Quantitative Economics 14 (2023), 609–650 1759-7331/20230609

Quantifying noise in survey expectations

Artūras Juodis
Amsterdam School of Economics, University of Amsterdam and Tinbergen Institute

Simas Kučinskas
School of Business and Economics, Humboldt University of Berlin

Expectations affect economic decisions, and inaccurate expectations are costly. Expectations can be wrong due to either bias (systematic mistakes) or noise (unsystematic mistakes). We develop a framework for quantifying the level of noise in survey expectations. The method is based on the insight that theoretical models of expectation formation predict a factor structure for individual expectations. Using data from professional forecasters, we find that the magnitude of noise is large (10%–30% of forecast MSE) and comparable to bias. We illustrate how our estimates can be applied to calibrate models with incomplete information and bound the effects of measurement error.

Keywords. Expectation formation, factor models, measurement error, noise, panel data, subjective expectations.

JEL classification. C53, D83, E70, G40.

> *Now, it's hard to say what there is more of, noise or bias. But one thing is very certain—that bias has been overestimated at the expense of noise. Virtually all the literature and a lot of public conversation is about biases. But in fact, noise is, I think, extremely important, very prevalent.*
>
> Daniel Kahneman[1]

Artūras Juodis: a.juodis@uva.nl
Simas Kučinskas: simas.kucinskas@hu-berlin.de

We would like to thank two anonymous referees for numerous suggestions that greatly improved this paper. For comments and feedback at various stages of the project, we thank Max Bruche, Spiros Bougheas, Ryan Chahrour, Peter Egger, Daniel Kahneman, Frank Kleibergen, Alexandre Kohlhas, Marcelo Pedroni, Florian Peters, Andreas Pick, Jane Ryngaert, Andrei Shleifer, Tom Stark, Martin Weidner, and Fabian Winkler as well as audiences at the Federal Reserve Board, Humboldt University, KOF-ETH-UZH, University of Nottingham, Dutch Central Bank, and IPDC2019 (Vilnius). Financial support from the Netherlands Organization for Scientific Research (NWO) under research grant 451-17-002 is gratefully acknowledged by the first author. A replication file is posted (Juodis and Kučinskas (2023)).

[1]Interview with Tyler Cowen at George Mason University (November 12, 2018), https://medium.com/conversations-with-tyler/tyler-cowen-daniel-kahneman-economics-bias-noise-167275de691f.

Available at http://qeconomics.org. https://doi.org/10.3982/QE1633

1. Introduction

Motivation

Expectations are a key driver of economic decisions. As a result, inaccurate expectations are costly.[2] There are two reasons why expectations may be wrong. The first possibility is *bias* (systematic mistakes). For example, Alice always expects to win the lottery, but sadly that never happens. The second possibility, however, is *noise* (unsystematic mistakes). That is like Bob's view of the economy, which depends on whether his favorite football team won or lost the game yesterday. Alice's and Bob's mistakes are equally detrimental to forecast accuracy, as formalized by the bias-variance decomposition (e.g., Hastie, Tibshirani, and Friedman (2009), p. 223). Yet noise has received little attention in prior research on subjective expectations. This paucity of research is especially surprising given the extant literature in psychology on noise in human decision making (Dawes, Faust, and Meehl (1989), Grove, Zald, Lebow, Snitz, and Nelson (2000), Kahneman, Sibony, and Sunstein (2021)).

What we do

We develop a general framework for quantifying the level of noise in survey expectations. The key insight is that existing models imply a factor structure for the cross-section of individual expectations. Imagine many people forecasting the same variable, say, inflation. Existing models have the property that individual forecasts can be written as a linear combination of some underlying factors plus a noise term. These factors capture macroeconomic news (e.g., changes in oil prices) as well as any common sentiment shocks that are orthogonal to fundamentals (e.g., animal spirits). Factor loadings may vary across forecasters, capturing the possibility that forecasters may disagree on how to interpret the available evidence.

The noise term is orthogonal to the common factors and uncorrelated across forecasters. Hence, our definition captures *idiosyncratic* noise in expectations. For instance, that is the variation in Bob's expectations about the economy that is driven by the performance of his favorite football team. Other people have different favorite teams, so such noise is idiosyncratic. (If everyone has the same favorite team, that is a sentiment shock, and hence a common factor.) In practice, we first estimate a factor model using panel data on individual forecasts. Noise is then given by the residuals from the factor model.

The key measurement challenge is to distinguish noise from heterogeneity. Consider two forecasters predicting inflation, and suppose that they make different forecasts. One possibility is that the forecasters disagree on how to interpret the available data. For example, they may have different views on the importance of a recent oil price shock. Alternatively, the forecasters may agree on how to interpret the data, but their forecasts

[2]For surveys, see Pesaran and Weale (2006, Section 5,) and Manski (2018). Coibion, Gorodnichenko, and Kamdar (2018) provide a review focused on inflation expectations in macro, while Barberis (2018, Chapters 4, 5, and 6) discusses the role of expectations in finance. Bachmann and Elstner (2015) and Barrero (2018) quantify welfare losses from inaccurate expectations.

could be noisy. With a single cross-section of forecasts, heterogeneity and noise are observationally equivalent. We overcome this measurement challenge by using panel data. With panel data, we can track forecasters over time, and in effect, estimate how they react to new information.

Noise recovered by our method can be due to either (i) inherent randomness in expectations; or (ii) measurement error. Potential mechanisms for (i) include imperfect information (Lucas (1973), Woodford (2003)), limited attention (Sims (2003), Maćkowiak and Wiederholt (2009)), and diagnostic expectations (Bordalo, Gennaioli, and Shleifer (2018)). Measurement error may arise if survey questions are not fully precise, or survey participants have difficulty conveying their true subjective expectations (Manski and Molinari (2010), Drerup, Enke, and von Gaudecker (2017)). The two potential explanations are observationally equivalent when only data on expectations are available.[3] However, our measures of noise provide empirically-informative upper bounds. Intuitively, since we estimate "(i) + (ii)," we can easily bound measurement error, or (ii). With additional (stronger) assumptions on the expectation-formation process, our estimates can be used to point-identify the magnitude of measurement error.

Empirical findings

We apply the new methodology to inflation forecasts from the Survey of Professional Forecasters (SPF). Our estimates indicate that noise is large and pervasive, equal to 10–30% of mean-squared forecast error and of a similar magnitude to bias. Importantly, we find that variation in the data is well captured by factor models with a small (1–3) number of factors.

Our framework has applications that go beyond quantifying the amount of noise. To show this, we first investigate the consequences of noise for forecast combination. Consensus (or average) forecasts tend to perform better than most individual forecasts, a phenomenon popularized as the *wisdom-of-the-crowd effect* (Surowiecki (2005)). Depending on the forecast horizon, we find that 20%–60% of the wisdom-of-the-crowd effect is due to the reduction in noise (with the rest due to a reduction in bias). Hence, to a large extent, the crowd is wise because it is less noisy than individual forecasters.

Next, we illustrate how our estimates can be used to (i) guide theory and calibrate models of expectation formation; and (ii) bound the effects of measurement error. For the first question, we demonstrate how our estimates can be used to develop novel tests of the noisy-information model (Lucas (1973), Woodford (2003), Sims (2003)). We find that the baseline noisy-information model underpredicts the amount of noise relative to what we find empirically. To be consistent with the data, the noisy-information model needs to attribute around half of our estimated noise to measurement error.

For the second application, we study whether negative micro-level estimates of the Coibion and Gorodnichenko (2015) regression may be explained by measurement error. Under Bayesian updating, the micro Coibion–Gorodnichenko coefficient should be zero. Recent empirical work, however, has obtained negative coefficients at the micro

[3]Even with data on actions, the two possibilities are difficult to tell apart; see Drerup, Enke, and von Gaudecker (2017).

level (Broer and Kohlhas (2018), Bordalo, Gennaioli, Ma, and Shleifer (2020)). Unfortunately, when expectations are measured with error, the Coibion–Gorodnichenko regression suffers from a nonclassical measurement-error problem that can spuriously yield negative estimates. To check robustness, we calculate how much measurement error would be necessary to generate the observed micro coefficients under Bayesian updating. We find that most of the noise that we estimate needs to be attributed to measurement error to explain the empirical findings. Given that noise is likely to arise from multiple reasons, not just measurement error, we conclude that these findings appear robust to empirically-realistic amounts of measurement error.

Related literature

The paper builds on an extensive literature on expectation formation. For comprehensive overviews, we refer to the surveys cited above (footnote 2). While the present paper is, to the best of our knowledge, the first to develop a method specifically designed to measure noise in subjective expectations, some existing techniques can be used for answering related questions. In particular, Davies and Lahiri (1995) decompose forecast errors to (i) an individual time-invariant bias term; (ii) shocks that are common to all forecasters; and (iii) shocks that are forecaster specific. The last component is closely related to our own measure of noise. As shown below (Section 2.2), however, the assumptions made by Davies and Lahiri are violated by prominent models of expectation formation; see also Crowe (2010). Gaglianone and Issler (2021) develop a factor-based approach to forecast combination.[4] Under certain assumptions made by the authors, individual expectations are driven by a single common factor. However, many popular theoretical models violate the single-factor assumption. In addition, we find that approximating the data well in our empirical application often requires more than one factor. The single-factor assumption is also made in the pioneering work of Figlewski (1983) on optimal forecast combination.

In a study that appeared after the initial version of our article was posted online, de Silva and Thesmar (2021) also develop a method to quantify noise in survey expectations. While our goal is to estimate idiosyncratic noise from forecaster-level data, the method proposed by de Silva and Thesmar uses consensus forecasts and, therefore, captures sentiment shocks. Unfortunately, the setup of de Silva and Thesmar (2021) rules out any form of nonlinear interaction between variables measurable with respect to private and public information sets. In addition, their approach requires the researcher to accurately approximate true conditional expectations. Nevertheless, the two papers are highly complementary and reach similar empirical conclusions.

[4] In contemporaneous and independent work, Herbst and Winkler (2019) use a factor model to study the structure of multivariate disagreement among professional forecasters. Consistent with our own empirical findings, Herbst and Winkler find that for most variables in the Survey of Professional Forecasters, disagreement is mostly due to the idiosyncratic component.

2. Methodology

2.1 *Framework*

We are interested in measuring noise in expectations about some common *target variable* x_t. Throughout, we assume that any deterministic component in x_t is removed, and x_t has mean zero. Following Kučinskas and Peters (2022), the target variable is assumed to follow

$$x_t = \sum_{j=1}^{M} \sum_{\ell=0}^{+\infty} \alpha_{j,\ell} \varepsilon_{j,t-\ell}, \tag{1}$$

for some square-summable coefficients $\alpha_{j,\ell}$. Here, $\boldsymbol{\varepsilon}_t = (\varepsilon_{1,t}, \varepsilon_{2,t}, \dots, \varepsilon_{M,t})^\top$ is a martingale difference sequence of structural shocks with $\mathbb{E}[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t^\top] = \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$. Without loss of generality, we assume that $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ is a diagonal matrix, and we normalize $\alpha_{j,0} = 1$.

We assume that the econometrician has access to a panel of individual-level expectations about the target variable, $\mathbb{F}_{i,t}[x_{t+1}]$, of the form

$$\mathbb{F}_{i,t}[x_{t+1}] = b_{i,0} + \sum_{j=1}^{M} \sum_{\ell=0}^{+\infty} a_{i,j,\ell+1} \varepsilon_{j,t-\ell} + \nu_{i,t}, \tag{2}$$

where $i = 1, \dots, N$ denotes a particular forecaster, $b_{i,0}$ is a forecaster-specific unconditional bias term, and $\nu_{i,t}$ is forecaster-specific zero-mean *noise*. To capture the notion that noise is idiosyncratic, we assume $\nu_{i,t}$ is orthogonal to the structural shocks $\varepsilon_{j,t}$ at all leads and lags and independent across forecasters. To be consistent with existing theory (see Section 2.2), we allow $\nu_{i,t}$ to be serially correlated. The $a_{i,j,\ell}$ coefficients capture how expectations respond to the underlying structural shocks. Importantly, these reactions are allowed to be heterogeneous across forecasters. Finally, we assume that $a_{i,j,\ell}$'s are almost surely square summable for every i, while $\nu_{i,t}$ is covariance stationary and satisfies standard regularity conditions for factor models (see, e.g., Bai (2003)).

As shown by Kučinskas and Peters (2022), the specification of expectation formation above nests all major models of expectations in the existing literature. Such generality is intuitive: Equation (1) is satisfied by the solution to any (linearized) economic model, while equation (2) allows for an unrestricted form of model misspecification as well as idiosyncratic noise on the part of forecasters. While the model specification assumes linearity, it is straightforward to allow for state and/or time dependence by assuming that $\alpha_{j,\ell}$ and $a_{i,j,\ell}$ are functions of some state variables.

Truncating the infinite sum in equation (2) at a finite lag K, we obtain an approximate finite-length dynamic-factor model:

$$\mathbb{F}_{i,t}[x_{t+1}] \approx b_{i,0} + \sum_{j=1}^{M} \sum_{\ell=0}^{K} a_{i,j,\ell+1} \varepsilon_{j,t-\ell} + \nu_{i,t},$$

which we can also write as a static-factor model with $M(K+1)$ factors:

$$\mathbb{F}_{i,t}[x_{t+1}] \approx b_{i,0} + \boldsymbol{\lambda}_i^\top \mathbf{f}_t + \nu_{i,t}, \tag{3}$$

where factor loadings, $\boldsymbol{\lambda}_i$, and factors, $\mathbf{f}_t$, are given by

$$\boldsymbol{\lambda}_i = (a_{i,1,1}, \dots, a_{i,1,1+K}, \dots, a_{i,M,1}, \dots, a_{i,M,1+K})^\top,$$
$$\mathbf{f}_t = (\varepsilon_{1,t}, \dots, \varepsilon_{1,t-K}, \dots, \varepsilon_{M,t}, \dots, \varepsilon_{M,t-K})^\top.$$

Equation (3) is the main theoretical justification for our empirical approach, demonstrating that individual expectations follow an approximate factor model. To estimate noise, we only need to take data on individual expectations and purge them of the factor structure. Importantly, our method does not require data on the realizations of the target variable—a key difference from measuring bias. This feature is useful in applied work, for example, when data are revised over time, or realizations are not observed.

2.2 *Illustration: Noisy-information model*

We now illustrate the general framework with the noisy-information model. While the baseline model is fairly simple, in the Appendix (Section B) we show that the model can be generalized to include richer information structures, more realistic dynamics of the target variable, behavioral biases (including diagnostic expectations), and strategic behavior without changing any of the economic conclusions.

Suppose that the variable of interest, x_t, is an autoregressive process with a persistence coefficient $\rho \in (-1, 1)$ and normally distributed disturbances:

$$x_t = \rho x_{t-1} + \sigma_\varepsilon \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.}\, \mathcal{N}(0, 1) \text{ and } \sigma_\varepsilon > 0.$$

Following the work of Lucas (1973) and Woodford (2003), among others,[5] we assume that forecasters are subject to an information friction. Each forecaster i only observes a noisy signal of x_t:

$$y_{i,t} = x_t + \sigma_{i,\omega}\omega_{i,t}, \quad \omega_{i,t} \sim \text{i.i.d.}\, \mathcal{N}(0, 1) \text{ and } \sigma_{i,\omega} > 0. \tag{4}$$

Here, $\omega_{i,t}$ is an idiosyncratic shock, which is i.i.d. across time and agents, and uncorrelated with ε_t at all leads and lags. The formulation above is standard. Our only generalization from most standard expositions is that we allow for the precision of the idiosyncratic signal to vary across forecasters.[6] Given that noisy information can be microfounded by rational inattention (Sims (1998, 2003)), such heterogeneity is natural. For example, professional macro forecasters are likely to pay closer attention to inflation than football players (i.e., macro forecasters have a lower $\sigma_{i,\omega}$ than football players).

