Measuring Biases in Expectation Formation*

Simas Kučinskas† Florian S. Peters‡

First Version: June 2018 This Version: May 2019

Abstract

We develop a framework for measuring biases in expectation formation. The basic insight is that under- and overreaction to new information is identified by the impulse response function of forecast errors. This insight leads to a simple and widely applicable measurement procedure. The procedure yields estimates of under- and overreaction to new information at different horizons. Our framework encompasses all major models of expectations, sheds light on existing approaches to measuring biases, and provides new empirical predictions. In an application to inflation expectations, we find that forecasters underreact to aggregate shocks but overreact to idiosyncratic shocks.

Keywords: expectation formation; bias; underreaction; overreaction.
JEL Classification: C53, D83, D84, E70, G40.

*For comments and suggestions at various stages of the project, we thank Peter Boswijk, Frank de Jong, Andreas Fuster, Hans-Martin von Gaudecker, Augustin Landier, Yueran Ma, Marco Ottaviani, Nicola Pavoni, Marcelo Pedroni, Guillermo Ordoñez, Stefano Rossi, Ioanid Rosu, Christian Stoltenberg, Paul Tetlock, Uwe Thümmel, Michael Weber, and Johannes Wohlfahrt as well as seminar audiences at the Bank of Lithuania, BEAM-ABEE, Bocconi, De Nederlandsche Bank, HEC Paris, KIT, RBFC 2018, T2M 2019, University of Amsterdam, University of Bonn, and Vrije Universiteit Amsterdam. We are especially grateful to Olivier Coibion for numerous comments, including the idea to explore using external instruments for identifying shock-specific bias coefficients. Replication files (with Python and STATA code) are available online.

†Humboldt University of Berlin (simas.kucinskas@hu-berlin.de, simaskucinskas.com).
‡University of Amsterdam (f.s.peters@uva.nl, florianpeters.org).

1 Introduction

There is ample evidence that subjective expectations deviate from simple forms of rational expectations.[1] However, there is little agreement on how subjective expectations are actually formed. The lack of consensus has led to a proliferation of models, some of them taking very different views on expectation formation. For instance, much research in macroeconomics has focused on models featuring *underreaction* to new information.[2] At the same time, many prominent models in finance exhibit *overreaction*.[3] Even some of the empirical evidence seems conflicting, with some findings supporting underreaction and others more consistent with overreaction.[4]

Part of the reason for the divergent views has to do with the lack of a general measurement framework. By this we mean clear definitions for what terms like under- and overreaction mean as well as empirical measures that are tightly linked to these theoretical concepts. In the absence of such a framework, empirical work has resorted to either reduced-form approaches that can be difficult to interpret, or methods relying on restrictive assumptions.

Our goal is to provide a measurement framework that fills this gap. We start with a natural definition of under- and overreaction in expectations:

> *An agent is said to <u>underreact</u> to a shock when forming expectations about some variable x_t if the agent perceives the impact of a shock to x_t to be smaller, in absolute terms, than it actually is. If the agent perceives the impact to be larger than it actually is, the agent is said to <u>overreact</u>.*

The definition can be cast in terms of *impulse response functions* (IRFs). Underreaction to news at some horizon is equivalent to the perceived IRF being smaller than than the actual IRF at a particular lag. We show that the definition is reasonable in the sense that models commonly thought to exhibit under- or overreaction indeed do so according to our definition.

1 For surveys, see Pesaran and Weale (2006, Section 5) and Manski (2018). Coibion, Gorodnichenko, and Kamdar (2018) provide a review focused on inflation expectations See also Gennaioli and Shleifer (2018).

2 Examples include sticky information (Mankiw and Reis, 2002), rational inattention (Sims, 2003; Maćkowiak and Wiederholt, 2009), imperfect information (e.g., Lucas, 1973; Woodford, 2003), and sparsity-based models of limited attention (Gabaix, 2014, 2017b, 2018).

3 Examples include diagnostic expectations (Gennaioli and Shleifer, 2010; Bordalo, Gennaioli, and Shleifer, 2018b), extrapolative expectations (e.g., Cutler, Poterba, and Summers, 1990; DeLong, Shleifer, Summers, and Waldmann, 1990; Barberis, Greenwood, Jin, and Shleifer, 2015), and overconfidence (e.g., Daniel, Hirshleifer, and Subrahmanyam, 1998; Odean, 1998).

4 An example is given by De Bondt and Thaler (1990) (overreaction) and Abarbanell and Bernard (1992) (underreaction) for the case of stock market analysts. See also Bouchaud, Krüger, Landier, and Thesmar (2018) and Bordalo, Gennaioli, La Porta, and Shleifer (2017).

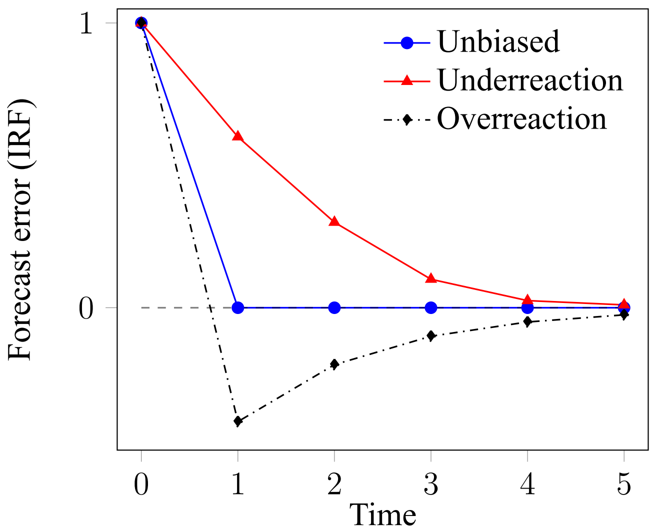


Figure 1: An example of impulse response functions (IRFs) of forecast errors corresponding to unbiased reaction to news as well as under- and overreaction.

The main insight of the paper is that there is a close link between under- and overreaction and the IRF of forecast errors. Forecast errors, by definition, are equal to the realized value less the forecast. The IRF of forecast errors, then, is given by the true IRF of the variable minus the IRF as it is perceived by the agent. The IRF of forecast errors therefore directly corresponds to our definitions of under- and overreaction. We formalize this notion by defining *bias coefficients* as the difference between the perceived and actual IRFs. The key theoretical result is that these bias coefficients are identified from the IRF of forecast errors.

The bias coefficients measure under- and overreaction to news at various horizons. As a result, our framework recovers the entire *term structure* of biases in expectations. This feature distinguishes our framework from previous approaches in the literature. As we show, failing to disentangle biases at various horizons can lead to incorrect conclusions about whether agents under- or overreact to new information.

To gain intuition about our approach, imagine an analyst predicting future inflation. Suppose that because of an oil price shock, inflation in the current quarter is higher than was expected by the analyst. If the analyst reacts to the shock in an unbiased way, the forecast error in the current quarter should not be predictive of forecast errors in the future. In the absence of any systematic bias, the only reason for making a forecast error in the first place is the realization of a shock, and shocks are by definition unpredictable. However, suppose that the analyst tends to underreact to news. Since inflation is persistent, a positive forecast error today implies that the forecast error next quarter is likely to again be positive. In contrast, if the analyst overreacts to the higher-than-expected inflation and makes a very high forecast, inflation in the next quarter will on average be lower than anticipated. With overreaction, a positive forecast error in the current quarter is associated with negative forecast errors in the future.

The IRFs of forecast errors that result from such expectations are plotted in Figure 1. With unbiased reaction to news, the initial shock to forecast errors dies out immediately, and the IRF is zero for all subsequent periods. With underreaction, the IRF is positive, and the shock decays slowly. With overreaction, on the other hand, the IRF is negative. More generally, the IRF of forecast errors may, for example, first be negative and then turn positive, corresponding to overreaction to recent news and underreaction to distant news.

This basic insight leads to a simple and widely applicable measurement procedure which boils down to estimating the IRF of forecast errors. The measurement procedure has several attractive features. First, it does not require precise knowledge of the true data-generating process. Second, the procedure is straightforward to apply and can be used in a variety of empirical settings. These include both experimental and observational data, individual- and consensus-level forecasts, and forecasts of various horizons. The IRF of forecast errors can be estimated using a variety of existing methods, depending on the application. The IRF then yields a set of estimated bias coefficients. These coefficients measure under- and overreaction to news at various horizons, thereby estimating the whole term structure of biases. Since the IRF can be estimated using flexible estimation techniques, biases can be measured without imposing strong parametric restrictions.

The data requirements for estimating biases depend on how many shocks are present. When a *single shock* is driving the variable being forecast, bias coefficients can be estimated with data on forecasts and realizations only. When *multiple shocks* are present, additional information is necessary for identification, as is standard in multivariate settings.

With multiple shocks present, we distinguish between two types of bias coefficients. *Composite bias coefficients* provide a summary measure of how the agent reacts to the multiple shocks that are present. *Shock-specific bias coefficients* measure the agent's reaction to a particular shock. Composite bias coefficients can be estimated with data on forecasts and realizations only. These coefficients are informative about under- and overreaction at various horizons. However, they may be difficult to interpret if the agent underreacts to some shocks while overreacting to others. In addition, they are silent on what information the agent is under- or overreacting to. For example, composite bias coefficients alone cannot say whether the agent underreacts to oil price shocks or monetary surprises when forming inflation expectations.

Identifying the shock-specific bias coefficients requires additional information. Fortunately, these additional informational requirements are fairly low. Formally, if we can identify how the variable being forecast responds to a particular shock, we can also identify the associated shock-specific bias coefficients. For example, suppose we can consistently estimate the effect of a monetary policy shock on inflation. Then, we can also identify biases in the way the agent reacts to that monetary policy shock when forming inflation expectations. This insight leads to an instrumental-variables procedure that can be used for estimating the shock-specific bias coefficients in practice.

The estimated bias coefficients can have multiple economic interpretations. Non-zero bias coefficients provide evidence of *statistical* bias and need not imply irrationality. For example, underreaction could arise from informational frictions, strategic behavior as well as psychological biases. That said, if we choose a particular model of expectations and postulate a process for the variable being forecast, it is straightforward to derive the implied bias coefficients. Comparing the estimated bias coefficients to their theoretical counterparts provides a natural test of the model. Bias coefficients also provide a natural set of moments to target for calibration exercises and structural estimation.

We illustrate the methodology using inflation forecasts from the Survey of Professional Forecasters. To estimate composite bias coefficients without imposing strong parametric assumptions, we employ local projections (Jordá, 2005). Composite bias coefficients indicate underreaction for both individual as well as consensus forecasts. The qualitative pattern for both sets of bias coefficients is very similar. However, bias coefficients estimated from the individual forecasts are attenuated towards zero, suggesting that expectations exhibit some forecaster-specific noise.

Looking at the individual forecasts more closely, we find that forecasters exhibit overreaction to *idiosyncratic* shocks. We identify idiosyncratic shocks by using deviations of individual forecasts from the consensus forecast as an instrument in our instrumental-variables procedure. The idiosyncratic shocks may represent forecaster-specific information, overconfidence, or capture expectations that are inherently noisy. Using the method of external instruments (Stock, 2008; Mertens and Ravn, 2013; Stock and Watson, 2018), we find short lived (around one quarter) but statistically significant overreaction. The fact that the reaction to idiosyncratic shocks is short lived, combined with the previous findings on underreaction as measured by composite bias coefficients, implies that forecasters underreact to *aggregate* shocks (i.e., shocks common to all forecasters such as monetary policy or oil price shocks). These findings are consistent with the recent work of Broer and Kohlhas (2018) and Bordalo, Gennaioli, Ma, and Shleifer (2018a) and highlight the need to distinguish how forecasters react to idiosyncratic and aggregate shocks.

Next, we illustrate how the procedure can be used to measure biases in reaction to specific shocks. For this purpose, we use the Romer and Romer (2004) measure of monetary surprises in an instrumental-variables estimation. We find that forecasters underreact to monetary policy shocks, and the pattern of underreaction is consistent with the impact of monetary policy shocks on actual inflation. Finally, we demonstrate the flexibility of our method by showing how to obtain time- and state-dependent estimates of biases. Our estimates indicate that forecasters are slow to react to time and state dependence in actual inflation, again suggesting underreaction to aggregate shocks.

The existing literature on expectations is voluminous, and we refer to the surveys cited above (Footnote 1) for comprehensive reviews. We are certainly not the first to study the predictability of forecast errors. A key result in the literature on forecast evaluation is that for optimal one-step-ahead forecasts, forecast errors are white noise (see, e.g., Diebold and Lopez, 1996).[5] Our contribution is to show that the structure of predictability in forecast errors is informative about how expectations are formed. In his pioneering work Muth (1961, pp. 321–322) already considered a model with potential under- and overreaction to current news that is a special case of our theoretical framework. However, Muth did not study how such biases may be estimated empirically. The paper that is most closely related to our work is Coibion and Gorodnichenko (2012). Their main empirical specification, in fact, is nested in our framework and can be understood as the reduced-form equation of our instrumental-variables procedure. We discuss certain advantages of our approach, especially the robustness against measurement error, in Section 2.3.1 after presenting the framework.

To our knowledge, the current paper is the first to provide an explicit framework for measuring under- and overreaction in expectations. However, applied work has employed various reduced-form approaches to measure similar phenomena. Our framework is helpful for understanding whether these approaches provide valid measures of under- and overreaction, as we show in Section 2.3.2. Procedures commonly used in practice—such as methods based on Mincer-Zarnowitz regressions, autocorrelations of forecast errors, or forecast revisions as proxies for news—turn out to recover under- and overreaction only under restrictive conditions. When these assumptions are violated, different methods may yield different conclusions about the dominant form of bias, an important concern in practice. A key problem with the existing approaches is that they do not estimate the entire term structure of biases. As a result, underreaction at some horizon can be conflated with overreaction at another. In some cases, reduced-form

[5] Here, optimality is taken to mean minimization of mean squared error. The result does not generalize to asymmetric loss functions; see Patton and Timmermann (2007) and references therein.

results that are commonly interpreted as indicating underreaction are in fact consistent with overreaction (and vice versa). For example, forecast errors may be positively autocorrelated even when the agent exhibits overreaction, provided that the agent overreacts to information at multiple lags.

Related issues have also been extensively studied outside economics, most prominently in psychology. Findings in psychology exhibit a similar tension between under- and overreaction. Early experimental studies of Bayesian updating found that subjects often do not update enough (Edwards, 1968). Studies on belief persistence show that people often hold on to incorrect beliefs (see, e.g., Nickerson, 1998, pp. 187–188). Conservatism bias and belief persistence are both forms of underreaction to new information. However, other well-known findings in psychology are more consistent with overreaction. For example, Kahneman and Tversky (1973) find that subjects fail to incorporate base rates and the reliability of the available information when making predictions. The famous hot-hand fallacy study of Gilovich, Vallone, and Tversky (1985) suggests that people overreact to noise.[6]

The methodology is developed in Section 2. In Section 3 we show how existing models can be mapped into our framework. In Section 4 we apply the method to data on inflation forecasts from the Survey of Professional Forecasters. Section 5 concludes.

2 Methodology

2.1 Framework

We want to measure biases in how an agent forms expectations about some variable x_t. To build intuition, we first study the situation in which x_t is driven by a single shock. Then, we generalize to the case with multiple shocks. Throughout, we assume that any deterministic component from x_t is already removed, and x_t is demeaned. In the Appendix (Appendix C), we show how the framework can be extended to account for multiple-step-ahead forecasts, heterogeneity in expectations, and state- and time-dependent biases. There, we also discuss the impact of measurement error on our procedure as well as how individual forecasts can be used to measure the reaction of ex-

[6] But see also Miller and Sanjurjo (2018). Other classic findings in psychology suggestive of overreaction include illusion of choice (Langer, 1975) and illusory correlation (for a review, see Chapman and Chapman, 1982); Andreassen (1987, p. 490) provides additional references highlighting the tension between under- and overreaction in sequential settings. Griffin and Tversky (1992) argue that the conflicting results can be reconciled if people focus too much on how diagnostic a piece of information is about a given hypothesis but place too little emphasis on the credence of that information. See Nisbett and Ross (1980, especially Chapters 5, 7, and 8) for further discussion. Benjamin (2018) provides a recent survey on related issues.

pectations to idiosyncratic shocks.

2.1.1 Single Shock

Suppose that x_t follows a linear stationary process

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

for some coefficients α_ℓ with $\alpha_0 = 1$ and a martingale difference sequence of shocks ε_t (so that $\mathbb{E}_t[\varepsilon_{t+1}] = 0$, with the expectation conditional on all information available at time t). The class of processes nested in Eq. (1) is already fairly general and nests all stationary ARMA processes with shocks that may exhibit conditional heteroskedasticity. However, the framework can be generalized naturally to handle non-stationary x_t as well as some forms of non-linearity and time-varying parameters, as shown in the Appendix.

We observe an agent making one-step ahead forecasts denoted by $\mathbb{F}_t[x_{t+1}]$. We assume that the forecasts are generated as

$$\mathbb{F}_t[x_{t+1}] = b_0 + \sum_{\ell=0}^{+\infty} a_{\ell+1} \varepsilon_{t-\ell}. \tag{2}$$

Here, b_0 is a time-invariant bias term, while the coefficients a_ℓ capture how subjective expectations react to past shocks. If $a_\ell \neq \alpha_\ell$, the subjective reaction to past shocks is different from the reaction of the true process. The formulation of expectations generalizes an early approach of Muth (1961).[7]

We say that the expectation formation process is *unbiased* if $\mathbb{F}_t[x_{t+1}] = \mathbb{E}_t[x_{t+1}]$ with probability one, i.e., the forecast coincides with the true conditional expectation almost surely. We define under- and overreaction by comparing the true reaction of x_t to a shock with how the agent perceives x_t to react to that shock. Formally, we say that the agent *underreacts* to a shock that arrived ℓ periods ago (i.e., $\varepsilon_{t-\ell}$) if the perceived response $a_{\ell+1}$ is smaller than the true response $\alpha_{\ell+1}$ in absolute value. The agent is said to *overreact* if the perceived response is greater than the actual response in absolute value. For example, the agent overreacts to current news, ε_t, if $|a_1| > |\alpha_1|$.

We can express the difference between the true conditional expectation and the ob-

[7] Muth studied a specification of expectations (his equation 3.18) that is a special case of our Eq. (2). In the notation of the present work, Muth allowed the subjective reaction to current news (a_1) to differ from the true reaction of the process (α_1).

served forecast as

$$\mathbb{E}_{t-1}[x_t] - \mathbb{F}_{t-1}[x_t] = -b_0 - \sum_{\ell=1}^{+\infty} \text{sgn}(\alpha_\ell) b_\ell \varepsilon_{t-\ell}, \tag{3}$$

where $b_\ell \equiv \text{sgn}(\alpha_\ell)(a_\ell - \alpha_\ell)$, $\ell \geq 1$ denote *bias coefficients*.[8] For the expectation formation process to be unbiased, all bias coefficients must evidently be zero. A negative bias coefficient b_ℓ indicates underreaction to news that arrived ℓ periods before the realization of x_t, while positive bias coefficients indicate overreaction.[9] Bias coefficients measure the deviation of the subjective model entertained by the agent from the true model generating x_t. Since the true model is typically unknown, even a fully Bayesian agent may exhibit non-zero bias coefficients.[10]

Outside experimental settings we are unlikely to know how exactly x_t is generated. As a result, we typically do not observe either the true conditional expectation or the shocks. The main insight of this paper is that the bias coefficients can be inferred from the behavior of observed forecast errors. Let $e_t \equiv x_t - \mathbb{F}_{t-1}[x_t]$ denote the *forecast error*. Since $x_t = \mathbb{E}_{t-1}[x_t] + \varepsilon_t$ by Eq. (1) and $\mathbb{E}_{t-1}[\varepsilon_t] = 0$, Eq. (3) implies

$$\underbrace{x_t - \mathbb{F}_{t-1}[x_t]}_{\text{forecast error, } e_t} = -b_0 - \sum_{\ell=1}^{+\infty} \text{sgn}(\alpha_\ell) b_\ell \varepsilon_{t-\ell} + \varepsilon_t. \tag{4}$$

But Eq. (4) is just the *impulse response function* (IRF) of forecast errors. As a result, the bias coefficients are identified—up to the sign—by the IRF of forecast errors. Some knowledge of the true process is needed to conclude whether the agent under- or overreacts, namely the sign of α_ℓ.[11] In many economic settings, even if the precise value of α_ℓ is not known, we have some prior knowledge about its sign. For simply testing whether expectations are unbiased, the sign of α_ℓ is not needed.

Figure 2 illustrates the main idea behind the measurement framework. The dashed blue line shows the IRF of the true process for x_t. The solid red line plots an example IRF of how the process may be perceived by the agent. As seen in the picture, the bias coefficients b_ℓ are equal to the difference between the two IRFs. Since forecast errors are just the difference between realized values and forecasts, the bias coefficients are in turn equal to the IRF of forecast errors. Our specification of expectations is flexible enough

8 The sign function, $\text{sgn}(\alpha_\ell)$, is equal to -1 if $\alpha_\ell < 0$ and 1 otherwise.

9 For the special case of $\alpha_\ell = 0$, any non-zero bias coefficient indicates overreaction.

10 To directly test whether observed beliefs are consistent with Bayesian updating, other methods may be better suited. See, for example, Augenblick and Rabin (2018) and Augenblick and Lazarus (2018).

11 Intuitively, suppose the IRF of forecast errors is positive at some lag. That could be either because (i) the agent underreacts to a shock while x_t responds positively to it; or (ii) the agent overreacts but x_t responds negatively. Both options imply the same response of forecast errors.

