ROBUST EXPECTATIONS ADAPTATION*

Te Bao[1] Peter Bossaerts[2,°]
Jiaoying Pei[2]

[1]Nanyang Technological University, Singapore [2]University of Cambridge, U.K.

This Version: July 2026

°Corresponding author: Faculty of Economics, University of Cambridge, Cambridge CB39DD, U.K.; Email: plb32@cam.ac.uk; Phone: +44.1223335235.

ABSTRACT

We amend adaptive expectations (AE) to make it robust in the face of nonstationarities such as those emerging in self-referential forecasting systems. We borrow insights from robust control in engineering and propose that the learning rate α in adaptive expectations is to be modulated in a way to minimize surprise relative to a reference model. As reference, we suggest the Kalman filter model recently used in a study examining how professional forecasters predict economic outcomes. We show how this prescribes changing α in the direction of autocovariance of prediction errors. We refer to the resulting forecasting model as *Robust Expectations Adaptation* (REA). Ours contrasts with the traditional prescription in reinforcement learning, which is to change α in the direction of the change in the size of the prediction error, the Pearce-Hall model, recently imported into the economics literature. Using almost 42,000 forecasts from experiments on self-referential economic markets, we discover that participants change α as in REA, but generally only if surprise is above the median level experienced. The Pearce-Hall model almost never fits the data.

JEL Codes: C91, D84, G41; **Keywords:** Adaptive Expectations, Forecasting, Kalman Filter, Self-Referential, Robust Expectations Adaptation, Model-Reference Based Adaptive Control

*We gratefully acknowledge financial support through a Tier 1 Grant from MOE of Singapore (RG121/23), the RIE2025 Industry Alignment Fund – Industry Collaboration Projects (IAF-ICP) (Award I2301E0026) administered by A*STAR and supported by Alibaba Group and NTU Singapore through the Alibaba-NTU Global e-Sustainability CorpLab (ANGEL), and the Leverhulme International Professorship at the University of Cambridge. We thank the following for valuable comments and discussions: Rava Azeredo da Silveira, John Duffy, Nobuyuki Hanaki, Harvey (Shijie) Huang, Konstantinos Ioannidis, Peiran Jiao, Jamie Lien, Kuang Pei, Luba Peterson, Stefan Trautmann, Rob Woods, Jubo Yan, participants of the International Conference in Computing and Economics, Society for Experimental Finance, Theoretical and Experimental Macroeconomics Workshop, 30th Theoretical and Experimental Macroeconomics Workshop, and seminar audience at Dongbei University of Finance and Economics.

1 Introduction

The traditional model for learning to forecast in economics is *adaptive expectations*. Mathematically, the model can be written as follows:

$$p^e_{t+1} = p^e_t + \alpha(p_t - p^e_t) \equiv p^e_t + \alpha\epsilon_t, \tag{1}$$

where p^e_t is the forecast of observation p_t in a time series of "objects" such as prices, indexed $t = 1, 2, ...$ $p_t - p^e_t$ is the prediction error, denoted ϵ_t. α is the learning rate (Cagan, 2000). In psychology, adaptive expectations is known as the *Rescorla Wagner Rule* (Rescorla and Wagner, 1972). It is also referred to as the "delta rule," which underlies *Prediction-Error Learning* in neuroscience (Montague et al., 1996; Schultz et al., 1997) and *Temporal Difference Learning* in machine learning (Sutton and Barto, 2018).

The updating rule can capture moving targets, i.e., situations where the variable to be forecast changes continuously. In such situations, the *Kalman Filter* provides a canonical example. There, α is referred to as the Kalman Gain. The Kalman filter has recently become the core model in studies of analysts' forecasting (Bordalo et al., 2020).

In economics, however, the object being forecast is often endogenous: the observation moves not only because of exogenous forces, but also because of the forecasts themselves (Marcet and Sargent, 1989). As a result, the forecast–observation system is said to be *self-referential*: inflation may be high, not only because of exogenous factors such as scarce production inputs, but also because high inflation expectations prompt more agents to bring forward purchases, thereby increasing demand and causing extra upward price pressure.

Self-referentiality leads to *nonstationarity*, by which we mean that the parameters of the data-generating process of the objects to be forecast change. It has long been appreciated that nonstationarity constitutes one of the main causes of poor out-of-sample performance of economic and financial forecasting models (Bossaerts and Hillion, 1999).

Here, we study how humans should and do adapt expectations to account for self-referentiality. We identify an approach, derived from control of dynamical systems in engineering, that produces *robustness* in expectations adaptation in the face of the nonstationarities arising from a self-referential system.

One way to capture nonstationarities is to track whether there are "regime switches" (Hamilton, 1989; Jansch-Porto et al., 2020). The idea is to focus on switchpoint detection, and once a switchpoint has been identified, to optimally adjust the learning rate α to the new "regime." However, the approach requires the forecaster to know beforehand what the possible "regimes" could be, and that there are sufficiently long and stable (stationary) episodes for renewed learning after switchpoints to have material impact. If the episodes are too short, there may be problems with policy convergence (Csáji and Monostori, 2008). A similar but simpler way is gain-scheduling, also known as heuristic switching model in economics. There, a strategy or an α is set offline *ex ante* to respond to different observed scenarios (Anufriev and Hommes, 2012). Such systems, however, can lead to poor performance or even failure when the environment changes in unanticipated ways (Tin and Poon, 2005).

In economics, the standard way to robustly deal with nonstationarities is to prepare for the worst (see, e.g., Hansen and Sargent, 2001). This is obviously not ideal, not only because the worst possible regime may rarely apply, but also because the worst regime may be unknown.

In psychology and neuroscience, robustness has been introduced by allowing the learning rate α to increase with *surprise*, i.e., with $|\epsilon_t|$. This resulting robust adaptive expectations technique is known as *Pearce-Hall Prediction-Error Learning* (Pearce and Hall, 1980). This mechanical rule is maladaptive, however, when outliers are salient and frequent (i.e., under leptokurtosis) and when they revert. In that case, the opposite adjustment to learning rates (i.e., reduction in α) is called for (see, e.g., d'Acremont and Bossaerts, 2016, for details). The Pearce-Hall rule has recently been studied in a macro-economic context, in Carvalho et al. (2023), building on the model initially developed by Marcet and Nicolini (2003). As we shall see, our data rejects the Pearce-Hall model.

In control theory, one proven way to obtain robustness in the face of unexpected nonstationarities is to *minimize surprise with respect to a reference model*. The resulting control technique, referred to as *Model-Reference Adaptive Control* (MRAC), has proven extremely useful in industrial applications in widely different contexts, such as robotics, aeronautics, or automotive control. In MRAC, hyper-parameters of the controller are modulated to ensure surprise is minimized. Here, surprise is defined relative to a reference model that represents the desired properties of the system and is used to supervise the adaptation of the controller (Nguyen, 2018; Bossaerts, 2018; Tin and Poon, 2005). In the context of forecasting, this translates into modulating the parameter α in the adaptive expectations model, so that surprise of forecasting errors is minimized relative to expected surprise in a reference model of the environment.

In MRAC, the reference model provides forecasts that are used to evaluate the outcomes from actual interaction with the environment. This contrasts with models such as Pearce-Hall where the past observation guides adaptation. MRAC was motivated by the concern that the past may provide bad guidance in nonstationary contexts. Indeed, the environment could have changed in ways such that even partial reliance on the past generates the worst possible performance. The latter was the case in d'Acremont and Bossaerts (2016), cited before. Instead, MRAC proposes that a carefully chosen reference model could generate more robust outcomes across a wide variety of possible environments.

We propose that economic agents use the Kalman filter model as reference to adjust the learning rate α in adaptive expectations. We refer to this type of modulation of α as *Robust Expectations Adaptation* (REA). We find that the optimal adjustment requires the forecaster to change α as a function of autocovariance of prediction errors. Using only 2 periods to estimate this autocovariance, we make this specific with the following adaptation rule:

$$\Delta\alpha_t \propto \epsilon_t \epsilon_{t-1}, \tag{2}$$

with positive constant of proportionality. That is, agents adjust the learning rate in proportion to the estimated autocovariance of prediction errors. The rule was proposed in the machine learning literature as well (though in an *ad hoc* way). There, it is known as the Delta-Bar-Delta Algorithm (Sutton, 1992).

The adaptation rule in (2) implies a departure from the conventional way α is selected. Rather than choosing α to minimize mean-squared error (MSE), as is conventional, REA has agents opt for the α that minimizes the autocovariance of their prediction errors. The intuition is that large autocorrelated errors signal that the model fails to capture structural features of the data — a failure that is especially costly when forecasting a nonstationary variable. In that case, the most recent error is not fully informative about news, since part of it is simply a predictable echo of past errors. Crucially, these two objectives need not coincide. Nursimulu and Bossaerts (2014) shows that estimating a reinforcement-learning model's learning rate by minimizing MSE leaves strongly autocorrelated errors in individual-level forecasts.

A major advantage of the adaptation rule in (2) is that it is independent of the parametrization of the Kalman filter model. Only the constant of proportionality depends on parameter values such as the observation error variance or the variance of state transitions in the model. Consequently, we can test whether agents use MRAC with the Kalman filter model as reference without having to be more specific about the model.

In the study cited previously, d'Acremont and Bossaerts (2016), it was shown that humans implemented this MRAC rule. Predictions reflected the use of the Kalman filter model as reference. The evidence suggested that the learning rate was changed in proportion to the autocovariance of prediction errors only when surprise ($|\epsilon_t|$) was sufficiently large.

We study almost 42,000 *price forecasts* from a series of experimental markets (Bao et al., 2012, 2013, 2017; Bao and Hommes, 2019; Bao et al., 2024). There, the nature of self-referentiality was controlled: in the *positive feedback* treatment, increased price expectations led to higher actual prices, while in the *negative feedback* treatment, increased price expectations caused lower prices.[1]

We study estimates of changes in the learning rate: their *direction* (up/down) and their *magnitude* (signed change). We reject the constant learning rate assumed in Bordalo et al. (2020). We reject the Pearce-Hall model as well because learning rate changes are found to be insensitive to the changes in size of the prediction error ($|\epsilon_t|$). In contrast, we find convincing evidence that the learning rate is modulated by the sign of the prediction error autocovariance ($\epsilon_t\epsilon_{t-1}$), as predicted by REA. But, as in d'Acremont and Bossaerts (2016), we discover that this modulation is only observed when surprise is sufficiently high. In particular, it emerges when surprise is above the median level agents experience individually in each experiment.

The remainder of this paper is organized as follows. Section 2 provides a more detailed literature review. Section 3 presents the theoretical model and derives the optimal robust expectations adaptation (REA) formula. Section 4 explains the data. Section 5 lays out the testable hypotheses. Section 6 presents the empirical results. Robustness of the results is discussed in Section 7. Section 8 provides concluding remarks.

2 Literature Review

We present the literature review in a systematic way, structured around topics relevant to the REA model.

2.1 Rational Expectations

There is an extensive literature in economics dedicated to understanding how human agents predict future outcomes. This is in large part motivated by the fact that economic systems are self-referential: predictions affect choices and choices affect economic outcomes (prices, trade volume, ...), making modeling economic outcomes challenging. We will discuss why this self-referentiality makes modeling economic outcomes challenging later.

[1]Some experiments in the cited literature required participants to predict quantity instead of price. Those experiments were not included for this study.