The Kalman filter equations imply that in the steady state

$$\mathbb{F}_{i,t}[x_{t+1}] = \rho\{G_i y_{i,t} + (1 - G_i)\,\mathbb{F}_{i,t-1}[x_t]\}, \tag{5}$$

[5]Angeletos and Lian (2016) provide an overview of this extensive literature.

[6]Similar models with heterogeneity in the precision of signals are studied by Coibion and Gorodnichenko (2012, pp. 130–131) and Dovern (2015, Section 5.1). See also Lahiri and Sheng (2008) and Baker, McElroy, and Sheng (2019).

where $G_i \in [0, 1]$ is the forecaster-specific steady-state Kalman gain, and $\mathbb{F}_{i,t}[x_{t+1}]$ denotes the one-step-ahead forecast of x_{t+1} made by forecaster i. A key feature of equation (5) is that the Kalman gain varies across forecasters, due to the heterogeneous precision of individual signals.

Iterating on equation (5), we find that $\mathbb{F}_{i,t}[x_{t+1}]$ follows an infinite-order dynamic-factor model:

$$\mathbb{F}_{i,t}[x_{t+1}] = \sum_{\ell=0}^{\infty} \lambda_{i,\ell} f_{t-\ell} + \nu_{i,t},$$

where the factor f_t, factor loadings $\lambda_{i,\ell}$, and forecaster-specific noise $\nu_{i,t}$ are

$$f_t = \varepsilon_t, \qquad \lambda_{i,\ell} = \rho^{\ell+1}\left[1 - (1 - G_i)^{\ell+1}\right],$$
$$\nu_{i,t} = \rho G_i \sum_{\ell=0}^{\infty} \left[\rho(1 - G_i)\right]^{\ell} \omega_{i,t-\ell}.$$

Hence, as expected from the general case, the noisy-information model predicts a factor structure for the cross-section of individual expectations. We note that the noise term, $\nu_{i,t}$, is serially correlated with a first-order autocorrelation coefficient of $\rho(1 - G_i)$. For this reason, it is important to allow for serially-correlated noise.

Noise versus sentiment shocks In the Appendix (Section B), we demonstrate that the cross-section of individual expectations continues to exhibit a factor structure in a generalized noisy-information model. The generalized model features sentiment shocks (Lorenzoni (2009), Angeletos and La'O (2013), Nimark (2014)). As the extension makes clear, sentiment shocks are soaked up by the factors. Hence, our measures of noise do not capture sentiment shocks. Some authors use a broader definition of noise that does include sentiment shocks (see, e.g., Chahrour and Jurado (2018)). For such broader definitions of noise, our estimates provide a lower bound. Measuring sentiment shocks requires different tools; see, for instance, Angeletos, Collard, and Dellas (2018).

Violations of simple common-factor structure We are not aware of any prominent models of expectation formation in which individual expectations do not exhibit a factor structure. In part, that is to be expected given the generality of the setup in Section 2.1. In addition, there are fundamental statistical reasons to expect that expectations can often be well approximated by factor models (Juodis (2020), Freeman and Weidner (2022), Menzel (2021)). That being said, it is straightforward to construct (arguably contrived) examples that violate the assumptions in Section 2.1. For that purpose, it is easiest to consider data-generating processes that can generate "weak factor structures" (Chudik, Pesaran, and Tosetti (2011)). As an example, suppose that instead of equation (4), the noisy signal exhibits a spatial MA(1) structure, with

$$y_{i,t} = x_t + \underbrace{\sigma_{i,\omega}\omega_{i,t} + \theta\sigma_{i-1,\omega}\omega_{i-1,t}}_{\equiv\widetilde{\omega}_{i,t}} = x_t + \widetilde{\omega}_{i,t},$$

for $i = 2, \ldots, N$ and appropriately extended for $i = 1$. Although this model generates a factor structure similar to the one discussed previously, factor loadings (through G_i) and forecaster-specific noise terms, $\nu_{i,t}$, are now spatially dependent. Such spatial dependence leaves unchanged most of the statistical properties of the estimation procedures that we consider. However, noise is no longer idiosyncratic, changing the economic interpretation of our estimates.

2.3 *Mean-squared error decomposition*

To compare the relative importance of noise and bias in the data, we provide a novel decomposition of mean-squared forecast errors (MSE). Straightforward algebra using equation (1) and equation (2) shows that forecast errors follow

$$x_{t+1} - \mathbb{F}_{i,t}[x_{t+1}] = -b_{i,0} - \sum_{j=1}^{M}\sum_{\ell=1}^{+\infty} \operatorname{sgn}(\alpha_{j,\ell}) b_{i,j,\ell}\varepsilon_{j,t+1-\ell} - \nu_{i,t} + \sum_{j=1}^{M}\varepsilon_{j,t+1},$$

where $b_{i,j,\ell} \equiv \operatorname{sgn}(\alpha_{j,\ell})(a_{i,j,\ell} - \alpha_{j,\ell})$ denote *bias coefficients* introduced by Kučinskas and Peters (2022). Positive (negative) values of these coefficients indicate overreaction (underreaction) to particular shocks. We now calculate that the mean-squared error of forecaster i is given by

$$\begin{aligned}\mathrm{MSE}_i &\equiv \mathbb{E}_i\big[\big(x_{t+1} - \mathbb{F}_{i,t}[x_{t+1}]\big)^2\big] \\ &= \underbrace{b_{i,0}^2}_{\text{uncond. bias}} + \underbrace{\sum_{j=1}^{M}\mathbb{V}\mathrm{ar}[\varepsilon_{j,t}]\left(\sum_{\ell=1}^{+\infty} b_{i,j,\ell}^2\right)}_{\text{conditional bias}} + \underbrace{\mathbb{V}\mathrm{ar}_i[\nu_{i,t}]}_{\text{noise}} + \underbrace{\sum_{j=1}^{M}\mathbb{V}\mathrm{ar}[\varepsilon_{j,t}]}_{\text{irreducible error}}. \end{aligned} \tag{6}$$

Here, $\mathbb{E}_i[\cdot]$ denotes the expectation conditional on forecaster-specific time-invariant characteristics (and similarly for $\mathbb{V}\mathrm{ar}_i[\cdot]$). The equation above shows that the mean-squared error of an individual forecaster can be decomposed into components due to (i) bias (from both unconditional and conditional bias); (ii) noise; and (iii) irreducible error.

To implement equation (6) empirically, we use the following approach. Unconditional bias $b_{i,0}^2$ and MSE_i can be consistently estimated from the data forecaster-by-forecaster. Our factor-based procedure produces an estimate of $\mathbb{V}\mathrm{ar}_i[\nu_{i,t}]$. Hence, given an estimate for the variance of irreducible error, we can back out the size of the conditional bias for every forecaster. To estimate the variance of irreducible error, we follow the approach put forward by Kučinskas and Peters (2022). First, define the consensus forecast as $\bar{\mathbb{F}}_t[x_{t+1}] = N^{-1}\sum_{i=1}^{N}\mathbb{F}_{i,t}[x_{t+1}]$. For large N, idiosyncratic noise vanishes at the consensus level, and the consensus mean-squared forecast error, $\mathrm{MSE}_{\text{cons.}}$, is given by

$$\mathrm{MSE}_{\text{cons.}} = \underbrace{(\bar{b}_0)^2}_{\text{uncond. bias}} + \underbrace{\sum_{j=1}^{M}\mathbb{V}\mathrm{ar}[\varepsilon_{j,t}]\left(\sum_{\ell=1}^{+\infty}\bar{b}_{j,\ell}^2\right)}_{\text{conditional bias}} + \underbrace{\sum_{j=1}^{M}\mathbb{V}\mathrm{ar}[\varepsilon_{j,t}]}_{\text{irreducible error}}, \tag{7}$$

where $\bar{b}_0 = N^{-1}\sum_{i=1}^{N} b_{i,0}$ and $\bar{b}_{j,\ell} = N^{-1}\sum_{i=1}^{N} b_{i,j,\ell}$ denote average bias coefficients. As before, we estimate the unconditional bias directly from the data on forecast errors. We then approximate the conditional bias term by estimating the univariate impulse-response function of forecast errors. Intuitively, if the consensus systematically under- or overreacts to news, its forecast errors are predictable. We purge the observed consensus forecast errors of this predictability, and thereby obtain an estimate of the irreducible error. We provide the details behind our empirical implementation in Section 3.2.

2.4 *Estimation*

To measure noise in the data, we employ established techniques for estimating factor models (see, e.g., Stock and Watson (2016)). While we make no contributions here, we outline the key ideas. We emphasize that the panel data set must contain forecasts of the *same* common variable x_{t+1} made by multiple forecasters.

Single factor The easiest case for estimation is when a single factor drives the observed forecasts, as in the noisy-information model of Section 2.2:

$$\mathbb{F}_{i,t}[x_{t+1}] = b_{i,0} + \psi_i(L) f_t + \nu_{i,t}$$

for some lag polynomial $\psi_i(L)$. As before, define the consensus, or average, forecast by $\bar{\mathbb{F}}_t[x_{t+1}] = N^{-1}\sum_{i=1}^{N} \mathbb{F}_{i,t}[x_{t+1}] \approx \bar{b}_0 + \bar{\psi}(L) f_t$ when N is large, where $\bar{b}_0$ and $\bar{\psi}(L)$ denote the cross-sectional averages of $b_{i,0}$ and $\psi_i(L)$. Then we can write

$$\mathbb{F}_{i,t}[x_{t+1}] \approx b_{i,0} + \psi_i(L)\left(\frac{\bar{\mathbb{F}}_t[x_{t+1}] - \bar{b}_0}{\bar{\psi}(L)}\right) + \nu_{it} \equiv a_{i,0} + \theta_i(L)\bar{\mathbb{F}}_t[x_{t+1}] + \nu_{i,t},$$

where $\theta_i(L) \equiv \psi_i(L)/\bar{\psi}(L)$, and $a_{i,0}$ is a composite forecaster-specific intercept. Hence, to estimate noise, we can simply regress individual forecasts on the consensus (and its lagged values) forecaster-by-forecaster:

$$\mathbb{F}_{i,t}[x_{t+1}] = \alpha_i + \sum_{\ell=0}^{K} \beta_{i,\ell}\bar{\mathbb{F}}_{t-\ell}[x_{t-\ell+1}] + \nu_{i,t}.$$

Here, K is the number of lagged consensus forecasts that are included in the regression. Then the estimated noise is given by the residuals of the regression. As discussed in Chudik and Pesaran (2015), for consistent estimation, K should grow with T, the time-series dimension of the data set (albeit at a rate slower than $T^{1/3}$).

This approach generalizes the pioneering work of Figlewski (1983) who estimated a special case of the regression above with $K = 1$. While the approach of using consensus forecasts is only valid when the true model has a single factor, it has the benefit of being very easy to implement in practice. In our empirical application, it also gives similar results to those from more sophisticated procedures.[7]

[7]The use of the cross-sectional averages as factor proxies has been popularized by Pesaran (2006) in the context of the so-called Common Correlated Effects (CCE) estimator. Theoretical and Monte Carlo results in Westerlund and Urbain (2015) and Juodis, Karabıyık, and Westerlund (2021) suggest that (when applicable) this method has better small sample behavior as compared to the principal components procedure.

Multiple factors To estimate models with multiple factors, we use an iterative maximum-likelihood procedure (Bai (2003, 2009)). Suppose that the model to be estimated is

$$\mathbb{F}_{i,t}[x_{t+1}] = b_{i,0} + \boldsymbol{\lambda}_i^\top \mathbf{f}_t + \nu_{it}$$

for a given number of factors R. We initialize the estimation procedure by some initial values for the factors, say, $\hat{\mathbf{f}}_t^{(0)}$. Then we estimate the factor loadings $\boldsymbol{\lambda}_i$ and $b_{i,0}$ by running time-series regressions

$$\mathbb{F}_{i,t}[x_{t+1}] = b_{i,0} + \boldsymbol{\lambda}_i^\top \hat{\mathbf{f}}_t^{(0)} + \nu_{i,t}$$

for each forecaster $i = 1, \ldots, N$ using only those time periods when the specific forecaster is active. With the estimated loadings and forecaster fixed effects at hand (say $\hat{\boldsymbol{\lambda}}_i^{(0)}$ and $\hat{b}_{i,0}^{(0)}$), we run cross-sectional regressions,

$$\mathbb{F}_{i,t}[x_{t+1}] - \hat{b}_{i,0}^{(0)} = \mathbf{f}_t^\top \hat{\boldsymbol{\lambda}}_i^{(0)} + \nu_{it},$$

for each time period $t = 1, \ldots, T$ (using only forecasters active at time t) to obtain estimates of the factors $\hat{\mathbf{f}}_t^{(1)}$. We then iterate this procedure until the change in the sum of squared residuals is lower than some tolerance level. The estimated noise is then given by the residuals from the time-series regressions. To increase the probability of locating the global rather than just a local minimum, we repeat the procedure for multiple initial starting values.

A key advantage of this simple iterative procedure in our empirical setting is that the panel data set we use is highly unbalanced (as is the case for most data sets on survey expectations).[8] Due to the unbalancedness, the estimates of factors and factor loadings obtained using the iterative procedure are not asymptotically efficient. In particular, more efficient estimates can be obtained by combining principal-component estimation with the EM algorithm (as suggested by Stock and Watson (2002) and Jin, Miao, and Su (2021)), or using matrix-completion methods (as suggested by Fernández-Val, Freeman, and Weidner (2021)). However, estimates obtained in such a way will generally result in larger estimates of noise. This conclusion follows from observing that our suggested procedure minimizes the sum of squared residuals (hence noise), while the other procedures do not. Thus, the estimates of noise obtained using our iterative procedure should be seen as lower bounds on the true level of noise.

3. Empirical results

For our empirical application, we use data from the Survey of Professional Forecasters (SPF) administered by the Federal Reserve Bank of Philadelphia. The survey is conducted quarterly and contains forecasts of multiple macro variables made by professional forecasters.

[8]This feature of the data effectively rules out estimation of the underlying dynamic-factor model using frequency domain techniques, as studied by Forni, Hallin, Lippi, and Zaffaroni (2015, 2017).

Table 1. Summary statistics.

	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$
			Initial Sample		
Total Forecasts	7712	7727	7723	7696	7297
Med. # Forecasts/Forecaster	16.00	16.00	16.00	16.00	14.00
Med. # Forecasts/Survey	36.00	37.00	37.00	36.00	35.00
Avg. Forecast	3.51	3.46	3.44	3.41	3.36
Min. Forecast	−12.74	−6.35	−6.66	−16.11	−19.45
Max. Forecast	29.95	32.11	32.29	31.16	29.75
Avg. Forecast Error	−0.04	0.01	0.04	0.06	0.04
Std. Dev. Forecast	2.56	2.37	2.27	2.26	2.23
Std. Dev. Forecast, Within	1.79	1.68	1.60	1.65	1.66
			Final Sample		
Total Forecasts	6007	5964	5924	5858	5552
Avg. Forecast	3.47	3.44	3.41	3.39	3.33
Min. Forecast	−12.74	−6.35	−6.66	−16.11	−19.45
Max. Forecast	29.95	32.11	32.29	31.16	29.75
Avg. Forecast Error	−0.01	0.05	0.07	0.11	0.11
Std. Dev. Forecast	2.65	2.45	2.34	2.33	2.32
Std. Dev. Forecast, Within	1.86	1.71	1.66	1.71	1.75

Note: Summary statistics for quarterly inflation forecasts. The initial sample period is 1968Q2–2019Q2. The final sample consists of all forecasters with at least 20 forecasts and all survey dates with at least 20 forecasters. Forecast horizon h runs from $h=0$ (nowcasts) to $h=4$ (1-year-ahead forecasts) quarters. *Std. dev. forecast, within* denotes the standard deviation of forecasts after the within transformation (subtracting the sample mean for each individual forecaster).