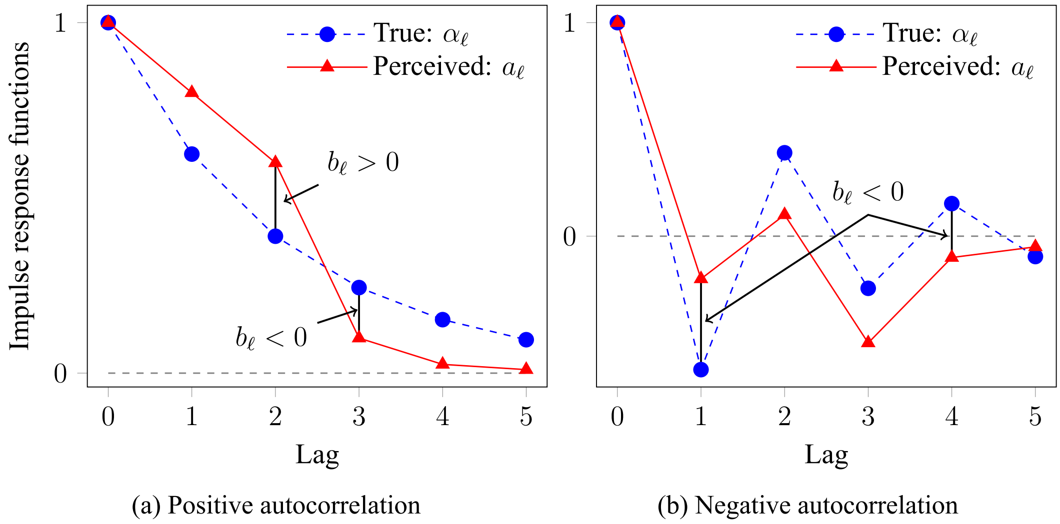


Figure 2: Measurement framework illustrated. The dashed blue line shows the true impulse response function (IRF) of the variable of interest. The solid red line plots the IRF of the process as it may be perceived by the agent (an example). *Left panel*: Positively autocorrelated process. *Right panel*: Negatively autocorrelated process.

to allow for both under- and overreaction at different lags. In the case of positively autocorrelated x_t (shown in the left panel) we have overreaction for $\ell \in \{1, 2\}$ and underreaction for $\ell \geq 3$.

The right panel of Figure 2 shows why we multiply the bias coefficients by $\text{sgn}(\alpha_\ell)$. We say that the agent overreacts whenever the perceived impulse response is larger than the true impulse response in absolute value. Multiplying by $\text{sgn}(\alpha_\ell)$ ensures that a positive bias coefficient indicates overreaction when the true impulse response is negative. For $\ell = 3$, for example, the perceived impulse response is smaller than the true impulse response, but larger in absolute value, and we classify this bias as overreaction.[12]

2.1.2 Multiple Shocks

We now consider the general case in which x_t is driven by $M \geq 1$ shocks

$$x_t = \sum_{i=1}^{M} \sum_{\ell=0}^{+\infty} \alpha_{i\ell} \varepsilon_{i,t-\ell}, \tag{5}$$

[12] A wrinkle arises when the perceived and actual impulse responses differ in sign. For example, suppose that $\alpha_1 = 0.25$ but $a_1 = -0.50$. According to our definition, the agent exhibits overreaction yet the bias coefficient is $b_1 = -0.75 < 0$. In such cases, one can multiply b_1 by (-1) to ensure that $b_1 < 0$ indeed indicates underreaction. The fundamental issue, though, is that it is challenging to define under- and overreaction when the perceived and actual signs of the impulse responses differ. Care in interpreting the bias coefficients is necessary in these circumstances.

where $\boldsymbol{\varepsilon}_t \equiv (\varepsilon_{1t}, \varepsilon_{2t}, \cdots, \varepsilon_{Mt})^\top$ is a martingale difference sequence with $\mathbb{E}[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t^\top] = \boldsymbol{\Sigma}$. Without loss of generality, we assume that $\boldsymbol{\Sigma}$ is diagonal. We generalize the previous specification of expectations in Eq. (2) to

$$\mathbb{F}_t[x_{t+1}] = b_0 + \sum_{i=1}^{M} \sum_{\ell=0}^{+\infty} a_{i,\ell+1} \varepsilon_{i,t-\ell}.$$

As before, b_0 is a time-invariant bias term, while $a_{i\ell}$'s capture the agent's subjective reaction to past shocks. Formally, we say that an agent *overreacts to shock* $\varepsilon_{i,t-\ell}$ if $|a_{i,\ell+1}| > |\alpha_{i,\ell+1}|$ and underreacts if $|a_{i,\ell+1}| < |\alpha_{i,\ell+1}|$.

Performing the same manipulations as in Section 2.1.1, we arrive at

$$x_t - \mathbb{F}_{t-1}[x_t] = -b_0 - \sum_{i=1}^{M} \sum_{\ell=1}^{+\infty} \operatorname{sgn}(\alpha_{i\ell}) b_{i\ell} \varepsilon_{i,t-\ell} + \sum_{i=1}^{M} \alpha_{i0} \varepsilon_{it}, \tag{6}$$

where $b_{i\ell} = \operatorname{sgn}(\alpha_{i\ell})(a_{i\ell} - \alpha_{i\ell})$ denote *shock-specific bias coefficients*. As before, the expectation formation process is unbiased if and only if all shock-specific bias coefficients are zero. The bias coefficients are again identified—up to the sign—from the IRF of forecast errors.

With a single shock present, only data on forecasts and realizations of x_t are necessary to estimate the bias coefficients. That is no longer true with multiple shocks. The reason is that the underlying shocks are not observed. Hence, to estimate the shock-specific bias coefficients, additional information is required. We emphasize that this data requirement is generic and not specific to our method. Estimating IRFs in multivariate settings requires additional information because of standard identification problems (see, e.g., Hamilton, 1994, Chapter 11).[13]

For this reason we consider two cases. First, we study the case when only data on forecasts and realizations of x_t are available. In this situation, it is not feasible to estimate the shock-specific bias coefficients. However, estimating the univariate IRF of forecast errors is nevertheless informative. Such estimation yields a set of *composite bias coefficients*, a summary measure of biases in how the agent reacts to the various shocks that are present.[14] Next, we consider the case in which additional information is available. We show that if the econometrician can identify how the variable of interest

[13] This remains the case if we observe forecasts and realizations of other variables in addition to x_t. That said, if forecasts of multiple variables are available, standard SVAR identification schemes may be used to estimate the shock-specific bias coefficients. In contract to the IV procedure that we pursue below, standard SVARs require shocks to be invertible to be consistent; see Stock and Watson (2018) and Plagborg-Møller and Wolf (2018a).

[14] In a previous version of the paper, we used "aggregate" and "individual" instead of "composite" and "shock-specific" bias coefficients.

responds to a particular shock, it is also possible to identify the associated *shock-specific* bias coefficients. Theoretically, the takeaway is that the informational requirements for measuring biases are the same as estimating the IRF of the variable of interest. Practically, the result provides a way to estimate shock-specific bias coefficients using the method of external instruments.

Case #1: Estimating *composite* bias coefficients. Provided that the forecast errors are stationary, by the Wold's Theorem we can write

$$x_t - \mathbb{F}_{t-1}[x_t] = -b_0 - \sum_{\ell=0}^{+\infty} \theta_\ell \nu_{t-\ell}, \tag{7}$$

for some square-summable coefficients θ_ℓ and a white noise series ν_t. The Wold shocks, ν_t, are innovations from a projection of forecast errors on all their past values and hence distinct from the structural shocks, ε_t. For Eq. (7) to be valid, it is only necessary to have stationary forecast errors, a much weaker requirement than x_t itself being stationary. We can similarly represent x_t as

$$x_t = \sum_{\ell=0}^{+\infty} \alpha_\ell \xi_{t-\ell},$$

with a slight abuse of notation (i.e., α_ℓ's are not equal to the $\alpha_{i\ell}$'s in Eq. (5)).

With this notation, we define *composite bias coefficients* as $b_\ell = \mathrm{sgn}(\alpha_\ell)\theta_\ell$. We say that (for $\alpha_\ell \neq 0$) the agent exhibits *overall overreaction* if $b_\ell > 0$ and exhibits *overall underreaction* if $b_\ell < 0$. (If $\alpha_\ell = 0$ and $\theta_\ell \neq 0$, the agent is said to exhibit overall overreaction.)

The proposition below summarizes the key properties of composite bias coefficients. Given the shock-specific bias coefficients, one can readily calculate the autocovariance function of forecast errors. The autocovariance function can then be inverted to find the implied composite bias coefficients. This way, composite bias coefficients provide a summary measure of the shock-specific bias coefficients.

Proposition 1. *For the expectation formation process to be unbiased, it is necessary (but not sufficient) for all composite bias coefficients to be zero. Composite bias coefficients,* $b_\ell = \mathrm{sgn}(\alpha_\ell)\theta_\ell$*, can be calculated from the individual bias coefficients* $\{b_{i\ell}\}$ *by solving*

$$\sum_{\ell=0}^{+\infty} \theta_\ell \theta_{\ell+k} = \frac{\gamma_k}{\gamma_0} \sum_{\ell=0}^{+\infty} \theta_\ell^2 \text{ for } k = 1, 2, \ldots, \tag{8}$$

where $\gamma_k = \gamma_k(\{b_{i\ell}\})$ *is the autocovariance function of forecast errors (explicit expression given in Eq. (A.1) in the Appendix).*

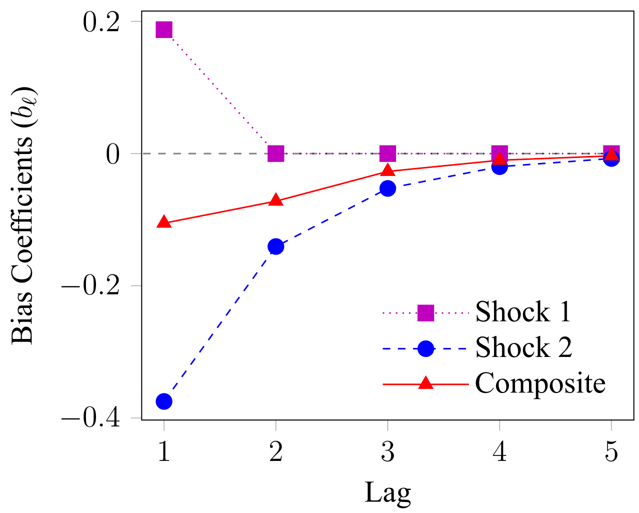


Figure 3: Numerical example of shock-specific and composite bias coefficients. The true process is given by $x_t = \sum_{\ell=0}^{+\infty} \rho^\ell (\varepsilon_{1,t-\ell} + \varepsilon_{2,t-\ell})$ with $\mathbb{V}\mathrm{ar}[\varepsilon_{1t}] = \mathbb{V}\mathrm{ar}[\varepsilon_{1t}] = 1.0$ and $\rho = 0.75$. The process for the forecast errors is given by $e_t = \sum_{\ell=0}^{5} (\lambda\rho)^\ell \varepsilon_{1,t-\ell} + (\varepsilon_{2t} - \theta\varepsilon_{2,t-1})$ with $\lambda = 0.50$ and $\theta = 0.25$. The composite bias coefficients are found by solving Eq. (8) numerically. The expectations formation process is a combination of diagnostic expectations (Bordalo, Gennaioli, and Shleifer, 2018b) and the sticky information model of Mankiw and Reis (2002); see Section 3.

Proof. In the Appendix. □

In general, the relationship between the shock-specific and composite bias coefficients is non-linear, as given in Eq. (8). If the autocovariances vanish at some finite lag, then the process for the forecast errors has a finite Wold representation, and Eq. (8) boils down to a system of nonlinear equations with a unique solution (see, e.g., Ansley, Spivey, and Wrobleski, 1977).

Figure 3 illustrates the relationship between the shock-specific and composite bias coefficients. In this numerical example, there are two shocks with equal variances. The agent overreacts to the first shock but underreacts to the second. The expectations formation process is a combination of diagnostic expectations (Bordalo, Gennaioli, and Shleifer, 2018b) and the sticky information model of Mankiw and Reis (2002); see Section 3 for a description of these models. The agent has diagnostic expectations with respect to the first shock but sticky information with respect to the second shock. The figure shows the two sets of shock-specific bias coefficients as well as the implied composite bias coefficients obtained by solving Eq. (8).

The composite bias coefficients lie between the two shock-specific bias coefficients. Although the agent exhibits overreaction to the first shock, all composite bias coefficients are negative. Intuitively, at the chosen parameter values, underreaction to the second shock turns out to quantitatively dominate overreaction to the first shock. The figure also highlights a potential pitfall when only data on forecasts and realizations is available. If the agent underreacts to one shock and simultaneously overreacts to another shock in a way that makes the biases "cancel out," composite bias coefficients may be close to zero even when shock-specific bias coefficients are not. As formally shown in Proposition 1, composite bias coefficients being zero is a necessary, but not sufficient, condition for the expectation formation process to be unbiased.

Case #2: Estimating *shock-specific* bias coefficients. Suppose that in addition to data on forecasts and realizations, we have an *instrumental variable* (IV) denoted by z_t. We assume that the instrument z_t allows us to identify the IRF of x_t to a shock in ε_{1t}, following the ideas in Stock and Watson (2018). Let

$$w_t = \sum_{i=1}^{M} \sum_{\ell=0}^{+\infty} \gamma_{i\ell} \varepsilon_{i,t-\ell}.$$

be another observed variable that will be instrumented by z_t. As a concrete example, we may wish to learn how the agent reacts to a monetary policy shock (ε_{1t}) when forming expectations about inflation (x_t). In this case, the instrument z_t may be a monetary surprise as measured by Romer and Romer (2004) or a shock estimated using an SVAR. The additional variable w_t could be the Federal Funds Rate. We only assume that the instrument allows us to consistently estimate the effects of a monetary policy shock on realized inflation. No further assumptions are made (e.g., on how the instrument is related to the agent's inflation expectations).

Formally, the external instrument is assumed to satisfy the following conditions.

Assumption 1. *External instrument z_t satisfies the following conditions:*

- *Relevance:* $\mathbb{E}[z_t \varepsilon_{1t}] = \phi \neq 0$*;*
- *Contemporaneous exogeneity:* $\mathbb{E}[z_t \varepsilon_{jt}] = 0$ *for* $j \neq 1$*;*
- *Lead-lag exogeneity:* $\mathbb{E}[z_t \varepsilon_{j,t+s}] = 0$ *for all* j *and* $s \neq 0$*;*
- *Normalization:* $\gamma_{10} = 1$.

The first three assumptions are familiar from standard IV estimation. The last assumption is necessary since, given that the shocks are unobservable, the scale of ε_{1t} is indeterminate. In the example above, the normalization corresponds to a choice of units under which the monetary policy shock is equal to the change in the Federal Funds Rate. In cases in which the normalization is not appropriate, the external instrument identifies the IRF up to the scale parameter $(1/\gamma_{10})$. Finally, to simplify notation, we assume that z_t has been demeaned so that $\mathbb{E}[z_t] = 0$.

Stock and Watson (2018) show that under Assumption 1 the instrument z_t identifies the IRF of x_t to a shock in ε_{1t}. Our key result is that the same instrument can also be used to identify the shock-specific bias coefficients. The proof is a straightforward modification of the arguments used by Stock and Watson.

Proposition 2. *Suppose that an instrument z_t satisfies Assumption 1. Then, the shock-specific bias coefficients can be consistently estimated by $\hat{b}_{1\ell} = -\operatorname{sgn}(\alpha_{1\ell})\hat{\beta}^{(\ell)}_{IV}$, where $\hat{\beta}^{(\ell)}_{IV}$ is the instrumental-variables estimate of $\beta^{(\ell)}$ in the regression $e_{t+\ell} = \beta^{(\ell)} w_t + u_{t+\ell}, \ell \geq 1$, with w_t instrumented by z_t.*

Proof. In the Appendix. □

A practical implication is that shock-specific bias coefficients can be consistently estimated using the method of external instruments (Stock, 2008; Mertens and Ravn, 2013; Stock and Watson, 2018). In this regression, forecast errors, $e_{t+\ell}$, are regressed on w_t, instrumenting w_t with the external instrument. In the previous example, the procedure amounts to regressing forecast errors for inflation on the lagged Federal Funds Rate, instrumenting the Federal Funds Rate with an appropriate instrument.

Several extensions to Proposition 2 are immediate. For example, if several instruments are available, they can be combined using a standard two-stage least squares estimator. In addition, control variables can be included in the regression in Proposition 2. Additional control variables may be helpful in obtaining more precise estimates. Moreover, an instrument may only be valid after conditioning on some other variables. For more on these points, see the discussion in Stock and Watson (2018).

An attractive feature of the IV procedure is its robustness to various types of measurement error, a feature originally emphasized by Mertens and Ravn (2013). Measurement error is likely to be especially important in the case of expectations. First, the instrument need not be perfectly correlated with the underlying shock ε_{1t}. The correlation only needs to be high enough for z_t to be a strong instrument. Less obviously, the IV approach is robust to measurement error in expectations. Suppose that instead of observing the true expectations $\mathbb{F}_t[x_{t+1}]$, we only observe $\mathbb{F}^*_t[x_{t+1}] = \mathbb{F}_t[x_{t+1}] + v_t$ where v_t is measurement error. Then, the IV estimator remains consistent as long as the measurement error is uncorrelated with the instrument at all leads and lags (even if measurement error itself is serially correlated). This assumption is fairly weak since in applications, the instrument will typically be based on a different dataset than the one from which data on expectations are drawn.

2.2 Estimation

Measuring the bias coefficients boils down to estimating the IRF of forecast errors. The best way to estimate the IRF depends on the application at hand, and many existing approaches may be employed. While our paper has nothing new to say about the estimation of IRFs, we briefly outline how bias coefficients can be estimated in the most empirically relevant scenarios.

Case #1: Estimating bias coefficients with known shocks. If the true shocks are observed, bias coefficients can be estimated by directly regressing forecast errors, $e_t = x_t - \mathbb{F}_{t-1}[x_t]$, on past shocks. This approach is likely to be especially relevant for experimental work.[15] For instance, in the case of a single shock, one may estimate

$$e_t = \mu + \beta_0 \varepsilon_t + \beta_1 \varepsilon_{t-1} + \cdots + \beta_K \varepsilon_{t-K} + u_t. \tag{9}$$

The estimated bias coefficients for $\ell \geq 1$ are then $\hat{b}_\ell = -\operatorname{sgn}(\alpha_\ell)\hat{\beta}_\ell$.

Case #2: composite bias coefficients. To estimate composite bias coefficients (or shock-specific bias coefficients if only one shock is present) only data on forecasts and realizations is required. In this case, one may use *local projections* (Jordá, 2005) to estimate the IRF flexibly. With local projections, for each $\ell = 1, 2, ..., L$, the following regression is estimated by least squares:

$$e_{t+\ell} = \beta_0^{(\ell)} + \underbrace{\beta_1^{(\ell)}}_{\text{coef. of interest}} e_t + \beta_2^{(\ell)} e_{t-1} + \cdots + \beta_K^{(\ell)} e_{t-K+1} + u_{t+\ell}. \tag{10}$$

Here, K is the number of lagged forecast errors included in the local projection. The estimated bias coefficients for $\ell \geq 1$ are then $\hat{b}_\ell = -\operatorname{sgn}(\alpha_\ell)\hat{\beta}_1^{(\ell)}$, where α_ℓ's are the coefficients of the univariate IRF of x_t, as in Section 2.1.2. The time-invariant bias b_0 is estimated by the sample average of e_t.

While we prefer local projections in our empirical application below, there are many ways to estimate an IRF. For example, we may fit a more parsimonious time series model (e.g., an AR(4) for quarterly forecast errors). This alternative may be especially useful when sample size is limited.[16] Alternatively, the IRF may be estimated by fitting a high-order moving average model using maximum likelihood. That said, there are important practical benefits to using local projections. First, it is immediate to extend Eq. (10) to cases in which we have multiple forecasters or want to pool multiple forecasts. Second, it is straightforward to adjust standard errors for clustering that occurs when individual forecasts are used (Keane and Runkle, 1990). Finally, it is easy to adapt Eq. (10) to

[15] There is a large experimental literature on expectation formation, going back to at least Schmalensee (1976). For a recent overview, see Assenza, Bao, Hommes, and Massaro (2014). A recent large-scale experimental study is provided by Ma, Landier, and Thesmar (2017).

[16] As shown by Plagborg-Møller and Wolf (2018b) local projections and (V)ARs estimate the same IRFs. However, the two methods implicitly make different choices on how to trade off bias and variance in finite samples. Hence, the choice between the two options necessarily depends on the application.

handle time and state dependence.

Case #3: Estimating shock-specific bias coefficients. The final major case of interest is when we have an instrument, z_t, that allows us to estimate the effect of some shock on the variable of interest, x_t. As shown in Proposition 2, in this case shock-specific bias coefficients can be estimated by a simple IV regression. In our empirical application, we also control for lagged values of forecast errors and the variable being instrumented (denoted by w_t) for additional precision. Thus, for $\ell = 1, 2, \ldots, L$ we estimate

$$e_{t+\ell} = \beta_0^{(\ell)} + \sum_{s=1}^{K_1} \beta_s^{(\ell)} w_{t+1-s} + \sum_{s=1}^{K_2} \gamma_s^{(\ell)} e_{t+1-s} + u_{t+\ell}, \tag{11}$$

instrumenting w_t with z_t. Here, α_ℓ denote the response of x_t to the particular shock in question, as in Section 2.1.2. The estimated bias coefficients are then $\hat{b}_\ell = -\operatorname{sgn}(\alpha_\ell)\hat{\beta}_1^{(\ell)}$ for $\ell \geq 1$. A special case is when $w_t = \mathbb{F}_t[x_{t+1}]$ and z_t is some proxy for a shock to expectations. This special case corresponds to an IV version of the classic Mincer and Zarnowitz (1969) regression.

Again, there is nothing special about using the method of external instruments. One can also employ a somewhat more parametric method such as the proxy-SVAR technique of Mertens and Ravn (2013) or use conventional SVARs.

2.3 Related Work

2.3.1 Coibion and Gorodnichenko (2012)

The paper that is most closely related to the present work is Coibion and Gorodnichenko (2012). Similarly to us, Coibion and Gorodnichenko study how forecast errors respond to shocks. The authors derive the response of forecast errors to shocks in a number of models and then estimate these IRFs empirically. These IRFs are interpreted as primarily capturing information rigidities. Our framework shows that the IRF of forecast errors identifies biases in expectation formation much more generally, including models that do not feature information rigidities.