One shortcut to simplify the analysis of economic systems is to assume rational expectations (Muth, 1961), but this would require agents to both predict future states with unbiased beliefs (Bossaerts, 2004) and have perfect foresight, i.e., know prices and volumes in each future state (Bossaerts et al., 2024). This has led economists to investigate alternative approaches, such as adaptive expectations or Kalman filters (to be discussed in the following). The assumption that humans form forecasts in self-referential systems according to rational expectations is misleading. Not only may forecasts be biased (see Bordalo et al. (2020)), but it appears that forecasts of prices of financial assets exhibit excessive volatility, just as prices themselves are excessively volatile (Nursimulu and Bossaerts, 2014).

2.2 Adaptive Expectations

As mentioned before, economists have studied alternative forecasting models besides rational expectations. Arguably the most influential one is adaptive expectations, whereby predictions are updated based on prediction errors (Cagan, 2000). Kalman filter analysis (to be discussed below) shows that such an adaptive updating rule may be optimal in a stationary Gaussian world, or a stationary world with quadratic loss functions.

Rescorla-Wagner Rule. In psychology and neuroscience, adaptive expectations is known as the *Rescorla-Wagner Rule* (Rescorla and Wagner, 1972). Some also refer to it as the "Delta Rule," which, if the learning rate can be adjusted flexibly, will emulate Bayesian learning (Nassar et al., 2010).

Prediction-Error Learning. In machine learning and computational neuroscience, delta-rule learning has become known as "prediction error learning." In its most sophisticated formulation, the prediction error derives from Bellman's equation, and hence, can be used not only to learn to forecast but also to adapt one's actions to become dynamically optimal. This then is the core of modern machine learning (Sutton and Barto, 2018). It deserves mention that such sophisticated "temporal difference" prediction errors have been found to be encoded in the dopaminergic system of the primate brain (Montague et al., 1996; Schultz and Dickinson, 2000).

Kalman Filter. A popular adaptive expectations rule to predict the position/value of a moving target is the *Kalman filter*. The target moves independently of the forecasts, so the system is autonomous, i.e., not self-referential.[2] Driving processes are Gaussian or the forecaster's loss function is quadratic. The Kalman filter has recently been applied successfully in modeling economic forecasts by professionals (see Bordalo et al., 2020). The pure Kalman filter, where learning rates are set optimally as a constant, generally predicts lower learning rates than inferred from professionals' forecasts. Then again, as mentioned before, the Kalman filter is not meant to apply to self-referential systems, to which we turn now.

[2]In more sophisticated versions, the target can be controlled by the predictor, but that does not make the system self-referential since the agent knows how her actions influence the target.

2.3 Self-Referentiality and Nonstationarity

In economics, forecasts determine outcomes, and hence, outcomes are endogenous. When there are prediction errors, they lead to updated forecasting and hence, changes in the data-generating process. Consequently, self-referential systems tend to be nonstationary – that is, the parameters of the data-generating process for the objects being forecast change over time – unless there are no mistakes in the forecasting rule, meaning expectations are "rational."[3] A large literature exists addressing whether and when economic systems can settle on rational expectations, starting from Marcet and Sargent (1989). The nonstationarities, caused by the mistakes, are not innocuous even if learning eventually converges: they invalidate standard tests of rational expectations models that ignore the transient mistakes (see Bossaerts, 1995).

Self-referentiality is especially important when cohort size is small, where only a few agents determine the outcomes. If the cohort size is larger, agents will be price takers and self-referentiality will not exist. As we shall see in Section 4, we analyze experiments conducted with a maximum of 9 participants to better study learning under self-referentiality.

Regime-Switching Model. Because of the foregoing discussion, it is no surprise that econometricians have been long interested in modeling nonstationarities in economic time series. Most statistics have been developed in the context of a stationary regime switching paradigm, illustrated by the influential Hamiltonian switching model (Hamilton, 1989). In machine learning, the problem of jointly learning to forecast and to select the optimal algorithm upon detecting a regime switch has become the object of much study recently (see, e.g., Jansch-Porto et al., 2020).

The problem with stationary regime switching models is that they do not always fit economic reality. Economic history generally does not repeat itself (the system is non-ergodic, and hence, there does not exist a stationary distribution with which to analyse the data statistically). This is not to say that regime switching models cannot be used successfully to understand forecasting in economics; a nice counterexample is Carvalho et al. (2023).

2.4 Surprise Based Learning

Surprise. Surprise has to be discussed due to its central importance in neuroscience, not only because neural signals correlating with surprise are ubiquitous (Preuschoff et al., 2008), but also because of its critical role in understanding psychopathology (Paulus et al., 2003). There are various ways to mathematically express surprise, from concepts more familiar to economists such as the size of risk prediction errors – the driving force behind GARCH models (Engle, 2001), to concepts from probability theory such as log posterior probabilities, or concepts from information theory such as entropy (see the discussion in Modirshanechi et al., 2022). Surprise simply measured as the size (absolute value or squared value) of the prediction error has

[3]The influential Lucas model of consumption/saving and asset pricing (Lucas, 1978) is perhaps the best example that it is possible, under rational expectations, for a self-referential economic system to be stationary. In an experimental setting, various aspects of the Lucas model have been found to be robust to small mistakes. Excess price volatility emerges, however (see Asparouhova et al., 2016). With simulations, one can also prove that mistakes lead to spuriously high risk premia (see Adam et al., 2016).

played a prominent role in understanding animal learning in psychology and neuroscience, as well as having influenced machine learning, as we now discuss.

Pearce-Hall Prediction-Error Learning. One influential model where surprise plays a crucial role to ensure robustness in the face of changes in the environment is the Pearce-Hall model. There, the learning rate in the Rescorla-Wagner adaptive expectations rule, or in more sophisticated Temporal Difference (machine) learning, increases in proportion to the size of the prediction error (surprise) (Pearce and Hall, 1980). This ensures that learning accelerates – prediction errors more distant in the past are downweighted – upon outliers.

It can be maladaptive, however, in some environments where reverting outliers are a feature of the environment. This is common in financial markets, where it has been estimated that 2/3 of outliers revert over the short run (Brogaard et al., 2014). As mentioned in the Introduction, the Pearce-Hall rule accomplishes exactly the opposite of what is required: rather than ignoring outliers, one reacts in a more extreme way. We test and reject this model in Section 6.

2.5 Reference-Model Based Learning

Model-Reference Adaptive Control. Control engineers have long been interested in how to ensure robustness in standard dynamic control techniques. The problem is that full optimization may lead to fragility – a small turbulence in the environment may lead to inferior choice or even instability. In situations where the environment fluctuates too much, there is no point in trying to optimize over the short epochs when there are no changes in the environment. One approach, namely, to plan for the worst environment, has become popular in economics (Hansen and Sargent, 2001; Wu and Sun, 2023), but has been deemed excessively pessimistic.

Since the 1950s, control engineers have advanced an approach referred to as Model-Reference Adaptive Control (MRAC). The approach can best be understood as adding a supervisory system to traditional feedback control. The goal is to adjust the parameters of the controller when the outcomes it generates through interaction with the environment are surprising. Surprise is measured here as the deviation of observed outcomes from those predicted by the supervisor using a *reference model* of the environment. Industrial applications exist, among others, in robot impedance control (Zhang and Wei, 2017). Theoretical analysis of MRAC instances focus on specification of the right reference model and assurance that control remains stable for a wide range of realistic environments. A classical textbook is Nguyen (2018).[4]

[4]Some may not immediately recognize MRAC in our analysis because the reference model is generally not made fully explicit in traditional accounts, instead limiting specification to that of a desired trajectory. In addition, the goal in early versions of MRAC was to eventually match the desired trajectory, entirely eliminating surprise. Later modeling acknowledges explicitly that desired trajectories may not be reachable, and attention has turned to bounding of tracking errors. While traditional MRAC analysis assumes away randomness, we explicitly allow for stochastics, thereby introducing new challenges. Stochastics have recently been introduced into engineering applications as well. For instance, Herzallah (2020) provides a fully probabilistic MRAC design, allowing the environment as well as the effect of the controller to be stochastic, and using KL divergence as a measure of surprise. Here, we use as surprise metric the – simpler – squared deviations from expectation. See also Bossaerts (2018); Berrada et al. (2024).

The idea of MRAC is analogous to that of a human supervisor monitoring and adapting an electronic autopilot. Camerer et al. (2024) provides an example of the autopilot supervision idea to modeling switches from habitual to goal-directed behavior. An application of MRAC to portfolio investment can be found in Berrada et al. (2024).

In sensorimotor control, reference models (called "teaching models") are known to make movements of limbs fast and flexible despite an ever-changing world, even if those movements can be shown to be sub-optimal (Shadmehr and Mussa-Ivaldi, 1994; Kawato, 1999; Franklin and Wolpert, 2011). Applications of MRAC can also be found in the software behind neural prosthetics (Musallam et al., 2004). Kawato (1999) hypothesized that higher cognition could also be driven by reference models.[5] The evidence presented in this paper appears to confirm his hypothesis.

Reference Model. MRAC obviates the need to have correct cause-and-effect predictions about future states and to update those when the environment changes, unlike in conventional probabilistic predictive processing models (Mansell et al., 2025). It does so by introducing a *reference model*, which can be viewed as an idealized world where control is easier to accomplish than in the real world. The agent then guides its interaction with ("control" of) the real world to minimize expected surprise relative to the reference model. In a sense, the agent wants outcomes that look like those in his idealized world, even if the agent could do better in particular circumstances. In motor learning, for instance, the reference model may imply moving a limb along a smooth, straight-line trajectory toward a target regardless of interfering forces. The resulting control may not be optimal: in certain situations, the agent may prefer to give in to those interfering forces, because they eventually lead to the same goal (trajectory endpoint) with less energy. But it may take time to realize that the agent has landed in such a situation. In the meanwhile, doubt about the true environment leads her not to give in to interfering forces because in other situations that may cause her to go completely astray (see Shadmehr and Mussa-Ivaldi, 1994).

Robust Expectations Adaptation. One can apply MRAC to forecasting. This was first done to explain the goal of certain brain regions in tracking surprise (Bossaerts, 2018). We assume that humans are adapted to stationary environments and therefore adopt a stationary Kalman filter as their reference model. Essentially, we hypothesize that humans fail to recognize nonstationarity unless told explicitly. Explicit evidence in favor of this hypothesis can be found in, e.g., Payzan-LeNestour and Bossaerts (2015). Instead, they behave as if the environment were stationary. We build on this to model forecasting prices in a market-like setting and test the model on data from a series of experiments. We refer to the resulting forecasting model as Robust Expectations Adaptation (REA).

Satisficing. The goal of MRAC is *not* to optimize, but to *ensure robustness*, which means that the "supervisor" will interfere only when the controller generates prediction errors (surprise) *different from* the reference model. In the model to be presented here, interference takes the form of adjustment of the learning rate. Note that MRAC also predicts interference when surprise is *smaller* than expected. In engineering,

[5] From p. 724 of Kawato (1999): "In the near future, the author expects major breakthroughs in the concepts and computational theories of internal models entering into cognitive domains such as communication, thinking, and consciousness, on the basis of their firm foundations in sensory motor integration."

where MRAC originated, this makes sense because too small a surprise (relative to the reference model) also signals that the environment may have changed, not only large surprise.

By contrast, the behavior of an MRAC agent reflects *satisficing* (Simon, 1955), featuring aversion to intervention upon small surprise. Satisficing comes from the argument that one may need to attend to other tasks that require more attention. As such, adaptation may be ill-advised when inattention is called for because of limited attention scope (Sims, 2003).[6]

Here, we account for the tension between the needs for attention and adaptation when analyzing our data, by investigating the extent to which surprise prompts participants adjusting their learning rates.

