t - F_{t-1}^i x_t\right)\right] + E_i\left[\kappa_t^i F_{t-1}^i x_t\right] + E_i\left[\kappa_t^i G_x^i \nu_t^i\right] = E_i\left[\kappa_t^i G_x^i\left(x_t - F_{t-1}^i x_t\right)\right]$$

where the last part of the derivation is based on the independency of the noise ν_t^i and on the forecast $F_{t-1}^i x_t$ being pre-determined at time t.

Hence, the possible cross-sectional correlation between κ_t^i and $F_t^i x_t$ could be assessed by looking on the expectation term of $E_i\left[\kappa_t^i G_x^i\left(x_t - F_{t-1}^i x_t\right)\right]$. Although an explicit expression for the correlation would be quite complex in general, it could be possible to draw some conclusions regarding the *sign* of the correlation. First, it should be noticed that the Kalman gain G_x^i should be positively correlated with the underlying persistence of the variable x, as it is evaluated by the individual - when the persistence is perceived higher by an individual, she would place a greater weight on the new signal she received about the fundamental. Thus, κ_t^i and G_x^i are positively correlated.

It follows that the sign of $E_i\left[\kappa_t^i G_x^i\left(x_t - F_{t-1}^i x_t\right)\right]$ will depend on the forecast error term $\left(x_t - F_{t-1}^i x_t\right)$. This is the ex-post error of the forecast from the previous period, which should contain both idiosyncratic noises before time t that are assumed not to be correlated with those of time t, as well as the innovation to x_t which was realized at time t. This innovation is a common shock which should be treated as a constant relative to the cross-sectional expectation term. Thus, in times of positive shocks to x_t, the evaluation of persistence (κ_t^i) and the forecast of the state variable ($F_t^i x_t$) will tend to be positively correlated across individuals and the opposite in times dominated by negative shocks to x_t. Therefore, when running (B.7) as a cross-sectional regression the estimated persistence could be biased in both directions.

Finally, the regression in equation (7) could be extended to additionally consider more general uncertainty about the model, by defining u_t^i as follows:

$$u_t^i = \kappa_t^i F_t^i x_t + e_t^i \qquad \text{(B.8)}$$

As before, the first error component is driven by uncertainty regarding the persistence in the underlying AR(1) process of the fundamental, while the second component, e_t^i, would account for uncertainty about the specification itself. For instance, some forecasters might base their forecast on an AR of a higher order or on a VAR specification rather than modelling with an AR(1). This could add an additional form of bias as in the general case of omitting relevant variables from a regression. Thus, the sign of the total bias in the persistence estimate will remain unclear.

Appendix C: Forward information model: Mathematical derivations

C.1. Deriving the optimal forecasts

In section 5.2., we consider the optimal forecasting for following state-space representation:

State:

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

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

Measurement:

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

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 be based on four useful signals:

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

All other noisy signals from the past are not useful anymore after x_t is perfectly observed.

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

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

where $\sum_{k=1}^4 W_k = 1$. The forecaster should choose the weights which minimize the expected squared error $E_t\{x_{t+1} - x_{t+1|t}\}^2$. The expected squared error could be expressed as:

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

Thus, the optimization problem could 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 $\sigma_\omega^2 = 1$, the FOC 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:

$$W_1 = \frac{\sigma_2^2\sigma_1^2(1+\sigma_2^2)}{m}$$
$$W_2 = \frac{\sigma_1^2(1+\sigma_2^2)}{m} \qquad \text{(C.4)}$$
$$W_3 = \frac{\sigma_2^2(1+\sigma_2^2)}{m}$$
$$W_4 = \frac{\rho^2\sigma_2^2\sigma_1^2}{m}$$

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 $x_{t+2|t}$. The same four signals used in the one-step ahead forecast, should also be used in $x_{t+2|t}$, 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. $\rho y_{t-1,t+1}^i$ (from the previous period), with ex-post forecast error equal to $(-\rho v_{t-1,t+1}^i + \omega_{t+2})$.
3. $\rho y_{t,t+1}^i$, with ex-post forecast error equal to $(-\rho v_{t,t+1}^i + \omega_{t+2})$.
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