The empirical procedure of Coibion and Gorodnichenko is nested in our framework. Their main regression specification (their Eq. (34)) regresses forecast errors on past shocks, as in our Eq. (9).[17] In practice, however, Coibion and Gorodnichenko do not observe these shocks directly but estimate them from observational data. As argued by Mertens and Ravn (2013) and Stock and Watson (2018), these estimates are best interpreted as imperfect measures of the true structural disturbances, i.e., instruments. From

[17] Coibion and Gorodnichenko include lagged values of forecast errors on the right-hand side and use simulation to obtain the IRFs, following the approach of Romer and Romer (2004).

this perspective, the main specification of Coibion and Gorodnichenko corresponds to the reduced-form equation of our instrumental-variables regression in Eq. (11). Indeed, if deeper lags of x_{1t} are not included in the IV regression (i.e., $K_1 = 1$), the reduced form associated with Eq. (11) is

$$e_{t+\ell} = \delta_0^{(\ell)} + \delta_1^{(\ell)} z_t + \sum_{s=1}^{K_2} \vartheta_s^{(\ell)} e_{t+1-s} + u_{t+\ell} \text{ for } \ell = 1, 2, \ldots, L.$$

Hence, the reduced form associated with Eq. (11) is exactly the specification employed by Coibion and Gorodnichenko (2012).[18] A key advantage of our IV method is that it remains consistent in the presence of measurement error, as discussed in Section 2.1.2. In contrast, if forecast errors are regressed on mismeasured shocks, the resulting estimates of bias coefficients are inconsistent.

2.3.2 Reduced-Form Approaches

A large literature tests forecast optimality by looking at whether forecast errors are predictable, as surveyed by Pesaran and Weale (2006, Section 5). Our main regressions, Eqs. (10) and (11), are special cases of such tests. A key advantage of the traditional predictability tests is that they are relatively assumption free. However, the traditional tests can reject the null hypothesis without being particularly informative about the alternative. In contrast, results of our regressions have a natural economic interpretation in terms of under- or overreaction to news at various lags.

In practice, applied work has used several predictability tests to obtain proxies for under- and overreaction. Our framework is helpful for understanding whether such estimates are valid measures of under- and overreaction. We show that this is the case only under restrictive assumptions. When these assumptions are violated, different methods may lead to different conclusions about the dominant form of bias. That is an important concern in practice. For instance, Capistrán and Timmerman (2009, Figure 5) document that most individual forecasters in the Survey of Professional Forecasters exhibit positively autocorrelated forecast errors. Positive autocorrelation, as we discuss below, is often interpreted as indicating underreaction. On the other hand, looking at the same dataset, Broer and Kohlhas (2018); Bordalo, Gennaioli, Ma, and Shleifer (2018a) find that for individual forecasters, forecast errors are negatively predicted by past revisions, suggesting overreaction.

[18] Coibion and Gorodnichenko (2012) include multiple lags of z_t in their specification (ε_t in their notation). For that, it is enough to include lagged values of z_t in the IV regression. Also, Coibion and Gorodnichenko estimate a single equation and then iterate to obtain the IRF while the specification above employs local projections.

To simplify exposition, we assume that there is a single shock present, as in Eqs. (1) and (2). In the main text, we discuss two main approaches based on (i) autocorrelations of forecast errors; and (ii) forecast revisions. In the Appendix (Appendix B), we discuss methods based on the Mincer and Zarnowitz (1969) regression that are also often used.

Autocorrelation of forecast errors. A common approach in applied work estimates the autocorrelation of forecast errors. Positively autocorrelated forecast errors are then interpreted as indicating underreaction (e.g., Abarbanell and Bernard, 1992; Ma, Sraer, and Thesmar, 2018). We can use our framework to see whether such an interpretation is warranted. We calculate that

$$\mathbb{C}\mathrm{ov}(e_t, e_{t-1}) \geq 0 \quad \Leftrightarrow \quad -\operatorname{sgn}(\alpha_1) b_1 + \sum_{\ell=2}^{+\infty} \operatorname{sgn}(\alpha_\ell) \operatorname{sgn}(\alpha_{\ell-1}) b_\ell b_{\ell-1} \geq 0.$$

Hence, positively autocorrelated forecast errors need not indicate overreaction. For example, if $\alpha_\ell \geq 0$ for all ℓ, and the agent exhibits overreaction at all lags, the first-order autocorrelation coefficient is positive whenever b_1 is small enough. The first-order autocorrelation measures the agent's reaction to current news only under the restrictive assumption that $b_\ell = 0$ for all $\ell \geq 2$.

Intuitively, what matters for the first-order autocorrelation is whether *adjacent* bias coefficients have the same sign, not what the sign of each bias coefficient is per se. Hence, if the agent overreacts to information at multiple lags, the first-order autocorrelation can well be positive. The result holds for autocorrelations of higher order, too. The key issue is that the approach does not measure the whole term structure of biases, leading to potentially incorrect inferences.

Forecast revisions as proxies for news. Another method for estimating under- and overreaction uses forecast revisions as a proxy for news. In this approach, one-step-ahead forecast errors are regressed on lagged forecast revisions. A positive slope coefficient is interpreted as evidence of underreaction, while a negative coefficient is taken to indicate overreaction (see, e.g., Broer and Kohlhas, 2018; Bordalo, Gennaioli, Ma, and Shleifer, 2018a). The approach of using forecast revisions to measure information rigidities was originally proposed by Coibion and Gorodnichenko (2015). Coibion and Gorodnichenko show that the regression discussed above recovers structural parameters of expectation formation in several models with information frictions. However, subsequent empirical work has at times interpreted the results of such regressions as directly indicating under- and overreaction, even outside the settings originally studied by Coibion and Gorodnichenko.

We now consider whether this approach estimates under- and overreaction generally.

Generalizing Eq. (2), suppose that h-step-ahead forecasts are given by $\mathbb{F}_t[x_{t+h}] = b_0^{(h)} + \sum_{\ell=0}^{+\infty} a_{\ell+h}^{(h)} \varepsilon_{t-\ell}$. Denote forecast revisions by $r_t \equiv \mathbb{F}_t[x_{t+1}] - \mathbb{F}_{t-1}[x_{t+1}]$. We find

$$\mathbb{C}\text{ov}(e_t, r_{t-1}) \geq 0 \quad \Leftrightarrow \quad -a_1^{(1)} \operatorname{sgn}(\alpha_1) b_1^{(1)} - \sum_{\ell=2}^{+\infty} b_\ell^{(1)} [b_\ell^{(1)} - b_\ell^{(2)}] \geq 0.$$

Hence, the method measures the reaction to current news only under the restrictive assumption that $b_\ell^{(1)} = b_\ell^{(2)}$ for all ℓ. This assumption is typically violated in existing models (outside simple forms of rational expectations). When the assumption is violated, the covariance between forecast errors and past revisions combines biases at various horizons. In addition, one may find a positive correlation between forecast errors and past forecast revisions even when the agent exhibits only overreaction. For example, that is the case whenever b_1^1 is sufficiently close to zero and $b_\ell^{(2)}$ is larger than $b_\ell^{(1)}$. Intuitively, the approach does not distinguish between biases in how the agent forms short- and longer-run expectations. In addition, the method is sensitive to measurement error in expectations. If expectations are measured with classical measurement error, measurement error leads to a mechanical negative correlation between forecast errors and past revisions. The reason is that $\mathbb{F}_{t-1}[x_t]$ is part of both the forecast error and the forecast revision, with opposite signs. The challenges above notwithstanding, if forecast revisions are valid instruments for some shocks, the key ideas in Coibion and Gorodnichenko (2015) can be readily accommodated in our framework following the methods in Section 2.1.2.[19]

3 Mapping Existing Models

We now show how existing models of expectations can be mapped into our framework.

The exercise leads to three key takeaways. First, our definitions of under- and overreaction are reasonable in the sense that models commonly thought to generate under- or overreaction in fact do so according to our definition. Second, the framework is flexible enough to accommodate all major models of expectations. Finally, the implied bias coefficients provide a useful lens for looking at models of expectations and deriving new empirical predictions.

To obtain closed-form expressions, we assume that x_t follows a stationary AR(1):

$$x_t = \rho x_{t-1} + \varepsilon_t, \rho \in (-1, 1), \tag{12}$$

[19] For example, Cascaldi-Garcia (2019) uses forecast revisions to construct an instrument for news shocks.

where, as in Section 2, ε_t is a martingale difference sequence. The simple AR(1) process is a reasonable first approximation for many economic time series. For more complicated processes, it can be more difficult to obtain the bias coefficients analytically. In those cases, it is still straightforward to obtain estimates by simulation.

To streamline the exposition, we discuss four major models of expectations in the main text (rational expectations, sticky information, diagnostic expectations, and adaptive expectations). In Appendix D, we show how to obtain bias coefficients for models of noisy information, diagnostic expectations with noisy information, misperceived law of motion, extrapolative expectations, adaptive learning, forecasting under adjustment costs, and asymmetric loss functions.

The results for the models discussed in the current section are summarized in Figure 4. As shown in the figure, existing models have sharp predictions for the structure of bias coefficients. The sticky information model by Mankiw and Reis (2002) implies that the bias coefficients are all negative and decay geometrically (dashed blue line). This finding is consistent with the standard view that sticky information is a model of underreaction. In contrast, diagnostic expectations of Bordalo, Gennaioli, and Shleifer (2018b) predict that the agent overreacts to current news but reacts rationally to all past news (dotted magenta line). Again, this result accords with the intuition that diagnostic expectations exhibit overreaction. Finally, we plot the bias coefficients implied by adaptive expectations (solid red line). At the chosen parameter values, adaptive expectations predict strong underreaction to current news but mild overreaction to news received further in the past. This last result is somewhat surprising since a common perception is that adaptive expectations respond to new information sluggishly. Hence, adaptive expectations may be expected to only generate underreaction. Intuitively, precisely because of the fact that adaptive expectations react to new information slowly, they end up overreacting to old news.

3.1 Rational Expectations

Rational expectations in the sense of Muth (1961) are given by

$$
\begin{aligned}
a_\ell &= \rho^\ell \\
b_\ell &= 0
\end{aligned}
$$

Here, the perceived response to a shock, a_ℓ, is identical to the true response of x_t. As a result, all bias coefficients are zero.

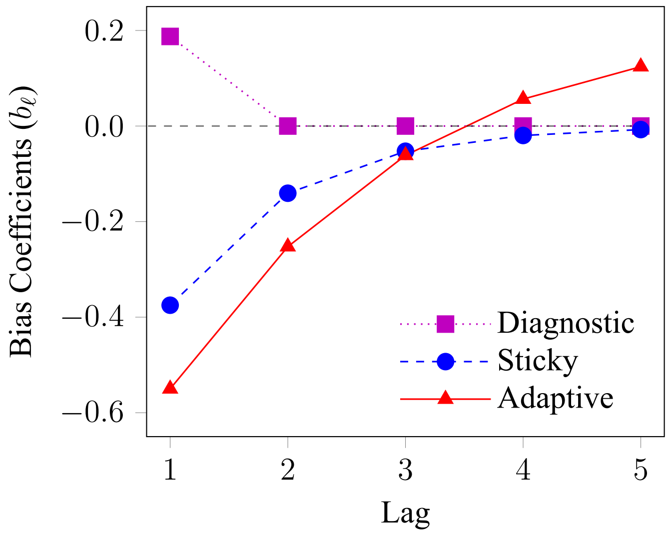


Figure 4: Bias coefficients for selected models of expectations. Positive bias coefficients indicate overreaction to news at a particular lag, and negative coefficients indicate underreaction. Unbiased reaction to news is given by a zero bias coefficient. The underlying process for x_t is an AR(1) with $x_t = 0.75x_{t-1} + \varepsilon_t$. The models shown are: (i) diagnostic expectations of Bordalo, Gennaioli, and Shleifer (2018b) with $\theta = 0.25$; (ii) sticky information model of Mankiw and Reis (2002) with $\lambda = 0.50$; (iii) adaptive expectations of Cagan (1956) and Nerlove (1958) with $\kappa = 0.20$.

3.2 Sticky Information

Consider the sticky information model of expectations proposed by Mankiw and Reis (2002). Each period a fraction $(1-\lambda) \in (0,1]$ of agents update their forecast to the full-information rational expectation. The remaining agents use information obtained in some previous period to form expectations that are rational conditional on their information set.[20] Given these assumptions, expectations at the aggregate (or consensus) level follow

$$\mathbb{F}_t[x_{t+1}] = (1-\lambda)\sum_{\ell=0}^{+\infty} \lambda^\ell \, \mathbb{E}_{t-\ell}[x_{t+1}]. \tag{13}$$

For the AR(1) model, we have that $\mathbb{E}_{t-\ell}[x_{t+1}] = \rho^{\ell+1}x_{t-\ell}$, and some algebra yields

$$\mathbb{F}_t[x_{t+1}] = \sum_{\ell=0}^{+\infty} \rho^{\ell+1}(1-\lambda^{\ell+1})\varepsilon_{t-\ell}. \tag{14}$$

[20] Reis (2006, Section 5) provides a microfoundation for the Poisson adjustment process. See also Carroll (2003).

As a result, we find that

$$a_\ell = \rho^\ell(1-\lambda^\ell)$$
$$b_\ell = -\operatorname{sgn}(\rho^\ell)(\lambda\rho)^\ell$$

As long as expectations do not adjust to news immediately ($\lambda > 0$), the sticky information model exhibits underreaction at all lags.[21]

As seen above, the bias coefficients depend on both (i) how people form expectations; and (ii) the data-generating process. Different processes for x_t will imply different bias coefficients, even if people form expectations in the same way. In the present example, the bias coefficients are larger in absolute value if the process is more persistent. Intuitively, underreaction is more severe when the process is highly persistent.

3.3 Diagnostic Expectations

Suppose that the agent has diagnostic expectations as in Bordalo, Gennaioli, and Shleifer (2018b) and overweights representative events. Bordalo, Gennaioli, and Shleifer (2018b, Proposition 1) show that in this case expectations follow

$$\mathbb{F}_t[x_{t+1}] = \mathbb{E}_t[x_{t+1}] + \theta\left\{\mathbb{E}_t[x_{t+1}] - \mathbb{E}_{t-1}[x_{t+1}]\right\}, \theta \geq 0, \tag{15}$$

where θ is a parameter capturing the extent to which the agent overweights representative events. The expression can be rewritten as $\mathbb{F}_t[x_{t+1}] = \mathbb{E}_t[x_{t+1}] + \rho\theta\varepsilon_t$. Therefore, diagnostic expectations imply that

$$a_\ell = \begin{cases} \rho(1+\theta) & \text{if } \ell = 1 \\ \rho^\ell & \text{if } \ell \geq 2 \end{cases} \quad \text{and} \quad b_\ell = \begin{cases} \theta|\rho| & \text{if } \ell = 1 \\ 0 & \text{if } \ell \geq 2 \end{cases}$$

Hence, diagnostic expectations predict overreaction to current news and unbiased reaction to all other news.

3.4 Adaptive Expectations

Finally, consider adaptive expectations of Cagan (1956) and Nerlove (1958):

$$\mathbb{F}_t[x_{t+1}] = \mathbb{F}_{t-1}[x_t] + \kappa\{x_t - \mathbb{F}_{t-1}[x_t]\} \text{ with } \kappa \in (0,1].$$

[21] Underreaction at all lags with sticky information extends to more general processes. To see this, consider a general x_t as in Eq. (1) and perform the same calculations as for the AR(1) case. The calculation shows that in the general case $b_\ell = -\operatorname{sgn}(\alpha_\ell)\lambda^\ell\alpha_\ell$.

	N	$\bar{x}$	σ_x	ρ_x	$\bar{e}$	RMSE	ρ_e	R^2_{adj}
Inflation: Consensus	196	3.45	2.59	0.83	-0.05	1.40	0.40	0.20
Inflation: Individual	7,475	3.77	2.72	0.82	0.02	1.96	0.32	0.16

Table 1: Summary statistics. Root mean-squared error (RMSE) is calculated as $\sqrt{\frac{1}{N}\sum e_t^2}$ where e_t is the forecast error. Persistence ρ_z is measured by the estimate of b in the regression $z_t = a + bz_{t-1} + v_t$. R^2_{adj} is the adjusted R-squared in the regression of forecast errors on the past four forecast errors e_{t-1}, $e_{t-2}, \ldots, e_{t-4}$.

Iterating we have that

$$\mathbb{F}_t[x_{t+1}] = \kappa \sum_{\ell=0}^{+\infty} (1-\kappa)^\ell x_{t-\ell} = \kappa \sum_{\ell=0}^{+\infty} \left(\frac{(1-\kappa)^{\ell+1} - \rho^{\ell+1}}{1-\kappa-\rho} \right) \varepsilon_{t-\ell}.$$

Hence, we obtain

$$a_\ell = \kappa \left[\frac{(1-\kappa)^\ell - \rho^\ell}{1-\kappa-\rho} \right] \quad \text{and} \quad b_\ell = \text{sgn}(\rho^\ell) \left[\frac{\kappa(1-\kappa)^\ell - (1-\rho)\rho^\ell}{1-\kappa-\rho} \right].$$

Inspecting the expressions above, it is immediate that adaptive expectations can generate both under- and overreaction to new information.

4 Application: Inflation Expectations

Our empirical application uses inflation forecasts from the Survey of Professional Forecasters (SPF), currently run by the Federal Reserve Bank of Philadelphia. This dataset has been used extensively in prior work and provides a natural testing ground for our method. To streamline the discussion, we focus on the findings and explain how we construct the dataset in Appendix E.

We study one-step ahead quarterly GDP deflator inflation forecasts. Summary statistics are provided in Table 1.[22] Both consensus- and individual-level forecasts are considered. We use median forecasts for the consensus to be consistent with prior work, but the median and mean forecasts are very similar. To avoid the possibility that our results are driven by data revisions, we use real-time data to measure realized inflation.

[22] The number of participants in the SPF has not been constant over time. As a result, estimates using the consensus- and individual-level datasets implicitly weight the data somewhat differently. The individual-level dataset implies a somewhat higher weight on observations coming from the earlier part of the sample.

4.1 Composite Bias Coefficients

The estimated composite bias coefficients are shown in Figure 5. As discussed in Section 2.1.2, composite bias coefficients measure the total reaction to shocks and do not disentangle between the reaction to specific shocks (e.g., we cannot say whether agents underreact to oil price shocks or monetary policy surprises). To calculate the bias coefficients, we estimate the univariate IRF of forecast errors using local projections, as in Eq. (10).[23] To be conservative, we show both the 65% and 95% confidence intervals calculated with Newey-West standard errors.

The top panel plots the bias coefficients estimated using consensus forecasts. We observe statistically significant negative bias coefficients for lag 1 ($p = 0.005$),[24] lag 3 ($p = 0.022$), and lag 4 ($p < 0.001$). The evidence suggests that participants in the SPF underreact to information that arrived up to one year ago. The bias coefficients are estimated fairly precisely. That is especially reassuring in light of the fact that we consider quarterly inflation forecasts, and quarterly inflation is rather volatile. In Appendix F, we show that virtually identical results obtain—with somewhat smaller standard errors—if instead of local projections we use maximum likelihood to fit a high-order moving average model (Figure A.3) or an AR(4) model for forecast errors (Figure A.5).

The magnitude of underreaction is substantial. The point estimates indicate that a positive 1σ shock to inflation in the current quarter leads the forecasters to underpredict inflation by roughly 0.30σ four quarters from now.

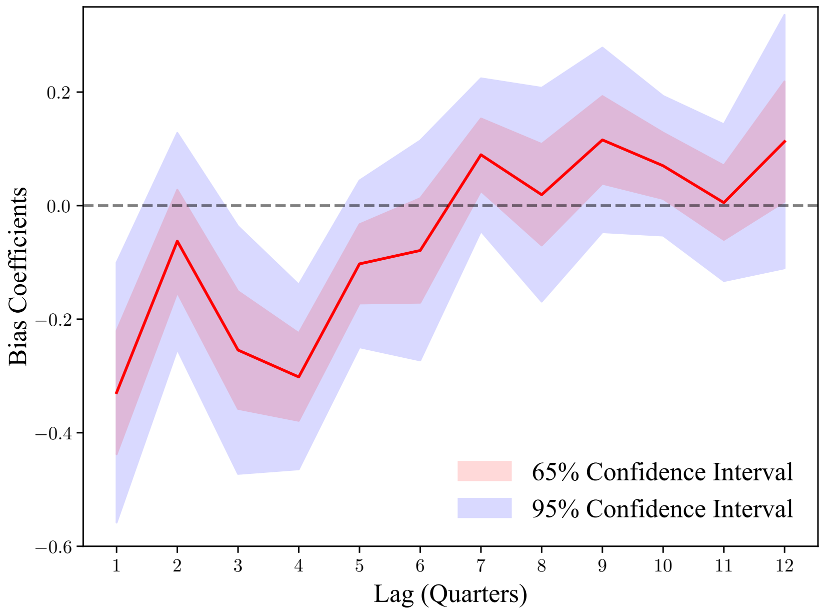
The bottom panel of Figure 5 performs the same exercise using individual forecasts.[25] A nice feature local projections is that it is straightforward to use them with panel data. We estimate a panel-data version of Eq. (10) including forecaster fixed effects. To account for the fact that the respondents are all forecasting the same variable, and the forecast errors may be correlated over time for a given respondent, we cluster the standard errors by both forecaster and quarter. The pattern of the estimated bias coefficients is very similar to that obtained using the consensus forecasts. The key difference is that the coefficients are smaller in absolute value. As a result, the bias coefficient at

---

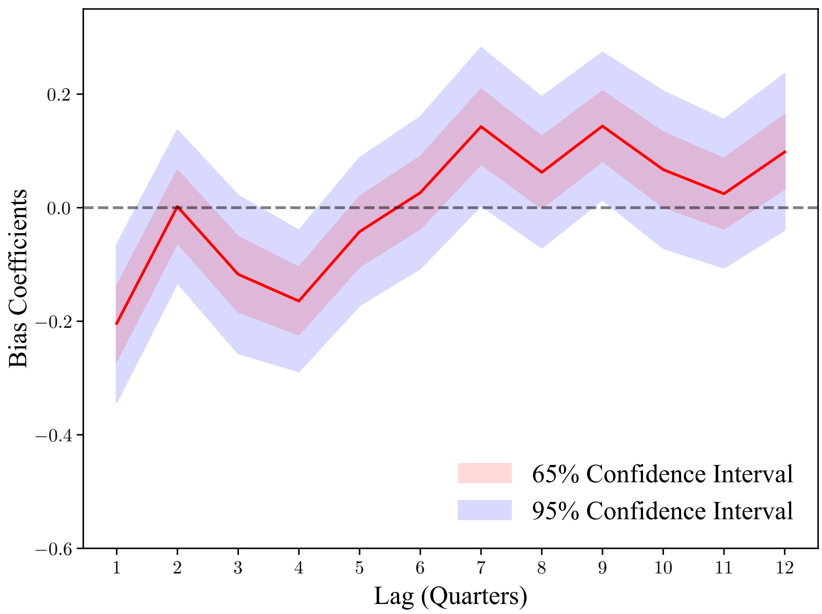
[23] As shown in Appendix F (Figure A.4), the univariate IRF of inflation is positive at all the relevant lags, so that $\text{sgn}(\alpha_\ell) = 1$. Hence, we obtain the bias coefficients by multiplying the IRF by (-1).