3 Theory

There are two modules in MRAC, the *controller* and the *reference model*. In our application, the controller is the entity that generates the actual forecasts. (Engineers often refer to the controller as the "plant.") The reference model is used as a supervisory tool: it generates its own predictions and its own prediction errors. If the actual prediction errors from the controller deviate too much from the prediction errors generated from the reference model, the supervising entity intervenes by adjusting the controller's prediction model to bring performance in line with that of the reference model.

Reference Model. Here, we propose REA as an implementation of MRAC with the standard Gaussian Kalman filter model as reference. The supervisor uses this model to regularly adapt the learning rate α of the controller, who uses standard adaptive expectations as in (1).

Specifically, the supervisor assumes a world in which observed scalar prices p are driven by an unobserved state x following an autoregressive process:

$$p_t = x_t + o_t,$$

$$x_{t+1} = ax_t + s_t,$$

where the errors (o_t, s_t) are independent of x_t, Gaussian, and uncorrelated over time, such that $E[o_t] = E[s_t] = E[o_t s_t] = 0$, $E[(o_t)^2] = v^2$ and $E[(s_t)^2] = w^2$.

In the reference world, the standard Kalman filter provides optimal forecasts p^*_{t+1}:[7]

$$p^*_{t+1} = E[p_{t+1}|p_t, p_{t-1}, ...] = a\left(p^*_t + \alpha^*(p_t - p^*_t)\right),$$

where α^* is a constant if we assume that the Kalman filter has reached its steady state. For example, if $a = 1$ (the state follows a unit root process, like in adaptive expectations), then $\alpha^* = 1$, and hence p^e_{t+1} and x_t are co-integrated. α is traditionally referred to as "Kalman Gain," but to tie our analysis more closely to

[6] Notice that our interpretation of Herbert Simon's "satisficing" is different from other accounts in the economics literature, which argue that it emerges because of optimal search when search is costly (see, e.g., Caplin et al., 2011).

[7] Forecasts are optimal in that they minimize any strictly convex loss function.

adaptive expectations in economics, we call it the "learning rate." α^* is then the optimal learning rate in the reference world.

At the steady state, the expected size of the prediction error under the optimal learning rate, measured as the squared value of $p_{t+1} - p^*_{t+1}$, is constant:

$$E[(p_{t+1} - p^*_{t+1})^2] = Z > 0.$$

Controller. The controller uses standard adaptive expectations with a learning rate α, as in (1). Allowing for $a \neq 1$, this means:

$$p^e_{t+1} = a\left(p^e_t + \alpha(p_t - p^e_t)\right). \tag{3}$$

The core idea of MRAC is that α should change as a function of *expected surprise*. The supervisor, interpreting the data through the lens of a standard Gaussian Kalman filter model and therefore expecting constant squared prediction errors Z, adjusts the learning rate α of the controller whenever the actual squared prediction errors deviate from expectations. This adjustment will be optimal, in the sense that it will minimize expected surprise.

As before, define the prediction error generated by the controller as follows:

$$\epsilon_{t+1} \equiv p_{t+1} - p^e_{t+1}.$$

Surprise at $t+1$ is defined as $(\Omega_{t+1})^2$, where

$$\Omega_{t+1} = (\epsilon_{t+1})^2 - Z. \tag{4}$$

The goal is to choose α to minimize *expected surprise*

$$\min_{\alpha} E[(\Omega_{t+1})^2]. \tag{5}$$

Here, the operator $E[\cdot]$ is to be interpreted as the expectations of future surprise based on past observations. We elaborate on these expectations later. Under the usual conditions for interchange of differentiation and integration, the following first-order condition obtains:

$$-4aE[\Omega_{t+1}\epsilon_{t+1}\epsilon_t] = 0. \tag{6}$$

or:

$$E[(\epsilon_{t+1})^3\epsilon_t] - ZE[\epsilon_{t+1}\epsilon_t] = 0.$$

If skewness of prediction errors is expected to be uncorrelated with past (signed) prediction errors,[8] i.e.,

[8]More general adaptation rules can be derived. For instance, if future prediction errors tend to be negatively skewed following positive prediction errors (Chen et al., 2001), as may be the case in forecasting stock prices, then mild negative autocovariance should be tolerated. Indeed, if $E[(\epsilon_{t+1})^3\epsilon_t] < 0$, then the first-order condition changes to $E[\epsilon_{t+1}\epsilon_t] = E[(\epsilon_{t+1})^3\epsilon_t]/Z < 0$. Alternatively, without imposing the assumption that $E[(\epsilon_{t+1})^3\epsilon_t] = 0$, (6) can be written as $E[((\epsilon_{t+1})^2 - Z)(\epsilon_{t+1}\epsilon_t)] = 0$. This implies that surprise is minimized

if

$$E[(\epsilon_{t+1})^3\epsilon_t] = 0,$$

then expected surprise is minimized by choosing α such that autocovariance of prediction errors is set to zero:

$$\text{Choose } \alpha \text{ such that : } \; E[\epsilon_{t+1}\epsilon_t] = 0.$$

The MRAC supervisor will therefore monitor autocovariance of the prediction errors generated by the controller, and adjust the learning rate in the direction of minimal autocovariance. If autocovariance is positive, the supervisor will infer that the controller does not learn fast enough (undershooting follows undershooting), and the supervisor will increase the learning rate α. Conversely, if autocovariance is negative (overshooting follows undershooting), then overreaction to prediction errors is inferred, and α will be decreased.

Note that this differs from learning models derived from minimizing mean-squared error (MSE): MSE penalizes the magnitude of errors but is silent on their temporal structure, whereas autocovariance targets precisely the predictable component of the errors. REA thus drives out predictable structure rather than merely shrinking error size.

The adaptation policy boils down to:

$$\Delta\alpha_t \propto E(\epsilon_t\epsilon_{t-1}). \tag{7}$$

Let us now clarify what the expectations operator $E[\cdot]$ in (7) refers to, namely, prediction errors as experienced by the supervisor over the recent past. The supervisor could estimate this autocovariance with as little data as the prediction error in the current or immediately preceding period, i.e., as $\epsilon_t\epsilon_{t-1}$. This is the estimate we will use in our empirical analysis[9].

As emphasized in the Introduction, a major advantage of the adaptation rule in (7) is that it is independent of the parametrization of the Kalman filter model. Only the constant of proportionality depends on the specific values an agent assigns to parameters such as v^2 or w^2. Consequently, we can test whether agents use MRAC with the Kalman filter model as reference without having to make more specific assumptions about the model.

As mentioned before, this learning rate adjustment rule is the one proposed in Sutton (1992) and referred to as the Delta-Bar-Delta algorithm in machine learning. Up until now, no specific motivation has been given, except that it appeared to be reasonable to produce forecasts so that prediction errors become uncorrelated. Here, we provide a justification for this adjustment rule in terms of MRAC.

The supervisor does not have to constantly monitor autocovariance. A more sensible rule that recognizes

---

by choosing α such that future autocovariance and surprise are expected to be orthogonal.

[9] As there are limited data points per participant-experiment, with a minimum of 20, this limits our ability to estimate autocovariance beyond two periods. For example, estimating autocovariance over 5 periods would require dropping 5 periods of data. This leads to a loss of valuable information in half of the experiments in the dataset, where adjustment happens in the first few periods (see right panel of Figure 2). We therefore use only 2 periods of observations. As we shall see in the results, these are sufficient to show evidence supporting REA.

bounds on cognitive effort would be to adjust the learning rate only if surprise reaches a minimum level. Estimating this cutoff point involves conducting kinked regression with the fat-tailed α resulting from small ϵ_t, which makes the estimation extremely unreliable. As a remedy, we split the data per participant-experiment in half to fairly compare the adjustment in learning under small and large surprise.

We hypothesize that participants adjust the learning rate only if surprise surpasses the median surprise experienced in each participant-experiment. This leads to the following amended rule:

$$\Delta\alpha_t \propto E(\epsilon_t\epsilon_{t-1}) \text{ when } \Omega_t \geq \text{median}(\Omega_t). \tag{8}$$

The amended adaptation policy could be useful in multitasking situations under limited attention, akin to Sims (2003). Not monitoring tasks that generate less than median expected surprise would be an effective way to ensure all attention goes to other tasks that generate more surprise than expected.

Remark that since the reference model has constant expected surprise (Z), there is no difference between surprise relative to the reference model and surprise as defined in the Pearce-Hall model (we will be more specific later on). The difference between Pearce-Hall and MRAC adaptation then boils down to the following: the Pearce-Hall agent increases the learning rate as surprise increases; the MRAC agent changes the learning rate in proportion to experienced autocovariance of prediction errors when encountering large surprises.

4 The Data

We study almost 42,000 forecasts from Learning-to-Forecast Experiments (LtFEs) in which participants were tasked with predicting a market variable we label the "price" – determined by aggregating individual forecasts. The idea was to emulate price formation in self-referential markets, where prices were calculated as a function of the average of individual forecasts. The experiments distinguished three major Treatments: (i) Positive Feedback, where increases in average price forecasts caused the outcome (price) to increase as well; (ii) Negative Feedback, where increases in average price forecasts caused the price to decrease; (iii) Mixed Feedback, where price formation was based on price forecasts by two groups, one whose price forecasts caused positive feedback, and another group whose price forecasts generated negative feedback.

A typical market in the LtFEs comprised 6 to 9 participants. Participants in each treatment made between 20 and 50 consecutive forecasts of the price.[10] Participant payoffs were a decreasing quadratic function of their prediction error. They had access to the entire history of their own past predictions and the realized market price, and hence, the history of personal prediction errors. This information was presented graphically and numerically, as illustrated in Figure 1. Time pressure was minimal: in early rounds, participants sometimes took up to 10 minutes to submit forecasts, but this dropped to five seconds or less in later rounds, depending on the treatment. In later experiments, forecast decision time was capped

[10]The cohort size was 6 in all experiments, except Experiments 11 and 12, where 3 additional suppliers (generating negative feedback) were introduced in the positive feedback condition, such that the total cohort size increased to 9. Each experiment in our dataset consists of 50 consecutive forecasts from each participant, except Experiments 1–3 and 13–15, which used a within-subjects design and include approximately 20 consecutive forecasts per participant in each treatment.

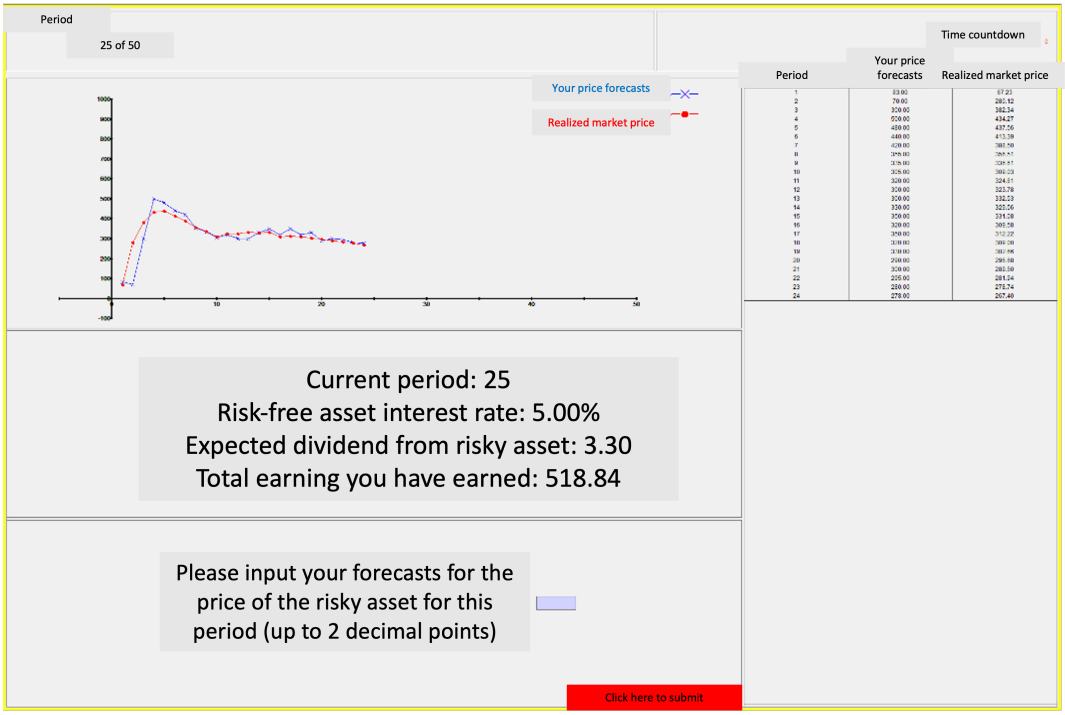

Figure 1. Sample computer interface from Bao et al. (2024). The market price is computed using the following formula: $p_t = \frac{1}{1+r}(\bar{p}_t^e + d) + e_t$, where $\bar{p}_t^e$ denotes the average forecast of the cohort, $r = 0.05$, $d = 3.3$, and $e_t \sim \mathcal{N}(0, 1)$. This function has a equilibrium point ($p_t = \bar{p}_t^e$) when $\bar{p}_t^e = 66$. To the right is a list of past periods, the participant's forecasts in those periods, and the prices obtained as a result of both the participant's forecasts and those of others in the cohort. While participants do not know the functional form of the pricing function – or that it is perturbed by Gaussian noise – they are given the values of the parameters (risk-free rate, dividend) which enter this function.