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

The expected squared error could therefore be expressed as follows:

$$E_t\{x_{t+2} - x_{t+2|t}\}^2 = E_t\{w_1(\rho\omega_{t+1} + \omega_{t+2}) + w_2(-\rho v_{t-1,t+1}^i + \omega_{t+2}) + w_3(-\rho v_{t,t+1}^i + \omega_{t+2}) + w_4 v_{t,t+2}^i\}^2 = (w_1)^2\rho^2 + (w_1 + w_2 + w_3)^2 + (w_2)^2\rho^2\sigma_2^2 + (w_3)^2\rho^2\sigma_1^2 + (w_4)^2\sigma_2^2$$

and the optimization problem is consequently

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

FOC 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 would be

$$w_1 = \frac{\sigma_2^2\sigma_2^2\sigma_1^2}{m} \qquad \text{(C.6)}$$
$$w_2 = \frac{\sigma_2^2\sigma_1^2}{m}$$

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

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

when moving further to forecast three steps-ahead, the optimal forecast should obey $x_{t+3|t} = \rho x_{t+2|t}$, which implies there is no further change in the optimal weights. This reason is that there are no forward signals referring to three-steps ahead horizon or beyond. To see that, we repeat the same above steps to derive $x_{t+3|t}$. The four available signals should be transformed to 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. $\rho^2 y_{t-1,t+1}^i$ (from the previous period), 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 should follow

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

Notice that in this case the expected squared error could be expressed as

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

Thus, the minimization of $E_t\{x_{t+3} - x_{t+3|t}\}^2$ would be equivalent to the minimization $E_t\{x_{t+3} - x_{t+3|t}\}^2$, producing the same optimal weights w_k derived above. The same reasoning holds for any horizon beyond two steps-ahead and, consequently, the relationship of $x_{t+h|t} = \rho x_{t+h-1|t}$ for any $h \geq 3$.

C.2. Kalman filter representation

In this section, we represent the optimal forecast in the Kalman filter framework. The filter should be of the following form:

$$\boldsymbol{x}_{t|t} = \boldsymbol{x}_{t|t-1} + \mathrm{G}\left(\boldsymbol{y}_t^{\boldsymbol{i}} - \boldsymbol{x}_{t|t-1}\right) \qquad \text{(C.8)}$$

where the gain matrix G is needed to be specified. Expanding the matrix notation, we have:

$$\begin{bmatrix} x_{t|t} \\ x_{t+1|t} \\ x_{t+2|t} \end{bmatrix} = \begin{bmatrix} x_{t|t-1} \\ x_{t+1|t-1} \\ x_{t+2|t-1} \end{bmatrix} + \begin{bmatrix} \mathrm{G}_{1,1} & \mathrm{G}_{1,2} & \mathrm{G}_{1,3} \\ \mathrm{G}_{2,1} & \mathrm{G}_{2,2} & \mathrm{G}_{2,3} \\ \mathrm{G}_{3,1} & \mathrm{G}_{3,2} & \mathrm{G}_{3,3} \end{bmatrix} \times \begin{bmatrix} y_{t,t}^i - x_{t|t-1} \\ y_{t,t+1}^i - x_{t+1|t-1} \\ y_{t,t+2}^i - x_{t+2|t-1} \end{bmatrix} \qquad \text{(C.9)}$$

Notice that each element 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, recalling that the filter algorithm is also based on minimization of the squared error.

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

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

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

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

$$x_{t+1|t} = x_{t+1|t-1} + W_1\rho\left(y_{t,t}^{i} - x_{t|t-1}\right) + W_3\left(y_{t,t+1}^{i} - x_{t+1|t-1}\right) + W_4\rho^{-1}\left(y_{t,t+2}^{i} - x_{t+2|t-1}\right) \qquad \text{(C.10)}$$