[24] At the time of responding to the survey, participants know the advance estimate of inflation in the previous quarter but not inflation in the current quarter (Federal Reserve Bank of Philadelphia, 2017, p. 21). As a result, interpretation of the bias coefficient in the first lag requires some care. On the one hand, forecasters have access to various real-time information on prices. On the other hand, they do not yet know the official number for inflation in the current quarter. In that sense, the one-step ahead forecast may really be a two-step ahead forecast. If that is the case, a non-zero bias coefficient at the first lag should not be interpreted as bias. See Keane and Runkle (1990) for further discussion.

[25] For recent papers that study individual forecasts in the SPF, see Fuhrer (2017), Bordalo, Gennaioli, Ma, and Shleifer (2018a) and Ryngaert (2018).



(a) Inflation: Consensus-Level Estimates



(b) Inflation: Individual-Level Estimates

Figure 5: Composite bias coefficients for one-quarter-ahead inflation forecasts. *Top panel*: Estimates using consensus (median) forecasts; Newey-West standard errors with $\max\{4, \ell - 1\}$ lags are used to calculate the confidence intervals. *Bottom panel*: Estimates using individual-level data (with forecaster fixed effects); standard errors clustered by both forecaster and quarter. Both sets of estimates are obtained by first using local projections (with $K = 4$) to estimate the univariate IRF of forecast errors:

$$x_{t+\ell} - \mathbb{F}_{t+\ell-1}[x_{t+\ell}] = \beta_0^{(\ell)} + \beta_1^{(\ell)}\{x_t - \mathbb{F}_{t-1}[x_{t-1}]\} + \cdots + \beta_4^{(\ell)}\{x_{t-3} - \mathbb{F}_{t-4}[x_{t-3}]\} + u_{t+\ell},$$

for $\ell = 1, 2, \ldots, 12$. The bias coefficients are then estimated by $\hat{b}_\ell = -\hat{\beta}_1^{(\ell)}$.

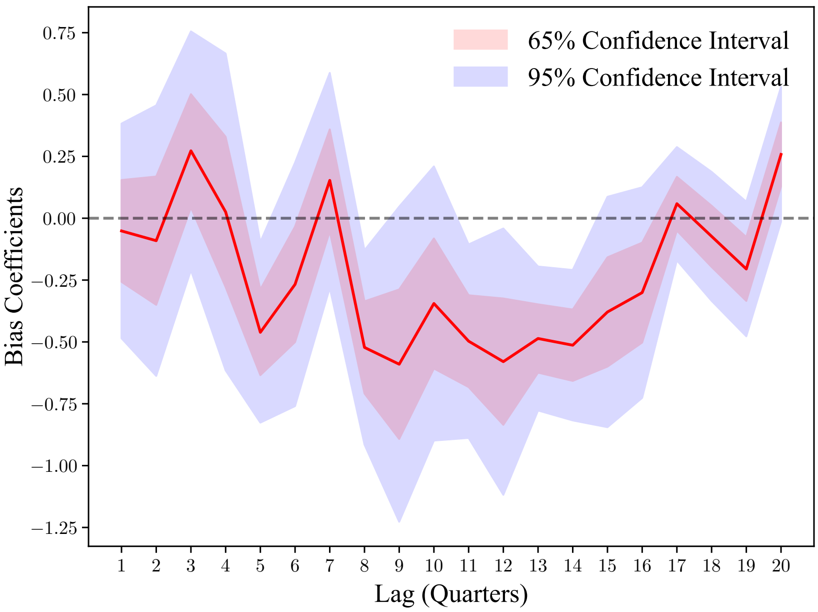


Figure 6: Biases in response to monetary policy shocks. The response of forecast errors to the monetary policy shock is estimated by

$$x_{t+\ell} - \mathbb{F}_{t+\ell-1}[x_{t+\ell}] = \beta_0^{(\ell)} + \sum_{s=1}^{4} \beta_s^{(\ell)} \mathrm{FFR}_{t+1-s} + \sum_{s=1}^{4} \gamma_s^{(\ell)} \{x_{t+1-s} - \mathbb{F}_{t-s}[x_{t-s+1}]\} + u_{t+\ell},$$

for $\ell = 1, 2, \ldots, 20$, instrumenting the Federal Funds Rate, FFR, at time t with the Romer and Romer (2004) measure of a monetary surprise at time t. The bias coefficients are then given by $\hat{b}_\ell = \hat{\beta}_1^{(\ell)}$.

lag 3 is no longer statistically significant at the 5% level ($p = 0.097$).

The fact that bias coefficients estimated using individual forecasts are attenuated towards zero suggests that expectations at the individual level contain some noise. There are two primary explanations for this result. One possibility is measurement error. As discussed in Section C.5, if observed expectations contain some measurement error, the estimated composite bias coefficients will biased towards zero, explaining the observed pattern. Another possibility is that expectations are measured without error but forecasters overreact to idiosyncratic signals. A model of this kind has recently been developed by Bordalo, Gennaioli, Ma, and Shleifer (2018a).[26] As we show in Appendix D.2, their model predicts exactly the same pattern that we find empirically. The standard noisy information model, while more noisy at the individual level, does not predict attenuation because expectations incorporate idiosyncratic signals optimally (see Appendix D.1 and Ryngaert, 2018).

In the Appendix, we also provide estimates of the bias coefficients using individual

[26] Broer and Kohlhas (2018) present a model based on overconfidence which similarly features overreaction at the individual level. See also da Silveira and Woodford (2019).

forecasts that are obtained from a simple AR(4) model, using iteration to obtain the IRF of forecast errors (Figure A.5). The resulting estimates are very similar. The only difference is that for deeper lags, bias coefficients are closer to zero, and the attenuation bias in individual forecasts is more pronounced.

4.2 Monetary Policy Shocks

We now illustrate how our IV procedure can be used to measure biases in how agents react to specific shocks. We focus on the effects of monetary policy on inflation expectations. As our measure of a monetary surprise, we use the Romer and Romer (2004) variable extended to a longer sample period by Wieland and Yang (2017). Since Greenbook forecasts, used to construct the Romer-Romer surprises, are made public with a five-year lag, our sample is reduced somewhat and stops at 2007Q4. Our specification regresses forecast errors on lagged forecast errors and lagged values of the Federal Funds Rate, instrumenting the Federal Funds Rate with the Romer-Romer monetary surprise, as in Eq. (11).[27]

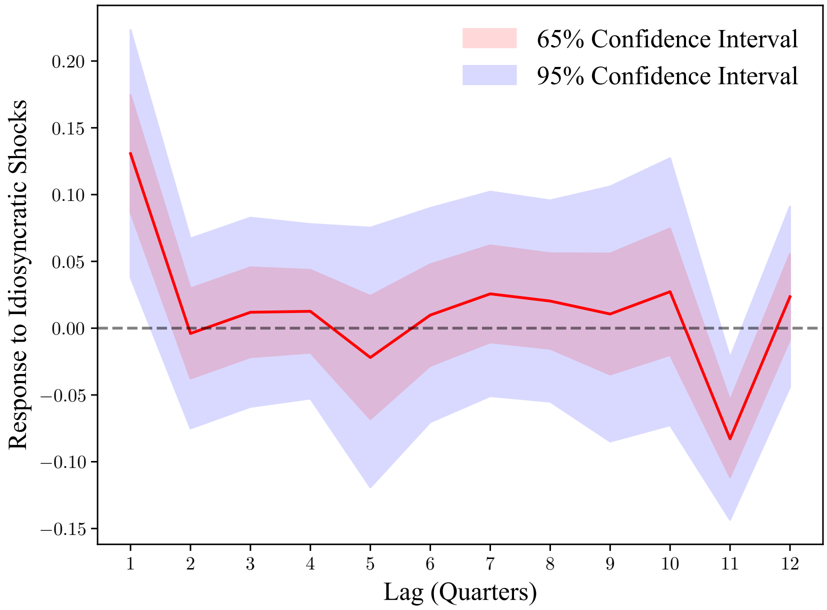
The results are shown in Figure 6. The results indicate statistically significant underreaction to monetary policy shocks that arrived 8 to 16 quarters ago. The timing of underreaction is consistent with the effects of monetary policy shocks on actual inflation. The response of inflation to monetary policy shocks is delayed, with the impact on inflation kicking in around two years after the initial shock (see Figure A.6 in the Appendix and Figures 4 and 5 in Romer and Romer, 2004). The magnitude of underreaction is again substantial. Bias coefficients are equal to roughly half of the actual response of inflation (see Figure A.6 in the Appendix).[28] Overall, these results are consistent with both our findings for the composite bias coefficients above, as well as those of Coibion and Gorodnichenko (2012) who document underreaction to other aggregate shocks, including technology, news, and oil shocks.

4.3 Idiosyncratic Shocks

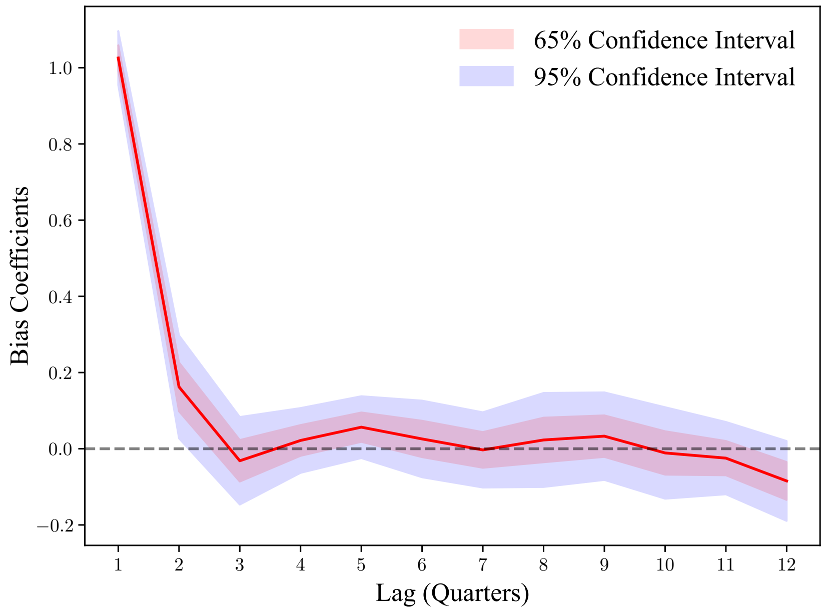
Next, we investigate how agents react to idiosyncratic shocks. By "idiosyncratic," we refer to shocks that are specific to each individual forecaster. These shocks could reflect forecaster-specific information as well as phenomena such as overconfidence. In

[27] Since monetary policy shocks are deflationary, we do not multiply the estimated IRF of forecast errors by (-1).

[28] The effects of Romer-Romer monetary policy surprises on inflation as well as other macro variables are quantitatively high. This fact helps to explain why the magnitude of underreaction that we find is is also substantial. See Coibion (2012) for more discussion on the magnitude of the effects of Romer-Romer monetary policy surprises.



(a) Reaction to Idiosyncratic Shocks



(b) Bias Coefficients w.r.t. Idiosyncratic Shocks

Figure 7: *Top panel*: estimated reaction to idiosyncratic shocks. The same regression specification as in the bottom panel of Figure 5 is used but with time fixed effects included. *Bottom panel*: Estimated bias coefficients with respect to idiosyncratic shocks. The estimates are obtained from

$$x_{t+\ell} - \mathbb{F}_{i,t+\ell-1}[x_{t+\ell}] = \mu_i^{(\ell)} + \sum_{s=1}^{4} \beta_s^{(\ell)}\, \mathbb{F}_{i,t+1-s}[x_{t+2-s}] + \sum_{s=1}^{4} \gamma_s^{(\ell)} \{x_{t+1-s} - \mathbb{F}_{i,t-s}[x_{t-s+1}]\} + u_{i,t+\ell},$$

for $\ell = 1, 2, \ldots, 12$, instrumenting $\mathbb{F}_{it}[x_{t+1}]$ with the deviation from consensus, $\mathbb{F}_{it}[x_{t+1}] - 1/N \sum_{j=1}^{N} \mathbb{F}_{jt}[x_{t+1}]$; standard errors clustered by both forecaster and quarter. The estimated bias coefficients are then $\hat{b}_\ell = -\hat{\beta}_1^{(\ell)}$.

Appendix C.4, we show that the way expectations react to idiosyncratic shocks can be identified from the IRF of deviations of individual forecasts from the consensus forecast. It is straightforward to obtain this IRF by estimating the univariate IRF of forecast errors, just as in the case of composite bias coefficients, but now including time fixed effects. The key identifying assumption for the procedure to be valid is that forecasters react to aggregate shocks homogeneously.

The results of this exercise are shown in the top panel of Figure 7. Forecasters indeed react to idiosyncratic shocks. However, these reactions are fairly small and short lived, lasting up to one quarter. Hence, underreaction exhibited by the composite bias coefficients stems from underreaction to aggregate shocks.

This empirical result is inconsistent with the basic noisy information model (see Appendix D.1). The basic noisy information model predicts exactly the same behavior for deviations from consensus as that for forecast errors. Our empirical results, in contrast, indicate that deviations from the consensus forecast are eliminated much more rapidly than are forecast errors. In the Appendix (Figure A.8), we show that very similar results obtain if instead of local projections, we use an AR(4) model with individual and time fixed effects.

A natural next question is whether reacting to idiosyncratic shocks constitutes under- or overreaction. In the benchmark noisy information model, idiosyncratic shocks are pure noise, and hence—according to our definition—such behavior implies overreaction.[29] In this class of models, the estimates in the top panel of Figure 7 can be directly interpreted as measures of overreaction to idiosyncratic shocks.

A potential worry with such an interpretation is that inflation may be correlated with the idiosyncratic shocks.[30] To address this concern, one would ideally have an IV for the idiosyncratic shocks. For example, Lagerborg (2017) uses local school shootings as an IV for consumer confidence; similar ideas could be employed in our context. However, given that information about individual participants in the SPF is very limited, we do not have an ideal IV. Hence, we use deviations from consensus as instrument, which as we show in the Appendix, identify idiosyncratic shocks in an infinitely large sample. The instrument is not ideal because in any finite sample, it is contaminated by the idiosyncratic shocks hitting all other agents. In addition, for any forecast that is above the consensus, there must be a forecast that lies below the consensus. Hence,

[29] In the benchmark noisy information model, agents observe a signal $y_{it} = x_t + \omega_{it}$ where ω_{it} is idiosyncratic noise which is orthogonal to x_t, the variable being forecast.

[30] As a concrete (although stylized) example, suppose that there are N forecasters, and actual inflation is equal to $x_{t+1} = \rho \sum_{i=1}^{N} \omega_{it}$, with each forecaster i only observing an idiosyncratic shock ω_{it}; the idiosyncratic shocks are i.i.d. across agents but potentially correlated across time. If forecasters use Bayesian updating to form expectations, their forecasts are given by $\mathbb{F}_{it}[x_{t+1}] = \rho\omega_{it}$. These expectations exhibit unbiased reaction to ω_{it}, despite reacting to idiosyncratic shocks.

on average, deviations from the consensus forecast are mechanically not predictive of actual inflation. Note, however, that in the special case in which the idiosyncratic shocks are pure noise (as in the baseline noisy information model), the IV procedure is still valid.

With these caveats in mind, we employ our instrumental variables procedure and regress forecast errors on lagged forecast errors as well as lagged forecasts, instrumenting the forecast at time t with the deviation from the consensus forecast at time t. The findings are shown in the bottom panel of Figure 7 and indicate overreaction to idiosyncratic shocks that lasts for around one quarter. The first bias coefficient is very close to one. The IV procedure, by construction, ensures that the idiosyncratic shock is such that the forecast at time t is higher by one unit. The fact that the estimated bias coefficient is close to one means that the increase in the forecast translates one-to-one into a larger forecast error. The second bias coefficient is also positive and statistically significant, implying that overreaction to idiosyncratic shocks is persistent. The magnitude of the bias coefficient is very close to the reaction found in the top panel of Figure 7, again consistent with idiosyncratic shocks translating into forecast errors one-for-one. However, the amount of overreaction is quantitatively not very large. This latter finding explains why we find that underreaction is the dominant feature of expectations when looking at the composite bias coefficients.[31]

4.4 State and Time Dependence

Our method can also be used to investigate whether biases in expectation formation vary across time and state of the economy. We generalize Eq. (10) to

$$e_{t+s} = \alpha_0 + \alpha_1 I_t + (1 - I_t) \sum_{i=1}^{K} \beta_{i,0}^{(s)} e_{t-i} + I_t \sum_{i=1}^{K} \beta_{i,1}^{(s)} e_{t-i} + u_{t+s},$$

where $I_t \in \{0, 1\}$ is a dummy variable equal to 1 if the economy is in a particular state at time t (such as a high-inflation state), and $s = 1, 2, \ldots, L$. Following Ramey and Zubairy (2017), we allow all coefficients in the regression to differ across the two states.

Table 2 collects the findings for a number of different states of the economy. The estimated bias coefficients indeed exhibit state dependence. Expectations display more underreaction in high-inflation periods, with a difference that is statistically significant. In addition, there is some evidence that expectations show less underreaction during

[31] We cannot rule out that the estimated positive bias coefficients with respect to idiosyncratic shocks are driven by autocorrelated measurement error. However, it is not clear why measurement error would have exactly this specific autocorrelation structure.

	Baseline	High Inflation	Recession	Great Moderation	Large FE
Lag 1	-0.33	-0.54	-0.16	-0.30	-0.27
	(0.12)	(0.14)	(0.29)	(0.09)	(0.15)
		[-2.27]	[0.80]	[0.15]	[0.50]
Lag 2	-0.06	-0.06	-0.19	0.10	-0.07
	(0.10)	(0.13)	(0.15)	(0.09)	(0.15)
		[-0.42]	[-0.87]	[1.46]	[-0.18]
Lag 3	-0.25	-0.32	-0.11	-0.15	-0.41
	(0.11)	(0.22)	(0.16)	(0.12)	(0.15)
		[-0.70]	[1.28]	[0.76]	[-1.90]
Lag 4	-0.30	-0.48	0.09	-0.19	-0.47
	(0.08)	(0.13)	(0.14)	(0.10)	(0.13)
		[-2.28]	[3.12]	[1.03]	[-1.12]

Table 2: State dependence in the composite bias coefficients. Standard errors in parentheses; t-statistics of a test of no state dependence in brackets. High absolute values of the t-statistics indicate evidence of state dependence with negative t-statistics indicating more underreaction. The same regression specification is used as in Figure 5 but with all coefficients allowed to differ across the states.

High Inflation: Realized inflation above the 75$^{\text{th}}$ percentile of sample values. *Recession*: The economy is in a recession as given by the NBER recession indicator. *Great Moderation*: Dummy variable that is equal to 1 between 1985-01-01 and 2006-12-31. *Large FE*: Forecast error above the 75$^{\text{th}}$ percentile of sample values.

recessions. However, state dependence in the estimated bias coefficients is primarily driven by state dependence in the behavior of actual inflation. Table A.1 in the Appendix shows that exactly the same qualitative pattern of state dependence is observed in actual inflation. Hence, the results from this section further corroborate the previous evidence of underreaction to aggregate shocks, suggesting that forecasters are slow to react to state dependence in actual inflation.

We can also perform the same exercise for time variation in biases. To do so, we estimate bias coefficients using a rolling-window regression. The results are given in Figure 8 which plots the first bias coefficient over time. Recall that the first bias coefficient measures how expectations react to current news. There is substantial time variation in the estimated bias coefficient. Once again, however, this time variation is driven by time variation in the persistence of actual inflation, as shown in the Appendix (Figure A.7).

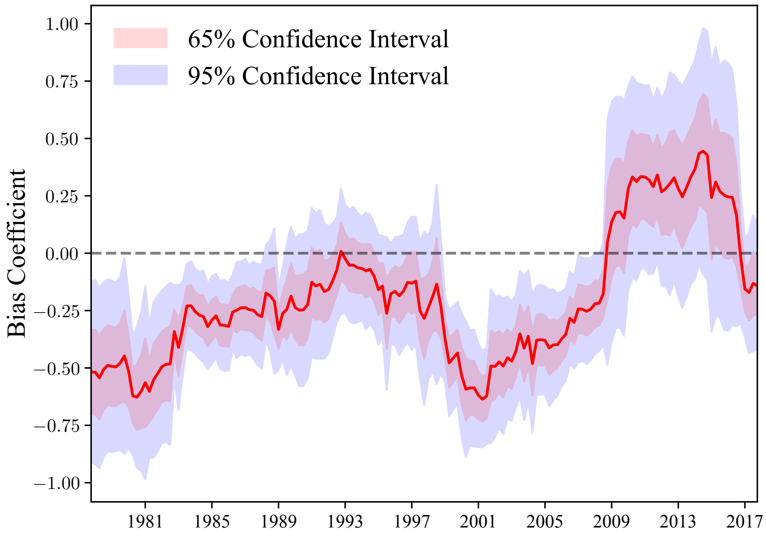


Figure 8: Rolling window estimates of the first composite bias coefficient, $\hat{b}_1$. A window of 32 quarters (8 years) is used for estimation. The regression specification is the same as in Figure 5.

4.5 Calibration Exercise

The estimated bias coefficients can be used to guide theory. To illustrate this point, we perform a simple calibration exercise. For a number of models of expectations, we choose their parameters to fit the estimated bias coefficients as closely as possible. Specifically, for each model, we choose its parameters to minimize the sum of squared deviations of the empirically estimated bias coefficients from the theoretically predicted bias coefficients. To obtain theoretical predictions, we assume that the true inflation process is an AR(1). Estimating the persistence parameter using least squares yields an estimate of $\hat{\rho} = 0.83$ (see Table 1).