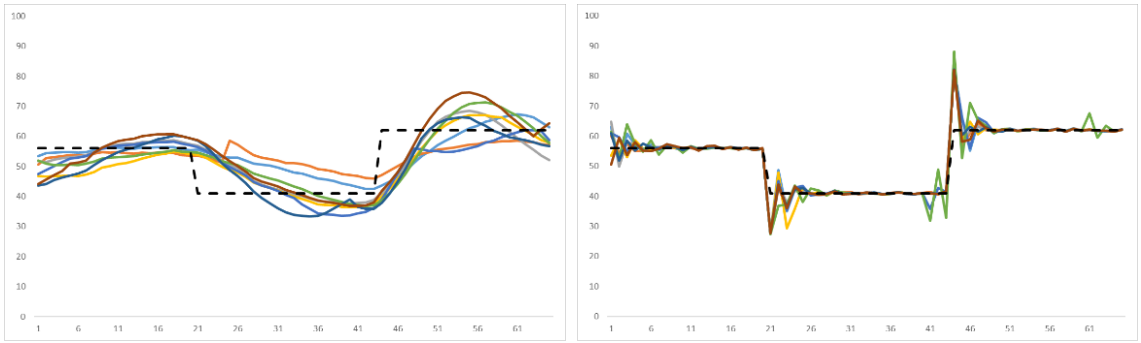

Figure 2. Sample price dynamics in positive (left panel) and negative feedback (right panel), representing markets from Experiments 1-3 (positive feedback), and markets from Experiments 13-15 (negative feedback), respectively. The dashed line shows the equilibrium price under rational expectations.

at one minute.

The dataset was constructed from experiments in five studies: Bao et al. (2012, 2013, 2017), Bao and Hommes (2019), and Bao et al. (2024). Considering each treatment in each study as a separate experiment, we thus have data from 18 experiments. The total number of price forecasts that we obtain is 41,490. They constitute forecasts generated by 801 individuals. Some of these individuals participated in more than one "experiment" because we define "experiment" in terms of treatment, not the usual experimental session, and many experimental sessions encompassed multiple treatments.

Across the five studies, there are 10 experiments with positive feedback markets, which we refer to as Experiments 1-10. Experiments 11 and 12 feature mixed feedback. Experiments 13 to 18 feature negative feedback. Appendix A provides detailed information on the 18 experiments. Sample experiment instructions can be found in Appendix B.

In the negative feedback treatment, the forecasts rapidly converged to the outcomes. With positive feedback, persistent bubbles and crashes emerged. The evidence suggested that these findings were robust to the size of the cohort (number of participants in a market).[11] Typical price dynamics in positive and negative feedback treatment can be found in Figure 2.

5 Testable Hypotheses

We have three testable hypotheses.

Hypothesis 1: α is constant. This is consistent with Bordalo et al. (2020), where the standard Kalman filter model is assumed. We expect to reject this hypothesis since REA predicts that the learning rate will change.

Hypothesis 2: α changes in the direction of the change in surprise ($|\epsilon_t|$). This is expected under the Pearce-Hall model (Pearce and Hall, 1980). We expect to reject this hypothesis since REA predicts that the learning rate will change as a function of autocovariance of predictions errors instead.

Hypothesis 3: α changes as a function of the autocovariance of the prediction errors ($\epsilon_t\epsilon_{t-1}$). We expect to find support for this hypothesis; however, we expect surprise ($|\epsilon_t|$) to modulate the impact of autocovariance.

[11] The results were also robust to changing from price to quantity or return forecasting, or changing forecast horizon; see Bao et al. (2021). We do not discuss results from these robustness tests, however.

6 Results

We start our analysis with the standard adaptive expectations model in economics, where the learning rate α is constant. Subsequently, we allow α to change as a function of the squared prediction error, as in the Pearce-Hall model in machine learning. We then explore the MRAC model, whereby α should change in the direction of autocovariance of prediction errors.

6.1 Imposing a Constant Learning Rate

Figure 3 plots the estimated learning rate α with 95% confidence intervals. The model is estimated on each of the 18 experiments separately using random-effects at the participant level. Mathematically, letting $p_t^{e,i}$ denote the price forecast of participant i in period t, and ϵ_t^i the prediction error $(= p_t - p_t^{e,i})$, we estimate:

$$p_{t+1}^{e,i} - p_t^{e,i} = \kappa_0^i + \kappa_1 \epsilon_t^i + \xi_{t+1}^i. \tag{9}$$

Here, the estimate of κ_1 allows us to infer the (constant) learning rate α, as in (1). Random-effects modeling is reflected in the superscript i to the parameter κ_0, (in principle, this parameter should equal zero). ξ_{t+1}^i is a mean-zero noise term. Confidence intervals are robust for clustering at the participant level.[12] In addition, confidence intervals have been adjusted for multiple hypothesis testing, to reflect the fact that 18 tests are displayed.[13]

In Experiments 1-10, feedback is positive. Imposing a constant α assumption, we find that α is either not significantly different from 1 or significantly above 1. This is consistent with evidence from the field. Bordalo et al. (2020) likewise shows that professional forecasters of macroeconomic time series exhibit excessively high learning rates. There, it is shown that the optimal Kalman filter fit to historical data would imply a substantially lower learning rate. The finding is also consistent with experiments in neuropsychology: in a model that allows the learning rate to change across epochs but remain fixed within an epoch, Lee et al. (2020) show that learning rates tend to be high.

We will not determine whether the level of the learning rate is inappropriate by fitting an optimal constant-gain Kalman filter to the data, as in Bordalo et al. (2020). Instead, we use autocorrelation of prediction errors as evidence that the adaptive expectations formulation with a constant learning rate is mis-specified.

We observe significant drift in prices in many experiments. This is evident, for instance, in the left panel of Figure 2. This drift implies that prediction errors were positively autocorrelated. Table 1 documents that positive autocorrelation appears in most of the positive-feedback experiments (Experiment 1 to 10). There, participants actually under-adjusted to forecast errors, implying that their learning rate was too small. Similarly, drift appeared in prices in the mixed-feedback experiments as well (Experiments 11 and

[12] In this paper, a participant is defined as an individual in a treatment. In some experiments, participants remained the same across treatments. For the purpose of cluster correction of standard errors, the same individual across different treatments is treated as different participants. Robustness tests demonstrate that the results would not change qualitatively if we had clustered at the cohort level. Tests available from the authors upon request.

[13] Bonferroni-corrected 95% confidence intervals are within 2.99 standard errors of the coefficient estimates.

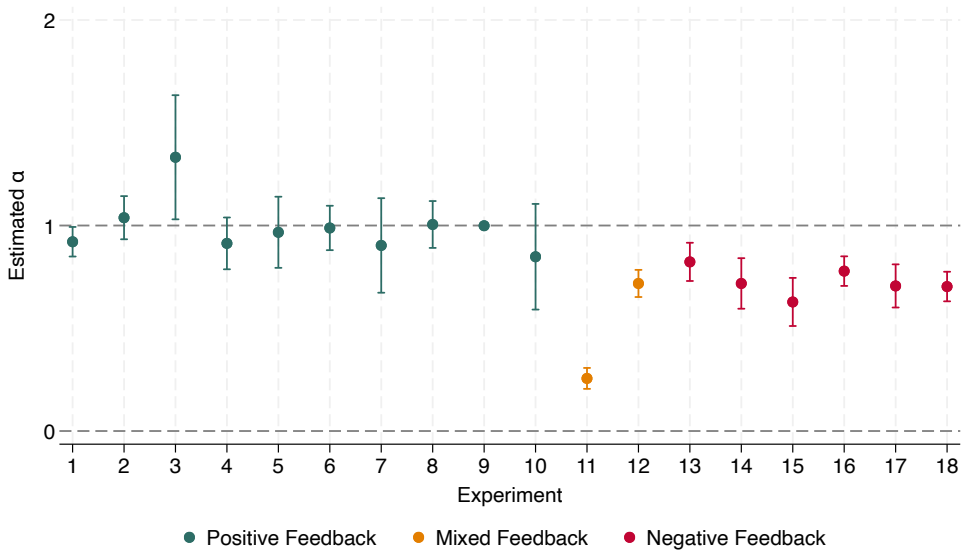

Figure 3. Estimates of κ_1 in (9), per experiment, with 95% confidence intervals Bonferroni-corrected for multiple hypothesis testing (18 tests). Standard errors are corrected for clustering at the participant level.

Table 1. Autocorrelation of prediction errors per experiment, estimated as the sample mean of $\epsilon_t\epsilon_{t-1}$ divided by the sample mean of ϵ_t^2, across all participants and periods. P-values based on standard errors computed as $1/\sqrt{T}$, where T is the number of valid $\epsilon_t\epsilon_{t-1}$ observations per experiment.

Experiment	**Sample Autocorrelation**	***p*-value**
Positive feedback		
1	0.101	0.002
2	0.003	0.917
3	0.150	< 0.001
4	0.103	< 0.001
5	0.052	< 0.001
6	0.020	0.169
7	0.108	< 0.001
8	0.094	< 0.001
9	0.002	0.918
10	0.123	< 0.001
Mixed feedback		
11	0.722	< 0.001
12	0.571	< 0.001
Negative feedback		
13	-0.246	< 0.001
14	-0.224	< 0.001
15	-0.187	< 0.001
16	-0.357	< 0.001
17	-0.218	< 0.001
18	-0.240	< 0.001

12). This translates into positive autocorrelation in the prediction errors; again see Table 1. There too, participants should have adopted to a higher learning rate. By contrast, in the negative-feedback experiments (Experiments 13 to 18) where αs are consistently below 1, prediction errors were negatively autocorrelated, as documented in Table 1. This is also evident in the example in the right panel of Figure 2, where prices overshoot at times. The negative autocorrelation implies that forecasts reacted too strongly to prediction errors.

We summarize as follows.

Result 1: We reject Hypothesis 1 that α is constant since prediction errors exhibit positive or negative autocorrelation, implying either under- or over-reaction.

Later we will directly reject Hypothesis 1 by allowing α to vary explicitly as a function of autocovariance of prediction errors. Such variability would be consistent with REA.