The lagged forecasts on the right-hand-side follow

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

$$x_{t+1|t-1} = x_{t+1|t-2} + w_1\rho^2\left(y_{t-1,t-1}^{i} - x_{t-1|t-2}\right) + w_3\rho\left(y_{t-1,t}^{i} - x_{t|t-2}\right) + w_4\left(y_{t-1,t+1}^{i} - x_{t+1|t-2}\right)$$

$$x_{t+2|t-1} = \rho x_{t+1|t-1}$$

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

$$x_{t+1|t} = W_1\rho y_{t,t}^{i} + [(1 - W_3 - W_4)w_4 - W_1W_4]y_{t-1,t+1}^{i} + W_3 y_{t,t+1}^{i} + W_4\rho^{-1}y_{t,t+2}^{i} + other\ signals \qquad \text{(C.11)}$$

where $other\ signals$ represent all the signals in the lagged forecasts, which are not the signal $y_{t-1,t+1}^{i}$. The signal $y_{t-1,t+1}^{i}$ is the only lagged signal which is still informative on 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 (C.4) and (C.6), it can be verified that the term $other\ signals$ in (C.11) is equal to zero. Furthermore, the coefficient on the lagged signal $y_{t-1,t+1}^{i}$ in (C.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 (C.3). Thus, the optimal forecast in (C.11), based on the Kalman filter, is the same optimal forecast derived in (C.3) by directly optimizing the weights to minimize the squared forecast error. In a similar way, it can be shown that the optimal forecast $x_{t+2|t}$ according to the Kalman filter is the same forecast derived in (C.5). Thus, our guessed solution for the gain matrix is verified again.

C.3. Patterns of regression properties across forecast horizons

The simulation results, as presented in figure 7, show that in a regression of $x_{t+h|t}$ on $x_{t+h-1|t}$ 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 $h = 3$.

$\boldsymbol{h = 1}$: Suppose that we run a cross-sectional regression of $x_{t+1|t}$ on $x_{t|t}$. From (26), it is clear that the coefficient on $x_{t|t}$ should be ρ. However, the OLS coefficient estimate would be zero. This is easily verified, when recalling that $x_{t|t} = x_t$ as in (19), due to the perfect signal $y^i_{t,t}$ in (18). so that there is no cross-sectional variation in $x_{t|t}$. It also follows that all the cross-sectional variation in $x_{t+1|t}$ is explained by the error term, as specified in (26), and the R-squared should therefore be 0.

In the simulation presented in figure 7, this limiting case was avoided by introducing some low degree of noise into the signal $y^i_{t,t}$ ($\sigma_0^2 = 0.2$), thereby, allowing some variation in $x_{t|t}$. As a consequence, simulated coefficient estimates and R-squared were positive but still very low (Especially the R-squared which is very close to zero, as described in panel B).

$\boldsymbol{h = 3}$: At the other extreme, if we regress $x_{t+3|t}$ on $x_{t+2|t}$, we should get a perfect fit where the OLS coefficient estimate is equal to ρ. As shown in (25), the optimal weights in the two forecasts $x_{t+3|t}$ and $x_{t+2|t}$ are the same, leading to the exact relationship of $x_{t+3|t} = \rho x_{t+2|t}$. 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 in (24). The coefficient on $x_{t+1|t}$ should be ρ, but the OLS estimate would be biased since $x_{t+1|t}$ is correlated with $Signal_k$ in the error term. The R-squared would be between 0 and 1.

It is also interesting to examine how regression properties, for $h = 2$, would 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 increase in the noise is obtained by moving to longer horizons. As shown in section 5.3 in the main text, by combining (C.3) and (C.5), the relation between the optimal forecasts with consecutive horizons $x_{t+2|t}$ and $x_{t+1|t}$ could be expressed as

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

Thus, a deviation from the simple state relation of $x_{t+2|t} = \rho x_{t+1|t}$ 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 $y^i_{t,t+2}$ become meaningless. As a consequence, forecasts at time t should only apply two signals: the perfect signal about realized x_t and the forward signal $y^i_{t,t+1}$. The one step-ahead forecast should therefore be

$$x_{t+1|t} = W_1^{lim} \rho x_t + W_3^{lim} y^i_{t,t+1} \qquad \text{(C.13)}$$

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

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

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:

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

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

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

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 variations in the optimal weights across consecutive horizons and we obtain the simple relation of $x_{t+2|t} = \rho x_{t+1|t}$.