The results from this exercise are shown in Table 3. Since the models all have a single parameter (except for rational expectations which have no free parameters), we do not adjust for model complexity. The model that best fits the data is a simple model in which forecasters think that the true persistence of inflation is smaller than it actually is (see Section D.3). The perceived level of persistence that provides the best fit is 0.61. This number is roughly 25% lower than the estimated persistence of inflation. As discussed by Gabaix (2017a, pp. 14–15), limited attention can naturally lead to such misperception.

The sticky information model also does well, with only a slightly worse fit than the misperception model. The estimated information stickiness parameter for this model is $\hat{\lambda} = 0.51$. The estimate is very close to that reported by Coibion and Gorodnichenko

Model	Parameter	SSR
Misperception	0.61	0.16
Sticky	0.51	0.18
Adjustment Cost	0.43	0.25
Rational	NA	0.32
Adaptive	0.17	0.36
Extrapolative	-0.31	0.55

Table 3: Calibration exercise for matching the empirically estimated bias coefficients. The sum of squared residuals (SSR) is calculated as $\sum_{\ell=1}^{12}[\hat{b}_\ell - b_\ell(\theta^*)]^2$ where $\hat{b}_\ell$ is the empirically estimated bias coefficient, and $b_\ell(\theta^*)$ is the theoretically predicted bias coefficient; θ^* denotes the parameter value that minimizes the sum of squared residuals. The estimated process for inflation is $x_t = 0.83x_{t-1} + \varepsilon_t$. The precise descriptions of the models are given in Section 3 and Appendix D.

(2015) who find $\hat{\lambda} = 0.54$ (s.e. $= 0.10$).[32] While Coibion and Gorodnichenko (2015) also use the SPF data, they consider one-year ahead forecasts, and their methodology estimates the level of stickiness by regressing forecast errors on past forecast revisions. The fact that we get a very similar number using a completely different methodology is reassuring. In the context of the sticky information model, our estimate implies that forecasters in the SPF update their information sets roughly twice a year on average. The noisy information model is observationally equivalent at the level of consensus forecasts, as shown in Appendix D.1. Hence, the noisy information model can match the estimated bias coefficients equally well as sticky information.[33]

The model with adjustment costs performs better than simple rational expectations. Finally, mechanical adaptive and extrapolative expectations perform worse than rational expectations. The finding is consistent with previous research that has documented that participants in the SPF are quite accurate.[34] It is therefore not surprising that their behavior is not very well described by mechanical models of expectations. The fact that extrapolative expectations perform especially poorly is interesting in light of the fact that extrapolative models explain inflation expectations well in laboratory experiments (e.g., Pfajfar and Žakelj, 2014).

[32] We use the Delta Method to calculate the standard error for $\hat{\lambda}$ from the estimates provided by Coibion and Gorodnichenko, i.e., s.e.$(\hat{\lambda}) = $ s.e.$(\hat{\beta})/(1+\hat{\beta})^2$.

[33] Bordalo, Gennaioli, Ma, and Shleifer (2018a) develop a model that introduces noisy information into the model of diagnostic expectations. Since their model nests the basic noisy information model (when the agent does not exhibit representativeness), this model does at least as well as the noisy information model in fitting the empirical bias coefficients. The model has an additional parameter, and hence comparing the fit requires adjusting for model complexity.

[34] For example, Croushore (2010) documents that the median forecast in the SPF performs better than simple time series models of inflation (Table 5), and that it is difficult to adjust forecasts for biases observed in the past to obtain higher forecasting accuracy in real time (Table 4). See also Ang, Bekaert, and Wei (2007).

5 Conclusions

Expectations play a key role in economics. However, views diverge on how expectations are formed. The divergence stems in part from the lack of a general framework for measuring biases in expectation formation. Definitions of theoretical concepts—such as under- and overreaction—are not always made explicit in applied work. As a result, interpreting the results of existing empirical tests can be difficult. Theory does not necessarily provide clear guidance either as there are many competing models of expectations.

This paper attempts to provide a general measurement framework that fills this gap. The key insight is that under- and overreaction to new information is identified by the IRF of forecast errors. The starting point is a definition of under- and overreaction that is natural and consistent with the existing theory. We then show that these biases can be measured using a simple and widely applicable measurement procedure. The measurement procedure boils down to estimating the IRF of forecast errors. The estimated bias coefficients directly measure under- and overreaction to news at various lags, characterizing the whole term structure of biases.

To be sure, our measurement procedure is not a silver bullet. One issue is that the notion of bias used in this paper is a statistical one. There may be multiple reasons for why expectations are biased according to this definition, including psychological as well as non-psychological ones. Distinguishing between competing explanations requires additional structure. A related challenge is that different models of expectations can yield similar predictions for bias coefficients. Predictions on additional moments of the data are then necessary to discriminate between competing explanations. Finally, identifying the IRF of forecast errors with respect to specific shocks requires valid and strong instruments.

In our empirical application we found evidence of forecasters underreacting to *aggregate* shocks but overreacting to *idiosyncratic* shocks. The fact that we find *both* under- and overreaction highlights the importance of a measurement framework that can accommodate both types of biases. It remains to be seen whether our empirical findings generalize to other contexts. Our own prior is that biases are likely to be context dependent. Hence, we would expect to see overreaction to aggregate shocks in other, especially financial market, contexts. The current framework may prove helpful for researchers investigating these issues.

References

Abarbanell, J. S., Bernard, V. L., 1992. Tests of analysts' overreaction/underreaction to earnings information as an explanation for anomalous stock price behavior. The Journal of Finance 47 (3), 1181–1207.

Andolfatto, D., Hendry, S., Moran, K., 2008. Are inflation expectations rational? Journal of Monetary Economics 55 (2), 406–422.

Andreassen, P. B., 1987. On the social psychology of the stock market: Aggregate attributional effects and the regressiveness of prediction. Journal of Personality and Social Psychology 53 (3), 490–496.

Ang, A., Bekaert, G., Wei, M., 2007. Do macro variables, asset markets, or surveys forecast inflation better? Journal of Monetary Economics 54 (4), 1163–1212.

Angeletos, G.-M., Lian, C., 2016. Incomplete information in macroeconomics. In: Taylor, J. B., Uhlig, H. (Eds.), Handbook of Macroeconomics. Vol. 2. pp. 1065–1240.

Ansley, C. F., Spivey, W. A., Wrobleski, W. J., 1977. On the structure of moving average processes. Journal of Econometrics 6 (1), 121–134.

Assenza, T., Bao, T., Hommes, C., Massaro, D., 2014. Experiments on expectations in macroeconomics and finance. In: Duffy, J. (Ed.), Experiments in Macroeconomics (Research in Experimental Economics). Vol. 17. Emerald Group Publishing Limited, Bingley, pp. 11–70.

Augenblick, N., Lazarus, E., 2018. Restrictions on asset-price movements under rational expectations: Theory and evidence. Working paper.

Augenblick, N., Rabin, M., 2018. Belief movement, uncertainty reduction, and rational updating. Working paper, 1–52.

Barberis, N., Greenwood, R., Jin, L., Shleifer, A., 2015. X-CAPM: An extrapolative capital asset pricing model. Journal of Financial Economics 115 (1), 1–24.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. Journal of Financial Economics 49 (3), 307–343.

Barrero, J. M., 2018. The micro and macro of managerial beliefs. Working paper.

Ben-David, I., Graham, J. R., Harvey, C. R., 2013. Managerial miscalibration. The Quarterly Journal of Economics 128 (4), 1547–1584.

Benjamin, D. J., 2018. Errors in probabilistic reasoning and judgment biases. Working paper.

Bordalo, P., Gennaioli, N., La Porta, R., Shleifer, A., 2017. Diagnostic expectations and stock returns. Working paper.

Bordalo, P., Gennaioli, N., Ma, Y., Shleifer, A., 2018a. Overreaction in macroeconomic expectations. Working paper.

Bordalo, P., Gennaioli, N., Shleifer, A., 2018b. Diagnostic expectations and credit cycles. The Journal of Finance 73 (1), 199–227.

Bouchaud, J.-P., Krüger, P., Landier, A., Thesmar, D., 2018. Sticky expectations and the profitability anomaly. The Journal of Finance, forthcoming.

Branch, W. A., Evans, G. W., 2006. A simple recursive forecasting model. Economics Letters 91 (2), 158–166.

Broer, T., Kohlhas, A. N., 2018. Forecaster (mis-)behavior. Working paper.

Cagan, P., 1956. The monetary dynamics of hyperinflation. In: Friedman, M. (Ed.), Studies in the Quantity Theory of Money. University of Chicago Press, Chicago, pp. 25–117.

Capistrán, C., Timmerman, A., 2009. Disagreement and biases in inflation expectations. Journal of Money, Credit and Banking 41 (2), 365–396.
Carroll, C. D., 2003. Macroeconomic expectations of households and professional forecasters. The Quarterly Journal of Economics 118 (1), 269–298.
Cascaldi-Garcia, D., 2019. Forecast revisions as instruments for news shocks. Working paper.
Chapman, L. J., Chapman, J., 1982. Test results are what you think they are. In: Kahneman, D., Slovic, P., Tversky, A. (Eds.), Judgment under Uncertainty. Cambridge University Press, Cambridge, Ch. 17.
Cochrane, J., 2005. Asset Pricing: Revised Edition. Princeton University Press, Princeton.
Coibion, O., 2012. Are the effects of monetary policy shocks big or small? American Economic Journal: Macroeconomics 4 (2), 1–32.
Coibion, O., Gorodnichenko, Y., 2012. What can survey forecasts tell us about information rigidities? Journal of Political Economy 120 (1), 116–159.
Coibion, O., Gorodnichenko, Y., 2015. Information rigidity and the expectations formation process: A simple framework and new facts. The American Economic Review 105 (8), 2644–2678.
Coibion, O., Gorodnichenko, Y., Kamdar, R., 2018. The formation of expectations, inflation and the Phillips Curve. Journal of Economic Literature, forthcoming.
Croushore, D., 2010. An evaluation of inflation forecasts from surveys using real-time data. The B.E. Journal of Macroeconomics 10 (1).
Cutler, D. M., Poterba, J. M., Summers, L. H., 1990. Speculative dynamics and the role of feedback traders. The American Economic Review: Papers & Proceedings 80 (2), 63–68.
da Silveira, R. A., Woodford, M., 2019. Noisy memory and over-reaction to news. Working paper.
Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and overreactions. The Journal of Finance 53 (6), 1839–1885.
De Bondt, W. F. M., Thaler, R. H., 1990. Do security analysts overreact? The American Economic Review: Papers & Proceedings 80 (2), 52–57.
DeLong, B. J., Shleifer, A., Summers, L. H., Waldmann, R. J., 1990. Positive feedback investment strategies and destabilizing rational speculation. The Journal of Finance 45 (2), 379–395.
Diebold, F. X., Lopez, J. A., 1996. Forecast evaluation and combination. In: Maddala, G., Rao, C. (Eds.), Handbook of Statistics (Statistical Methods in Finance). Vol. 14. Elsevier, Amsterdam, Ch. 8, pp. 241–268.
Dominitz, J., Manski, C. F., 2011. Measuring and interpreting expectations of equity returns. Journal of Applied Econometrics 26, 352–370.
Drerup, T., Enke, B., von Gaudecker, H. M., 2017. The precision of subjective data and the explanatory power of economic models. Journal of Econometrics 200 (2), 378–389.
Edwards, W., 1968. Conservatism in human information processing. In: Kleinmuntz, B. (Ed.), Formal Representation of Human Judgement. John Wiley and Sons, New York.
Evans, G. W., Honkapohja, S., 2001. Learning and Expectations in Macroeconomics. Princeton University Press, Princeton.
Evans, G. W., Honkapohja, S., 2009. Learning and macroeconomics. Annual Review of Economics 1, 421–449.
Federal Reserve Bank of Philadelphia, 2017. Survey of Professional Forecasters: Docu-

mentation. Link to website.
Fuhrer, J., 2017. Intrinsic expectations persistence: Evidence from professional and household survey expectations. Working paper.
Fuster, A., Laibson, D., Mendel, B., 2010. Natural expectations and macroeconomic fluctuations. Journal of Economic Perspectives 24 (4), 67–84.
Gabaix, X., 2014. A sparsity-based model of bounded rationality. The Quarterly Journal of Economics, 1661–1710.
Gabaix, X., 2017a. Behavioral inattention. Working paper.
Gabaix, X., 2017b. Behavioral macroeconomics via sparse dynamic programming. Working paper.
Gabaix, X., 2018. A behavioral New Keynesian model. Working paper.
Gennaioli, N., Shleifer, A., 2010. What comes to mind. The Quarterly Journal of Economics 125 (4), 1399–1433.
Gennaioli, N., Shleifer, A., 2018. A Crisis of Beliefs: Investor Psychology and Financial Fragility. Princeton University Press, Princeton.
Giglio, S., Maggiori, M., Stroebel, J., Utkus, S., 2019. Five facts about beliefs and portfolios. Working paper.
Gilovich, T., Vallone, R., Tversky, A., 1985. The hot hand in basketball: On the misperception of random sequences. Cognitive Psychology 17, 295–314.
Goodwin, R. M., 1947. Dynamical coupling with especial reference to markets having production lags. Econometrica 15 (3), 181–204.
Graham, J. R., Harvey, C. R., 2001. Expectations of equity risk premia, volatility and asymmetry from a corporate finance perspective. Working paper.
Greenwood, R., Shleifer, A., 2014. Expectations of returns and expected returns. Review of Financial Studies 27 (3), 714–746.
Griffin, D., Tversky, A., 1992. The weighing of evidence and the determinants of confidence. Cognitive Psychology 24, 411–435.
Hamilton, J. D., 1994. Time Series Analysis. Princeton University Press, Princeton.
Hoff, P. D., 2009. A First Course in Bayesian Statistical Methods. Springer.
Huo, Z., Pedroni, M., 2018. A single-judge solution to beauty contests. Working paper.
Jeong, J., Maddala, G. S., 1991. Measurement errors and tests for rationality. Journal of Business and Economic Statistics 9 (4), 431–439.
Jordá, Ò., 2005. Estimation and inference of impulse responses by local projections. The American Economic Review 95 (1), 161–182.
Kahneman, D., Tversky, A., 1973. On the psychology of prediction. Psychological Review 80 (4), 237–251.
Keane, M. P., Runkle, D. E., 1990. Testing the rationality of price forecasts: New evidence from panel data. The American Economic Review 80 (4), 714–735.
Kohlhas, A. N., Walther, A., 2018. Asymmetric attention. Working paper.
Koop, G., Pesaran, M., Potter, S. M., 1996. Impulse response analysis in nonlinear multivariate models. Journal of Econometrics 74 (1), 119–147.
Lagerborg, A., 2017. Confidence and local activity: An iv approach. Working paper.
Langer, E. J., 1975. The illusion of control. Journal of Personality and Social Psychology 32, 311–328.
Lucas, R. E., 1973. Some international evidence on output-inflation tradeofs. American Economic Review 63 (3), 326–334.

Ma, Y., Landier, A., Thesmar, D., 2017. New experimental evidence on expectations formation. Working Paper.

Ma, Y., Sraer, D. A., Thesmar, D., 2018. The aggregate cost of systematic forecast errors. Working paper.

MacKinnon, J. G., Smith, A. A., 1998. Approximate bias correction in econometrics. Journal of Econometrics 85 (2), 205–230.

Maćkowiak, B., Wiederholt, M., 2009. Optimal sticky prices under rational inattention. The American Economic Review 99 (3), 769–803.

Malmendier, U., Nagel, S., 2016. Learning from inflation experiences. The Quarterly Journal of Economics 131 (1), 53–87.

Mankiw, G. N., Reis, R., 2002. Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips Curve. The Quarterly Journal of Economics 117 (4), 1295–1328.

Manski, C. F., 2004. Measuring expectations. Econometrica 72 (5), 1329–1376.

Manski, C. F., 2018. Survey measurement of probabilistic macroeconomic expectations: Progress and promise. NBER Macroeconomics Annual 32, forthcoming.

Marinovic, I., Ottaviani, M., Sorensen, P., 2013. Forecasters' objectives and strategies. In: Handbook of Economic Forecasting. Vol. 2. Elsevier B.V., Ch. 12, pp. 689–720.

Mertens, K., Ravn, M. O., 2013. The dynamic effects of personal and corporate income tax changes in the United States. American Economic Review 103 (4), 1212–1247.

Miller, J. B., Sanjurjo, A., 2018. Surprised by the hot hand fallacy? A truth in the law of small numbers. Econometrica, forthcoming.

Mincer, J. A., Zarnowitz, V., 1969. The evaluation of economic forecasts. In: Mincer, J. A. (Ed.), Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance. National Bureau of Economic Research, Ch. 1, pp. 3–46.

Muth, J. F., 1960. Optimal properties of exponentially weighted forecasts. Journal of the American Statistical Association 55 (290), 299–306.

Muth, J. F., 1961. Rational expectations and the theory of price movements. Econometrica 29 (3), 315–335.

Nerlove, M., 1958. Adaptive expectations and cobweb phenomena. The Quarterly Journal of Economics 72 (2), 227–240.

Nickerson, R. S., 1998. Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology 2 (2), 175–220.

Nisbett, R., Ross, L., 1980. Human Inference: Strategies and Shortcomings of Social Judgement. Prentice Hall, Englewood Cliffs.

Odean, T., 1998. Volume, volatility, price, and profit when all traders are above average. The Journal of Finance 53 (6), 1887–1934.

Orphanides, A., Williams, J. C., 2005. Imperfect knowledge, inflation expectations, and monetary policy. In: Bernanke, B. S., Woodford, M. (Eds.), The Inflation-Targeting Debate. University of Chicago Press, Chicago, pp. 201–246.

Ottaviani, M., Norman, P. S., 2006. The strategy of professional forecasting. Journal of Financial Economics 81, 441–466.

Patton, A. J., Timmermann, A., 2007. Properties of optimal forecasts under asymmetric loss and nonlinearity. Journal of Econometrics 140 (2), 884–918.

Pesaran, M. H., Weale, M., 2006. Survey expectations. In: Elliott, G., Granger, C. W. J., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Vol. 1. Elsevier, Amster-

dam, Ch. 14, pp. 715–776.
Pfajfar, D., Žakelj, B., 2014. Experimental evidence on inflation expectation formation. Journal of Economic Dynamics & Control 44, 147–168.
Plagborg-Møller, M., Wolf, C. K., 2018a. Instrumental variable identification of dynamic variance decompositions. Working paper.
Plagborg-Møller, M., Wolf, C. K., 2018b. Local projections and VARs estimate the same impulse responses. Working paper.
Ramey, V., Zubairy, S., 2017. Government spending multipliers in good times and in bad: Evidence from U.S. historical data. Journal of Political Economy 126 (2), 696277.
Reis, R., 2006. Inattentive producers. Review of Economic Studies 73 (3), 793–821.
Romer, C. D., Romer, D. H., 2004. A new measure of monetary shocks: Derivation and implications. The American Economic Review 94 (4), 1055–1084.
Ryngaert, J., 2018. What do (and don't) forecasters know about U.S. inflation? Working paper.
Schmalensee, R., 1976. An experimental study of expectation formation. Econometrica 44 (1), 17–41.
Sims, C. A., 2003. Implications of rational inattention. Journal of Monetary Economics 50 (3), 665–690.
Stock, J. H., 2008. Recent developments in structural VAR modeling. In: NBER Summer Institute Minicourse 2008: What's New in Econometrics—Time Series. National Institute for Economic Research, www.nber.org/minicourse_2008.html, Cambridge, MA.
Stock, J. H., Watson, M. W., 2018. Identification and estimation of dynamic causal effects in macroeconomics using external instruments. Economic Journal 128 (610), 917–948.
Vissing-Jorgensen, A., 2004. Perspectives on behavioral finance: Does 'irrationality' disappear with wealth? Evidence from expectations and actions. NBER Macroeconomics Annual 18, 139–208.
Wieland, J. F., Yang, M.-J., 2017. Financial dampening. Working paper.
Woodford, M., 2003. Imperfect common knowledge and the effects of monetary policy. In: Aghion, P., Frydman, R., Stiglitz, J., Woodford, M. (Eds.), Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps. Princeton University Press, Princeton, pp. 25–58.

Appendix A Proofs

Proof of Proposition 1

For necessity, suppose that some composite bias cofficients are not zero but—to obtain a contradiction—all shock-specific bias coefficients are zero. In that case, Eq. (6) implies that $x_t - \mathbb{F}_{t-1}[x_t] = \sum_{i=1}^{M} \alpha_{i0}\varepsilon_{it}$. But then all composite bias coefficients are zero, a contradiction. To see that zero composite bias coefficients are not sufficient for the expectation formation process to be unbiased, suppose that $x_t = 0$ with probability one but $\mathbb{F}_t[x_t] = \xi_t$ where ξ_t is a non-degenerate i.i.d. shock. Then, the agent overreacts to ξ_t at the individual level, but the composite bias coefficients are all zero since ξ_t is an i.i.d. shock.