6.2 Pearce-Hall Adaptation of Learning Rates

To test to what extent αs increase with the size of the prediction error, as prescribed by the Pearce-Hall Model, we first estimate learning rates, as follows. The approach may seem crude, but any estimation errors should be absorbed by the error term in the regressions.

We proceed as follows. Consider the controller's updating equation (3), where we add subscripts t to (i) the autoregressions a, to reflect that self-referentiality may cause nonstationarity in the outcome generating process, and (ii) the learning rate α, to reflect that α may be updated by the supervisor:

$$p^e_{t+1} = a_{t+1}\left(p^e_t + \alpha_{t+1}(p_t - p^e_t)\right) = a_{t+1}\left(p^e_t + \alpha_{t+1}\epsilon_t\right).$$

We can solve for α_{t+1}, and after taking first differences, express the change in the learning rate as a function of forecast updates and prediction errors:

$$\alpha_{t+1} - \alpha_t = \frac{p^e_{t+1} - p^e_t}{a_{t+1}\epsilon_t} - \frac{p^e_t - p^e_{t-1}}{a_t\epsilon_{t-1}} - \left(\frac{1 - a_{t+1}}{a_{t+1}\epsilon_t}p^e_{t+1} - \frac{1 - a_t}{a_t\epsilon_{t-1}}p^e_t\right).$$

We assume that participants apply an autoregression coefficient a that equals 1. This is implicitly assumed in the standard adaptive expectations updating equation (1).[14] It allows us to estimate the change in α as the change in the ratio of forecast updates and prediction errors:

$$\alpha_{t+1} - \alpha_t \approx \frac{p^e_{t+1} - p^e_t}{\epsilon_t} - \frac{p^e_t - p^e_{t-1}}{\epsilon_{t-1}}.$$

Directly estimating the learning rate from the observed prediction errors is problematic, however, sinceϵ_t is often close to zero — and, though unlikely, may even be exactly zero[15]. Hence, the estimate of the learning

[14]It also means that participants perceive the future as a martingale, as any Bayesian would – Bayesian posterior beliefs form a martingale under the Bayesian's own beliefs (Doob, 1949).
[15]The prediction error is zero in 98 out of our 41,490 observations.

rate would diverge, and the distribution of estimates of changes in the learning rate will become fat-tailed. We could accommodate fat tails in regressions of those estimates on explanatory variables by means of M-Estimation (robust estimation).

Robust estimation is one way we will try to overcome issues with our way of estimating the change in the learning rate. Another way is to resort to discretization and explain the *direction* of changes in learning rates (positive, negative, no change) using ordered logit estimation.

We implemented both approaches, but argue that the discretization results should be more reliable. As we shall see, qualitatively no differences emerge.

We thus estimate the learning rates in two ways, one discrete, Model D, preferred, and the other one continuous, Model C, more problematic because of fat tails in the distribution of the regressor.

(D) Direction of change Y_{t+1}:

$$Y_{t+1} = \begin{cases} +1 & \text{if } \frac{p^e_{t+1}-p^e_t}{\epsilon_t} - \frac{p^e_t-p^e_{t-1}}{\epsilon_{t-1}} > 0 \\ 0 & \text{if } \frac{p^e_{t+1}-p^e_t}{\epsilon_t} - \frac{p^e_t-p^e_{t-1}}{\epsilon_{t-1}} = 0 \\ -1 & \text{if } \frac{p^e_{t+1}-p^e_t}{\epsilon_t} - \frac{p^e_t-p^e_{t-1}}{\epsilon_{t-1}} < 0 \end{cases}$$

We explain Y_{t+1} using ordered logit, allowing for random effects (intercept) at the participant level.

(C) Change itself:

$$\Delta_{t+1} = \frac{p^e_{t+1} - p^e_t}{\epsilon_t} - \frac{p^e_t - p^e_{t-1}}{\epsilon_{t-1}}.$$

We explain Δ_{t+1} using M-Estimation, allowing for fixed effects (intercept) at the participant level.

We discuss here the results based on Model D, postponing analysis of results based on Model C in a section on robustness (Section 7).

Figure 4 displays results from logit estimation of Y^i_{t+1} (direction of change in α) as a function of the Pearce-Hall driver, i.e., the change in surprise $|\epsilon^i_t|$.[16] We report the results for (simple) logit rather than for ordered logit since including the middle state ($Y_{t+1} = 0$) did not generate significantly different results, and logit estimation is easier to interpret.

Mathematically, we therefore report results for the following model:

$$Y^i_{t+1} = \begin{cases} -1 & \text{if } z^i_t + \eta^i_t < \mu^i, \\ +1 & \text{if } \mu^i \leq z^i_t + \eta^i_t, \end{cases} \tag{10}$$

where

$$z^i_t = \lambda 1_{\{|\epsilon^i_t| \geq |\epsilon^i_{t-1}|\}}$$

and

$$\eta^i_t$$

is noise that follows a logistic distribution with mean 0 and a scale parameter to be estimated.[17] The variable

[16] Subscripts i are added to reflect the panel structure of the data: one time series of forecasts per participant i.

[17] The scale parameter is effectively allowed to change across participants since we report standard errors

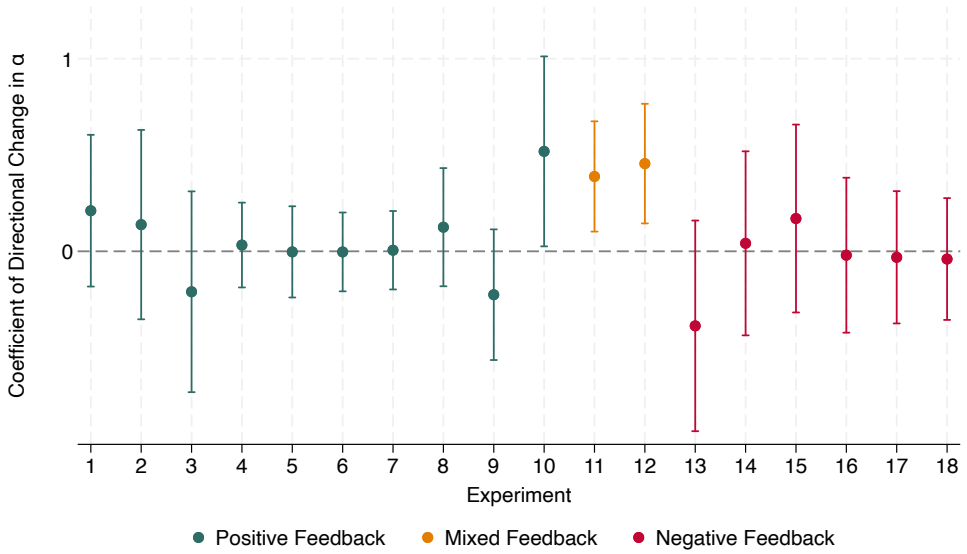

Figure 4. Shown are estimates of logit estimation of λ in (10), per experiment, with 95% confidence intervals Bonferroni-corrected for multiple hypothesis testing (18 tests). Logit estimation allows for random effects at the participant level (intercept); standard errors are corrected for clustering at the participant level.

1_E denotes a dummy variable for the event E. Note that the event "drop in surprise" $\{|\epsilon_t^i| < |\epsilon_{t-1}^i|\}$ provides the baseline. Note also that we allow for random effects at the participant level, captured by the (random) parameters μ^i.

Panel A of Table A.2 - Table A.4 list the parameter estimates, standard errors and corresponding p-values on which Figure 4 is based. Confidence intervals (95%) are adjusted for multiple hypothesis testing using Bonferroni correction.

We observe that except for one positive feedback experiment, and the two mixed positive/negative feedback experiments, the change in size of prediction error has no impact on the learning rate.

So, overall we fail to find support for the Pearce-Hall Model of learning rate adjustment to universally explain participants' strategy. This may seem to contradict recent evidence in Carvalho et al. (2023), who find that the learning rate in inflation forecasting increases substantially, but only after episodes of large prediction errors. Their model is one of regime switches, not unlike those applied to restless bandit problems, as in Payzan-LeNestour and Bossaerts (2015). In contrast, the Pearce-Hall model continuously alters the learning rate as a function of surprise, not just when there is evidence of a regime switch. It implements what is referred to in computational neuroscience and computer science as "model-free reinforcement learning" (Lee et al., 2020).

Result 2: We reject Hypothesis 2 that α changes in the direction of the change in surprise.

that allow for clustering at the participant level.

6.3 Robust Expectations Adaptation (REA)

We now test REA, implementing MRAC with the Gaussian Kalman filter as the reference model. We do so by using ordered logit modeling with Y_{t+1} as dependent variable (the direction of change of α). As explanatory variables, we construct dummy variables for (i) the magnitude of surprise and (ii) the sign of prediction error autocovariance. Binary *magnitude* categories are constructed by using a median split of surprise as experienced per participant. The categories are referred to as "low" (below-median $|\epsilon_t|$) and "high" surprise (above-median $|\epsilon_t|$)[18]. *Autocovariance* is estimated simply as the product of two previously experienced prediction errors, as proposed before. That is, autocovariance in trial t is estimated as

$$\epsilon_t \epsilon_{t-1}.$$

The logit regression also includes an interaction term between the surprise and autocovariance categorical variables. This interaction is what allows us to test the central prediction of Hypothesis 3: that adjustment in the direction of autocovariance is modulated by surprise rather than operating uniformly.

Mathematically, after adding subscripts i to denote participant-specific variables, we estimate the following model:

$$Y^i_{t+1} = \begin{cases} -1 & \text{if } z^i_t + \eta^i_t < \mu^i_1, \\ 0 & \text{if } \mu^i_1 \leq z^i_t + \eta^i_t < \mu^i_2, \\ +1 & \text{if } \mu^i_2 \leq z^i_t + \eta^i_t, \end{cases} \tag{11}$$

where

$$z^i_t = \gamma 1_{\{\epsilon^i_t \epsilon^i_{t-1} > 0\}} + \beta 1_{\{|\epsilon^i_t| < \text{median}(|\epsilon^i|)\}} + \delta 1_{\{|\epsilon^i_t| < \text{median}(|\epsilon^i|)\}} 1_{\{\epsilon^i_t \epsilon^i_{t-1} > 0\}}$$

and

$$\eta^i_t$$

is noise that follows a logistic distribution with a scale parameter independent of i (but standard errors are corrected for clustering at the participant level). Note that the joint events "high surprise" $\{|\epsilon^i_t| \geq \text{median}(|\epsilon^i|)\}$ and "non-positive autocovariance" $\{\epsilon^i_t \epsilon^i_{t-1} \leq 0\}$ provide the baseline. Note also that we allow for random effects at the participant level, captured by the (random) parameters μ^i_1 and μ^i_2.

Figure 5 shows how the autocovariance/surprise categories affect the probability that $Y_{t+1} = -1$ (a decrease in learning rate α), $Y_{t+1} = 0$ (no change in α), and $Y_{t+1} = +1$ (an increase in α). For instance, in Experiment 4, when surprise (absolute value of prediction error) is large and autocovariance is positive (dark blue), it is unlikely that the learning rate decreases (probability ≈ 0.25); instead it is highly likely that the learning rate increases (probability ≈ 0.75).

The probabilities displayed in Figure 5 are obtained as follows. First,

$$\text{P}(Y_{t+1} = -1 | \text{surprise above median}, \text{autocov} > 0) = \Lambda(\eta_t < \overline{\mu}_1 - \gamma),$$

where Λ denotes the logistic distribution function and $\overline{\mu}_1$ denotes the estimate of the mean value of the

[18] We also conduct analyses using the 25th percentile as the cutoff instead of the median; the results are consistent with those reported here.