We focus on inflation forecasts. Inflation expectations play a central role in modern economic models (Coibion, Gorodnichenko, and Kamdar (2018)). In addition, the data set has been used extensively in prior research, and hence provides an ideal testing ground for a new methodology. We study forecasts of quarterly inflation, with inflation derived from the GDP price deflator. In the main text, we describe the empirical findings and discuss data construction in the Appendix (Section A).

Table 1 provides summary statistics. The data set contains forecasts at horizons ranging from nowcasts (current quarter) up to 1-year-ahead forecasts. For most forecast horizons, our initial sample contains around 7700 individual forecasts. However, the data set is highly unbalanced, with the median forecaster participating in only 14–16 survey waves. To allow for reasonably precise estimation of noise, we restrict the sample to (i) forecasters that provide at least 20 nowcasts of inflation; and (ii) dates on which at least 20 forecasters submit nowcasts.[9] We ensure that these two restrictions hold simultaneously by iteratively removing observations that do not satisfy these conditions. Whenever we look at other subsamples of the data, we always make sure that the data set only contains forecasters with 20 nowcasts and dates with at least 20 nowcasts.

Our final sample has around 6000 individual forecasts. Comparing the initial and final samples, the summary statistics look very similar, with comparable average fore-

[9] Ideally, we would impose a more stringent requirement, given that the precision of factor estimates is proportional to the square root of the number of individuals. However, due to data limitations, satisfying the joint requirement for, say, 30 and 30 is not feasible; see Table C.8 in the Appendix.

casts, average forecast errors, and standard deviations of forecasts. Hence, sample selection effects seem unlikely to have a major impact on our results.

3.1 *Baseline estimates*

Estimated level of noise. Our baseline results are given in the Table 2. The left panel of the table shows the results from a dynamic-factor model with a single factor, as discussed in Section 2.4; K denotes the number of lagged consensus forecasts used in the regressions. The right panel shows the results using the iterative procedure of Bai (2009); R denotes the number of factors in the static-factor model representation. The table reports the average standard deviation of noise (averaged across forecasters).[10]

Depending on the forecast horizon and model specification, the standard deviation of noise ranges between 0.60 and 1.10. The magnitudes are economically significant. As shown in the Table 1, the within-forecaster standard deviation of forecasts is around 1.70. Hence, a substantial part of within-forecaster variation in the observed forecasts is due to noise. For example, if we take the standard deviation of noise to be 0.85 (midpoint of our estimates), then our estimates attribute 25% of the within-forecaster variance to noise (i.e., $(0.85/1.70)^2$). Especially given that we use data from professional forecasters, the amount of noise is substantial.

Number of factors Our estimates of noise depend on the number of factors included in the model, with more factors leading to a lower estimated level of noise. We therefore

Table 2. Noise: baseline estimates.

	Single Factor (Consensus)				Multiple Factors			
	$K=0$	$K=1$	$K=2$	$K=3$	$R=1$	$R=2$	$R=3$	$R=4$
$h=0$	1.11	1.07	1.03	1.00	1.10	0.95	0.84	0.73
	(0.07)	(0.06)	(0.06)	(0.06)	(0.06)	(0.04)	(0.04)	(0.03)
$h=1$	0.99	0.95	0.92	0.89	0.99	0.83	0.73	0.64
	(0.07)	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.04)	(0.03)
$h=2$	0.99	0.95	0.92	0.89	0.97	0.82	0.72	0.62
	(0.07)	(0.07)	(0.06)	(0.06)	(0.07)	(0.05)	(0.04)	(0.03)
$h=3$	1.03	0.99	0.97	0.94	1.02	0.84	0.72	0.60
	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)	(0.05)	(0.04)	(0.03)
$h=4$	1.08	1.04	0.99	0.96	1.08	0.87	0.73	0.62
	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.06)	(0.04)	(0.03)

Note: Standard deviations of noise (averages across forecasters). Cross-sectional standard errors (shown in parentheses) are obtained by regressing $\hat{\sigma}_i = \alpha + u_i$ and using the standard error for α, where $\hat{\sigma}_i$ denotes the estimated standard deviation of noise for forecaster i. Forecast horizon h runs from $h=0$ (nowcasts) to $h=4$ (1-year-ahead forecasts) quarters.

[10]In our calculation of standard errors, we make the implicit assumption that the data has a random-coefficient structure in that the underlying population values for the standard deviations vary across forecasters. Thus, the leading term in the asymptotic distribution of $N^{-1}\sum_{i=1}^{N}\hat{\sigma}_i$, where $\hat{\sigma}_i$ denotes the estimated standard deviation of noise, is solely determined by that variation; see, for example, Okui and Yanagi (2019).

Table 3. Optimal number of factors.

	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$
Raw Forecasts	1	1	1	1	3
Adjusted for Time F.E.	2	1	1	3	2

Note: Number of factors selected by the pseudo growth ratio (GR) criterion of Ahn and Horenstein (2013) using the least-squares objective function. We calculate the statistics for $R \in \{1, \dots, 5\}$ and report the number of factors R with the highest statistic. *Raw forecasts* show the number of factors selected when raw forecasts are used; *Adjusted for time F.E.* gives the number of factors selected after subtracting the consensus forecast from each individual forecast.

next study the number of factors supported by the data. To do so, we use the (pseudo) growth-ratio criterion of Ahn and Horenstein (2013) calculated on the basis of the sum of squared residuals. As shown in Table 3, when raw forecasts are used, the procedure typically selects a single factor. However, the procedure may be subject to the one-factor bias (see, e.g., Ahn and Horenstein (2013), p. 1209).[11] Hence, we also show the results when the consensus forecast is subtracted from the individual forecasts. Consistent with a one-factor bias, the procedure now often selects a higher factor. However, the maximum number of factors selected is still only 3, and the procedure never selects the maximum value of factors that we allow (5). While there is some remaining uncertainty about how many factors should be selected, the estimates of noise in Table 2 are fairly robust to the precise number of factors chosen. For most forecasting horizons, going from $R=1$ to $R=3$ (the maximum selected by the growth-ratio criterion) only decreases the estimated standard deviation of noise by around 25%.[12]

Factor estimates Differently from much prior work using factor models, we are not directly interested in the factor estimates. Nevertheless, Figure 1 plots the estimated factors across different forecasting horizons. For each forecasting horizon, we plot the estimated factor from the corresponding model with $R=1$ (estimated using the iterative maximum likelihood procedure).[13] As seen in the graph, the estimated factors are strongly correlated with realized inflation. For nowcasts ($h=0$), the correlation between the first factor and realized inflation is 0.91. The correlation declines at longer forecasting horizons, with a correlation of 0.70 for 1-year-ahead forecasts, as forecasters react less to transitory movements in inflation. The estimated factors become less volatile as the forecasting horizon increases, and forecasters filter out short-term shocks.

[11] Intuitively, the one-factor bias is caused by the fact that in a model such as $z_{i,t} = g_t + \lambda_i f_t + \varepsilon_{i,t}$, the first eigenvalue of the covariance matrix of $z_{i,t}$ can be made arbitrary large by increasing the value of g_t. This substantially complicates the consistent estimation of the true number of factors.

[12] Ideally, we would also like to formally test for the number of factors in our setup. However, the procedure of Onatski (2009) is not applicable due to the unbalanced data set we consider, while the weighted Cross-section Dependence (CD) statistic of Juodis and Reese (2022) will lack power to reject any null hypothesis because of the relatively short time-series dimension.

[13] Given the highly unbalanced nature of our data, the "first" (or "most important") factor is no longer well-defined for models with multiple factors. However, given that we find that a single-factor is typically selected by the Ahn and Horenstein (2013) procedure, any reasonable definition of the "first" factor would arguably look similar to the single-factor estimates that we report.

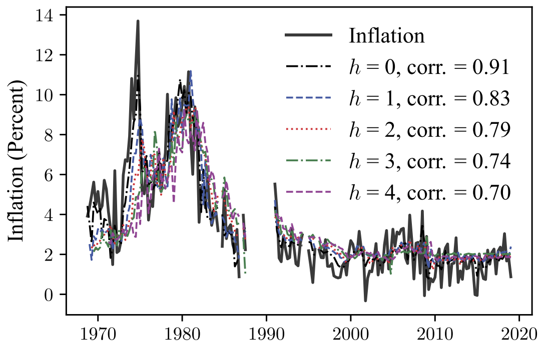

FIGURE 1. Estimated factors. *Note:* Estimated factors for h-quarter-ahead forecasts and realized inflation. Factor estimates from a single-factor model estimated using the iterative maximum–likelihood procedure shown. For easier interpretability, we rotate the factors by regressing realized inflation on the first factor (for each forecasting horizon separately) and then plotting the fitted values. The legend shows the correlation of the respective factor with realized inflation.

3.2 *Noise versus bias*

To compare the relative importance of noise and bias, we now implement the decomposition of mean-squared forecast errors developed in Section 2.3. We first use consensus forecasts to quantify the size of irreducible error (Kučinskas and Peters (2022)). Intuitively, we use the predictability of forecast errors to remove the conditional-bias component from the consensus MSE, and attribute the remainder to irreducible error. Empirically, the consensus forecast in the SPF is arguably the best available predictor of inflation (Ang, Bekaert, and Wei (2007)). Hence, the consensus MSE should already be a good proxy for the variance of the irreducible error. To be conservative, we go a step further and remove the observed (ex post) predictability of consensus forecast errors. The remaining details are provided in Appendix D (Section D.2).

The results are shown in Table 4. For short-term forecasts, noise is as important as bias, with both noise and bias accounting for around one-third of the MSE at the forecaster level (with the remaining third due to irreducible error). For three-quarter and

TABLE 4. Noise versus bias.

	Individual				Consensus		
	MSE	Bias	Noise	Irreducible	MSE	Bias	Irreducible
$h = 0$	3.53	37.3%	33.0%	29.7%	1.17	10.0%	90.0%
$h = 1$	4.46	31.4%	34.1%	34.5%	1.99	22.8%	77.2%
$h = 2$	5.12	33.7%	30.0%	36.3%	2.55	27.0%	73.0%
$h = 3$	6.42	57.4%	11.7%	30.9%	2.98	33.5%	66.5%
$h = 4$	7.50	53.1%	17.3%	29.6%	3.55	37.5%	62.5%

Note: Baseline estimates of noise and bias as a fraction of mean-squared forecast error. Forecast horizon h runs from $h = 0$ (nowcasts) to $h = 4$ (1-year-ahead forecasts) quarters. The iterative maximum-likelihood procedure with the number of factors chosen by the growth-ratio statistic of Ahn and Horenstein (2013) (adjusted for time fixed effects) is used to estimate the level of noise. The amount of irreducible error is obtained using the method proposed by Kučinskas and Peters (2022).

four-quarter-ahead forecasts, we find that noise matters slightly less, explaining around 10% of the MSE. This latter result is primarily driven by the fact that for these horizons, the Ahn and Horenstein (2013) growth-ratio criterion selects a higher number of factors, thereby reducing our estimated level of noise. At the consensus level, idiosyncratic noise washes out. Interestingly, we find that bias increases with the forecast horizon at the consensus level due to stronger underreaction to news (i.e., conditional bias) at longer horizons. Overall, we conclude that (i) around 10–30% of the observed MSE at the forecaster level is due to noise; and (ii) noise and bias are similarly important quantitatively, especially at shorter forecasting horizons.

3.3 *Robustness tests*

Our estimates of noise are just residuals from a factor model. Hence, a potential worry is that we may overestimate the amount of noise if the factor model is misspecified. To avoid this problem, we make conservative measurement choices to obtain lower bounds on the amount of noise whenever possible. For example, we do not perform degrees-of-freedom adjustments for our estimates of noise. Given that many of the time-series regressions have a small number of observations, this choice leads to a substantial underestimation of noise; see Chen, Fernández-Val, and Weidner (2021, Section 3) for the intuition.

While it is not possible to fully rule out misspecification, in this section we perform a number of model specification tests to help alleviate potential concerns. In addition, we investigate whether the estimates of noise are similar across alternative specifications. Overall, the results seem fairly robust across a range of assumptions. These findings suggest that our baseline results are not driven by specific modeling choices or sample restrictions.

3.3.1 *Predictability tests* If the baseline model is well specified, our estimates of noise should not predict actual inflation. That is a testable implication, which can be used to check the specification of our model. Specifically, in a regression of the form

$$x_{t+h} = \alpha_i + \gamma_i \hat{\nu}_{i,t}^{(h)} + u_{i,t},$$

we should have $\gamma_i = 0$, where $\hat{\nu}_{i,t}^{(h)}$ denotes our estimate of noise for h-step-ahead forecasts made by forecaster i at time t. In contrast, if there is some remaining information in our noise estimates that is helpful for predicting actual inflation, we would have $\gamma_i \neq 0$.

Table 5 presents the results of these predictability regressions, estimated for each forecaster separately. Conducting inference on γ_i requires calculating standard errors that are robust to the generated-regressors problem. The generated-regressors problem arises because we use residuals on the right-hand side of the regression, and these residuals are calculated by first estimating forecaster fixed effects and factor loadings. In the Appendix (Section D.1), we derive the corrected standard errors and use a Newey–West-type estimator that is robust to heteroskedasticity and autocorrelation.

The results indicate that for most forecasters and forecasting horizons, we cannot reject the null hypothesis of no predictability. The empirical frequency of rejections at

TABLE 5. Is noise predictive of actual inflation?

	N	% $t \geq 1.96$	% $t \leq -1.96$	% $\|t\| \geq 1.96$	Med. R^2	P90 R^2
$h = 0$	141	3.55	1.42	4.96	0.01	0.05
$h = 1$	141	2.13	0.00	2.13	0.01	0.09
$h = 2$	141	0.71	2.84	3.55	0.01	0.09
$h = 3$	141	2.84	1.42	4.26	0.02	0.10
$h = 4$	141	2.13	0.71	2.84	0.02	0.12

Note: Results from predictability regressions $x_{t+h} = \alpha_i + \gamma_i \hat{\nu}_{i,t}^{(h)} + u_{i,t}$ where $\hat{\nu}_{i,t}^{(h)}$ denotes our estimates of noise, and h is forecast horizon. The regression is run for all forecasters and forecast horizons. The table shows the fraction of regressions with t-statistics above 1.96, below -1.96, and above 1.96 in absolute value. The final two columns provide the median and 90th percentile of the regression R^2's. The standard errors used in the calculation of the t-statistics are adjusted for generated regressors and robust to heteroskedasticity and autocorrelation, with 4 Newey–West lags.

a nominal level of 5% is close to 5%. In addition, rejections do not appear to be concentrated among positive (or negative) values of t-statistics. Even looking at the raw R-squared's of these regressions, we obtain very small numbers, indicating no predictability. The median R-squared's for these regressions are in the range of 1–2%. Even the 90th percentiles of the estimated R-squared's are low, at 5–12% across the forecast horizons. We conclude that our estimates of noise do not seem predictive of actual inflation, as should be the case if our model is well specified.