To see that Eq. (8) holds, note that

$$
\begin{aligned}
\gamma_0 &= \mathbb{V}\mathrm{ar}[e_t] = \mathbb{V}\mathrm{ar}[\nu_t] \sum_{\ell=0}^{+\infty} \theta_\ell^2 \\
\gamma_k &= \mathbb{C}\mathrm{ov}[e_t, e_{t-k}] = \mathbb{V}\mathrm{ar}[\nu_t] \sum_{\ell=0}^{+\infty} \theta_\ell \theta_{\ell+k}
\end{aligned}
$$

Dividing the second equation by the first and rearranging yields Eq. (8). The explicit expressions for γ_k in terms of shock-specific bias coefficients are

$$
\begin{aligned}
\gamma_0 &= \sum_{i=1}^{M} \mathbb{V}\mathrm{ar}[\varepsilon_{it}] \left(\alpha_{i0}^2 + \sum_{\ell=1}^{+\infty} b_{i\ell}^2 \right) \\
\gamma_k &= \sum_{i=1}^{M} \mathbb{V}\mathrm{ar}[\varepsilon_{it}] \Bigg(-\alpha_{i0}\,\mathrm{sgn}(\alpha_{ik}) b_{ik} \\
&\qquad + \sum_{\ell=1}^{+\infty} \mathrm{sgn}(\alpha_{i\ell})\,\mathrm{sgn}(\alpha_{i,\ell+k}) b_{i\ell} b_{i,\ell+k} \Bigg) \text{ for } k \geq 1.
\end{aligned}
\tag{A.1}
$$

Proof of Proposition 2

Calculate that

$$
\begin{aligned}
\mathbb{E}[z_t e_{t+\ell}] &= -\,\mathrm{sgn}(\alpha_{1\ell}) b_{1\ell} \phi \text{ for } \ell \geq 1 \\
\mathbb{E}[z_t w_t] &= \underbrace{\gamma_{10}\phi = \phi}_{\text{normalization } \gamma_{10}=1}
\end{aligned}
$$

As a result, since $\phi \neq 0$ by Assumption 1, we have that

$$\frac{\mathbb{E}[z_t e_{t+\ell}]}{\mathbb{E}[z_t w_t]} = -\operatorname{sgn}(\alpha_{1\ell}) b_{1\ell}.$$

But the instrumental-variables estimator of $\beta^{(\ell)}$ is given by

$$\hat{\beta}^{(\ell)}_{IV} = \left(\frac{1}{T-\ell} \sum_{t=1}^{T-\ell} z_t w_t \right)^{-1} \left(\frac{1}{T-\ell} \sum_{t=1}^{T-\ell} z_t e_{t+\ell} \right) \xrightarrow{p} \frac{\mathbb{E}[z_t e_{t+\ell}]}{\mathbb{E}[z_t w_t]}$$

under standard regularity conditions under which a Law of Large Number applies to the terms in the parentheses. Since $\operatorname{sgn}(x)^2 = 1$, the result follows.

Appendix B Other Reduced-Form Approaches

Consider the standard Mincer and Zarnowitz (1969) regression which estimates

$$x_t = \alpha + \beta \mathbb{F}_{t-1}[x_t] + u_t,$$

and then tests the joint hypothesis that $\alpha = 0$ and $\beta = 1$.[35] The test for $\beta = 1$ can be intuitively thought of as a test of under- and overreaction. If the agent overreacts to new information, we expect the forecasts to be too extreme. As a result, one may expect to see $\beta < 1$ with overreaction, and vice versa for underreaction (see, for example, De Bondt and Thaler, 1990).

We can use our framework to check whether this intuition is in fact correct. Given that plim $\hat{\beta} = \mathbb{C}\text{ov}(x_t, \mathbb{F}_{t-1}[x_t]) / \mathbb{V}\text{ar}(\mathbb{F}_{t-1}[x_t])$, we find

$$\text{plim}\, \hat{\beta} \leq 1 \quad \Leftrightarrow \quad \sum_{\ell=1}^{+\infty} a_\ell \, \text{sgn}(\alpha_\ell) b_\ell \geq 0. \tag{A.2}$$

Hence, the basic intuition is only partially justified. The approach suffers from two problems, both stemming from the fact that the regression cannot disentangle how the agent reacts to information at different lags. First, even if expectations are biased, the estimated slope coefficient may be close to one. The reason is that the key determinant of the slope coefficient is the entire sum in Eq. (A.2). If some bias coefficients are positive and some negative, the whole sum may be zero even if the shock-specific bias coefficients are not. The second shortcoming is that the method is not able to determine the horizon at which the agent exhibits bias. In contrast, our approach directly identifies the whole term structure of biases and is not vulnerable to these problems. Finally, the Mincer-Zarnowitz regression is vulnerable to measurement error (see, e.g., Jeong and Maddala, 1991). If expectations are measured with classical measurement error, the estimated slope coefficient is biased towards zero. Hence, the regression may suggest that the agent exhibits biases even when no such biases exist.

A closely related empirical approach regresses forecast errors on the lagged value of the variable being predicted (see, e.g., Barrero, 2018; Bordalo, Gennaioli, and Shleifer, 2018b; Kohlhas and Walther, 2018):

$$x_t - \mathbb{F}_{t-1}[x_t] = \delta + \gamma x_{t-1} + u_t.0 \tag{A.3}$$

A negative estimate of γ is then interpreted as indicating overreaction or overextrapola-

[35] An equivalent test regresses forecast errors on $\mathbb{F}_{t-1}[x_t]$, with a negative coefficient interpreted as evidence for overreaction.

tion. Intuitively, if an agent observes a high value of x_t and exaggerates that information when making the forecast, the subsequent realization x_{t+1} will on average be lower than predicted, giving rise to a negative forecast error. We calculate that

$$\text{plim } \hat{\gamma} \leq 0 \quad \Leftrightarrow \quad \sum_{\ell=1}^{+\infty} \text{sgn}(\alpha_\ell)\alpha_{\ell-1}b_\ell \geq 0.$$

Hence, interpreting $\hat{\gamma} \leq 0$ as indicating overreaction makes an implicit assumption on the data-generating process for x_t. If the signs of α_ℓ and $\alpha_{\ell-1}$ are the same, then the estimated coefficient is indeed negative whenever the agent exhibits overreaction to news at all lags. In that case, still, the method is not able to disentangle the horizon at which the agent is overreacting. However, if the process for x_t exhibits reversals, some of the adjacent α_ℓ coefficients will differ in sign, leading to potentially incorrect conclusions about under- and overreaction. Finally, if the agent exhibits both under- and overreaction to information at different lags, the estimated coefficient may be close to zero even if the agent exhibits biases. An advantage of the approach relative to the basic Mincer-Zarnowitz regression is that it is robust to classical measurement error in expectations.

A special case in which the regression in Eq. (12) provides a valid measure of overreaction is given when the variable of interest is pure noise, $x_t = \varepsilon_t$. In that case, plim $\hat{\gamma} = -b_1$, and the regression directly estimates the first bias coefficient. In fact, more efficient estimates can be obtained by directly regressing forecasts on lagged values of x_t, in effect estimating Eq. (2). This type of regression is commonly used in the literature studying extrapolative expectations of stock market returns (see, e.g., Graham and Harvey, 2001; Vissing-Jorgensen, 2004; Ben-David, Graham, and Harvey, 2013; Dominitz and Manski, 2011; Greenwood and Shleifer, 2014). Classic asset pricing theory predicts that stock prices follow martingales at short horizons (see, e.g. Cochrane, 2005, p. 22), and $x_t = \varepsilon_t$ is therefore a reasonable first approximation for stock returns. This theory-implied restriction is likely to yield more precise estimates than unrestricted estimation of the IRF of forecast errors. Our framework highlights that stock returns must be uncorrelated over the relevant time horizon for the regression to yield a meaningful measure of overreaction. If stock returns exhibit momentum, some degree of extrapolation is justified.

Appendix C Extensions to the Basic Framework

This appendix develops several extensions to the basic framework: multiple-step-ahead forecasts, heterogeneity, state- and time-dependent biases, distinguishing biases with respect to aggregate and idiosyncratic shocks, and measurement error.

C.1 Multiple-Step-Ahead Forecasts

Suppose that the agent makes h-step ahead forecasts with $h \geq 1$ denoting the forecast horizon. To simplify notation, suppose that x_t is driven by a single shock, as given by Eq. (1). We assume that the forecasts are generated as

$$\mathbb{F}_t[x_{t+h}] = b_0 + \sum_{\ell=0}^{+\infty} a_{\ell+h}\varepsilon_{t-\ell}.$$

Performing the same calculations as in Section 2.1 shows that

$$x_t - \mathbb{F}_{t-h}[x_t] = -b_0 - \sum_{\ell=h}^{+\infty} \text{sgn}(\alpha_\ell)b_\ell\varepsilon_{t-\ell} + \sum_{\ell=0}^{h-1} \alpha_\ell\varepsilon_{t-\ell}.$$

Comparing the equation above to Eq. (4), we observe an additional term stemming from the multiple-step ahead nature of forecasts. Even if subjective expectations react to news in an unbiased way, forecast errors are mechanically autocorrelated—up to lag $(h-1)$—if the underlying process is autocorrelated at these lags.

The methodology again boils down to estimating the IRF of forecast errors. Differently from the one-step-ahead case, the first $(h-1)$ impulse responses need to be discarded. The remaining impulse responses are converted to bias coefficients, with $b_{\ell+h}$ giving the biased reaction to news that arrived ℓ periods ago.

C.2 Heterogeneity and Aggregation

Existing research has documented that expectations are heterogeneous across individuals (e.g., Manski, 2004, Section 5). Our method can easily be applied to subsamples of the population. For example, we may estimate the bias coefficients for young and old forecasters. When the method is applied to the whole population, it recovers an average of the bias coefficients of the individual forecasts, as we now show.

Suppose, as in Section 2.1.2, that $x_t = \sum_{i=1}^{M}\sum_{\ell=0}^{+\infty} \alpha_{i\ell}\varepsilon_{i,t-\ell}$. The forecast of forecaster f is generated as

$$\mathbb{F}_t^{(f)}[x_{t+1}] = b_0^{(f)} + \sum_{i=1}^{M} \sum_{\ell=0}^{+\infty} a_{i,\ell+1}^{(f)} \varepsilon_{i,t-\ell}, \quad f = 1, 2, \ldots, N.$$

Denote the consensus (average) forecast by $\mathbb{F}_t[x_{t+1}] \equiv \frac{1}{N}\sum_f \mathbb{F}_t^{(f)}[x_{t+1}]$. Then, the consensus forecast is given by

$$\mathbb{F}_t[x_{t+1}] = \bar{b}_0 + \sum_{i=1}^{M} \sum_{\ell=0}^{+\infty} \bar{a}_{i,\ell+1} \varepsilon_{i,t-\ell},$$

where $\bar{b}_0 \equiv \frac{1}{N}\sum_f b_0^{(f)}$ and $\bar{a}_j^{(\ell)} \equiv \frac{1}{N}\sum_f a_{i\ell}^{(f)}$. Applying the results from Section 2.1.2, the shock-specific bias coefficients of the consensus forecast, $\bar{b}_{i\ell}$, are equal to

$$\bar{b}_{i\ell} = \operatorname{sgn}(\alpha_{i\ell})(\bar{a}_{i\ell} - \alpha_{i\ell}) = \operatorname{sgn}(\alpha_{i\ell}) \left[\frac{1}{N} \sum_{f=1}^{N} \left\{ a_{i\ell}^{(f)} - \alpha_{i\ell} \right\} \right] = \frac{1}{N} \sum_{f=1}^{N} b_{i\ell}^{(f)},$$

where $b_{i\ell}^{(f)} \equiv \operatorname{sgn}(\alpha_{i\ell})[a_{i\ell}^{(f)} - \alpha_{i\ell}]$. Therefore, with heterogeneity, the bias coefficients of the consensus forecast equal the average bias coefficients of the individual forecasts.

The linear aggregation result only holds for the *shock-specific* bias coefficients, in the sense defined in Section 2.1.2. For the composite bias coefficients, aggregation is non-linear. Therefore, the composite bias coefficients of the consensus forecast are *not* given by the average composite bias coefficients of the individual forecasts.

C.3 State and Time Dependence

It is straightforward to generalize our framework to allow for state and time dependence in both the data-generating process for the variable being predicted as well as how expectations are formed.

Instead of being time invariant, suppose that the true reactions $\alpha_{i\ell}$ are given by a function of a vector of state variables $\mathbf{s}_t$, so that $\alpha_{i\ell} = \alpha_{i\ell}(\mathbf{s}_t)$, with $\alpha_{i\ell}$ a deterministic function mapping the state variables to the true reaction of the process.[36] Various types of state dependence can be achieved by choosing specific state variables $\mathbf{s}_t$ and functional forms for $\alpha_{i\ell}(\cdot)$. Similarly, let the perceived reaction now be $a_{i\ell} = a_{i\ell}(\mathbf{s}_t)$. In that

[36] If $\alpha_{i\ell}(\cdot)$ is not a deterministic function, then bias coefficients become random variables. In these cases, it may be desirable to estimate the whole distribution of bias coefficients, instead of just their first moment. See Koop, Pesaran, and Potter (1996) for a discussion of the issues.

case, the bias coefficients are also state dependent and given by

$$b_{i\ell}(\mathbf{s}_t) = \text{sgn}[\alpha_{i\ell}(\mathbf{s}_t)][a_{i\ell}(\mathbf{s}_t) - \alpha_{i\ell}(\mathbf{s}_t)].$$

Hence, state dependence does not affect the basic theory in any substantial way. The only difference is that estimating state-dependent bias coefficients in practice now requires controlling for the underlying state variables appropriately. The best way to do that depends on the application at hand. An explicit example of state-dependent bias coefficients is provided in Section D.7 of the Appendix which derives bias coefficients for a model in which the agent makes forecasts to minimize an asymmetric loss function, and the shocks follow a GARCH process, as in Patton and Timmermann (2007).

C.4 Aggregate and Idiosyncratic Shocks

Theoretical models commonly distinguish between information that is available to every agent in the economy and information that is agent specific. Such information structures are often used in models of incomplete information in macroeconomics and finance, as surveyed by Angeletos and Lian (2016).

Our multiple-shocks framework in Section 2.1.2 nests such information structures. In addition, when data on individual forecasts is available, we can use our measurement procedure to estimate how agents react to idiosyncratic shocks. This exercise is helpful for dissecting the sources of bias in expectations.

For ease of exposition, suppose that there is a single aggregate shock ε_t driving the variable of interest, as in Section 2.1.1. However, in addition to the aggregate shock, each agent i has an associated idiosynratic shock, ω_{it}. We assume that individual forecasts are generated as

$$\mathbb{F}_{it}[x_{t+1}] = b_0 + \sum_{\ell=0}^{+\infty} a_{\ell+1}^{(\varepsilon)} \varepsilon_{t-\ell} + \sum_{\ell=0}^{+\infty} a_{\ell+1}^{(\omega)} \omega_{i,t-\ell}.$$

A concrete example of a model that gives rise to such expectations is given by the noisy information model (see Appendix D.1). In this model, each agent only observes a noisy signal of x_t given by $y_{it} = x_t + \omega_{it}$ with ω_{it} denoting normally distributed noise. Then, the agent uses Bayesian updating to form expectations about x_{t+1}. Provided that the precision of signals, $\mathbb{V}\text{ar}[\omega_{it}]^{-1}$, is the same for each agent, the individual expectations can be represented in the form above.

Defining the average or consensus forecast by $\mathbb{F}_t[x_{t+1}] = \mathbb{E}\left\{\mathbb{F}_{it}[x_{t+1}]\right\}$ where the expectation is taken across agents, deviations from consensus are given by

$$\underbrace{\mathbb{F}_{it}[x_{t+1}] - \mathbb{F}_t[x_{t+1}]}_{\text{deviation from consensus forecast}} = \sum_{\ell=0}^{+\infty} a_{\ell+1}^{(\omega)} \omega_{i,t-\ell}.$$

As a result, reactions to idiosyncratic shocks can be obtained from the IRF of deviations from consensus. The result generalizes straightforwardly to the case with many aggregate shocks. By the Frisch-Waugh-Lovell Theorem, this IRF can be estimated by simply including time fixed effects in the local projections regression in Eq. (10) (when individual forecasts are used).

The assumption that the agents react identically to the aggregate shocks is key for the result above to hold. If the reactions to the aggregate shocks were heterogeneous across agents (e.g., agents observe signals with different levels of precision), deviations from consensus would arise both from idiosyncratic shocks as well differences in the loadings on the aggregate shocks. In that case, dynamic factor models may be used to measure reactions to idiosynratic shocks.

C.5 Measurement Error

Expectations are typically measured with some error. Measurement error can arise for many reasons, including survey design imperfections and limited attention of survey participants.[37] The magnitude of measurement error is often substantial, especially for individual-level data. For example, Giglio, Maggiori, Stroebel, and Utkus (2019) elicit expected stock market returns in a sample of individual investors using two independent measurement methods. They find a correlation between the two measures of 0.43, suggesting substantial measurement error.

The effects of measurement error depend on whether composite or shock-specific bias coefficients are being estimated. Measurement error is not problematic for estimating the shock-specific bias coefficients using the method of external instruments. As discussed in Section 2.1.2, this method provides consistent estimates of the bias coefficients as long as measurement error is uncorrelated with the instrument. Of course, in finite samples measurement error typically reduces precision of the estiamtes. This robustness to measurement error is a key advantage of the IV procedure.

[37] A closely related possibility is that subjective expectations are inherently noisy, such as in the model of da Silveira and Woodford (2019). The empirical predictions of measurement error and inherently noisy expectations are often very similar. Noise in expectations may also indicate that expectations are not very well formed and only weakly related to actual behavior; see Drerup, Enke, and von Gaudecker (2017) for more discussion and empirical evidence on this latter point.

Now consider the case of estimating composite bias coefficients. Estimating composite bias coefficients boils down to estimating the univariate IRF of forecast errors. Classical measurement error tends to mask any existing predictability of the forecast errors, biasing the estimated bias coefficients towards zero. As a result, the empirically estimated composite bias coefficients are likely to represent a lower bound on the true composite bias coefficients. To the extent that measurement error is idiosyncratic, measurement error is less problematic for consensus forecasts.

We now provide explicit formulas for the attenuation bias caused by measurement error for two models of expectations (sticky information and diagnostic expectations) when calculating composite bias coefficients. As in Section 3, x_t follows a stationary AR(1) process.

First, suppose that the true expectations are generated by the sticky information model of Mankiw and Reis (2002), implying that the true forecast errors follow

$$e_t = \lambda\rho e_{t-1} + \varepsilon_t.$$

However, instead of observing the true forecast $\mathbb{F}_t[x_{t+1}]$, we can only observe

$$\mathbb{F}^*_t[x_{t+1}] = \mathbb{F}_t[x_{t+1}] + v_t,$$

where v_t is white noise measurement error with variance σ_v^2 (and independent of ε_t). The observed forecast error is then equal to $e^*_t = e_t - v_{t-1}$. Now write

$$e^*_t - \lambda\rho e^*_{t-1} = \varepsilon_t - v_{t-1} + \lambda\rho v_{t-2}. \tag{A.4}$$

The right-hand side of Eq. (A.4) is the sum of an MA(1) process and white noise, and therefore also an MA(1) process (see, e.g., Hamilton, 1994, pp. 102–105). Denote the resulting process as $\xi_t + \theta\xi_{t-1}$ for some parameters θ and σ_ξ^2 to be determined. For the representation to be valid, the autocovariances must match, namely

$$\sigma_\varepsilon^2 + [1 + (\lambda\rho)^2]\sigma_v^2 = (1 + \theta^2)\sigma_\xi^2$$
$$-\lambda\rho\sigma_v^2 = \theta\sigma_\xi^2$$

Substituting out σ_ξ and rearranging leads to a quadratic equation in θ:

$$(-\lambda\rho)\theta^2 - \theta\left\{\frac{\sigma_\varepsilon^2}{\sigma_v^2} + \left[1 + (\lambda\rho)^2\right]\right\} - \lambda\rho = 0.$$

The equation has two real solutions. Picking the solution associated with the invertible representation (i.e., with $|\theta| < 1$) yields

$$\theta = \frac{\left\{\frac{\sigma_\varepsilon^2}{\sigma_v^2} + [1 + (\lambda\rho)^2]\right\} - \sqrt{\left\{\frac{\sigma_\varepsilon^2}{\sigma_v^2} + [1 + (\lambda\rho)^2]\right\}^2 - 4(\lambda\rho)^2}}{-2\lambda\rho}. \tag{A.5}$$

All in all, the observed forecast errors follow an ARMA(1, 1) process with

$$e_t^*(1 - \lambda\rho L) = (1 + \theta L)\xi_t,$$

where L is the lag operator. As a result, the Wold representation of e_t^* is

$$e_t^* = \xi_t + (\lambda\rho + \theta)\sum_{\ell=1}^{+\infty}(\lambda\rho)^{\ell-1}\xi_{t-\ell}.$$

Therefore, measurement error leads to an attenuation bias. The attenuation bias can be substantial if measurement error is large (i.e., signal-to-noise ratio, $\sigma_\varepsilon/\sigma_v$, is small). For example, suppose that $\lambda = 0.50$, $\rho = 0.75$, $\sigma_\varepsilon = 0.25$, and $\sigma_v = 0.15$. Then, the true bias coefficient b_1 is equal to $(-\lambda\rho) = -0.375$. In contrast, the bias coefficient in the process with measurement error is equal to

$$-(\lambda\rho + \theta) \approx -(0.375 - 0.097) = -0.278.$$

The attenuation bias is roughly 26% in relative terms.

Now consider the case of diagnostic expectations. In that case, the true forecast errors follow an MA(1) process with $e_t = \varepsilon_t - \theta\rho\varepsilon_{t-1}$, implying that the observed forecast errors are given by $e_t^* = \varepsilon_t - \theta\rho\varepsilon_{t-1} - v_{t-1}$. The right-hand side again follows an MA(1) process but with different parameters. Write $e_t^* = \xi_t + \psi\xi_{t-1}$ for some parameters ψ and σ_ξ^2. Similar calculations to those performed earlier show that

$$\psi = \frac{\left\{\frac{\sigma_v^2}{\sigma_\varepsilon^2} + [1 + (\theta\rho)^2]\right\} - \sqrt{\left\{\frac{\sigma_v^2}{\sigma_\varepsilon^2} + [1 + (\theta\rho)^2]\right\}^2 - 4(\theta\rho)^2}}{-2\theta\rho}.$$

To gauge the size of the attenuation bias, suppose that $\theta = 0.50$, $\rho = 0.75$, $\sigma_\varepsilon = 0.25$, and $\sigma_v = 0.15$. With these parameters, the true bias coefficient is equal to $b_1 = \theta\rho = 0.375$. However, the bias coefficient from the process with measurement error (i.e., $-\psi$) is equal to approximately 0.268. In relative terms, the attenuation bias is roughly 29%.

Appendix D Mapping Existing Models

D.1 Noisy Information

We now analyze a model in which agents are rational and understand the structure of the model but do not observe the underlying state perfectly. Models of this type include the rational inattention model of Sims (2003) and the imperfect information model studied by Woodford (2003).