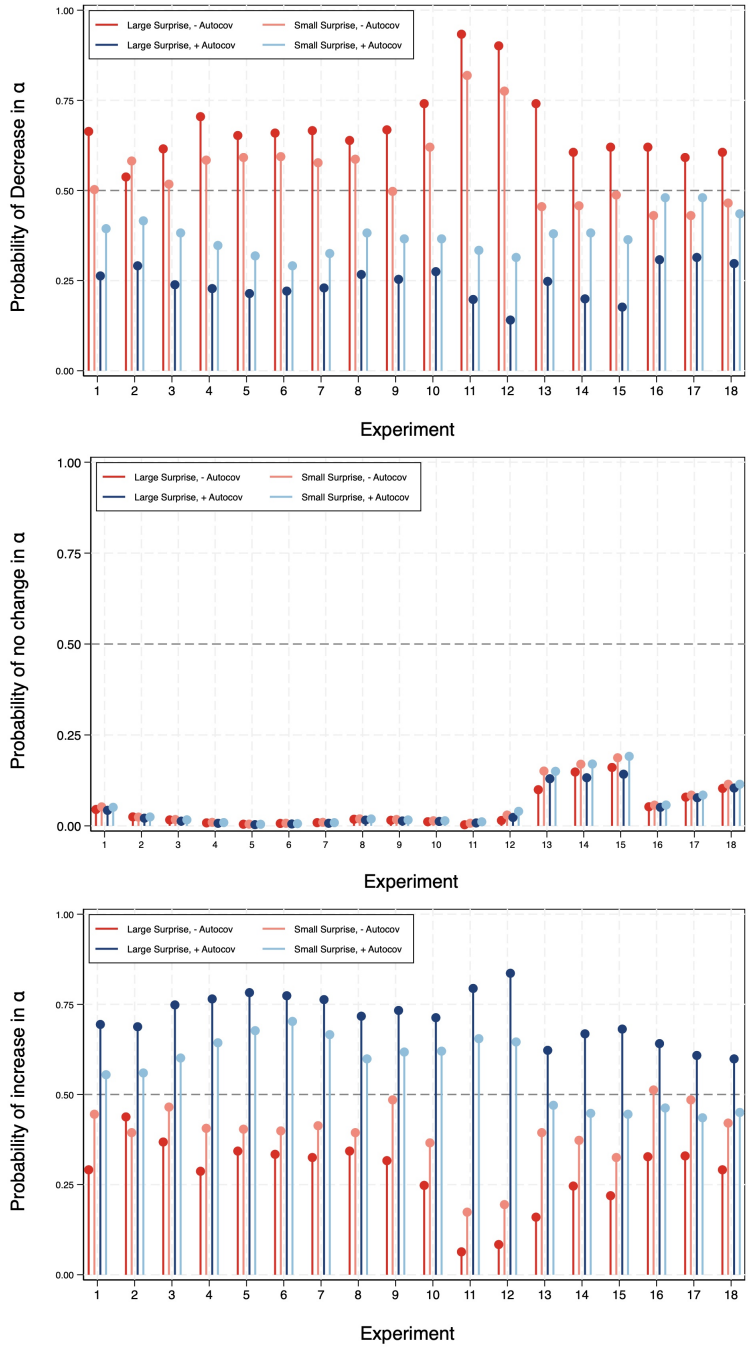

Figure 5. Estimated probabilities are color-coded per conditioning event: dark blue (large surprise, + autocov), dark red (large surprise, – autocov), light blue (small surprise, + autocov) and light red (small surprise, – autocov).

incidental parameters $\{\mu_1^i,\ i = 1, 2, ..., I\}$. Second,

$$\mathrm{P}(Y_{t+1} = +1|\text{surprise above median, autocov} > 0) = \Lambda(\overline{\mu}_2 - \gamma \leq \eta_t).$$

Analogously, if surprise is large and autocovariance is negative (dark red), then the learning rate is likely to decrease (probability ≈ 0.7) while it is unlikely to increase (probability ≈ 0.3). These probabilities are computed from the estimated parameters by computing

$$\mathrm{P}(Y_{t+1} = -1|\text{surprise above median, autocov} \leq 0) = \Lambda(\eta_t < \overline{\mu}_1)$$

and

$$\mathrm{P}(Y_{t+1} = +1|\text{surprise above median, autocov} \leq 0) = \Lambda(\overline{\mu}_2 \leq \eta_t).$$

Inspection of Figure 5 reveals that the ordering of the effect of the autocovariance/surprise categories supports our hypothesis that in all experiments, α changes in the direction of autocovariance, though the effect is attenuated when surprise is below the median an individual experiences. There is only one exception: in Experiment 2, the attenuation effect of the size of surprise does not emerge for decreases in α (negative autocovariance) and increase in α (positive autocovariance).

Table 2 lists the numerical estimation results on which the probabilities displayed in Figure 5 are based. p-values are Bonferroni-corrected, and standard errors are corrected for clustering at the participant level.[19]

Specifically, we estimate (i) $\gamma > 0$, so when surprise is large ($1_{\{|\epsilon_t^i| < \text{median}\}} = 0$), positive autocovariance of prediction errors increases the chance that the learning rate α increases (while decreasing the chance that the learning rate decreases); (ii) $\delta < 0$, so that low surprise reduces the chance of a change in the learning rate with respect to autocovariance, relative to high surprise.

Altogether, the evidence is fully consistent with Robust Expectations Adaptation (REA) following the prescription of MRAC: participants choose learning rates α that change in the direction of the autocovariance of prediction errors when encountering large surprises ($\gamma > 0$). This ensures robustness because it minimizes expected surprise with respect to a reference model, the Gaussian Kalman filter model.

The evidence suggests, however, that learning rates are modulated more intensely when surprise (size of the prediction error) is larger ($\delta < 0$). This is consistent with the idea that adaptation does not take place continuously, but only when surprise is sufficiently large. As such, our findings also confirm the experimental results in a very different forecasting task, d'Acremont and Bossaerts (2016). There, frequent outliers (leptokurtosis) complicated forecasting; here, prediction was made difficult because of self-referentiality.

Result 3: We find support for Hypothesis 3, and hence, REA, because we find evidence that α is modulated by the sign of the prediction error autocovariance ($\epsilon_t \epsilon_{t-1}$). However, this modulation is only observed when surprise is sufficiently high.

Result 3 reinforces Result 1 (that α is not constant).

[19] We also re-tabulate Table 2 using standard errors clustered at the cohort level. The results are presented in Table A.5. Overall, the findings remain similar.

Table 2. Ordered-logit modeling of direction of change in learning rate α, Robust Expectations Adaptation (REA). Listed are estimates of the parameters of the equation displayed in (11), per Experiment. Robust standard error clustered at participant level in parentheses. p values in brackets; p values for γ, δ and β are Bonferroni-corrected for multiple hypothesis testing (54 tests).

Panel A: Positive feedback

Experiment	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
γ	1.71	1.04	1.63	2.09	1.93	1.92	1.90	1.58	1.78	2.02
	(0.20)	(0.15)	(0.17)	(0.11)	(0.09)	(0.10)	(0.10)	(0.11)	(0.15)	(0.24)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
δ	-1.27	-0.37	-1.08	-1.12	-0.80	-0.65	-0.86	-0.75	-1.24	-0.98
	(0.27)	(0.25)	(0.25)	(0.15)	(0.14)	(0.14)	(0.12)	(0.18)	(0.21)	(0.25)
	[0.000]	[1.000]	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]	[0.002]	[0.000]	[0.004]
β	0.67	-0.18	0.40	0.53	0.26	0.28	0.38	0.22	0.71	0.56
	(0.22)	(0.20)	(0.19)	(0.12)	(0.10)	(0.11)	(0.09)	(0.15)	(0.14)	(0.20)
	[0.118]	[1.000]	[1.000]	[0.000]	[0.446]	[0.517]	[0.001]	[1.000]	[0.000]	[0.228]
$\bar{\mu}_1$	0.68	0.15	0.47	0.87	0.63	0.66	0.69	0.57	0.70	1.05
	(0.17)	(0.12)	(0.13)	(0.08)	(0.06)	(0.07)	(0.06)	(0.09)	(0.10)	(0.19)
	[0.000]	[0.189]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$\bar{\mu}_2$	0.89	0.25	0.54	0.91	0.65	0.69	0.73	0.65	0.77	1.11
	(0.17)	(0.13)	(0.13)	(0.08)	(0.06)	(0.07)	(0.07)	(0.08)	(0.11)	(0.19)
	[0.000]	[0.047]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	894	1,077	994	4,598	4,590	4,594	4,590	2,283	2,302	1,152
Number of Participants	48	48	48	96	96	96	96	48	48	24

Panel B: Mixed and negative feedback

Experiment	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
γ	4.04	4.02	2.16	1.82	2.03	1.30	1.15	1.29
	(0.35)	(0.24)	(0.19)	(0.19)	(0.22)	(0.22)	(0.15)	(0.20)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
δ	-1.84	-2.00	-1.85	-1.51	-1.52	-1.50	-1.35	-1.17
	(0.36)	(0.27)	(0.24)	(0.24)	(0.23)	(0.29)	(0.18)	(0.26)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
β	1.13	0.97	1.23	0.60	0.54	0.77	0.65	0.57
	(0.36)	(0.26)	(0.19)	(0.15)	(0.16)	(0.18)	(0.11)	(0.13)
	[0.086]	[0.010]	[0.000]	[0.004]	[0.035]	[0.001]	[0.000]	[0.000]
$\bar{\mu}_1$	2.64	2.21	1.05	0.43	0.49	0.49	0.37	0.43
	(0.33)	(0.22)	(0.15)	(0.13)	(0.18)	(0.15)	(0.10)	(0.11)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.005]	[0.001]	[0.000]	[0.000]
$\bar{\mu}_2$	2.69	2.39	1.66	1.12	1.27	0.72	0.71	0.89
	(0.33)	(0.23)	(0.14)	(0.14)	(0.13)	(0.14)	(0.09)	(0.14)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	2,160	2,586	910	1,088	998	1,150	2,012	1,726
Number of Participants	45	54	48	48	48	24	42	36

7 Robustness

We now investigate the robustness of our results by changing the nature of both the dependent and independent variables, properly adjusting estimation protocol if this is needed.

7.1 From Ordered Logit to Logit

We re-estimate (11) after excluding the middle outcome—i.e., cases where there is no change in α. This simplification is justified, as the middle state accounts for only 3% of the observations.

The resulting model for Y^i_{t+1} is a binary specification, with z defined as in Equation (11).

$$Y^i_{t+1} = \begin{cases} -1 & \text{if } z^i_t + \eta^i_t < \mu^i, \\ +1 & \text{if } \mu^i \leq z^i_t + \eta^i_t. \end{cases} \tag{12}$$

The estimation results are presented in Panel A of Tables A.6, A.7, and A.8, corresponding to the positive, mixed, and negative feedback experiments, respectively. The parameter estimates remain consistent with those from the previous ordered logit specification: we find $\gamma > 0$ in all experiments. We therefore conclude that our inference is unaltered when switching from ordered logit to standard logit estimation. We conclude that participants' behavior is consistent with REA. Essentially, this means that participants impute robustness in applying adaptive expectations by changing the learning rate to minimize expected surprise relative to the Kalman filter model. Further, we find $\delta < 0$ in all but Experiment 2, indicating that modulation following REA occurs only in tasks requiring more attention—namely, when surprise is sufficiently high.

7.2 Continuous Measurement of Surprise

We have been focusing on how the learning rate changes in response to a surprise that is above or below the median experienced by a participant in an experiment. In this section, we replace discretized surprise with its continuous analogue. This estimation is based on a simplified model that excludes the middle state.