3.3.2 *Further tests* Appendix C contains a battery of additional robustness checks. In particular, we show that the results remain very similar if we (i) orthogonalize our estimates w.r.t. realized inflation; (ii) redefine noise to be uncorrelated over time; (iii) redefine noise to be both uncorrelated over time and orthogonal to realized inflation; (iv) remove potential outliers; and (v) reestimate the model on different subsamples. Finally, we document that (vi) for a given forecaster, the level of noise is strongly positively correlated across different forecast horizons. This last result suggests that our estimates of noise capture an economically meaningful forecaster-specific trait.

4. Applications: Putting noise to work

We now provide several applications of our estimates. Our goal is to show that noise matters for understanding a variety of substantive economic questions.

4.1 *Wisdom of the crowd*

Consensus (or average) forecasts are typically substantially more accurate than most individual forecasts, an effect popularized as the "wisdom of the crowd" by Surowiecki (2005). If individual forecasts are noisy, then consensus forecasts may be more accurate simply because they average away the idiosyncratic noise.[14] Alternatively, the reduction

[14]The use of cross-sectional averaging as a way to reduce measurement error and endogeneity is well known in the literature on pseudo panel data models; see, for example, Verbeek (2008). For example, Juodis (2018) uses this insight to estimate factor-augmented panel data models from repeated cross-sections.

in mean-squared error could stem primarily from a reduction in bias. Our estimates can be used to quantify the sources of the wisdom-of-the-crowd effect.

The MSE decomposition results in Table 4 imply that 20–60% of the wisdom-of-the-crowd effect is explained by the reduction in idiosyncratic noise (with the rest of the effect due to a reduction in bias). For instance, for one-quarter-ahead forecasts ($h = 1$), 61.6% of the improvement in the consensus forecast can be attributed to the reduction in noise ($0.341 \times 4.46/(4.46 - 1.99)$). Similarly, for nowcasts ($h = 0$) and two-quarter ahead ($h = 2$) forecasts, 50% or more of the improvement is due to noise. Noise plays a smaller role for three-step-ahead forecasts, with slightly over 20% of the wisdom-of-the-crowd effect attributable to a reduction in noise.

4.2 *Are micro Coibion–Gorodnichenko regressions robust?*

Our estimates of noise can be used to check whether empirical findings are robust to realistic levels of measurement error in expectations. We illustrate this idea using micro-level (Coibion and Gorodnichenko (2015), CG for short) regressions as a case study.

Empirical approach In an influential contribution, Coibion and Gorodnichenko showed that information rigidities can be identified by regressing consensus forecast errors on lagged revisions:

$$x_{t+h} - \bar{\mathbb{F}}_t[x_{t+h}] = \alpha + \beta_c\big\{\bar{\mathbb{F}}_t[x_{t+h}] - \bar{\mathbb{F}}_{t-1}[x_{t+h}]\big\} + u_t. \tag{8}$$

Here, $\bar{\mathbb{F}}_t[x_{t+h}]$ denotes the average (consensus) forecast. Recently, the same method has also been applied to *individual* data on expectations (e.g., see Bordalo et al. (2020), Broer and Kohlhas (2018), Bouchaud, Krüger, Landier, and Thesmar (2019)). With individual data, one estimates

$$x_{t+h} - \mathbb{F}_{i,t}[x_{t+h}] = \alpha + \beta_m\big\{\mathbb{F}_{i,t}[x_{t+h}] - \mathbb{F}_{i,t-1}[x_{t+h}]\big\} + u_{i,t}, \tag{9}$$

where i denotes an individual forecaster. The existing empirical work has found that the *micro estimate* $\hat{\beta}_m$ is typically smaller than the *consensus estimate* $\hat{\beta}_c$, and the micro estimate is often negative. This finding is inconsistent with standard models of information rigidities (as well as full-information rational expectations).

A potential concern with the micro estimates is that they are sensitive to measurement error in expectations. If $\mathbb{F}_{i,t}[x_{t+h}]$ is measured with error, measurement error enters both the left- and right-hand sides of equation (9). This leads to a nonclassical measurement-error problem. The measurement-error problem can generate a negative estimate of β_m with high probability even when the true coefficient is zero.

Our estimates of noise can be used to investigate whether empirical estimates of β_m are robust to realistic amounts of measurement error. Suppose that instead of observing the true expectations, we only observe

$$\mathbb{F}_{i,t}[x_{t+h}] = \mathbb{F}^*_{i,t}[x_{t+h}] + \eta_{i,t},$$

where $\mathbb{F}^*_{i,t}[x_{t+h}]$ denotes the true (unobserved) expectations of an agent i, and $\eta_{i,t}$ is measurement error. We assume that the measurement error is uncorrelated across forecasters, uncorrelated with the target variable x_t at all leads and lags, and has finite variance σ^2_η. However, we allow measurement error to be serially correlated and denote $\phi \equiv \mathbb{C}\mathrm{orr}(\eta_{i,t}, \eta_{i,t-1})$.

Suppose that the true forecast errors are unpredictable by past forecast revisions at the micro level, that is,

$$\mathbb{C}\mathrm{ov}\big(x_{t+h} - \mathbb{F}^*_{i,t}[x_{t+h}], \mathbb{F}^*_{i,t}[x_{t+h}] - \mathbb{F}^*_{i,t-1}[x_{t+h}]\big) = 0.$$

However, since the true expectations are unobserved, an analyst that uses the observed expectations to run the regression equation (9) would in effect be estimating

$$x_{t+h} - \mathbb{F}^*_{i,t}[x_{t+h}] - \eta_{i,t} = \alpha + \beta_m\big\{\mathbb{F}^*_{i,t}[x_{t+h}] - \mathbb{F}^*_{i,t-1}[x_{t+h}] + \eta_{i,t} - \eta_{i,t-1}\big\} + u_{i,t}.$$

Therefore, the probability limit of the OLS estimator in the presence of measurement error is

$$\mathrm{plim}_{N,T\to\infty}\,\hat{\beta}_m = -\frac{\sigma^2_\eta(1-\phi)}{\mathbb{V}\mathrm{ar}[r_{i,t}]}, \tag{10}$$

where $r_{i,t} \equiv \mathbb{F}_{i,t}[x_{t+h}] - \mathbb{F}_{i,t-1}[x_{t+h}]$ denotes the observed forecast revision. Hence, the estimated value of β_m is negative as long as there is some measurement error ($\sigma^2_\eta > 0$), and measurement error is not a unit root process ($|\phi| < 1$).

We can now invert equation (10) for σ^2_η to ask what amount of measurement error is necessary to generate the empirically observed value of $\hat{\beta}_m$ under the null hypothesis that the true coefficient is zero, that is,

$$\widetilde{\sigma}^2_\eta = \frac{(-\hat{\beta}_m)\widehat{\mathbb{V}\mathrm{ar}}[r_{i,t}]}{1-\phi}. \tag{11}$$

Here, $\widehat{\mathbb{V}\mathrm{ar}}[r_{i,t}]$ denotes the sample variance of forecast revisions. We can then compare the value of $\widetilde{\sigma}^2_\eta$ in equation (11) to our independent estimates of noise. If $\widetilde{\sigma}^2_\eta$ in equation (11) is larger than our estimate of noise, that would suggest robustness to measurement error: More measurement error is necessary to explain the empirical finding (assuming that the true value is zero) than is plausible given the amount of noise in the data.[15]

Results The results of this exercise (using one- and two-quarter ahead forecasts) are given in Table 6. The first two rows of the table show the estimated slope coefficients from Coibion–Gorodnichenko regressions using consensus as well as individual forecasts. Consistent with previous findings in the literature, the consensus estimates are all positive, while the estimates from the micro data are all negative. We next calculate the amount of measurement error that would be necessary to explain the negative micro-level estimates, under the assumption that the true coefficient is zero. The resulting estimates of measurement error are fairly high, at around 0.90–1.00 percentage points.

[15]Measurement error can only explain negative values of $\hat{\beta}_m$. Of course, we recognize that our simple calibration exercise does perform proper statistical inference as we compare point estimates and not their population values.

TABLE 6. Micro CG estimates: robustness to measurement error.

	$h = 0$	$h = 1$	$h = 2$	$h = 3$
Consensus	0.39	0.35	0.84	1.15
Individual	−0.36	−0.34	−0.33	−0.32
Revisions (S.D.)	1.56	1.41	1.48	1.58
Implied Measurement Error (S.D.), $\phi = 0$	0.94	0.82	0.85	0.90
Implied Measurement Error (S.D.), $\phi = 0.20$	1.05	0.92	0.95	1.01
Estimated Noise (S.D.)	0.95	0.99	0.97	0.72

Note: The table investigates the robustness of micro-level estimates of the Coibion and Gorodnichenko (2015) regression to measurement error. The first row shows the estimated slope coefficient in the Coibion–Gorodnichenko regression when consensus forecasts are used (for $h = 1$), as in equation (8). The second row shows the results when individual forecasts are used (for $h = 1$), as in equation (9). The third row gives the sample standard deviation of forecast revisions. The fourth and fifth rows give the standard deviation of measurement error that is necessary to generate the empirical micro estimates, assuming that the true coefficient is zero, using the formula in equation (11). The fourth row assumes that measurement error is serially uncorrelated ($\phi = 0$), while the fifth row assumes that the first-order autocorrelation of measurement error is $\phi = 0.20$. The final row gives our estimates of the standard deviation of noise, with the number of factors selected by the growth-ratio statistic of Ahn and Horenstein (2013) after subtracting the consensus forecast from the individual forecasts, as in Table 3.

Economically, roughly one-third of the variation in forecast revisions would need to be attributed to measurement error to explain the observed negative micro-level coefficients. In addition, the amount of measurement error required is very close to our estimated levels of noise. Since noise is likely to arise for multiple reasons, measurement error being just one of them, we conclude that the micro-level estimates appear to be robust to realistic amounts of measurement error.

4.3 *Using noise to estimate and test economic models*

Our estimates of noise can be used to estimate and test theoretical models of expectations, as we now illustrate using the noisy-information model introduced in Section 2.2. We also show how, under additional assumptions on the expectation-formation process, our estimates can be leveraged to quantify the level of measurement error.

The noisy-information model posits a link between the amount of noise in individual signals—and thereby expectations—and persistence in forecast errors. Intuitively, if forecast errors are persistent, the model suggests that agents are subject to a strong information friction. Vice versa, for a given level of noise, the model makes a prediction for how persistent forecast errors should be. We can therefore test the model by checking if these two predictions are consistent with each other. To the best of our knowledge, this prediction of the noisy-information model has not previously been tested, arguably because past research could not quantify the level of noise in expectations.

Noise-based estimate of information frictions In the noisy-information model of Section 2.2, forecast errors follow an AR(1) process with heterogeneous (forecaster-specific) coefficients:

$$x_{t+1} - \mathbb{F}_{i,t}[x_{t+1}] = \underbrace{\rho(1 - G_i)}_{\phi_i}\left\{x_t - \mathbb{F}_{i,t-1}[x_t]\right\} + (\varepsilon_{t+1} - \rho G_i \omega_{i,t}). \tag{12}$$

If the information friction is greater (i.e., G_i is lower), forecast errors are more persistent. Hence, following Ryngaert (2018), we can back out the level of information frictions from the persistence of forecast errors. To do, we regress forecast errors $e_{i,t} \equiv x_{t+1} - \mathbb{F}_{i,t}[x_{t+1}]$ on lagged forecast errors forecaster-by-forecaster; denote the estimated persistence of forecast errors by $\hat{\phi}_i$. Then we estimate the Kalman gain as $\hat{G}_i = 1 - \hat{\phi}_i/\hat{\rho}$, where $\hat{\rho}$ is the estimated persistence of inflation. This approach is standard in the literature (Coibion and Gorodnichenko (2012), Ryngaert (2018)). The only difference here is that, since we allow the information friction to vary across forecasters, we estimate equation (12) forecaster-by-forecaster.

At the same time, the variance of noise in this model equals

$$\mathbb{V}\text{ar}_i[\nu_{i,t}] = \frac{\rho^2 G_i^2 \sigma_{i,\omega}^2}{1 - \rho^2(1 - G_i)^2}. \tag{13}$$

Our estimates of noise provide us with a value for the left-hand side. As discussed above, the Kalman gain, G_i, can be estimated from the persistence of forecast errors. Therefore, we can calculate the value of $\sigma_{i,\omega}^2$ that is necessary for the model to generate the observed level of noise. However, $\sigma_{i,\omega}^2$ and G_i are not independent—less informative signals decrease the Kalman gain. Hence, we can plug in the obtained value for $\sigma_{i,\omega}^2$ to the standard filtering formula to produce another estimate of the Kalman gain. If the noisy-information model provides a good approximation to the observed expectations, the two Kalman gain estimates should, on average, be close to each other.

Since our estimation leverages individual-level data, it is important to allow for potential measurement error. In particular, measurement error attenuates the estimated persistence of forecast errors toward zero. Hence, failing to account for measurement error may underestimate the level of information frictions. For this reason, we show our results for a range of assumptions on what fraction of noise in expectations is due to measurement error, and we use measurement-error-adjusted estimates of forecast-error persistence. The full details are provided in Appendix D (Section D.3).

Results Figure 2 shows the estimated Kalman gain parameters using the two alternative identification schemes. Under the assumption that forecasts are measured without error, the standard estimate of the (average) Kalman gain is $\hat{G} \approx 0.77$. The standard estimate is somewhat higher than the estimate based on the amount of noise in expectations. Without measurement error, the noise-based estimator yields a value of $\hat{G} \approx 0.65$. Intuitively, at $\hat{G} \approx 0.77$, the noisy-information model predicts that the level of noise in expectations should be lower than what we measure in the data. For this reason, the noise-based estimate is lower. In turn, the noise-based estimate predicts a higher persistence for forecast errors than is the case in the data. For the two estimates to be consistent, one needs to attribute around 50% of the noise in expectations to measurement error.

The Kalman gain estimates we obtain are slightly higher than those in the existing literature.[16] Based on the findings by Ryngaert (2018), we conjecture that part of the

[16]For example, Coibion and Gorodnichenko (2015) find $\hat{G} \approx 0.46$ using a different estimation strategy that uses forecast revisions; Ryngaert (2018) finds values of $\hat{G}$ between 0.40 and 0.50.

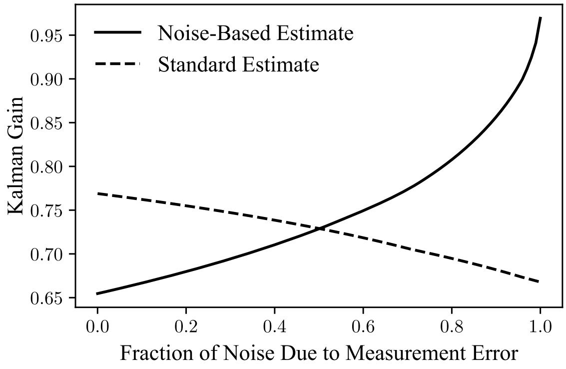

FIGURE 2. Noise-based estimate of information frictions. *Note:* Estimated average Kalman gain for one-quarter-ahead forecasts, as the fraction of noise attributed to measurement error is varied. The *standard* estimate backs out the Kalman gain from the autocorrelation of forecast errors. The alternative estimate identifies the Kalman gain from the estimated level of noise. For a given estimate of the variance of noise, $\widehat{\mathbb{V}\mathrm{ar}}_i[\nu_{i,t}]$, we attribute a fraction α to measurement error and then use $(1-\alpha)\widehat{\mathbb{V}\mathrm{ar}}_i[\nu_{i,t}]$ to calculate the implied value for the variance of idiosyncratic shocks, $\sigma^2_{i,\omega}$, and the Kalman gain.

difference may be due to a behavioral bias that is not captured in the baseline noisy-information model (e.g., underestimated persistence of actual inflation). However, testing this hypothesis is challenging and goes beyond the current paper. In particular, if the information friction is heterogeneous across forecasters, we can no longer use data from individual forecasters in panel regressions to estimate the perceived persistence, as in Ryngaert (2018).