Suppose the true process for x_t is an AR(1) but each agent i only observes

$$y_{it} = x_t + \omega_{it}.$$

Here, ω_{it} is a normally distributed mean-zero noise term which is i.i.d. across time and agents and uncorrelated with ε_t at all leads and lags. The Kalman filter equations imply

$$\mathbb{F}_{it}[x_{t+1}] = \rho \left\{ G y_{it} + (1 - G) \, \mathbb{F}_{i,t-1}[x_t] \right\}, \tag{A.6}$$

where $G \in [0, 1]$ is the Kalman gain.

From Eq. (A.6) and the fact that x_t follows an AR(1), forecast errors at the individual level also follow an AR(1) with

$$x_t - \mathbb{F}_{i,t-1}[x_t] = \rho(1 - G) \left\{ x_{t-1} - \mathbb{F}_{i,t-2}[x_{t-1}] \right\} + \varepsilon_t - \rho G \omega_{i,t-1}. \tag{A.7}$$

According to our definitions, at the individual level, the agent overreacts to the idiosyncratic noise, ω_{it}, and underreacts to the aggregate shock ε_t.

We now calculate bias coefficients for the consensus forecast. Assuming that there is a continuum of agents and making the usual Law of Large Numbers assumption, we can integrate over agents to find that

$$x_t - \mathbb{F}_{t-1}[x_t] = \rho(1 - G) \left\{ x_{t-1} - \mathbb{F}_{t-2}[x_{t-1}] \right\} + \varepsilon_t.$$

Here, $\mathbb{F}_t[x_{t+1}] = \mathbb{E}\{\mathbb{F}_{it}[x_{t+1}]\}$ is the average (consensus) forecast, with the expectation taken across agents. Hence, at the consensus level forecast errors again follow an AR(1). Therefore,

$$\begin{aligned} a_\ell &= \rho^\ell - [(1 - G)\rho]^\ell \\ b_\ell &= -\operatorname{sgn}(\rho^\ell)[(1 - G)\rho]^\ell \end{aligned}$$

Hence, as long as the signal is not perfectly revealing of the state ($G \neq 1$), the noisy information model predicts underreaction at the consensus level. Since idiosyncratic shocks wash out in the aggregate, the consensus forecast does not overreact to ω_{it}. The noisy information model predicts identical bias coefficients as the sticky information model when $\lambda = 1 - G$.

We note that deviations from the consensus forecast also follow an AR(1) with the same persistence parameter. Specifically, we have that

$$\mathbb{F}_{it}[x_{t+1}] - \mathbb{F}_t[x_{t+1}] = \rho(1 - G)\left\{\mathbb{F}_{i,t-1}[x_t] - \mathbb{F}_{t-1}[x_t]\right\} + \rho G\omega_{it}.$$

Hence, in the basic noisy information model, forecast errors at the individual level, forecast errors at the consensus level, and deviations from the consensus all follow an AR(1) process with the same persistence parameter $\rho(1 - G)$.

The results of this section generalize to models in which forecasters interact strategically (such as trying to stay close to the consensus forecast). In particular, Huo and Pedroni (2018) show that the solution to a wide class of models with strategically interacting agents and noisy information is given by a single-agent forecasting problem with a modified information structure. In this modified information structure, private signals are more noisy, in accordance to the degree of strategic complementarity in the model. Hence, in a model with strategically interacting agents, the only change would be in the value of G, the Kalman gain parameter.

D.2 Diagnostic Expectations With Noisy Information

Bordalo, Gennaioli, Ma, and Shleifer (2018a) combine noisy information and diagnostic expectations. As in Appendix D.1, suppose that the agents do not observe x_t directly but only receive a noisy signal. Each agent i observes $y_{it} = x_t + \omega_{it}$ where ω_{it} is normally distributed noise. If the agent used Bayesian updating, the agent would form expectations as

$$\mathbb{F}^*_{it}[x_{t+1}] = \rho\left\{Gy_{it} + (1 - G)\,\mathbb{F}^*_{i,t-1}[x_t]\right\},$$

where G is the Kalman gain, just as in Section D.1. However, the agent is subject to the representativeness heuristic and overweights representative events. Under representativeness, Bordalo, Gennaioli, Ma, and Shleifer (2018a, Proposition 1) show that expectations follow

$$\mathbb{F}_{it}[x_{t+1}] = \mathbb{F}^*_{it}[x_{t+1}] + \rho\theta G\left\{y_{it} - \mathbb{F}^*_{i,t-1}[x_t]\right\}.$$

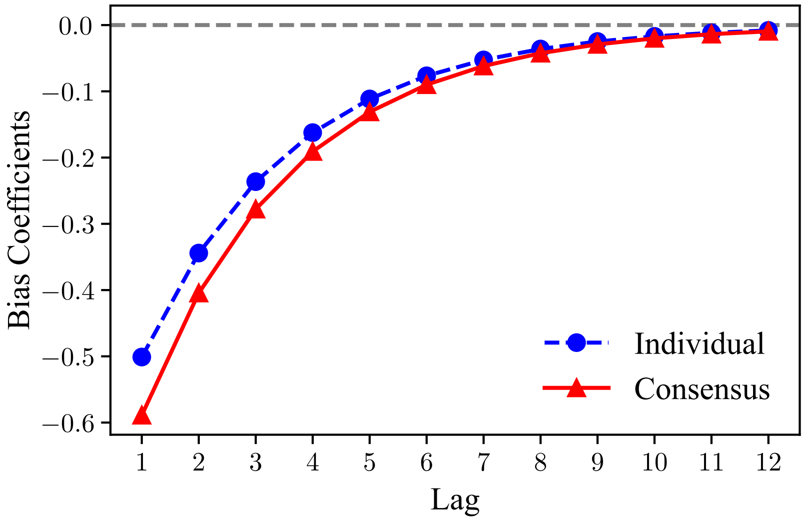


Figure A.1: Composite bias coefficients for the mixed noisy information and diagnostic expectations model of Bordalo, Gennaioli, Ma, and Shleifer (2018a) at the individual and consensus level. Parameter values are $\rho = 0.85$, $\sigma_\varepsilon^2 = 1.005$, $\mathbb{V}\text{ar}[\omega_{it}] = 3.18^2\sigma_\varepsilon^2$, and $\theta = 0.60$.

Denoting $e_{it}^* = x_t - \mathbb{F}_{i,t-1}^*[x_t]$, we can express forecast errors as

$$\begin{aligned} e_{i,t+1} &= e_{i,t+1}^* - \rho\theta G e_{it}^* - \rho\theta G\omega_{it} \\ &= \rho[1-(1+\theta)G]e_{it}^* + \varepsilon_{t+1} - \rho(1+\theta)G\omega_{it}, \end{aligned} \tag{A.8}$$

where the second equality uses Eq. (A.7). Subtracting $\rho(1-G)e_{it}$ from both sides and using Eqs. (A.7) and (A.8), we have

$$e_{i,t+1} - \rho(1-G)e_{it} = \varepsilon_{t+1} - \rho\theta G\varepsilon_t - \rho G(1+\theta)\left(\omega_{it} - \frac{\rho\theta}{1+\theta}\omega_{i,t-1}\right).$$

The right-hand side is a sum of two independent MA(1) processes and hence also an MA(1) process (see, e.g., Hamilton, 1994, pp. 102–105). All in all, then, forecast errors follow an ARMA(1, 1) process. The ARMA(1, 1) representation then yields expressions for the composite bias coefficients. We omit the resulting lengthy formulas.

Consensus forecasts, $\mathbb{F}_t[x_{t+1}] \equiv \mathbb{E}\{\mathbb{F}_{it}[x_{t+1}]\}$, with the expectation taken across agents i, also follow an ARMA(1, 1) process. The only difference is that idiosyncratic shocks, ω_{it}, wash out in the consensus forecast. Therefore, forecast errors at the consensus level follow

$$e_{i,t+1} - \rho(1-G)e_{it} = \varepsilon_{t+1} - \rho\theta G\varepsilon_t.$$

A numerical example of the resulting bias coefficients is provided in Figure A.1. The parameter values are based on the estimates provided by Bordalo, Gennaioli, Ma, and Shleifer (2018a, Table 8). They find that for a broad range of macroeconomic expectations, the representativeness parameter θ is around 0.60. For GDP deflator inflation, they estimate the standard deviation of noise in the measurement equation to be roughly three times larger than the standard deviation of the true shocks. Following the parametrization in Coibion and Gorodnichenko (2012), we set $\rho = 0.85$ and $\sigma_\varepsilon^2 = 1.005$ to approximate the behavior of actual GDP deflator inflation. The optimal steady-state Kalman gain parameter is then calculated using values for the variances of the shocks and persistence of the process, implying a value of $G \approx 0.19$.[38] The predicted bias coefficients are consistent with our empirical findings in Section 4. Bias coefficients of the individual forecasts are attenuated toward zero with respect to the bias coefficients of the consensus forecast, just as in the data.

D.3 Misperceived Law of Motion

Suppose that the agent misperceives the true persistence of the process and makes forecasts as

$$\mathbb{F}_t[x_{t+1}] = \hat{\rho} x_t, \hat{\rho} \in (-1, 1),$$

with $\hat{\rho}$ potentially different from ρ. Examples of models with misperceived laws of motion abound in the literature, with two prominent cases given by Barberis, Shleifer, and Vishny (1998) and Fuster, Laibson, and Mendel (2010); see also Gabaix (2017a, pp. 14–15). In the present case,

$$\mathbb{F}_t[x_{t+1}] = \hat{\rho} \sum_{\ell=0}^{+\infty} \rho^\ell \varepsilon_{t-\ell}$$

and therefore

$$\begin{aligned} a_\ell &= \hat{\rho} \rho^{\ell-1} \\ b_\ell &= \operatorname{sgn}(\rho^\ell) \left[\hat{\rho} \rho^{\ell-1} - \rho^\ell \right] \end{aligned}$$

[38] The steady-state Kalman gain is $G = \Sigma/(\Sigma + \sigma_\omega^2)$ where $\sigma_\varepsilon^2 = \mathbb{V}\text{ar}[\varepsilon_t]$, $\sigma_\omega^2 = \mathbb{V}\text{ar}[\omega_{it}]$ and

$$\Sigma = \frac{-(1-\rho^2)\sigma_\omega^2 + \sigma_\varepsilon^2 + \sqrt{[(1-\rho^2)\sigma_\omega^2 - \sigma_\varepsilon^2]^2 + 4\sigma_\omega^2\sigma_\varepsilon^2}}{2}.$$

When $\rho \neq 0$, we can write

$$b_\ell = \operatorname{sgn}(\rho^\ell)\rho^\ell \left(\frac{\hat{\rho} - \rho}{\rho}\right).$$

If $\rho > 0$, the agent overreacts to news whenever $\hat{\rho} > \rho$ and underreacts otherwise.

D.4 Extrapolative Expectations

We now consider pure extrapolative expectations

$$\mathbb{F}_t[x_{t+1}] = x_t + \gamma(x_t - x_{t-1}),$$

as in Goodwin (1947, p. 191). The parameter γ could be either positive or negative, with a positive γ representing extrapolation or trend following, while a negative γ could capture contrarian expectations.

Substituting in the expression for x_t, we calculate that

$$\mathbb{F}_t[x_{t+1}] = (1+\gamma)\varepsilon_t + \sum_{\ell=1}^{+\infty} \left\{(1+\gamma)\rho^\ell - \gamma\rho^{\ell-1}\right\} \varepsilon_{t-\ell},$$

and so we find that

$$a_\ell = \begin{cases} 1+\gamma & \text{if } \ell = 1 \\ (1+\gamma)\rho^{\ell-1} - \gamma\rho^{\ell-2} & \text{if } \ell \geq 2 \end{cases},$$

and

$$b_\ell = \begin{cases} \operatorname{sgn}(\rho)(1+\gamma-\rho) & \text{if } \ell = 1 \\ \operatorname{sgn}(\rho^\ell)\left\{(1+\gamma)\rho^{\ell-1} - \gamma\rho^{\ell-2} - \rho^\ell\right\} & \text{if } \ell \geq 2 \end{cases}$$

Suppose expectations are of the trend-following type ($\gamma > 0$). Then, if x_t is positively autocorrelated, the agent always overreacts to current news ($b_1 > 0$). However, extrapolative expectations may well lead to underreaction to past news.

D.5 Adaptive Learning

Many models in economics feature agents that form expectations by estimating econometric models; for recent overviews, see Evans and Honkapohja (2001) and Evans and Honkapohja (2009). Such models lead to a time-varying process of expectation formation. As a result, they are nested in the class of models studied in Section C.3. While deriving the time-varying bias coefficients is straightforward, we may also be interested in calculating the *average* bias coefficients, especially for comparison with empirical estimates. Calculating the average bias coefficients analytically is challenging. Nevertheless, it is straightforward to estimate them using Monte Carlo simulation, as we now demonstrate.

The true process for x_t is again an AR(1). However, we no longer assume that the agent knows the true parameter values governing the process. Instead, the agent estimates the parameters using least squares. We assume that the first value that is observed is x_1. As in Orphanides and Williams (2005) or Malmendier and Nagel (2016), the agent perceives the process to be

$$x_t = b_0 + b_1 x_{t-1} + u_t,$$

with parameter vector $\mathbf{b} = (b_0, b_1)^\top$ to be estimated.

Denoting the data for period t by $\mathbf{X}_t = (1, x_t)^\top$, the agent estimates the parameters using the following recursion:

$$\begin{aligned} \mathbf{b}_t &= \mathbf{b}_{t-1} + \gamma_t \mathbf{R}_t^{-1} \mathbf{X}_{t-1}(x_t - \mathbf{X}_{t-1}^\top \mathbf{b}_{t-1}) \\ \mathbf{R}_t &= \mathbf{R}_{t-1} + \gamma_t(\mathbf{X}_{t-1}\mathbf{X}_{t-1}^\top - \mathbf{R}_{t-1}) \end{aligned} \tag{A.9}$$

Here, $\mathbf{b}_t$ is the current estimate of the parameter vector $\mathbf{b}$, $\mathbf{R}_t$ is the current estimate of the second moment matrix of $\mathbf{X}_{t-1}$, and $\{\gamma_t\}$ is a *sequence of gains*.

We choose initial values for $\mathbf{b}_t$ and $\mathbf{R}_t$ to ensure that when the sequence of gains is chosen appropriately, $\mathbf{b}_t$ equals the standard least-squares estimator. Specifically, we start the recursion at $t = 3$ with

$$\begin{aligned} \mathbf{b}_3 &= \frac{1}{2}\mathbf{R}_3^{-1}(\mathbf{X}_1 x_2 + \mathbf{X}_2 x_3) \\ \mathbf{R}_3 &= \frac{1}{2}(\mathbf{X}_1\mathbf{X}_1^\top + \mathbf{X}_2\mathbf{X}_2^\top) \end{aligned}$$

Then, for $t \geq 4$ we use the recursion in Eq. (A.9); see Evans and Honkapohja (2001, pp. 32–33) for more details.[39]

The model with adaptive learning is effectively a time-varying version of the model in Appendix D.3 in which the agent misperceives the law of motion for x_t. With adaptive learning, the perceived persistence of x_t is determined by the observed data and varies over time. In the notation of Section C.3, bias coefficients depend on a state variable $\mathbf{s}_t$ that contains all past values of x_t, i.e., $\mathbf{s}_t = (x_t, x_{t-1}, \ldots, x_1)^\top$. The dependence of

[39] Recall that the first available value of x_t is x_1. Hence, data for estimation is available for $t \geq 2$.

expectations on $\mathbf{s}_t$ depends on the exact form of learning and the chosen gain sequence. At any rate, the results from Appendix D.3 carry through to the present setting, the only difference being that the perceived persistence of x_t is now a function of $\mathbf{s}_t$.

In practice, we may also be interested in calculating the *average* bias coefficients. Doing so analytically is challenging. As a result, we now show how Monte Carlo simulation can be used to estimate the average bias coefficients. We investigate two different specifications for the sequence of gains, $\{\gamma_t\}$. First, we consider standard *least-squares learning* with $\gamma_t = 1/(t-1)$. With this choice for γ_t, $\mathbf{b}_t$ coincides with the usual least-squares estimator. As is well known, such least-squares learning is optimal—in the sense of being implied by Bayesian updating—under certain conditions. Specifically, if the true shocks are i.i.d. normally distributed, and the prior distribution of $\mathbf{b}$ is normal with $\mathbb{V}\text{ar}[\mathbf{b}]^{-1} = \mathbf{0}$ (diffuse prior), then the posterior distribution of $\mathbf{b}$ is also normal, with its mean given by the least-squares estimate (see, e.g., Hoff, 2009, pp. 154–155).

Second, we consider *constant-gain learning* with $\gamma_t = \gamma$. Constant-gain learning discounts past data and leads to perpetual learning. This form of learning is optimal when parameters undergo certain forms of structural change (see, e.g., Branch and Evans, 2006, and references therein). In addition, as shown by Malmendier and Nagel (2016, Section V.A.), if people overweight information from their own lifetimes, *average* expectations can be closely approximated by constant-gain learning, even if individual expectations are generated by adaptive learning with a decreasing gain sequence. We use two values for the gain parameter, informed by existing empirical evidence. First, we investigate $\gamma = 0.018$. Malmendier and Nagel (2016) find that constant-gain learning with $\gamma = 0.018$ can closely approximate average inflation expectations in a model with learning from experience. Second, we consider $\gamma = 0.0345$. Branch and Evans (2006) document that forecasts generated with this value of the gain parameter closely match both inflation and GDP forecasts from the Survey of Professional Forecasters.

Results from the simulation are shown in Figure A.2. We set $\rho = 0.85$, $\sigma_\varepsilon^2 = 1.005$ and simulate the model for $T = 150$ periods, following the Monte Carlo simulation study in Coibion and Gorodnichenko (2012). These parameters are chosen to approximate the behavior of GDP deflator inflation.[40] In each simulation, we estimate the bias coefficients by regressing forecast errors on the true shocks, as in Eq. (9). Since the true bias coefficients are time varying, such estimation yields a time average of the true bias coefficients. The Monte Carlo estimate is then obtained by averaging across 10,000 replications.

[40] We use a burn-in period of 1,000 periods for simulations with constant-gain learning. For least-squares learning, we do not use a burn-in period. Since the least-squares estimator is consistent, in a large sample the estimates converge to the true values, and bias coefficients would be close to zero with a large burn-in period.

For standard least-squares learning, the agent underreacts to recent news but overreacts to news that arrived more than 5 periods ago. Underreaction to recent news is likely related to the fact that the least-squares estimate of ρ, the AR(1) persistence parameter, is biased towards zero in small samples (see, e.g., MacKinnon and Smith, 1998, and references therein). Hence, we expect an agent with least-squares learning, on average, to behave similarly in finite samples to an agent who misperceives ρ to be smaller than it actually is, as in Appendix D.3. We emphasize that in the present case, non-zero bias coefficients is a small-sample phenomenon. In a large enough sample, least-squares learning would converge to the true (constant) parameter values, and bias coefficients would tend to zero.[41]

For constant-gain learning, average bias coefficients exhibit a hump-shaped pattern, mostly exhibiting overreaction to new information. This finding is intuitive, given that constant-gain learning in effect discards past data, while the true process has time-invariant coefficients. As a result, the agent thinks that incoming data is more informative than it actually is.

In both cases the magnitude of biases is fairly small (in the range of $[-0.04, 0.03]$) and significantly smaller than biases we estimate in the data for inflation expectations (Section 4). To be fair for these models, however, we note that the postulated process for x_t in the simulations is time invariant. In the data, inflation has undergone important structural shifts, and hence a process with time-invariant coefficients may be not be an appropriate benchmark. That is especially the case for constant-gain learning which can be microfounded as an optimal response to certain types of structural change (Muth, 1960). The simulation results nevertheless indicate that agents who use adaptive learning to form expectations are fairly sophisticated, and their biases may be small quantitatively.

D.6 Adjustment Costs

We now consider a model in which the agent has rational expectations but faces a cost in adjusting forecasts from one period to the next. We interpret the adjustment cost as a stand-in for reputational costs or career concerns. For example, forecasters who change their forecasts by large amounts may be perceived as having lower forecasting ability.

Similarly to Coibion and Gorodnichenko (2015, p. 2660), suppose that in each period t the agent makes a forecast of x_{t+1}. The agent wishes to minimize the mean-squared

[41] Relatedly, Andolfatto, Hendry, and Moran (2008) document that time-varying parameters can lead to spurious rejections of the rational expectations hypothesis in small samples for tests using the Mincer-Zarnowitz regression.