Mathematically, we estimate (12) but define z^i_t as follows:

$$z^i_t = \gamma 1_{\{\epsilon^i_t \epsilon^i_{t-1} > 0\}} + \beta |\epsilon^i_t| + \delta |\epsilon^i_t| 1_{\{\epsilon^i_t \epsilon^i_{t-1} > 0\}}. \tag{13}$$

The results are reported in Panel B of Table A.6, A.7, and A.8, corresponding to the positive, mixed, and negative feedback experiments, respectively. We find evidence of REA in all experiments because learning rates increase when there is positive autocovariance ($\gamma + \delta > 0$, given that $|\epsilon^i_t|$ is almost never zero and therefore is strictly positive[20]). Further, in 16 out of 18 experiments, we find $\delta > 0$, suggesting that adjustments in the learning rate in the direction of autocovariance occur more frequently when participants experience larger surprise. This pattern is consistent with the efficient allocation of attention.

[20] Among the 41,490 observations, surprise is zero in only 98 instances.

7.3 Heterogeneity in Response to Autocovariance in Prediction Errors

Up to this point, we have forced γ to be constant across participants in our estimation. This assumes that all participants share the same propensity to increase their learning rate in response to autocovariance of prediction errors. This is a strong assumption. If this propensity differs across participants, estimates of other parameters may be biased.

In this section, we examine whether allowing for heterogeneity in γ changes our inference. Estimates are obtained within a *linear* regression framework, as standard statistical packages do not support slope heterogeneity in logistic models. If the parameter estimates remain consistent regardless of whether slope heterogeneity is allowed in a linear regression, it would suggest that our results are robust to heterogeneity.

We proceed as follows. Building on the discrete explanatory variable of increase/decrease in surprises as in (12), and additionally allowing for participant-level fixed effects at the level of the intercept, we first estimate the following model, in which slope homogeneity is assumed:

$$Y^i_{t+1} = \mu^i + z^i_t + \eta^i_t \tag{14}$$

where

$$z^i_t = \gamma 1_{\{\epsilon^i_t \epsilon^i_{t-1} > 0\}} + \beta 1_{\{|\epsilon^i_t| < \text{median}(|\epsilon^i|)\}} + \delta 1_{\{|\epsilon^i_t| < \text{median}(|\epsilon^i|)\}} 1_{\{\epsilon^i_t \epsilon^i_{t-1} > 0\}}$$

and

$$\eta^i_t$$

is mean-zero noise, with variance to be estimated and fixed across i.

We then relax the assumption of slope (γ) homogeneity. We add superscripts "i" to γ and estimate:

$$z^i_t = \gamma^i 1_{\{\epsilon^i_t \epsilon^i_{t-1} > 0\}} + \beta 1_{\{|\epsilon^i_t| < \text{median}(|\epsilon^i|)\}} + \delta 1_{\{|\epsilon^i_t| < \text{median}(|\epsilon^i|)\}} 1_{\{\epsilon^i_t \epsilon^i_{t-1} > 0\}}.$$

Estimation results for the models without (Panel A) and with (Panel B) slope heterogeneity are reported in Tables A.9, A.10, and A.11. These tables correspond to the positive, mixed, and negative feedback experiments, respectively. First, we find that the estimates and standard errors are consistent across the two panels, suggesting that allowing for slope heterogeneity in the linear regression has little impact on the results. Second, compared to the ordered-logit estimates with random intercepts but constant γ, the linear models with fixed intercepts yield similar patterns: in 17 out of 18 experiments, we observe $\gamma > 0$ or $\text{mean}(\gamma^i) > 0$, in addition to $\delta < 0$. We therefore conclude that our results are robust to heterogeneity in intensity of adjustment of the learning rate with respect to autocovariance.

7.4 Continuous Measurement of Changes in the Learning Rate (Model C)

The analyses so far treated updates in the learning rate α as a discrete variable, capturing only the direction of change (increase or decrease). We now switch to the continuous version of the estimated change in the

learning rate and estimate the following model:

$$\Delta^i_{t+1} = \mu^i + z^i_t + \eta^i_t \tag{15}$$

where

$$z^i_t = \gamma 1_{\{\epsilon^i_t \epsilon^i_{t-1} > 0\}} + \beta 1_{\{|\epsilon^i_t| < \text{median}(|\epsilon^i|)\}} + \delta 1_{\{|\epsilon^i_t| < \text{median}(|\epsilon^i|)\}} 1_{\{\epsilon^i_t \epsilon^i_{t-1} > 0\}}$$

and

$$\eta^i_t$$

is noise with mean zero and variance to be estimated. We implement robust regression (M-estimation) with fixed effects on the intercept, to mitigate the influence of outliers in the dependent variable.

The results are presented in Panel A of Tables A.12, A.13, and A.14, corresponding to positive, mixed, and negative feedback treatments, respectively. Across all but one experiment (Experiment 2), we find evidence of REA because learning rates increase when autocovariance is larger ($\gamma > 0$). However, we lose the interaction effect with surprise ($\delta = 0$) except in five experiments in the negative feedback treatments, suggesting that the propensity of learning rate adjustment remains constant regardless of the magnitude of surprise.

The reduced explanatory power of the interaction term is likely because the model attempts to explain a continuous dependent variable using only three binary independent variables. For example, surprise is measured discretely based on a median split per participant. In the following, we re-run the estimation allowing for continuous changes in surprise, as in (13). That is, we estimate the following:

$$\Delta^i_{t+1} = \mu^i + z^i_t + \eta^i_t \tag{16}$$

where

$$z^i_t = \gamma 1_{\{\epsilon^i_t \epsilon^i_{t-1} > 0\}} + \beta |\epsilon^i_t| + \delta |\epsilon^i_t| 1_{\{\epsilon^i_t \epsilon^i_{t-1} > 0\}}$$

and

$$\eta^i_t$$

is noise with mean zero and variance to be estimated. We continue with robust regression (M-estimation) with fixed effects on the intercept, to mitigate the influence of outliers in the dependent variable.

The results are presented in Panel B of the same tables. The estimation demonstrates slightly reduced explanatory power compared to when surprise is discretized. We find evidence of REA ($\gamma + \delta > 0$, given that $|\epsilon^i_t|$ is almost never zero and therefore is strictly positive) in all experiments.

Only one of them exhibit an increasing propensity for REA adjustment when surprises are large ($\delta > 0$). However, it is important to note that M-estimation only addresses fat-tail issues in the dependent variables but not in the independent variables (i.e., the magnitude of surprises). Consequently, the inference drawn from Panel B may still be misleading.

7.5 Pearce-Hall Re-visited with Continuous Recording of Learning Rate Changes

Finally, we revisit the Pearce-Hall model by examining whether an increase in surprise leads participants to increase their learning rates when these rates are measured continuously (Model C). The results, based on M-estimation, are presented in Panel B of Tables A.2, A.3, and A.4. We observe that Pearce-Hall model provides no explanatory power in any experiment. Worse, the estimated coefficient γ has the opposite sign (negative instead of positive) from what is expected in 7 out of 18 experiments. We therefore conclude that the poor fit of the Pearce-Hall model is a robust finding.

8 Conclusion

Analyzing almost 42,000 forecasts in 18 experiments with self-referential price determination, we find that the data are consistent with a key approach from engineering that ensures robustness in control. In this approach, the standard adaptive expectations protocol is "supervised" and its learning rate is adapted to minimize expected surprise. Expected surprise is measured against a reference model, which we took to be the standard (Gaussian) Kalman filter model. Expected surprise minimization can then be reached by changing the learning rate in proportion to autocovariance of prediction errors.

Our data supports this conjecture, though learning rate adaptation generally occurs only if surprise is above the median level experienced by a participant. This is consistent with the observation from a previous experiment in a different setting (forecasting a target whose movements are subject to leptokurtosis): participants change the learning rate only if surprise is sufficiently high (d'Acremont and Bossaerts, 2016; Bossaerts, 2018). This is also consistent with theory on limited attention (Sims, 2003): the supervisor pays attention to the learning rate only when it is deemed necessary, given that there are competing demands on attention; while the standard adaptive expectations protocol can be executed without attention.

We discover several additional features in our forecasting data.

- When fitting the traditional adaptive expectations model with a constant learning rate, we find that estimated learning rates are close to or above 1 in settings with positive feedback, and below 1 in settings with negative feedback. The autocorrelation of prediction errors, however, reveals that these learning rates were nonetheless mis-specified: prediction errors are positively autocorrelated under positive and mixed feedback, indicating under-reaction; and negatively autocorrelated under negative feedback, indicating over-reaction. High learning rates are consistent with the findings of Bordalo et al. (2020), who attribute such deviations from optimal learning to belief distortions. Instead, we explain them as the result of concern for robustness: agents adapt the learning rate optimally, but only when the level of surprise they experience exceeds a threshold.
- The neuropsychology literature has likewise reported excessive learning rates; see, e.g., Lee et al. (2020). By contrast, in an experiment with regular regime switches, Behrens et al. (2007) observed that learning rates were (Bayes-)optimal. Their participants, however, were told of possible contingency shifts and were given extensive practice in adjusting to them. Payzan-LeNestour and Woodford (2022)

demonstrate that humans need to be told about the presence of contingency shifts; otherwise, they apply a constant learning rate. The latter article does not report whether that learning rate was excessive, however.

- Payzan-LeNestour and Woodford (2022) show that humans exhibit "outlier blindness," i.e., that they have difficulty acting correctly upon an outlier event. REA demonstrates that human behavior is adapted to outlier blindness: (i) outliers lead to surprise, and surprise forces humans to get out of their "autopilot;" (ii) expectations adaptation upon surprise is robust because expected surprise is measured relative to a reference model, consistent with MRAC. In other words: yes, humans exhibit outlier blindness, but, then again, the brain appears to be able to recruit a supervisory system that ensures that the blindness is dealt with in a robust way.
- Surprise can be modeled in other ways besides the distance between the prediction error and a reference model's expectation of this prediction error (Modirshanechi et al., 2022). We chose to define surprise as the squared deviation of the squared prediction error relative to its expectation under the reference model. This definition was also used in Bossaerts (2018). In an investments context, the definition can be derived as a unique globally robust extension of a locally quadratic regret criterion (Berrada et al., 2024). Our definition is also in line with non-stochastic MRAC applications, where surprise is modeled as the squared tracking error (Nguyen, 2018).[21]
- The implications of REA for portfolio analysis have been explored elsewhere; see Berrada et al. (2024). Surprising behavior emerges, such as mean-variance optimization for a large set of reference risk-reward trade-offs, as well as willingness to invest in risky assets even if there is no reward for risk.

We assume the standard Kalman filter as the reference model and change the learning rate flexibly to ensure expected surprise is minimized. This leads to the prescription that the learning rate has to be changed in the direction of the autocovariance of prediction errors. This aligns with the prescription of the Delta-Bar-Delta learning algorithm (Sutton, 1992). It is worth noting that, when another reference model is used to define surprise, a different prescription may ensue.

We do not, however, have a theory of the nature of the reference model that is being used, and how quickly it changes. We used the Kalman filter since it provides a simple representation of an ever-changing world, and because it has recently been the focus of analysis in the economics literature on macro-forecasting (Bordalo et al., 2020). A possible avenue for future research is to consider alternative forms of reference models and test whether they generate the wide variety of learning algorithms observed empirically. Future work should focus on identification of the reference model and how it changes over time.