Our estimates can also be used to calibrate information structures in macro models with incomplete information. Our estimates imply that the signal-to-noise ratio, $\sigma^2_\varepsilon/\sigma^2_\omega$, the key parameter in such models, implicit in our estimates is 0.81.[17] For comparison, Melosi (2014) estimates a DSGE model in which firms receive noisy signals about monetary policy and productivity. Melosi finds that the signal-to-noise ratio for the productivity signal is around 0.66, while it is only 0.09 for the signal about monetary policy (his Table 2). Since professional forecasters have strong incentives to pay attention to inflation—just as firms have strong incentives to pay attention to productivity—it is reassuring that the two estimates are fairly similar.

5. Conclusions

This paper proposes a method for quantifying noise in survey expectations. The basic idea is to estimate a factor model using panel data on individual expectations, purge expectations of the common factor structure, and thereby recover noise.

In an empirical application to inflation expectations, we found that noise is large and pervasive. In our data, noise is as important as bias for explaining forecast accuracy, at least at short forecasting horizons. These findings lend strong support to the view by

[17] To estimate σ^2_ε, we fit an AR(1) model to actual inflation and calculate the variance of the residuals.

Daniel Kahneman cited at the beginning of this paper that "*noise is extremely important.*" Given that noise has so far received relatively little attention in the literature on expectations, our findings also accord with Kahneman's statement that "*bias has been overestimated at the expense of noise.*"

Since noise can lead to substantial forecast errors, practitioners may wish to use forecasting methods that are less susceptible to noise. Noise can be reduced by combining forecasts from different statistical models. Forecasting firms may also ask multiple employees to produce independent forecasts. Averaging such forecasts would help diminish human-induced noise. Our findings also caution against favoring judgmental forecasts over statistical algorithms. A key advantage of algorithms is that they always provide the same output for a given set of inputs. Humans are more noisy.

The present paper also has implications for researchers using micro data on expectations. Our results suggest that it is important to be mindful of noise in individual expectations. Some of that noise may be caused by measurement error. While measurement error in expectations has been the subject of much research in applied micro,[18] it appears to have received less attention in macro and finance. Our paper provides an easy way to check whether empirical results are robust to realistic amounts of measurement error.

Do people act upon the noise in their expectations? Addressing this question lies beyond the scope of the present paper. However, there are reasons to speculate that the answer may be "yes." One reason is the existing evidence on expectation formation from incentivized lab experiments (e.g., He and Kučinskas (2019), Landier, Ma, and Thesmar (2019)). This literature finds that even the best-fitting empirical models of expectations have fairly low R-squared's at the individual level. In the field, the evidence on volatility and large volume of trade in financial markets is suggestive of an important role played by noisy expectations (Shiller (1980), Kyle (1985), Black (1986), Augenblick and Lazarus (2018)). More directly, Bailey, Cao, Kuchler, and Stroebel (2018) document that households' investment decisions are affected by house prices experienced by their geographically distant friends. Producing direct evidence on the relationship between noise and actions may be worthy of future research.

Appendix: Contents

- Section A: detailed description of the data used in our empirical analysis;
- Section B: extensions to the basic noisy information model;
- Section C: additional robustness and specifications tests of our baseline model;
- Section D: relevant proofs and implementation details for the empirical exercises reported in the main text.

[18] For example, see Manski (2004), Manski and Molinari (2010), Drerup, Enke, and von Gaudecker (2017), Giustinelli, Manski, and Molinari (2018).

Appendix A: Data appendix

Data construction largely follows the procedure in Kučinskas and Peters (2022). We download data for individual responses from the website of the Federal Reserve Bank of Philadelphia (https://www.philadelphiafed.org/surveys-and-data/pgdp). The downloaded file contains forecasts of the GDP deflator index for the past quarter (*PGDP1*), current quarter (*PGDP2*), and the next four quarters (*PGDP3* up to *PGDP6*); see the Federal Reserve Bank of Philadelphia (2017, pp. 20–22). We use forecasts for all available forecast horizons in our analysis.

First, we calculate forecasts of quarterly inflation rates, annualized, from the forecasts of the price deflator. These are calculated as

$$\mathbb{F}_{i,t}[x_{t+h}] = 100\left[\left(\frac{PGDP(h+2)}{PGDP(h+1)}\right)^4 - 1\right] \quad \text{for } h \in \{0, 1, \dots, 4\}.$$

For realizations, we use the Real-Time Data Set for Macroeconomists, which is also provided by the Philadelphia Fed (https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/real-time-data-set-for-macroeconomists). We use the first-release data for "*Price Index for GNP/GDP (P)*." In 1995Q4, the first-release data for inflation is not available. In this period, we use the second-release data.

To match forecasts and actuals, we align the forecasts to the date for which they were made. For example, the one-quarter ahead forecast made in the 1970Q1 survey is matched with the actual inflation reported for 1970Q2.

Appendix B: Extending the noisy-information model

The baseline noisy-information model in Section 2.2 may seem fairly restrictive. However, nothing in the arguments in that section hinges on the special properties of the setup. In particular, the baseline noisy-information model can be generalized in a straightforward manner to account for both richer dynamics of the target variable as well as other drivers of expectations such as behavioral biases and strategic behavior.

We first extend the model to allow for richer dynamics in the target variable and then discuss how the model can capture strategic behavior and behavioral biases. Suppose the target variable, $\mathbf{x}_t$, is generated by

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t,$$

where $\mathbf{x}_t = (x_{1,t}, x_{2,t}, \dots, x_{M,t})^\top$ is an M-dimensional vector, $\mathbf{A}$ is an $(M \times M)$ matrix with all eigenvalues within the unit circle, and $\boldsymbol{\varepsilon}_t$ is a vector of normally distributed shocks with mean zero and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$. Without loss of generality, we can assume that $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ is diagonal, so that $\boldsymbol{\varepsilon}_t$ can be interpreted as a structural shock to $\mathbf{x}_t$. Finally, to avoid degenerate cases, we assume that $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ is strictly positive definite.

As before, the agents do not observe $\mathbf{x}_t$ directly. Instead, they observe noisy signals. First, forecasters observe a private (idiosyncratic) signal of $\mathbf{x}_t$ given by

$$\mathbf{s}_{i,t}^{(i)} = \mathbf{x}_t + \boldsymbol{\omega}_{i,t}.$$

The idiosyncratic shocks $\boldsymbol{\omega}_{i,t}$ are normally distributed, uncorrelated across forecasters, have mean zero and covariance matrix $\boldsymbol{\Sigma}_{i,\boldsymbol{\omega}}$. The covariance matrix $\boldsymbol{\Sigma}_{i,\boldsymbol{\omega}}$ is allowed to vary across forecasters. Second, forecasters observe a public signal—common to all forecasters—given by

$$\mathbf{s}_t^{(p)} = \mathbf{x}_t + \mathbf{v}_t.$$

Again, $\mathbf{v}_t$ is normally distributed, has mean zero and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{v}}$. The idiosyncratic shocks $\boldsymbol{\omega}_{i,t}$ and shocks to public signals $\mathbf{v}_t$ are uncorrelated at all leads and lags. Collecting all signals to a single vector

$$\underbrace{\mathbf{s}_{i,t}}_{(2M\times 1)} = \begin{pmatrix} \mathbf{s}_{i,t}^{(i)} \\ \mathbf{s}_t^{(p)} \end{pmatrix},$$

we can cast the model into standard state-space form as

$$\begin{aligned} \mathbf{x}_t &= \mathbf{A}\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t; \\ \mathbf{s}_{i,t} &= \mathbf{H}\mathbf{x}_t + \mathbf{w}_{i,t}, \end{aligned} \tag{14}$$

where

$$\mathbf{H} = \begin{pmatrix} \mathbf{I}_r \\ \mathbf{I}_r \end{pmatrix}, \qquad \mathbf{w}_{i,t} = \begin{pmatrix} \boldsymbol{\omega}_{i,t} \\ \mathbf{v}_t \end{pmatrix}, \quad \text{and} \quad \mathbb{V}\text{ar}_i[\mathbf{w}_{i,t}] \equiv \boldsymbol{\Sigma}_{i,\mathbf{w}} = \begin{pmatrix} \boldsymbol{\Sigma}_{i,\boldsymbol{\omega}} & \mathbf{0}_{M\times M} \\ \mathbf{0}_{M\times M} & \boldsymbol{\Sigma}_{\mathbf{v}} \end{pmatrix}.$$

Provided that all eigenvalues of $\mathbf{A}$ lie inside the unit circle (Hamilton (1994, Proposition 13.1)), in the steady state, one-step-ahead forecasts are given by

$$\mathbb{F}_{i,t}[\mathbf{x}_{t+1}] = \mathbf{A}\big\{\mathbf{G}_i\mathbf{s}_{i,t} + (\mathbf{I}_r - \mathbf{G}_i\mathbf{H})\,\mathbb{F}_{i,t-1}[\mathbf{x}_t]\big\}, \tag{15}$$

where $\mathbf{G}_i$ is the $(M \times 2M)$ steady-state Kalman gain matrix defined by

$$\begin{aligned} \mathbf{G}_i &= \mathbf{P}_i\mathbf{H}^\top\big(\mathbf{H}\mathbf{P}_i\mathbf{H}^\top + \boldsymbol{\Sigma}_{i,\mathbf{w}}\big)^{-1}; \\ \mathbf{P}_i &= \mathbf{A}\big[\mathbf{P}_i - \mathbf{P}_i\mathbf{H}^\top\big(\mathbf{H}\mathbf{P}_i\mathbf{H}^\top + \boldsymbol{\Sigma}_{i,\mathbf{w}}\big)^{-1}\mathbf{H}\mathbf{P}_i\big]\mathbf{A}^\top + \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}, \end{aligned}$$

with $\mathbf{P}_i$ an $(M \times M)$ matrix.

We can now iterate on equation (15) in the same way as in the baseline model.[19] Performing the same steps as before, we can arrive at the approximate factor model representation in equation (3). The only economic difference, relative to the baseline model, is that (i) factors now contain not only the structural shocks $\boldsymbol{\varepsilon}_t$ but also the shock to the public signal, $\boldsymbol{\nu}_t$; and (ii) noise now combines shocks to private signals about all variables in $\mathbf{x}_t$. The shock to the public signal, $\boldsymbol{\nu}_t$, can be interpreted as capturing animal spirits or market sentiment (Angeletos and La'O (2013)). Part (ii) implies that noisy expectations about, say, inflation can partly stem because GDP growth is observed with noise.

[19]Hamilton (1994, Proposition 13.2) shows that the eigenvalues of $(\mathbf{A} - \mathbf{A}\mathbf{G}_i\mathbf{H})$ are all strictly within the unit circle provided that at least one of $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ or $\boldsymbol{\Sigma}_{i,\mathbf{w}}$ is strictly positive definite. Hence, we can take limits when iterating on equation (15).

Strategic behavior

Huo and Pedroni (2020, Theorem 1) show that a model with information frictions (as above) but with strategic behavior is equivalent to a model without strategic behavior in which the private signals are discounted. Hence, the model above accounts for the possibility that forecasters are strategic such as wishing to be close to the average forecast or facing tournament incentives (Ottaviani and Norman (2006)). The only difference in a model with strategic concerns would be a different covariance matrix for the idiosyncratic shocks $\mathbf{\Sigma}_{i,\boldsymbol{\omega}}$.

Misperceived parameters

Nothing in the arguments above relied on the *actual* properties of the data-generating process. Hence, we can interpret $\mathbf{A}$, $\mathbf{\Sigma}_{i,\boldsymbol{\omega}}$, etc., as *perceived* parameters of the data-generating process. These parameters may well be different from the *actual* parameters of the model, denoted by $\mathbf{A}^*$, $\mathbf{\Sigma}^*_{i,\boldsymbol{\omega}}$, etc. Various biases can be captured by model misspecification. For example, in the univariate case with $M = 1$, $A > A^*$ represents an *extrapolation bias*. Ignoring mean reversion, as in the natural-expectations model of Fuster, Laibson, and Mendel (2010), can similarly be captured by a deviation of $\mathbf{A}$ from $\mathbf{A}^*$. For that, one could specify the true model as an AR(2) process (via $\mathbf{A}^*$) but postulate that the agents misperceive the process for x_t to be an AR(1) by setting the relevant parameters in $\mathbf{A}$ to zero. Similarly, misperceptions in $\mathbf{\Sigma}_{i,\boldsymbol{\omega}}$ or $\mathbf{\Sigma}_{\boldsymbol{\nu}}$ can capture an overprecision bias (Daniel, Hirshleifer, and Subrahmanyam (1998), Odean (1998), Broer and Kohlhas (2018)).

Diagnostic expectations

It is also straightforward to extend the model to allow for diagnostic expectations (Bordalo, Gennaioli, and Shleifer (2018), Bordalo et al. (2020)). First, rewrite equation (15) as

$$\mathbb{F}_{i,t}[\mathbf{x}_{t+1}] = \mathbf{A}\,\mathbb{F}_{i,t-1}[\mathbf{x}_t] + \mathbf{A}\mathbf{G}_i \underbrace{\left\{\mathbf{s}_{i,t} - \mathbf{H}\,\mathbb{F}_{i,t-1}[\mathbf{x}_t]\right\}}_{\text{surprise}}.$$

Here, the term in the curly brackets denotes the expectational "surprise," that is, the difference between the realized signal and its $t-1$ expectation. We can then define diagnostic expectations (abusing notation slightly) as $\mathbb{F}^*_{i,t}[x_{t+1}]$ as

$$\mathbb{F}^*_{i,t}[x_{t+1}] = \mathbb{F}_{i,t}[\mathbf{x}_{t+1}] + \underbrace{\mathbf{\Theta}\mathbf{A}\mathbf{G}_i\left\{\mathbf{s}_{i,t} - \mathbf{H}\,\mathbb{F}_{i,t-1}[\mathbf{x}_t]\right\}}_{\text{diagnostic component}},$$

where $\mathbf{\Theta} = \text{diag}(\theta_1, \theta_2, \ldots, \theta_M)$ and $\theta_m \geq 0$ is the diagnosticity parameter for the mth element for $\mathbf{x}_t$ for $m = 1, \ldots, M$. It is straightforward to check that this specification nests the models analyzed in (Bordalo, Gennaioli, and Shleifer (2018), Bordalo et al. (2020)) as special cases.

All in all, the model above captures a wide variety of existing models of expectations, including a general signal structure, information frictions, behavioral biases, and strategic behavior.