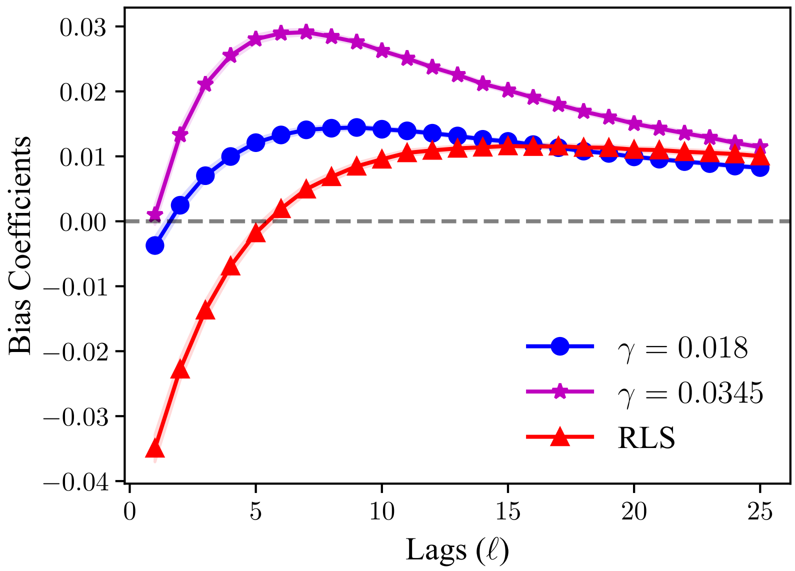


Figure A.2: Estimated average bias coefficients for selected models of adaptive learning, estimated by Monte Carlo simulation with 10,000 replications. The variable being predicted follows an AR(1) model $x_t = \rho x_{t-1} + \varepsilon_t$, $t = 1, 2, \ldots, 150$ with $\rho = 0.85$ and $\sigma_\varepsilon^2 = 1.005$. Results shown for constant-gain learning ($\gamma = 0.018$ and $\gamma = 0.0345$) and recursive least squares (RLS). A burn-in period of 1,000 periods is used for constant-gain learning. Shaded areas indicate Monte Carlo confidence bounds at the 0.1% significance level.

error of the prediction but faces a quadratic adjustment cost.[42] Denoting the current value of x_t as x and the previous forecast by F, the Bellman equation of the agent is given by

$$V(x, F) = \min_{F'} \frac{1}{2} \mathbb{E}[(\rho x + \tilde{\varepsilon} - F')^2] + \frac{\alpha}{2}(F' - F)^2 + \delta\, \mathbb{E}[V(\rho x + \tilde{\varepsilon}, F')],$$

where F' is the current period's forecast, $\alpha \geq 0$ is the weight on the adjustment cost, $\delta \in (0, 1)$ is a discount factor, and we have used tildes to denote random variables. The first-order condition is

$$-(\rho x - F') + \alpha(F' - F) + \delta\, \mathbb{E}[V_F(\rho x + \tilde{\varepsilon}, F')] = 0.$$

The envelope condition is just $V_F(x, F) = -\alpha(F' - F)$. Therefore, the optimal forecast

[42] There are no game-theoretic considerations in the present setting. Coibion and Gorodnichenko (2012, pp. 126–129) study a dynamic game in which forecasters have an incentive to stay close to the consensus forecast. They show that forecast errors follow an AR(1) process, implying geometrically decaying bias coefficients. Ottaviani and Norman (2006) study the effects of different incentive structures (reputational signaling and winner-take-all tournaments) and derive implications for observed forecasts; see also Marinovic, Ottaviani, and Sorensen (2013).

is given by

$$(F')^* = \frac{\rho x + \alpha(1-\delta)F}{1+\alpha(1-\delta)}.$$

If the agent is fully patient ($\delta = 1$) or there is no adjustment cost ($\alpha = 0$), the forecast coincides with the true conditional expectation. In the other extreme, if $\alpha \to +\infty$, then it is optimal to never change the forecast.

Defining $\phi \equiv \alpha(1-\delta)/[1+\alpha(1-\delta)]$, the optimal forecasting rule is

$$\mathbb{F}_t[x_{t+1}] = (1-\phi)\rho x_t + \phi\, \mathbb{F}_{t-1}[x_t].$$

Performing similar manipulations to those in Section 3.4, we arrive at

$$\mathbb{F}_t[x_{t+1}] = (1-\phi)\rho \sum_{\ell=0}^{+\infty} \left[\frac{\phi^{\ell+1} - \rho^{\ell+1}}{\phi - \rho}\right] \varepsilon_{t-\ell}.$$

Hence, the bias coefficients are equal to

$$\begin{aligned} a_\ell &= (1-\phi)\rho \left[\frac{\phi^\ell - \rho^\ell}{\phi - \rho}\right] \\ b_\ell &= \text{sgn}(\rho^\ell)(a_\ell - \rho^\ell) \end{aligned}$$

Inspecting the expressions above, it is clear that the model can generate both under- and overreaction.

D.7 Asymmetric Loss Function

Finally, consider the setting studied by Patton and Timmermann (2007, pp. 898–899). Suppose that the agent has the linex loss function

$$L(e_t; a) = \frac{1}{a^2}[\exp(ae_t) - ae_t - 1],$$

where a is a parameter controlling asymetry in losses. If $a \to 0$, the loss function is proportional to e^2, and hence symmetric around zero. If $a > 0$, the agent dislikes positive forecast errors (i.e., x_t is higher than predicted) more than negative forecast errors, and vice versa when $a < 0$.

The true process for x_t is now given by

$$x_t = \rho x_{t-1} + \varepsilon_t$$
$$\varepsilon_t = \sigma_t v_t, v_t \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$$
$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

with $\omega > 0$ and $\alpha, \beta \geq 0$. In contrast to before, the shocks ε_t now follow a GARCH(1, 1) process and are therefore conditionally heteroskedastic. To ensure that the $\{\varepsilon_t\}$ process is stationary, assume that $\alpha + \beta < 1$.

Some calculus shows that the optimal forecast is given by

$$\mathbb{F}_t[x_{t+1}] = \rho x_t + \frac{a}{2}\sigma_{t+1}^2.$$

Iterating the law of motion for σ_t^2, we have

$$\sigma_t^2 = \frac{\omega}{1-\beta} + \alpha \sum_{\ell=0}^{+\infty} \beta^\ell \varepsilon_{t-1-\ell}^2,$$

and therefore

$$\begin{aligned}\mathbb{F}_t[x_{t+1}] &= \rho x_t + \frac{a}{2}\left(\frac{\omega}{1-\beta} + \alpha\sum_{\ell=0}^{+\infty}\beta^\ell\varepsilon_{t-\ell}^2\right)\\ &= \frac{a\omega}{2(1-\beta)} + \sum_{\ell=0}^{+\infty}\left(\rho^{\ell+1} + \frac{a\alpha}{2}\beta^\ell\varepsilon_{t-\ell}\right)\varepsilon_{t-\ell}.\end{aligned}$$

Since GARCH is a nonlinear process, the expectation formation process is also nonlinear. The model is therefore nested in the setup studied in Section C.3, and we have

$$a_\ell(\mathbf{s}_t) = \rho^\ell + \frac{a\alpha}{2}\beta^{\ell-1}\varepsilon_{t-1-\ell}$$
$$b_\ell(\mathbf{s}_t) = \text{sgn}(\rho^\ell)\frac{a\alpha}{2}\beta^{\ell-1}\varepsilon_{t-1-\ell}$$

with the state variable $\mathbf{s}_t$ given by all past shocks, i.e., $\mathbf{s}_t = (\varepsilon_t, \varepsilon_{t-1}, \dots)^\top$. Depending on the realized values of the shocks, the model can generate both under- and overreaction. This result again illustrates how the first-order autocorrelation can be misleading as a measure of underreaction. Patton and Timmermann (2007, Eq. (23), p. 899) show that for this model, the first-order autocorrelation of forecast errors is always positive. Nevertheless, the model can in fact generate both under- and overreaction.

Appendix E Data Appendix

We download data for individual inflation forecasts from the website of the Federal Reserve Bank of Philadelphia (link). The downloaded file contains forecasts of GDP deflator inflation for the past quarter (*PGDP1*), current quarter (*PGDP2*), and the next four quarters (*PGDP3* up to *PGDP6*), see Federal Reserve Bank of Philadelphia (2017, pp. 20–22).

To construct consensus inflation forecasts, we first calculate the median forecast of *PGDP2* and *PGDP3* in each quarter. Then, we calculate annualized quarter-on-quarter inflation forecasts as

$$100\left[\left(\frac{PGDP3}{PGDP2}\right)^4 - 1\right].$$

The approach follows the standard practice in the Survey of Professional Forecasters. To calculate individual inflation forecasts, we directly use the equation above.

For realizations, we use the Real-Time Data Set for Macroeconomists which is also provided by the Philadelphia Fed (link). 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.

In addition, we use data from the FRED database (link) for the Federal Funds Rate (code DFF, link) and the NBER recession indicator (code USRECQ, link). To convert the Federal Funds Rate to a quarterly frequency, we take an average of the daily data.

For the Romer-Romer monetary-surprise measure, we use the extended dataset provided by Wieland and Yang (2017) (link). Wieland and Yang extend the original series of Romer and Romer (2004) up to the end of 2007.

Appendix F Additional Empirical Results

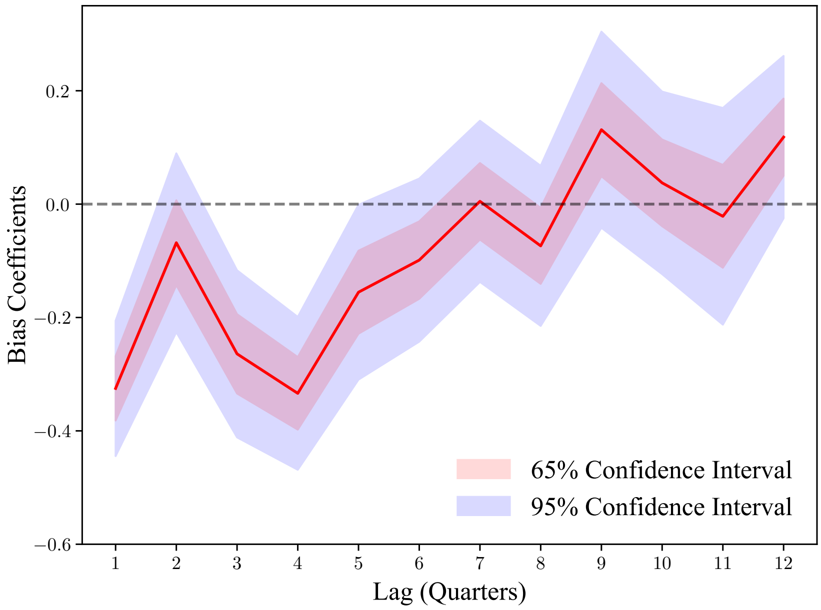


Figure A.3: Bias coefficients for one-quarter ahead inflation forecasts: maximum likelihood estimates. The estimation uses consensus (median) forecasts. The impulse response function of the forecast errors is obtained by estimating

$$x_t - \mathbb{F}_{t-1}[x_t] = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_{12} \varepsilon_{t-12}, \ \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$$

by maximum likelihood. The bias coefficients are then given by $\hat{b}_\ell = -\hat{\theta}_\ell$.

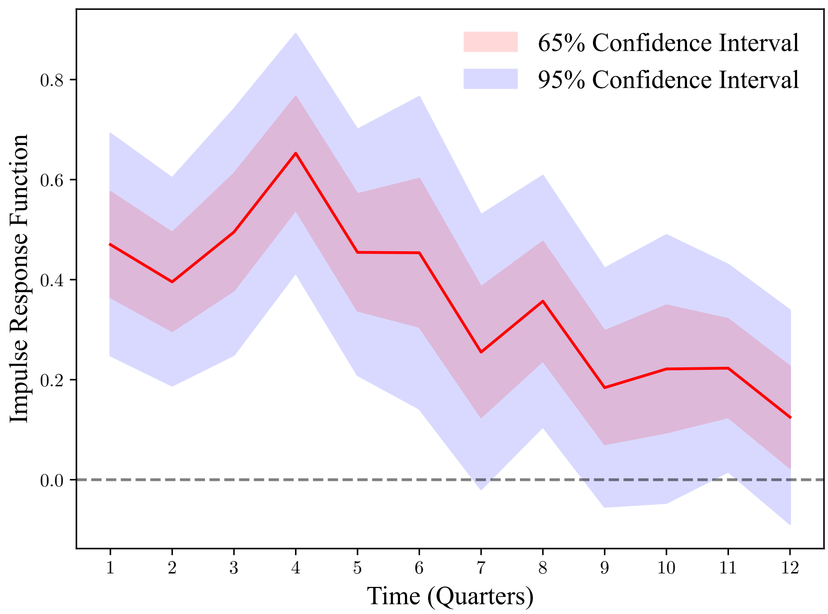
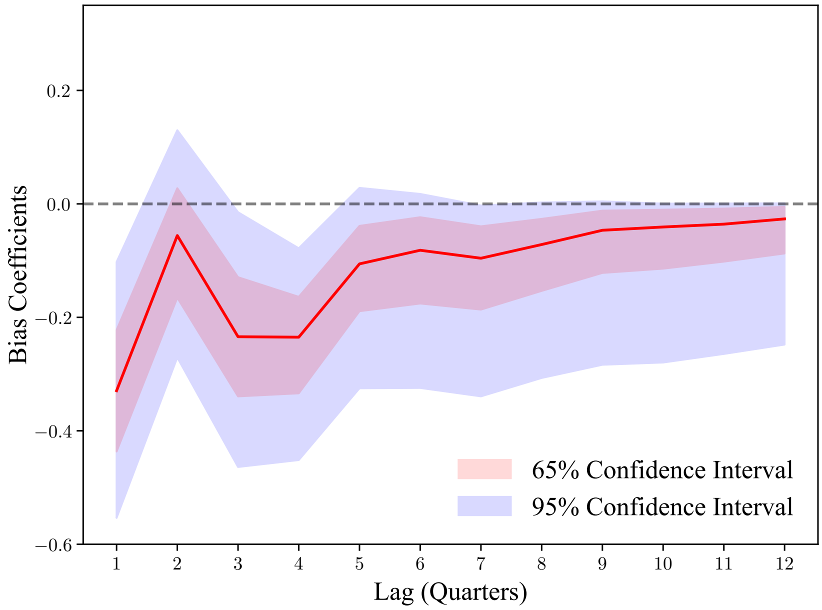
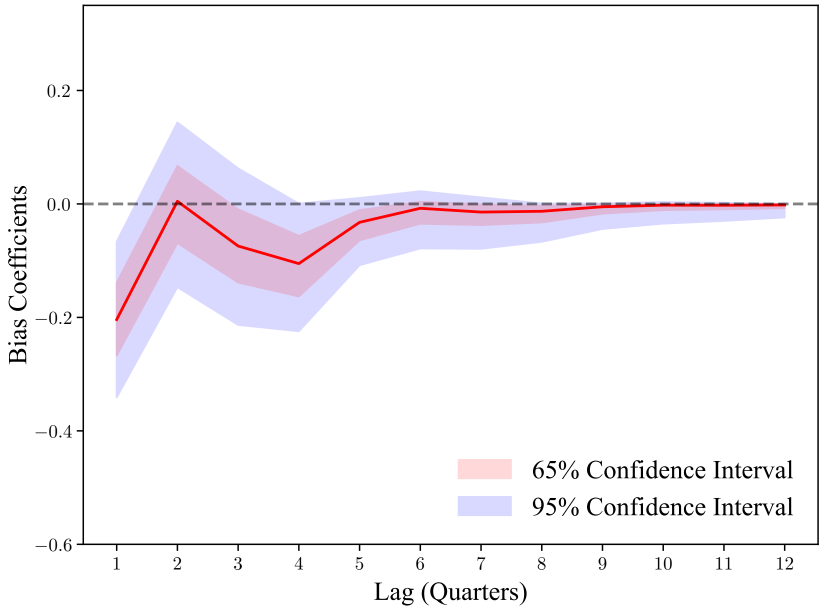


Figure A.4: Univariate impulse response function: quarterly GDP deflator inflation. The estimation uses local projections; Newey-West standard errors with max$\{4, \ell-1\}$ lags are used to calculate the confidence intervals where $\ell = 1, 2, \ldots, 12$ denotes the horizon of the local projection.



(a) Inflation: Consensus-Level Estimates



(b) Inflation: Individual-Level Estimates

Figure A.5: Composite bias coefficients for one-quarter-ahead inflation forecasts estimated using an AR(4) model. The model used to estimate the IRF of forecast errors is

$$x_t - \mathbb{F}_{t-1}[x_t] = \alpha + \beta_1\{x_{t-1} - \mathbb{F}_{t-2}[x_{t-1}]\} + \cdots + \beta_4\{x_{t-4} - \mathbb{F}_{t-5}[x_{t-4}]\} + u_t.$$

The IRF is then obtained by iteration. Confidence intervals are generated with a parametric bootstrap.

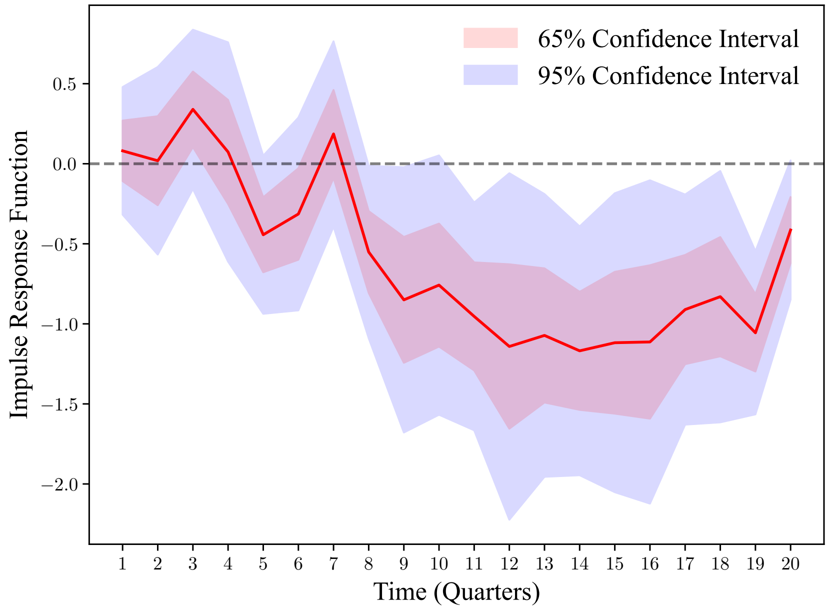


Figure A.6: Response of inflation to monetary policy shocks. The response is estimated by $\hat{\gamma_1}^{(\ell)}$ in the regression

$$x_{t+\ell} = \beta_0^{(\ell)} + \sum_{s=1}^{4} \gamma_s^{(\ell)} \mathrm{FFR}_{t+1-s} + \sum_{s=1}^{4} \beta_s^{(\ell)} x_{t+1-s} + u_{t+\ell},$$

for $\ell = 1, 2, \ldots, 12$, instrumenting the Federal Funds Rate, FFR, at time t with the Romer and Romer (2004) measure of a monetary surprise at time t.

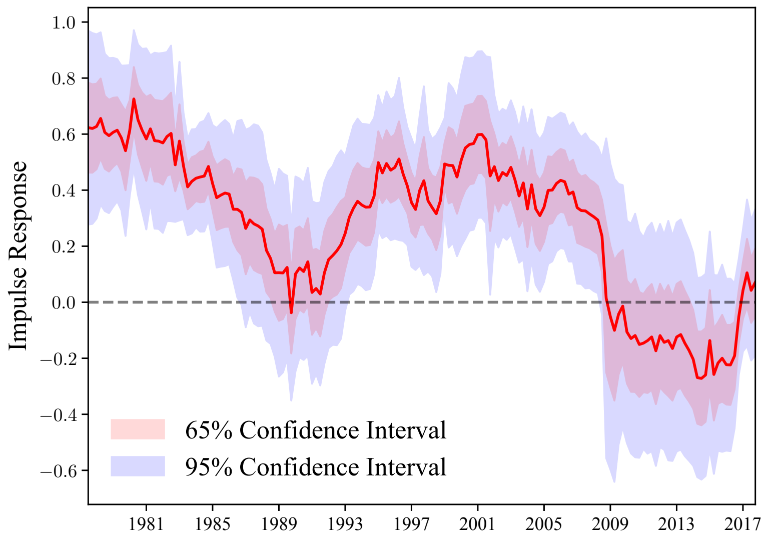


Figure A.7: Rolling window estimates of the univariate IRF of quarterly GDP deflator inflation (first lag). The estimation uses local projections and a window of 32 quarters (8 years); Newey-West standard errors with 4 lags are used to calculate the confidence intervals.

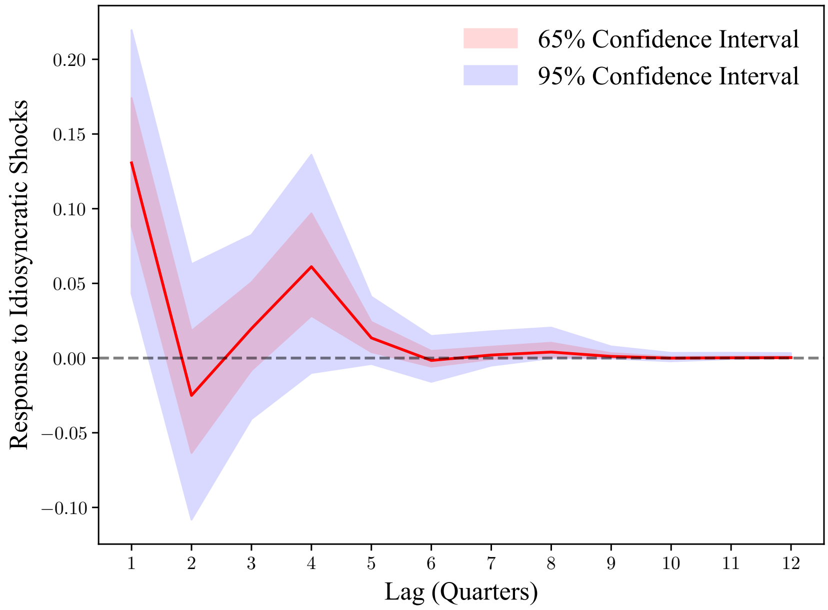


Figure A.8: Reaction of one-quarter-ahead inflation forecasts to idiosyncratic shocks, estimated using an AR(4) model. We first estimate

$$x_t - \mathbb{F}_{i,t-1}[x_t] = \alpha_i + \gamma_t + \beta_1\{x_{t-1} - \mathbb{F}_{i,t-2}[x_{t-1}]\} + \cdots + \beta_4\{x_{t-4} - \mathbb{F}_{i,t-5}[x_{t-4}]\} + u_{it}.$$

The IRF is then obtained by iteration. Confidence intervals are generated with a parametric bootstrap.

	Baseline	High Inflation	Recession	Great Moderation	Large FE
Lag 1	0.47	0.64	0.41	0.36	0.39
	(0.11)	(0.14)	(0.24)	(0.09)	(0.14)
		[2.60]	[-0.39]	[-0.80]	[-0.59]
Lag 2	0.40	0.32	0.51	0.07	0.37
	(0.11)	(0.17)	(0.17)	(0.12)	(0.15)
		[0.65]	[0.83]	[-2.38]	[0.08]
Lag 3	0.50	0.46	0.34	0.30	0.61
	(0.12)	(0.22)	(0.24)	(0.08)	(0.18)
		[0.87]	[-1.06]	[-1.39]	[1.43]
Lag 4	0.65	0.58	0.20	0.43	0.72
	(0.12)	(0.16)	(0.14)	(0.09)	(0.19)
		[1.50]	[-4.51]	[-1.51]	[0.42]

Table A.1: State dependence in the univariate IRF of quarterly GDP deflator inflation. Standard errors in parentheses; t-statistics of a test of no state dependence in brackets. High absolute values of the t-statistics indicate evidence of state dependence, with negative t-statistics indicating less persistence. Definitions of the different states are provided in Table 2.