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 (C.4) and (C.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 (C.4) and (C.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

$$\begin{aligned} m &= \sigma_2^2\sigma_1^2(1+\sigma_2^2) + \sigma_1^2(1+\sigma_2^2) + \sigma_2^2(1+\sigma_2^2) + \rho^2\sigma_2^2\sigma_1^2 \\ &= (2+\rho^2)\sigma_1^2\sigma_2^2 + (\sigma_2^2)^2 + (\sigma_2^2)^2\sigma_1^2 + \sigma_2^2 + \sigma_1^2 \end{aligned}$$

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 (C.4) and (C.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

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

$\underline{k = 3}$:

From (C.4) and (C.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 $k = 1$.

$\underline{k = 4}$:

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

$$\frac{\partial(w_4 - W_4)}{\partial\sigma_2^2} = \frac{\partial[1 - (w_1 - W_1) - (w_2 - W_2) - (w_3 - W_3)]}{\partial\sigma_2^2}$$
$$= -\frac{\partial(w_1 - W_1)}{\partial\sigma_2^2} - \frac{\partial(w_2 - W_2)}{\partial\sigma_2^2} - \frac{\partial(w_3 - W_3)}{\partial\sigma_2^2} < 0$$

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

C.4. Patterns of regression properties when changing persistence

Changes in persistence

Another empirical pattern, documented in section 4 and confirmed by the simulation results in figure 7, is the co-movement tendency of the coefficient and R-squared of our cross-sectional. 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. We first look at the two extreme cases of $\rho = 0$ and $\rho = 1$.

$\boldsymbol{\rho = 0}$: When there is no persistence in the state process, there should not be any persistence in the forecasts as well. A forecast $x_{t+\mathrm{h}|t}$ would only rely on signals referring to $t + \mathrm{h}$, and would not be correlated with $x_{t+\mathrm{h}-1|t}$ (nor with $x_{t+\mathrm{h}+1|t}$). In terms of our above equations, the optimal forecasts would now be:

$$x_{t|t} = x_t \qquad \text{(C.15)}$$

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

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

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

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 $x_{t+\mathrm{h}|t}$ on $x_{t+\mathrm{h}-1|t}$, the coefficient-estimate and R-squared should be zero. In terms of equation (C.14), for example, the whole cross-sectional variation in $x_{t+2|t}$ is due to the regression error term, driven by the differences in the optimal weights $(w_k - W_k)$, which equal to $-W_2$, $-W_3$ and w_4 for $k = 2,3,4$, respectively (according to (C.15)).

$\boldsymbol{\rho = 1}$: When the fundamental follows a random walk, the general results obtained above would hold. This is not a limiting case where the coefficient and R-squared should converge to 1. Rather, the regression properties would depend on the forecast horizon in the same way described above. Specifically, for the middle horizon $h = 2$, as implied by equation (C.14), the coefficient on $x_{t+1|t}$ should be $\rho = 1$, but the OLS estimate would be biased due to the correlation of $x_{t+1|t}$ with the error term, 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 $x_{t+2|t}$ and $x_{t+1|t}$ as specified in (C.12) would vary when changing the persistence of the state process. Specifically, we show that the difference between optimal weights placed on the same signals in $x_{t+2|t}$ and $x_{t+1|t}$, which is $(w_k - W_k)$, tends to diminish when the persistence ρ increases. This explains the tendency of the R-squared and ρ to move in the same direction, as documented for the SPF forecasts and in the simulation results.

From our previous results, we can notice that for $k = 1,2,3$, the difference in the weights can be expressed as $w_k - W_k = a_k m^{-1}$, where only m is a function of ρ. Hence, it is sufficient to show that the derivative of m^{-1} with respect to ρ is positive:

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

Thus, the negative difference $w_k - W_k$, $k = 1,2,3$, is getting closer to zero when 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 sum of the weights w_k and W_k should equal one.

Overall, this implies that the fit of a regression of $x_{t+2|t}$ to $x_{t+1|t}$, will tend to be higher with higher degree of persistence, since the disturbance term in (C.12) would diminish due to the diminishing difference between the optimal weights applied by the two forecasts.