Appendix C: Additional robustness checks

C.1 *Realized inflation and noise*

In Section 3.3.1, we saw that our estimates of noise are not predictive of future inflation. As an additional robustness check, we perform the following exercise. We estimate an extended model in which we also include the *actual* (or realized) inflation as a right-hand side variable:

$$\mathbb{F}_{i,t}[x_{t+h}] = b_{i,0} + \delta_i x_{t+h} + \boldsymbol{\lambda}_i^\top \mathbf{f}_t + \nu_{i,t}.$$

Given that we estimate the model by least squares, this specification automatically ensures that the estimated noise terms $\hat{\nu}_{i,t}$ are orthogonal to realized inflation. The results of this exercise are shown in Table C.1.

While the estimates of noise are smaller, as they should be, the differences are very small. This finding further suggests that our baseline noise estimates do not have predictive power for actual inflation.

C.2 *Serial correlation in noise*

A potential concern with our procedure is that the estimated noise levels $\nu_{i,t}$ may be serially correlated. Indeed, as shown in Table C.2, noise is serially correlated, with first-order autocorrelation coefficients fluctuating around 0.08–0.25 depending on the specification and forecast horizon. The concern with serial correlation has to do both with

Table C.1. Noise: orthogonalized w.r.t. realizations.

	Single Factor (Consensus)				Multiple Factors			
	$K=0$	$K=1$	$K=2$	$K=3$	$R=1$	$R=2$	$R=3$	$R=4$
$h=0$	1.09	1.05	1.00	0.97	1.09	0.93	0.81	0.71
	(0.07)	(0.06)	(0.06)	(0.06)	(0.06)	(0.04)	(0.04)	(0.03)
$h=1$	0.96	0.92	0.89	0.86	0.96	0.80	0.71	0.61
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.04)	(0.04)	(0.03)
$h=2$	0.97	0.93	0.90	0.87	0.95	0.80	0.70	0.60
	(0.07)	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.04)	(0.03)
$h=3$	1.00	0.97	0.94	0.92	0.99	0.82	0.70	0.59
	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)	(0.05)	(0.04)	(0.03)
$h=4$	1.05	1.00	0.96	0.93	1.05	0.86	0.69	0.60
	(0.08)	(0.08)	(0.08)	(0.07)	(0.08)	(0.06)	(0.04)	(0.03)

Note: Estimates of noise orthogonalized w.r.t. actual realizations of inflation. Standard deviations of noise (averages across forecasters) with standard errors in parentheses shown. The estimation procedure is the same as in Table 2, except that realized inflation is included as a right-hand side variable.

Table C.2. Average autocorrelation of noise: baseline estimates.

	Single Factor (Consensus)				Multiple Factors			
	$K=0$	$K=1$	$K=2$	$K=3$	$R=1$	$R=2$	$R=3$	$R=4$
$h=0$	0.11	0.13	0.10	0.11	0.12	0.10	0.07	0.07
$h=1$	0.14	0.15	0.14	0.12	0.15	0.14	0.11	0.08
$h=2$	0.18	0.18	0.18	0.16	0.21	0.14	0.10	0.08
$h=3$	0.22	0.22	0.21	0.18	0.25	0.22	0.19	0.13
$h=4$	0.22	0.25	0.24	0.21	0.22	0.20	0.15	0.10

Note: Average autocorrelation of noise for the baseline noise estimates. The estimation procedure is the same as in Table 2.

model specification as well as with how noise should be defined conceptually. Conceptually, one may argue that noise should be serially uncorrelated. As shown in Section 2.2, the standard noisy-information model predicts that noise should be serially correlated. For that reason, our baseline estimation allows for serially correlated noise. However, it is not unreasonable to *define* noise by saying it should be serially uncorrelated. From a model-specification perspective, serially correlated noise may indicate that the dynamics of individual forecasts are not captured well enough.

We first note that the estimated levels of serial correlation are not large enough to substantially change the results. Using a simple AR(1) approximation, a back-of-the-envelope calculation suggests that removing an autocorrelation of 0.20 would reduce the estimated variance of noise by only 4%.[20]

To further investigate this potential issue, however, we expand our baseline model to include the lagged forecast as a right-hand side variable, that is,[21]

$$\mathbb{F}_{i,t}[x_{t+h}] = b_{i,0} + \delta_i \mathbb{F}_{i,t-1}[x_{t+h-1}] + \boldsymbol{\lambda}_i^\top \mathbf{f}_t + \nu_{i,t}.$$

The model above is a factor-augmented AR(1), and the lagged forecast term should soak up most of the serial correlation that is present in our baseline estimates of noise. The resulting estimates of noise are provided in Table C.3. These estimates are somewhat smaller than our baseline results, but overall they are fairly similar. Conceptually, these results are reassuring in that they suggest that irrespective of whether we require noise to be serially uncorrelated or not, the magnitude of noise is going to be fairly similar.

C.3 *More stringent definition of noise*

Building on the previous section, we next estimate noise levels using a more stringent definition of noise. Specifically, we require estimated noise to be orthogonal to both (i) realized inflation (so that noise is not predictive of realized inflation); and (ii) lagged

[20]Suppose that noise follows $\nu_{i,t} = \rho\nu_{i,t-1} + \omega_{i,t}$. Then $\mathbb{V}\mathrm{ar}_i[\nu_{i,t}] = \mathbb{V}\mathrm{ar}_i[\omega_{i,t}]/(1-\rho^2)$ and so $\mathbb{V}\mathrm{ar}_i[\omega_{i,t}] = \mathbb{V}\mathrm{ar}_i[\nu_{i,t}](1-\rho^2)$. Hence, for a given estimate of $\mathbb{V}\mathrm{ar}_i[\nu_{i,t}]$, accounting for an autocorrelation level of $\rho = 0.20$ would only reduce the variance by a factor of $1-0.20^2 = 0.96$.

[21]Note that estimating this equation requires lagged forecasts, which reduces our sample somewhat and makes the resulting sample slightly different from our baseline sample.

Table C.3. Noise: orthogonalized w.r.t. lagged forecast.

	Single Factor (Consensus)				Multiple Factors			
	$K=0$	$K=1$	$K=2$	$K=3$	$R=1$	$R=2$	$R=3$	$R=4$
$h=0$	1.02	0.96	0.91	0.87	0.99	0.81	0.66	0.57
	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)	(0.04)	(0.03)	(0.02)
$h=1$	0.87	0.82	0.78	0.74	0.88	0.69	0.56	0.47
	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.04)	(0.03)	(0.02)
$h=2$	0.88	0.85	0.81	0.76	0.85	0.69	0.55	0.46
	(0.07)	(0.06)	(0.06)	(0.06)	(0.05)	(0.04)	(0.03)	(0.02)
$h=3$	0.89	0.85	0.82	0.77	0.84	0.69	0.56	0.44
	(0.07)	(0.07)	(0.07)	(0.07)	(0.06)	(0.04)	(0.03)	(0.02)
$h=4$	0.88	0.82	0.77	0.72	0.85	0.66	0.52	0.41
	(0.08)	(0.07)	(0.07)	(0.06)	(0.06)	(0.04)	(0.03)	(0.02)

Note: Estimates of noise orthogonalized w.r.t. lagged forecast. Standard deviations of noise (averages across forecasters) with standard errors in parentheses shown. The estimation procedure is the same as in Table 2, except that the lagged forecast is included as a right-hand side variable.

forecast (so that the possibility of serially correlated noise is reduced). We operationalize this definition by estimating

$$\mathbb{F}_{i,t}[x_{t+h}] = b_{i,0} + \gamma_{1,i} x_{t+h} + \gamma_{2,i} \mathbb{F}_{i,t-1}[x_{t+h-1}] + \boldsymbol{\lambda}_i^\top \mathbf{f}_t + \nu_{i,t}.$$

That is, we include realized inflation and lagged forecast as control variables in our estimation. As expected from the previous results that included realized inflation and lagged forecasts separately, the resulting estimates of noise are lower but not by much. As shown in Table C.4, the resulting estimates of noise fairly comparable to those in Table C.3.

Table C.4. Noise: more stringent definition.

	Single Factor (Consensus)				Multiple Factors			
	$K=0$	$K=1$	$K=2$	$K=3$	$R=1$	$R=2$	$R=3$	$R=4$
$h=0$	0.98	0.92	0.87	0.83	0.93	0.77	0.64	0.53
	(0.06)	(0.06)	(0.05)	(0.05)	(0.04)	(0.03)	(0.03)	(0.02)
$h=1$	0.84	0.79	0.75	0.71	0.82	0.67	0.53	0.45
	(0.06)	(0.05)	(0.05)	(0.05)	(0.04)	(0.03)	(0.02)	(0.02)
$h=2$	0.85	0.82	0.77	0.73	0.82	0.66	0.53	0.43
	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)	(0.03)	(0.03)	(0.02)
$h=3$	0.86	0.83	0.79	0.74	0.81	0.67	0.54	0.42
	(0.07)	(0.07)	(0.07)	(0.07)	(0.05)	(0.04)	(0.03)	(0.02)
$h=4$	0.84	0.79	0.73	0.68	0.83	0.62	0.49	0.38
	(0.07)	(0.07)	(0.07)	(0.06)	(0.06)	(0.04)	(0.03)	(0.02)

Note: Estimates of noise using a more stringent definition of noise. Standard deviations of noise (averages across forecasters) with standard errors in parentheses shown. The definition ensures that the estimated noise is orthogonal to both (i) realized inflation; and (ii) lagged forecast.

C.4 *Alternative subsamples: Outliers, IDs, and larger cross-sections*

Next, we reestimate our baseline model on different subsamples to investigate the potential effects of certain data issues in the SPF on our results. We also investigate the robustness of our results to the choice of the minimum number of participants in the SPF survey waves.

A natural concern is that our estimates of noise may be biased upward if outliers are present in the sample. These outliers could be, for example, the result of data collection or input errors. Looking at the minimum and maximum values of forecasts in Table 1, there are no obviously "crazy" values. In part, this is to be expected, given that the respondents of the survey are professional forecasters. This feature of the data set is very different from survey data on household expectations. However, some forecasts are fairly extreme, with the values for 1-year-ahead forecasts, for example, ranging from the low of −19.45% to the high of 29.75%. We note that these values need not be outliers but could, in fact, be due to noise.

To investigate the potential effect of outliers, we use the Tukey's rule and remove all observations that fall outside $[q_1 - 1.5 \cdot \text{IQR}, q_3 + 1.5 \cdot \text{IQR}]$ where q_i is the ith quartile, and IQR is the interquartile range of all forecasts for that particular horizon. Table C.5 shows that the absolute levels of noise goes down somewhat once outliers are removed and are in the range of 0.40–0.85.

Another potential concern is that our estimates of noise may be contaminated by measurement error in forecaster IDs. Specifically, the Survey of Professional Forecasters assigns an ID to each forecaster. In theory, this ID should be unique to each forecaster. However, as originally pointed out by Stark (1997) and discussed by Federal Reserve Bank of Philadelphia (2017, p. 34), sometimes certain forecaster IDs reappear in the data set after a long period of nonparticipation in the survey. Such reappearances suggest that some forecaster IDs may be reassigned to other survey participants. Such

TABLE C.5. Is noise driven by outliers?

	Single Factor (Consensus)				Multiple Factors			
	$K=0$	$K=1$	$K=2$	$K=3$	$R=1$	$R=2$	$R=3$	$R=4$
$h=0$	0.87	0.84	0.80	0.77	0.85	0.76	0.67	0.61
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
$h=1$	0.74	0.72	0.69	0.67	0.72	0.65	0.57	0.50
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)
$h=2$	0.72	0.69	0.67	0.65	0.70	0.61	0.53	0.47
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
$h=3$	0.70	0.67	0.65	0.63	0.68	0.60	0.52	0.46
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)
$h=4$	0.71	0.65	0.62	0.59	0.68	0.60	0.50	0.43
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)

Note: Estimates of noise with outliers removed by Tukey's rule. Standard deviations of noise (averages across forecasters) with standard errors in parentheses shown. Specifically, we calculate the first and third quartiles of individual forecasts, denoted by q_1 and q_3, with the resulting interquartile range given by $\text{IQR} = q_3 - q_1$. Then we remove all forecasts that fall outside $[q_1 - 1.5 \cdot \text{IQR}, q_3 + 1.5 \cdot \text{IQR}]$. Otherwise, the procedure is the same as in Table 2.

TABLE C.6. Noise estimates: “No gaps” subsample.

	Single Factor (Consensus)				Multiple Factors			
	$K=0$	$K=1$	$K=2$	$K=3$	$R=1$	$R=2$	$R=3$	$R=4$
$h=0$	1.12	1.09	1.06	1.04	1.12	0.94	0.83	0.73
	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.06)	(0.05)	(0.04)
$h=1$	0.99	0.94	0.92	0.89	0.99	0.82	0.70	0.62
	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.06)	(0.05)	(0.04)
$h=2$	0.91	0.89	0.86	0.83	0.90	0.73	0.64	0.56
	(0.08)	(0.08)	(0.08)	(0.07)	(0.08)	(0.05)	(0.04)	(0.03)
$h=3$	0.99	0.96	0.94	0.92	0.99	0.80	0.69	0.57
	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.07)	(0.05)	(0.04)
$h=4$	1.03	0.99	0.94	0.89	1.06	0.91	0.68	0.58
	(0.11)	(0.11)	(0.10)	(0.09)	(0.10)	(0.07)	(0.05)	(0.04)

Note: Estimates of noise for a subsample of forecasters with “no gaps” in survey participation. Standard deviations of noise (averages across forecasters) with standard errors in parentheses shown. Specifically, we remove all forecaster IDs for which the maximum length of time between two forecasts is more than 5 years.

measurement error in forecaster IDs may problematic for our purposes since it can confound noise with differences in factor loadings and lead to an overestimation of noise.

To investigate this concern, we drop forecasters that reappear in the sample after an absence of 5 years or more. The results are shown in Table C.6. The estimated noise levels are very similar in this restricted “no gaps” subsample.

Finally, we investigate the robustness of our estimates to the choice of the minimum cross-sectional dimension. Table C.7 shows the results once we impose a minimum cross-sectional dimension of at least 30 at all survey dates; we keep the minimum number of time series observations to 20. As shown in Table C.8, this sample restriction reduces the effective sample size substantially. However, our estimates of noise remain

TABLE C.7. Noise: larger cross-section.

	Single Factor (Consensus)				Multiple Factors			
	$K=0$	$K=1$	$K=2$	$K=3$	$R=1$	$R=2$	$R=3$	$R=4$
$h=0$	0.65	0.59	0.55	0.53	0.67	0.58	0.48	0.41
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
$h=1$	0.63	0.58	0.55	0.52	0.63	0.54	0.45	0.39
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
$h=2$	0.69	0.65	0.61	0.59	0.67	0.58	0.49	0.41
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)
$h=3$	0.72	0.67	0.62	0.58	0.89	0.65	0.55	0.43
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
$h=4$	0.80	0.76	0.72	0.69	0.86	0.73	0.67	0.58
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)

Note: Estimates of noise after imposing a minimum cross-sectional dimension of 30. Standard deviations of noise (averages across forecasters) with standard errors in parentheses shown. The estimation procedure is the same as in Table 2.

Table C.8. Sample sizes with different sample selection criteria.

$N_{\min}$ $T_{\min}$	0	10	20	30	40
0	7712	7245	6370	5149	4225
10	7703	7236	6361	5140	4216
20	7526	7063	6007	2380	0
30	6809	4566	1632	0	0
40	3761	0	0	0	0

Note: The number of nowcasts available in the data set after imposing that each forecaster makes at least $T_{\min}$ forecasts and each survey has at least $N_{\min}$ forecasters. Both conditions are imposed simultaneously. To do so, we apply the two conditions iteratively until the sample size no longer changes.

fairly similar. The key difference is that the amount of noise for shorter forecasting horizons is reduced.