We further derive the OLS coefficient estimate and show that it should increase in ρ, despite being a biased estimate of persistence. Using equations (C.3) – (C.6), we obtain

$$\beta_{OLS} = \frac{Cov(x_{t+2|t}, x_{t+1|t})}{Var(x_{t+1|t})} = \frac{W_2 w_2 \rho^3 \sigma_2^2 + W_3 w_3 \rho \sigma_1^2 + W_4 w_4 \rho^{-1} \sigma_2^2}{{W_2}^2 \rho^2 \sigma_2^2 + {W_3}^2 \sigma_1^2 + {W_4}^2 \rho^{-2} \sigma_2^2} =$$

$$= \rho + \frac{W_2(w_2 - W_2)\rho^3 \sigma_2^2 + W_3(w_3 - W_3)\rho\sigma_1^2 + W_4(w_4 - W_4)\rho^{-1}\sigma_2^2}{{W_2}^2 \rho^2 \sigma_2^2 + {W_3}^2 \sigma_1^2 + {W_4}^2 \rho^{-2} \sigma_2^2}$$

$$= \rho + \frac{-\sigma_1^2\sigma_1^2(1+\sigma_2^2)\rho^3\sigma_2^2 - \sigma_2^2\sigma_2^2(1+\sigma_2^2)\rho\sigma_1^2 + (\sigma_2^2\sigma_1^2 + \sigma_1^2 + \sigma_2^2)\rho^2\sigma_2^2\sigma_1^2\rho^{-1}\sigma_2^2}{\sigma_1^2(1+\sigma_2^2)\sigma_1^2(1+\sigma_2^2)\rho^2\sigma_2^2 + \sigma_2^2(1+\sigma_2^2)\sigma_2^2(1+\sigma_2^2)\sigma_1^2 + \rho^2\sigma_2^2\sigma_1^2\sigma_2^2\sigma_1^2\sigma_2^2} \qquad \text{(C.16)}$$

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

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

The bias in the coefficient estimate obeys

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

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

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

If $\beta_{OLS} > \rho$, it is easy to see that the derivative is positive. The only negative term is $-2\rho\sigma_1^2(1+\sigma_2^2)$, which is lower than the expression in the denominator under the assumption of $\sigma_2^2 > \sigma_1^2 > 1$.

In the typical case, where the coefficient estimate is biased downward ($\beta_{OLS} < \rho$), we can write the derivative in the following form:

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

where $0 < \frac{(\rho-\beta_{OLS})}{\rho} < 1$, and, as a result, the expressions in both squared brackets are positive, making the whole derivative positive as well.

C.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 (C.3), we obtain:

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

Taking the average across agents, would drop all the idiosyncratic terms. Hence, we get:

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

where the upper bar in $\bar{x}_{t+1|t}$ denotes the cross-sectional average.

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

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

Thus, when running a regression of forecast error on forecast revision, the expected OLS coefficient estimate would be:

$$\frac{Cov(x_{t+1} - \overline{x}_{t+1|t}, \overline{x}_{t+1|t} - \overline{x}_{t+1|t-1})}{Var(\overline{x}_{t+1|t} - \overline{x}_{t+1|t-1})} = \frac{E\left[(x_{t+1} - \overline{x}_{t+1|t})(\overline{x}_{t+1|t} - \overline{x}_{t+1|t-1})\right]}{E\left[(\overline{x}_{t+1|t} - \overline{x}_{t+1|t-1})^2\right]}$$
$$= \frac{E\left[\left(W_1\omega_{t+1} + W_4(-\rho^{-1}\omega_{t+2})\right)\left(w_1\rho\omega_t + (w_1 + w_2 + w_3 - W_1)\omega_{t+1} + W_4\rho^{-1}\omega_{t+2}\right)\right]}{E\left[(w_1\rho\omega_t + (w_1 + w_2 + w_3 - W_1)\omega_{t+1} + W_4\rho^{-1}\omega_{t+2})^2\right]}$$
$$= \frac{W_1(w_1 + w_2 + w_3) - (W_1)^2 - (W_4)^2\rho^{-2}}{(w_1)^2\rho^2 + (w_1 + w_2 + w_3 - W_1)^2 + (W_4)^2\rho^{-2}}$$

where we use the independency of the shock to the fundamental, and the standardization of its variance (σ^2_ω=1).

We can further verify that the expected coefficient is positive by showing that the numerator is positive. Using our expressions for the optimal weights from (C.4) and (C.6) we obtain

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

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