C.5 *Is noisiness a forecaster-specific trait?*

As a final robustness test, we check if, for a given forecaster, the level of noise is positively correlated across different forecasts. Economically, it is natural to expect that "noisiness" is a forecaster-specific trait, with some forecasters more noisy than others. In contrast, if our estimates of noise are an artefact of certain modeling choices, one may expect the level of noise to be uncorrelated across different forecasts.

To empirically test this, we calculate the standard deviation of noise for all forecaster-horizon pairs, and calculate the correlation matrix of these measures. Table C.9 documents that for a given forecaster, our measure of noise is strongly positively correlated across different forecast horizons. The smallest correlation coefficient is 0.55, and most coefficients lie between 0.60 and 0.80. We conclude that noisiness appears to be forecaster specific. This finding also suggests that our estimates of noise are not just capturing measurement error, as it is less immediate why measurement error should exhibit such within-forecaster correlation (rather than being something that is common to all forecasters).

Table C.9. Correlation between noise at various horizons.

	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$
$h=0$	1.00	0.78	0.74	0.70	0.67
$h=1$	0.78	1.00	0.91	0.55	0.57
$h=2$	0.74	0.91	1.00	0.62	0.56
$h=3$	0.70	0.55	0.62	1.00	0.61
$h=4$	0.67	0.57	0.56	0.61	1.00

Note: Correlation matrix of noise estimates in inflation forecasts at different horizons. From our estimates of noise $\hat{\nu}_{i,t}$, we calculate the standard deviations of noise for each forecaster i and forecast horizon h, denoted by $\hat{\sigma}_{i,h}$. We then report the correlation matrix of $\hat{\sigma}_{i,h}$. Forecast horizon h runs from 0 (nowcasts) to 4 (1-year-ahead forecast). The number of factors used for estimating noise levels is chosen by the growth-ratio statistic of Ahn and Horenstein (2013) after subtracting the consensus forecast from the individual forecasts, as shown in Table 3.

Appendix D: Implementation details

D.1 *Standard errors for predictability regressions*

Consider the following DGP for individual forecasts:

$$x^e_{i,t} = b_{i,0} + \boldsymbol{\lambda}_i^\top \mathbf{f}_t + \nu_{i,t}, \tag{16}$$

where for notational simplicity we let $x^e_{i,t} \equiv \mathbb{F}_{i,t}[x_{t+h}]$ and drop the horizon superscript for the noise term $\nu_{i,t}$. After estimating the common component $\boldsymbol{\lambda}_i^\top \mathbf{f}_t$, we define the corresponding residual as $\hat{\nu}_{i,t} = x^e_{i,t} - \hat{\delta}_i - \hat{\boldsymbol{\lambda}}_i^\top \hat{\mathbf{f}}_t$. We wish to use $\hat{\nu}_{i,t}$ in the predictive regression of the form:

$$x_t = \alpha_i + \gamma_i \hat{\nu}_{i,t} + \text{error}.$$

Since the sample mean of $\hat{\nu}_{i,t}$ is zero, the OLS estimator of γ_i is given by

$$\hat{\gamma}_i = \left(\frac{1}{T}\sum_{t=1}^T \hat{\nu}^2_{i,t}\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^T \hat{\nu}_{i,t} x_t\right).$$

Assuming that $\nu_{i,t}$ are uncorrelated with x_t, the OLS estimator is consistent for the true value $\gamma_i = 0$ as both N and T increase. However, due to the estimation of $b_{i,0}$ and $\boldsymbol{\lambda}_i$ in the first step regression inference on $\hat{\gamma}_i$ is complicated by the generated regressors problem.

Careful inspection of Theorem 3 in Bai (2003) shows that under the assumption that both $\sqrt{T}/N \to 0$ and $\sqrt{N}/T \to 0$, the asymptotic variance of the estimator $\hat{\gamma}_i$ is driven by

$$\frac{1}{\sqrt{T}}\sum_{t=1}^T x_t \nu_{i,t} - \begin{pmatrix} \sum_t x_t & \sum_t \mathbf{f}_t x_t \end{pmatrix} \times \begin{pmatrix} 1 & \sum_t \mathbf{f}'_t \\ \sum_t \mathbf{f}_t & \sum_t \mathbf{f}_t \mathbf{f}'_t \end{pmatrix}^{-1} \times \frac{1}{\sqrt{T}} \begin{pmatrix} \sum_t \nu_{i,t} \\ \sum_t \mathbf{f}_t \nu_{i,t} \end{pmatrix}.$$

In particular, we need to evaluate the variance-covariance matrix $\boldsymbol{\Sigma}_i$ of

$$\begin{pmatrix} \frac{1}{\sqrt{T}}\sum_{t=1}^T x_t \nu_{i,t} \\ \frac{1}{\sqrt{T}}\sum_{t=1}^T \nu_{i,t} \\ \frac{1}{\sqrt{T}}\sum_{t=1}^T \mathbf{f}_t \nu_{i,t} \end{pmatrix} = \frac{1}{\sqrt{T}}\sum_{t=1}^T \mathbf{z}_t \nu_{i,t},$$

where

$$\mathbf{z}_{i,t} \equiv (x_t, 1, \mathbf{f}_t)',$$

is an $(R+2)$-dimensional vector of "common covariates." As $\mathbf{f}_t$ are unobserved, to operationalize the estimation of the variance-covariance matrix we replace it with the factor estimates, that is,

$$\hat{\mathbf{z}}_{i,t} = (x_t, 1, \widehat{\mathbf{f}}_t)'.$$

Here, $\widehat{\mathbf{f}}_t$ is the R-dimensional vector of estimated common factors in equation (16). Then, under the assumption that $\nu_{i,t}$'s are serially uncorrelated, a consistent estimator of $\boldsymbol{\Sigma}_i$ is given by

$$\widehat{\boldsymbol{\Sigma}}_i = \frac{1}{T}\sum_{t=1}^{T} \hat{\mathbf{z}}_t \hat{\mathbf{z}}_t' (\hat{\nu}_{i,t})^2.$$

The asymptotic variance-covariance matrix of $\hat{\gamma}_i$ can then be estimated as

$$\widehat{\boldsymbol{\Sigma}}_{\gamma_i} = \hat{\xi}_i^2 \hat{\mathbf{v}}' \widehat{\boldsymbol{\Sigma}}_i \hat{\mathbf{v}},$$

where

$$\hat{\xi}_i = \left(\frac{1}{T}\sum_{t=1}^{T} \hat{\nu}_{i,t}^2 \right)^{-1},$$

while the "bread part" $\hat{\mathbf{v}}$ is defined as

$$\hat{\mathbf{v}} = \begin{pmatrix} 1 \\ -\begin{pmatrix} 1 & \frac{1}{T}\sum_{t=1}^{T} \widehat{\mathbf{f}}_t' \\ \frac{1}{T}\sum_{t=1}^{T} \widehat{\mathbf{f}}_t & \frac{1}{T}\sum_{t=1}^{T} \widehat{\mathbf{f}}_t \widehat{\mathbf{f}}_t' \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{T}\sum_{t=1}^{T} x_t \\ \frac{1}{T}\sum_{t=1}^{T} \widehat{\mathbf{f}}_t x_t \end{pmatrix} \end{pmatrix},$$

which is an $(R+2)$-dimensional vector. The corresponding standard error of $\hat{\gamma}_i$ is given by $\sqrt{\widehat{\boldsymbol{\Sigma}}_{\gamma_i}/T}$.

If the data set is balanced, the "bread part" $\hat{\mathbf{v}}$ is common to all forecasters. If the dataset is unbalanced, as in our empirical application, $\hat{\mathbf{v}}$ should be calculated separately for each individual i. The reason is that only the values of $x_{i,t}^e$'s for the dates on which the forecaster is active are used in estimating the factor loadings and individual fixed effects. Therefore, when calculating the corrected standard errors in practice, the summations are performed only over the dates on which the corresponding forecaster is active.

If we relax the no-serial-correlation assumption, then the only thing that requires modification is the formula for $\widehat{\boldsymbol{\Sigma}}_{\gamma_i}$. A simple Newey–West type of the estimator for the long-run variance is given by

$$\widehat{\boldsymbol{\Sigma}}_i = \frac{1}{T}\sum_{t=1}^{T}\hat{\mathbf{z}}_t\hat{\mathbf{z}}_t'(\hat{\nu}_{i,t})^2 + \frac{1}{T}\sum_{\ell=1}^{L}\sum_{t=\ell+1}^{T} w_\ell\big(\hat{\mathbf{z}}_{i,t}\hat{\mathbf{z}}_{i,t-\ell}' + \hat{\mathbf{z}}_{t-\ell}\hat{\mathbf{z}}_t'\big)(\hat{\nu}_{i,t}\hat{\nu}_{i,t-\ell}),$$

with $w_\ell = 1 - \ell/(L+1)$ for some prespecified L.

D.2 *MSE decomposition*

In this section, we provide the empirical implementation details for the MSE decomposition in Section 3.2. First, we estimate the variance of the irreducible error following the method proposed by Kučinskas and Peters (2022). We start by simplifying the decomposition in Eq. (7) by assuming that there is only one structural shock present ($M = 1$), as in the model of Section 2.2, yielding

$$\text{MSE}_{\text{cons.}} = (\bar{b}_0)^2 + \mathbb{V}\text{ar}[\varepsilon_{1,t}]\left(\sum_{\ell=1}^{+\infty}\bar{b}_{1,\ell}^2\right) + \mathbb{V}\text{ar}[\varepsilon_{1,t}].$$

Note that $M = 1$ should be thought of as an approximation that lumps all structural shocks into one composite measure; see the discussion of composite bias coefficients in Kučinskas and Peters (2022). We estimate the unconditional bias, $\bar{b}_0$, by the sample average of consensus forecast errors, and we estimate the bias coefficients $\bar{b}_{1,\ell}$ from the univariate impulse-response function of forecast errors (estimated with local projections with four lags of forecast errors as controls), as in Kučinskas and Peters (2022).

Next, we estimate the variance of irreducible error as

$$\widehat{\mathbb{V}\text{ar}[\varepsilon_{1,t}]} = \frac{\widehat{\text{MSE}}_{\text{cons.}} - \hat{\bar{b}}_0^2}{1 + \sum_{\ell=1}^{L}\hat{\bar{b}}_{1,\ell}^2},$$

where $\widehat{\text{MSE}}_{\text{cons.}}$ denotes the sample mean-squared forecast error of the consensus forecast, L is the number of bias coefficients we estimate, and variables with hats denote sample estimates. In our application, the truncation parameter is set to $L = 12$. Note that this approach of estimating irreducible error does not require us to take a stand on the true data-generating process. The formula above is very intuitive: We obtain the magnitude of irreducible error by subtracting the effects of biases (both conditional and unconditional) from the MSE of the consensus forecast.

With the estimate of $\widehat{\mathbb{V}\text{ar}[\varepsilon_{1,t}]}$ at hand, we can directly estimate the conditional bias component in equation (6) by subtracting the estimates of the unconditional-bias term, the variance of irreducible error, and the variance of noise from the forecaster-level MSE. Finally, we average equation (6) across forecasters.

D.3 *Noisy-information model*

In this section, we provide the details behind the empirical implementation of the exercise in Section 4.3. To model measurement error, we assume that instead of observing the true expectations, denoted by $\mathbb{F}^*_{i,t}[x_{t+1}]$, we only observe

$$\mathbb{F}_{i,t}[x_{t+1}] = \mathbb{F}^*_{i,t}[x_{t+1}] + \eta_{i,t},$$

where $\eta_{i,t}$ is i.i.d. measurement error with mean zero and variance $\sigma^2_{i,\eta}$. (We use asterisks to denote the true, but unobserved, values.) The level of measurement error is allowed to vary across forecasters. Since $e_{i,t+1} = e^*_{i,t+1} - \eta_{i,t}$, in a regression of

$$e_{i,t+1} = c_i + \phi_i e_{i,t} + u_{i,t},$$

estimated by OLS, we have that for each $i = 1, \dots, N$,

$$\operatorname{plim}_{T\to\infty} \hat{\phi}_i = \frac{\mathbb{C}\text{ov}_i\big[e^*_{i,t+1}, e^*_{i,t}\big]}{\mathbb{V}\text{ar}_i\big[e^*_{i,t+1}\big] + \sigma^2_{i,\eta}} = \rho(1 - G_i)\frac{\mathbb{V}\text{ar}_i\big[e^*_{i,t+1}\big]}{\mathbb{V}\text{ar}_i\big[e^*_{i,t+1}\big] + \sigma^2_{i,\eta}}, \tag{17}$$

where the second-equality uses equation (12). Hence, with measurement error, the OLS estimate of ϕ_i is biased downwards, and the Kalman gain estimate is biased upwards (i.e., the information friction is underestimated).

We use the attenuation-bias formula in equation (17) to adjust the empirical OLS estimates for measurement error. In particular, let

$$\kappa_i \equiv \frac{\mathbb{V}\text{ar}_i\big[e^*_{i,t+1}\big]}{\mathbb{V}\text{ar}_i\big[e^*_{i,t+1}\big] + \sigma^2_{i,\eta}}.$$

We discuss the strategy for estimating κ_i below. However, for a given consistent estimator of κ_i, say $\hat{\kappa}_i$, we clearly can consistently estimate the Kalman gain by

$$\hat{G}_i^{\text{stand.}} = 1 - \frac{\hat{\phi}_i}{\hat{\rho}\hat{\kappa}_i}, \tag{18}$$

where $\hat{\rho}$ denotes a consistent estimator of ρ. We call equation (18) the "standard" estimator of the Kalman gain to distinguish it from our new "noise-based" estimator below.

We are now ready to state the complete empirical procedure:

1. Estimate the variance of noise, $\mathbb{V}\text{ar}_i[\nu_{i,t}]$, forecaster-by-forecaster using a factor model, yielding an estimate $\widehat{\mathbb{V}\text{ar}}_i[\nu_{i,t}]$.

2. Obtain estimates $\hat{\rho}$ and $\hat{\sigma}^2_\varepsilon$ using an OLS regression of x_{t+1} on x_t.

3. Assume that a fraction $\alpha \in [0, 1]$ of the forecaster-specific noise estimated in step (1) is due to measurement error.

   (a) Standard approach:

i. Estimate the persistence of forecast errors forecaster-by-forecaster:

$$e_{i,t+1} = c_i + \phi_i e_{i,t} + u_{i,t}.$$

Denote the resulting OLS estimate by $\hat{\phi}_i$.

ii. Calculate forecaster-specific measurement error via $\hat{\sigma}^2_{i,\eta} = \alpha \widehat{\mathbb{V}\text{ar}}_i[\nu_{i,t}]$.

iii. Estimate the attenuation factor as $\hat{\kappa}_i = \{\widehat{\mathbb{V}\text{ar}}_i[e_{i,t+1}] - \hat{\sigma}^2_{i,\eta}\}/\widehat{\mathbb{V}\text{ar}}_i[e_{i,t+1}]$.

iv. Obtain the "standard estimate" of the Kalman gain, $\hat{G}_i^{\text{stand.}}$, via equation (18).

v. Since this procedure does not guarantee that $\hat{G}_i^{\text{stand.}} \in [0, 1]$, as it should in theory, we set values $\hat{G}_i^{\text{stand.}} > 1$ to 1, and values $\hat{G}_i^{\text{stand.}} < 0$ to 0.

vi. We obtain our final estimate $\hat{G}^{\text{stand.}}$ by taking an average of $\hat{G}_i^{\text{stand.}}$ across forecasters.

(b) Noise-based approach:

i. Use equation (13) and the assumed level of α to estimate the variance of idiosyncratic shocks as

$$\hat{\sigma}^2_{i,\omega} = \frac{(1-\alpha)\widehat{\mathbb{V}\text{ar}}_i[\nu_{i,t}]\left(1 - \hat{\rho}^2\left(1 - \hat{G}_i^{\text{stand.}}\right)^2\right)}{\hat{\rho}^2\left(\hat{G}_i^{\text{stand.}}\right)^2},$$

using the estimated $\hat{G}_i^{\text{stand.}}$.

ii. Calculate the "noise-based" estimate of the Kalman gain as

$$\hat{G}_i^{\text{noise}} = \frac{\widehat{\Sigma}_i}{\widehat{\Sigma}_i + \hat{\sigma}^2_{i,\omega}},$$

where

$$\widehat{\Sigma}_i = \frac{-(1-\hat{\rho}^2)\hat{\sigma}^2_{i,\omega} + \hat{\sigma}^2_{\varepsilon} + \sqrt{\left[(1-\hat{\rho}^2)\hat{\sigma}^2_{i,\omega} - \hat{\sigma}^2_{\varepsilon}\right]^2 + 4\hat{\sigma}^2_{i,\omega}\hat{\sigma}^2_{\varepsilon}}}{2}.$$

iii. We obtain our final estimate $\hat{G}^{\text{noise}}$ by taking an average of $\hat{G}_i^{\text{noise}}$ across forecasters.

References

Ahn, Seung C. and Alex R. Horenstein (2013), "Eigenvalue ratio test for the number of factors." *Econometrica*, 81, 1203–1227. [621, 622, 623, 627, 639]

Ang, Andrew, Geert Bekaert, and Min Wei (2007), "Do macro variables, asset markets, or surveys forecast inflation better?" *Journal of Monetary Economics*, 54, 1163–1212. [622]

Angeletos, George-Marios, Fabrice Collard, and Harris Dellas (2018), "Quantifying confidence." *Econometrica*, 86, 1689–1726. [615]

Angeletos, George-Marios and Jennifer La'O (2013), "Sentiments." *Econometrica*, 81, 739–779. [615, 632]

Angeletos, George-Marios and Chen Lian (2016), "Incomplete information in macroeconomics." In *Handbook of Macroeconomics*, Vol. 2 (John B. Taylor and Harald Uhlig, eds.), 1065–1240, Elsevier. [614]

Augenblick, Ned and Eben Lazarus (2018), "Restrictions on asset-price movements under rational expectations: Theory and evidence." Working paper. [630]

Bachmann, Rüdiger and Steffen Elstner (2015), "Firm optimism and pessimism." *European Economic Review*, 79, 297–325. [610]

Bai, Jushan (2003), "Inferential theory for factor models of large dimensions." *Econometrica*, 71, 135–171. [613, 618, 640]

Bai, Jushan (2009), "Panel data models with interactive fixed effects." *Econometrica*, 77, 1229–1279. [618, 620]

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel (2018), "The economic effects of social networks: Evidence from the housing market." *Journal of Political Economy*, 126, 2224–2276. [630]

Baker, Scott R., Tucker McElroy, and Xuguang S. Sheng (2019), "Expectation formation following large unexpected shocks." Review of Economics and Statistics. [614]

Barberis, Nicholas (2018), "Psychology-based models of asset prices and trading volume." In *Handbook of Behavioral Economics* (Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds.). Elsevier, North-Holland, Amsterdam. [610]

Barrero, Jose Maria (2018), "The micro and macro of managerial beliefs." Working paper. [610]

Black, Fischer (1986), "Noise." *The Journal of Finance*, 41, 528–543. [630]

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer (2020), "Overreaction in macroeconomic expectations." *American Economic Review*, 110, 2748–2782. https://www.aeaweb.org/articles?id=10.1257/aer.20181219. [612, 625, 633]

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2018), "Diagnostic expectations and credit cycles." *The Journal of Finance*, 73, 199–227. [611, 633]

Bouchaud, Jean-Philippe, Philipp Krüger, Augustin Landier, and David Thesmar (2019), "Sticky expectations and the profitability anomaly." *The Journal of Finance*, 74, 639–674. [625]

Broer, Tobias and Alexandre N. Kohlhas (2018), "Forecaster (mis-)behavior." Working paper. [612, 625, 633]

Chahrour, Ryan and Kyle Jurado (2018), "News or noise? The missing link." *American Economic Review*, 108, 1702–1736. [615]

Chen, Mingli, Ivan Fernández-Val, and Martin Weidner (2021), "Nonlinear factor models for network and panel data." *Journal of Econometrics*, 220, 296–324. [623]

Chudik, Alexander, M. Hashem Pesaran, and Elisa Tosetti (2011), "Weak and strong cross-section dependence and estimation of large panels." *The Econometrics Journal*, 14, 45–90. http://dx.doi.org/10.1111/j.1368-423X.2010.00330.x. [615]

Chudik, Alexander and M. Hashem Pesaran (2015), "Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors." *Journal of Econometrics*, 188, 393–420. [617]

Coibion, Olivier and Yuriy Gorodnichenko (2012), "What can survey forecasts tell us about information rigidities?" *Journal of Political Economy*, 120, 116–159. [614, 628]

Coibion, Olivier and Yuriy Gorodnichenko (2015), "Information rigidity and the expectations formation process: A simple framework and new facts." *The American Economic Review*, 105, 2644–2678. [611, 625, 627, 628]

Coibion, Olivier, Yuriy Gorodnichenko, and Rupal Kamdar (2018), "The formation of expectations, inflation and the Phillips Curve." *Journal of Economic Literature*, 56, 1447–1491. [610, 619]

Crowe, Christopher (2010), "Consensus forecast and inefficient information aggregation." Working paper. [612]

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam (1998), "Investor psychology and security market under- and overreactions." *The Journal of Finance*, 53, 1839–1885. [633]

Davies, Anthony and Kajal Lahiri (1995), "A new framework for analyzing survey forecasts using three-dimensional panel data." *Journal of Econometrics*, 68, 205–227. [612]

Dawes, Robyn M., David Faust, and Paul E. Meehl (1989), "Clinical versus actuarial judgment." *Science*, 243, 1668–1674. [610]

de Silva, Tim and David Thesmar (2021), "Noise in expectations: Evidence from analyst forecasts." Working paper. [612]

Dovern, Jonas (2015), "A multivariate analysis of forecast disagreement: Confronting models of disagreement with survey data." *European Economic Review*, 80, 16–35. [614]

Drerup, Tilman, Benjamin Enke, and Hans Martin von Gaudecker (2017), "The precision of subjective data and the explanatory power of economic models." *Journal of Econometrics*, 200, 378–389. [611, 630]

Federal Reserve Bank of Philadelphia (2017), "Survey of professional forecasters: Documentation." https://www.philadelphiafed.org/-/media/research-and-data/real-time-center/survey-of-professional-forecasters/spf-documentation.pdf?la=en. [631, 637]

Fernández-Val, Ivan, Hugo Freeman, and Martin Weidner (2021), "Low-rank approximations of nonseparable panel models." *Econometrics Journal*, 24, C40–C77. [618]

Figlewski, Stephen (1983), "Optimal price forecasting using survey data." *The Review of Economics and Statistics*, 65, 13–21. [612, 617]

Forni, Mario, Marc Hallin, Marco Lippi, and Paolo Zaffaroni (2015), "Dynamic factor models with infinite-dimensional factor spaces: One-sided representations." *Journal of Econometrics*, 185, 359–371. [618]

Forni, Mario, Marc Hallin, Marco Lippi, and Paolo Zaffaroni (2017), "Dynamic factor models with infinite-dimensional factor space: Asymptotic analysis." *Journal of Econometrics*, 199, 74–92. [618]

Freeman, Hugo and Martin Weidner (2022), "Linear panel regressions with two-way unobserved heterogeneity." Journal of Econometrics, (forthcoming). [615]

Fuster, Andreas, David Laibson, and Brock Mendel (2010), "Natural expectations and macroeconomic fluctuations." *Journal of Economic Perspectives*, 24, 67–84. [633]

Gaglianone, Wagner Piazza and João Victor Issler (2021), "Microfounded forecasting." Working paper. [612]

Giustinelli, Pamela, Charles F. Manski, and Francesca Molinari (2018), "Tail and center rounding of probabilistic expectations in the health and retirement study." Working paper. [630]

Grove, William M., David H. Zald, Boyd S. Lebow, Beth E. Snitz, and Chad Nelson (2000), "Clinical versus mechanical prediction: A meta-analysis." *Psychological Assessment*, 12, 19–30. [610]

Hamilton, James D. (1994), *Time Series Analysis*. Princeton University Press, Princeton. [632]

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman (2009), *The Elements of Statistical Learning*, second edition. Springer. [610]

He, Simin and Simas Kučinskas (2019), "Complexity and expectation formation." Working paper. [630]

Herbst, Edward and Fabian Winkler (2019), "A factor structure of disagreement." Working paper. [612]

Huo, Zhen and Marcelo Pedroni (2020), "A single-judge solution to beauty contests." *American Economic Review*, 110, 526–568. https://www.aeaweb.org/articles?id=10.1257/aer.20170519. [633]

Jin, Sainan, Ke Miao, and Liangjun Su (2021), "On factor models with random missing: Em estimation, inference, and cross validation." *Journal of Econometrics*, 222, 745–777. https://www.sciencedirect.com/science/article/pii/S0304407620302815. [618]

Juodis, Artūras (2018), "Pseudo panel data models with cohort interactive effects." *Journal of Business and Economic Statistics*, 36, 47–61. [624]

Juodis, Artūras (2020), "This shock is different: Estimation and inference in misspecified two-way fixed effects regressions." Mimeo. [615]

Juodis, Artūras, Hande Karabıyık, and Joakim Westerlund (2021), "On the robustness of the pooled cce estimator." *Journal of Econometrics*, 220, 325–348. [617]

Juodis, Artūras, and Simas Kučinskas (2023), "Supplement to 'Quantifying noise in survey expectations'." *Quantitative Economics Supplemental Material*, 14, https://doi.org/10.3982/QE1633. [609]

Juodis, Artūras and Simon Reese (2022), "The incidental parameters problem in testing for remaining cross-section correlation." *Journal of Business and Economic Statistics*, 40, 1191–1203. [621]

Kahneman, Daniel, Olivier Sibony, and Cass R. Sunstein (2021), *Noise: A Flaw in Human Judgment*. Little, Brown. [610]

Kučinskas, Simas and Florian S. Peters (2022), "Measuring under- and overreaction in expectation formation." Review of Economics and Statistics, (forthcoming). [613, 616, 622, 631, 642]

Kyle, Albert S. (1985), "Continuous auctions and insider trading." *Econometrica*, 53, 1315–1335. [630]

Lahiri, Kajal and Xuguang Sheng (2008), "Evolution of forecast disagreement in a Bayesian learning model." *Journal of Econometrics*, 144, 325–340. [614]

Landier, Augustin, Yueran Ma, and David Thesmar (2019), "Biases in expectations: Experimental evidence." Working paper. [630]

Lorenzoni, Guido (2009), "A theory of demand shocks." *American Economic Review*, 99, 2050–2084. [615]

Lucas, Robert E. (1973), "Some international evidence on output-inflation tradeofs." *American Economic Review*, 63, 326–334. [611, 614]

Maćkowiak, Bartosz and Mirko Wiederholt (2009), "Optimal sticky prices under rational inattention." *The American Economic Review*, 99, 769–803. [611]

Manski, Charles F. (2004), "Measuring expectations." *Econometrica*, 72, 1329–1376. [630]

Manski, Charles F. (2018), "Survey measurement of probabilistic macroeconomic expectations: Progress and promise." NBER Macroeconomics Annual, 32, (forthcoming). [610]

Manski, Charles F. and Francesca Molinari (2010), "Rounding probabilistic expectations in surveys." *Journal of Business and Economic Statistics*, 28, 219–231. [611, 630]

Melosi, Leonardo (2014), "Estimating models with dispersed information." *American Economic Journal: Macroeconomics*, 6, 1–31. [629]

Menzel, Konrad (2021), "Bootstrap with cluster-dependence in two or more dimensions." *Econometrica*, 89, 2143–2188. [615]

Nimark, Kristoffer P. (2014), "Man-bites-dog business cycles." *The American Economic Review*, 104, 2320–2367. [615]

Odean, Terrance (1998), "Volume, volatility, price, and profit when all traders are above average." *The Journal of Finance*, 53, 1887–1934. [633]

Okui, Ryo and Takahide Yanagi (2019), "Panel data analysis with heterogeneous dynamics." *Journal of Econometrics*, 212, 451–475. [620]

Onatski, Alexei (2009), "Testing hypotheses about the number of factors in large factor models." *Econometrica*, 77, 1447–1479. [621]

Ottaviani, Marco and Peter Sørensen Norman (2006), "The strategy of professional forecasting." *Journal of Financial Economics*, 81, 441–466. [633]

Pesaran, M. Hashem (2006), "Estimation and inference in large heterogeneous panels with a multifactor error structure." *Econometrica*, 74, 967–1012. [617]

Pesaran, M. Hashem and Martin Weale (2006), "Survey expectations." In *Handbook of Economic Forecasting*, Vol. 1 (Graham Elliott, Clive W. J. Granger, and Allan Timmermann, eds.), 715–776, Elsevier, Amsterdam. [610]

Ryngaert, Jane (2018), "What do (and don't) forecasters know about U.S. inflation?" Working paper. [628, 629]

Shiller, Robert (1980), "Do stock prices move too much to be justified by subsequent changes in dividends?" The American Economic Review. [630]

Sims, Christopher A. (1998), "Stickiness." *Carnegie-Rochester Conference Series on Public Policy*, 49, 317–356. [614]

Sims, Christopher A. (2003), "Implications of rational inattention." *Journal of Monetary Economics*, 50, 665–690. [611, 614]

Stark, Tom (1997), "Macroeconomic forecasts and microeconomic forecasters in the survey of professional forecasters." Working paper. [637]

Stock, James H. and Mark W. Watson (2002), "Macroeconomic forecasting using diffusion indexes." *Journal of Business and Economic Statistics*, 20, 147–162. [618]

Stock, James H. and Mark W. Watson (2016), *Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics*, 1 edition, Vol. 2. Elsevier. [617]

Surowiecki, James (2005), *The Wisdom of Crowds*. Anchor. [611, 624]

Verbeek, Marno (2008), "Pseudo-panels and repeated cross-sections." In *The Econometrics of Panel Data* (Laszlo Matyas and Patrick Sevestre, ed.), 369–383, Springer. [624]

Westerlund, Joakim and Jean-Pierre Urbain (2015), "Cross-sectional averages versus principal components." *Journal of Econometrics*, 185, 372–377. [617]

Woodford, Michael (2003), "Imperfect common knowledge and the effects of monetary policy." In *Knowledge, Information, and Expectations in Modern Macroeconomics:*

in Honor of Edmund S. Phelps (Philippe Aghion, Roman Frydman, Joseph Stiglitz, and Michael Woodford, eds.), 25–58, Princeton University Press, Princeton. [611, 614]

Co-editor Tao Zha handled this manuscript.

Manuscript received 14 May, 2020; final version accepted 18 December, 2022; available online 20 December, 2022.