In the empirical examination, we have provided only crude evidence in favor of MRAC. Because the data we analyzed here came from an experiment that was designed with other goals in mind, we at best have only confirmatory evidence that, in forecasting in nonstationary, self-referential systems, humans act *as if* minimizing expected surprise relative to a simple Kalman filter model. It is not clear whether this behavior is generic. We already pointed out that MRAC behavior may be maladaptive if the environment is sufficiently stationary for the agent to fully learn to optimize. As an example, we argued that sensorimotor control that insists on always minimizing surprise relative to a linear trajectory may not lead to energy minimization.

[21] Probabilistic versions of MRAC have recently been proposed whereby the distance between outcomes in the reference model and in the real world is defined in terms of KL divergence; see, e.g., Herzallah (2020).

Future research should develop experimental paradigms to explore this apparent tension between MRAC's "satisficing" behavior and the more standard optimization on which economic theory has been built.

References

Adam, K., Marcet, A., and Nicolini, J. P. (2016). Stock market volatility and learning. *The Journal of finance*, 71(1):33–82.

Anufriev, M. and Hommes, C. (2012). Evolutionary selection of individual expectations and aggregate outcomes in asset pricing experiments. *American Economic Journal: Microeconomics*, 4(4):35–64.

Asparouhova, E., Bossaerts, P., Roy, N., and Zame, W. (2016). 'lucas' in the laboratory. *Journal of Finance*, 71:2727–2780.

Bao, T., Duffy, J., and Hommes, C. (2013). Learning, forecasting and optimizing: An experimental study. *European Economic Review*, 61:186–204.

Bao, T., Füllbrunn, S., Pei, J., and Zong, J. (2024). Reading the market? expectation coordination and theory of mind. *Journal of Economic Behavior & Organization*, 219:510–527.

Bao, T. and Hommes, C. (2019). When speculators meet suppliers: Positive versus negative feedback in experimental housing markets. *Journal of Economic Dynamics and Control*, 107:103730.

Bao, T., Hommes, C., and Makarewicz, T. (2017). Bubble formation and (in) efficient markets in learning-to-forecast and optimise experiments. *The Economic Journal*, 127(605):F581–F609.

Bao, T., Hommes, C., and Pei, J. (2021). Expectation formation in finance and macroeconomics: A review of new experimental evidence. *Journal of Behavioral and Experimental Finance*, 32:100591.

Bao, T., Hommes, C., Sonnemans, J., and Tuinstra, J. (2012). Individual expectations, limited rationality and aggregate outcomes. *Journal of Economic Dynamics and Control*, 36(8):1101–1120.

Behrens, T. E., Woolrich, M. W., Walton, M. E., and Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nature neuroscience*, 10(9):1214–1221.

Berrada, T., Bossaerts, P., and Ugazio, G. (2024). Investments and asset pricing in a world of satisficing agents. *Swiss Finance Institute Research Paper*, 24(05).

Bordalo, P., Gennaioli, N., Ma, Y., and Shleifer, A. (2020). Overreaction in macroeconomic expectations. *American Economic Review*, 110(9):2748–2782.

Bossaerts, P. (1995). The econometrics of learning in financial markets. *Econometric Theory*, 11(1):151–189.

Bossaerts, P. (2004). Filtering returns for unspecified biases in priors when testing asset pricing theory. *The Review of Economic Studies*, 71(1):63–86.

Bossaerts, P. (2018). Formalizing the function of anterior insula in rapid adaptation. *Frontiers in Integrative Neuroscience*, 12:61.

Bossaerts, P., Fattinger, F., van den Bogaerde, F., and Yang, W. (2024). Asset pricing in a world of imperfect foresight. *Available at SSRN 3610634*.

Bossaerts, P. and Hillion, P. (1999). Implementing statistical criteria to select return forecasting models: what do we learn? *Review of Financial Studies*, 12(2):405–428.

Brogaard, J., Hendershott, T., and Riordan, R. (2014). High-frequency trading and price discovery. *The Review of Financial Studies*, 27(8):2267–2306.

Cagan, P. (2000). Phillips' adaptive expectations formula. *AWH Phillips: Collected Works in Contemporary Perspective, Cambridge University Press, New York*.

Camerer, C., Xin, Y., and Zhao, C. (2024). A neural autopilot theory of habit: Evidence from consumer purchases and social media use. *Journal of the Experimental Analysis of Behavior*, 121(1):108–122.

Caplin, A., Dean, M., and Martin, D. (2011). Search and satisficing. *American Economic Review*, 101(7):2899–2922.

Carvalho, C., Eusepi, S., Moench, E., and Preston, B. (2023). Anchored inflation expectations. *American Economic Journal: Macroeconomics*, 15(1):1–47.

Chen, J., Hong, H., and Stein, J. C. (2001). Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices. *Journal of financial Economics*, 61(3):345–381.

Csáji, B. C. and Monostori, L. (2008). Adaptive stochastic resource control: a machine learning approach. *Journal of Artificial Intelligence Research*, 32:453–486.

d'Acremont, M. and Bossaerts, P. (2016). Neural mechanisms behind identification of leptokurtic noise and adaptive behavioral response. *Cerebral Cortex*, 26(4):1818–1830.

Doob, J. L. (1949). Application of the theory of martingales. *Le calcul des probabilites et ses applications*, pages 23–27.

Engle, R. (2001). Garch 101: The use of arch/garch models in applied econometrics. *Journal of economic perspectives*, 15(4):157–168.

Franklin, D. W. and Wolpert, D. M. (2011). Computational mechanisms of sensorimotor control. *Neuron*, 72(3):425–442.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384.

Hansen, L. P. and Sargent, T. J. (2001). Acknowledging misspecification in macroeconomic theory. *Review of Economic Dynamics*, 4(3):519–535.

Herzallah, R. (2020). A fully probabilistic design for tracking control for stochastic systems with input delay. *IEEE Transactions on Automatic Control*, 66(9):4342–4348.

Jansch-Porto, J. a. P., Hu, B., and Dullerud, G. (2020). Convergence guarantees of policy optimization methods for markovian jump linear systems. *arXiv preprint*, abs/2002.04090. Provides theoretical convergence analysis of policy optimization in MJLS.

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current opinion in neurobiology*, 9(6):718–727.

Lee, S., Gold, J. I., and Kable, J. W. (2020). The human as delta-rule learner. *Decision*, 7(1):55.

Lucas, R. E. J. (1978). Asset prices in an exchange economy. *Econometrica*, 46(6):1429–1445.

Mansell, W., Gulrez, T., and Landman, M. (2025). The prediction illusion: perceptual control mechanisms that fool the observer. *Current Opinion in Behavioral Sciences*, 62:101488.

Marcet, A. and Nicolini, J. P. (2003). Recurrent hyperinflations and learning. *American Economic Review*, 93(5):1476–1498.

Marcet, A. and Sargent, T. J. (1989). Convergence of least squares learning mechanisms in self-referential linear stochastic models. *Journal of Economic theory*, 48(2):337–368.

Modirshanechi, A., Brea, J., and Gerstner, W. (2022). A taxonomy of surprise definitions. *Journal of mathematical psychology*, 110:102712.

Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *Journal of neuroscience*, 16(5):1936–1947.

Musallam, S., Corneil, B., Greger, B., Scherberger, H., and Andersen, R. A. (2004). Cognitive control signals for neural prosthetics. *Science*, 305(5681):258–262.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica: Journal of the Econometric Society*, pages 315–335.

Nassar, M. R., Wilson, R. C., Heasly, B., and Gold, J. I. (2010). An approximately bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *Journal of Neuroscience*, 30(37):12366–12378.

Nguyen, N. T. (2018). *Model-reference adaptive control*. Springer.

Nursimulu, A. and Bossaerts, P. (2014). Excessive volatility is also a feature of individual level forecasts. *Journal of Behavioral Finance*, 15(1):16–29.

Paulus, M. P., Frank, L., Brown, G. G., and Braff, D. L. (2003). Schizophrenia subjects show intact success-related neural activation but impaired uncertainty processing during decision-making. *Neuropsychopharmacology*, 28(4):795–806.

Payzan-LeNestour, E. and Bossaerts, P. (2015). Learning about unstable, publicly unobservable payoffs. *The Review of Financial Studies*, 28(7):1874–1913.

Payzan-LeNestour, E. and Woodford, M. (2022). Outlier blindness: A neurobiological foundation for neglect of financial risk. *Journal of Financial Economics*, 143(3):1316–1343.

Pearce, J. M. and Hall, G. (1980). A model for pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological review*, 87(6):532.

Preuschoff, K., Quartz, S. R., and Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *The Journal of Neuroscience*, 28(11):2745–2752.

Rescorla, R. A. and Wagner, A. R. (1972). A theory of classical conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In Black, A. H. and Prokasy, W. F., editors, *Classical Conditioning II: Current Research and Theory*, pages 64–99. Appleton-Century-Crofts, New York.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599.

Schultz, W. and Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience*, 23(1):473–500.

Shadmehr, R. and Mussa-Ivaldi, F. A. (1994). Adaptive representation of dynamics during learning of a motor task. *Journal of neuroscience*, 14(5):3208–3224.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, pages 99–118.

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690.

Sutton, R. S. (1992). Adapting bias by gradient descent: An incremental version of delta-bar-delta. In *AAAI*, volume 92, pages 171–176. Citeseer.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Tin, C. and Poon, C.-S. (2005). Internal models in sensorimotor integration: perspectives from adaptive control theory. *Journal of Neural Engineering*, 2(3):S147.

Wu, Z. and Sun, K. (2023). Distributionally robust optimization with wasserstein metric for multi-period portfolio selection under uncertainty. *Applied Mathematical Modelling*, 117:513–528.

Zhang, D. and Wei, B. (2017). A review on model reference adaptive control of robotic manipulators. *Annual Reviews in Control*, 43:188–198.

Appendices

A Additional Tables

Table A.1. Data Sources

Experiment	Treatment	Source	Details
Panel A: Positive Feedback			
1	Fundamental value (FV) = 56	Bao et al. (2012), JEDC	N = 960; Cohort size = 6; 8 cohorts; T = 20; $E(\lvert\epsilon^i\rvert) = 0.97$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 8.03$
2	FV = 41		N = 1,104; Cohort size = 6; 8 cohorts; T = 23; $E(\lvert\epsilon^i\rvert) = 0.57$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 2.46$
3	FV = 62		N = 1,056; Cohort size = 6; 8 cohorts; T = 22; $E(\lvert\epsilon^i\rvert) = 0.74$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 1.19$
4	High Theory of Mind (ToM)	Bao et al. (2024), JEBO	N = 4,800; Cohort size = 6; 16 cohorts; T = 50; $E(\lvert\epsilon^i\rvert) = 13.56$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 1,807.10$
5	Medium High ToM		N = 4,800; Cohort size = 6; 16 cohorts; T = 50; $E(\lvert\epsilon^i\rvert) = 16.25$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 2,317.83$
6	Medium Low ToM		N = 4,800; Cohort size = 6; 16 cohorts; T = 50; $E(\lvert\epsilon^i\rvert) = 16.04$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 3,259.39$
7	Low ToM		N = 4,800; Cohort size = 6; 16 cohorts; T = 50; $E(\lvert\epsilon^i\rvert) = 26.80$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 4,450.37$
8	Price prediction	Bao et al. (2017), EJ	N = 2,400; Cohort size = 6; 8 cohorts; T = 50; $E(\lvert\epsilon^i\rvert) = 1.27$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 4.09$
9	Price prediction & quantity decision		N = 2,400; Cohort size = 6; 8 cohorts; T = 50; $E(\lvert\epsilon^i\rvert) = 7.67$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 19,607.31$
10	Demand side only	Bao and Hommes (2019), JEDC	N = 1,200; Cohort size = 6; 4 cohorts; T = 50; $E(\lvert\epsilon^i\rvert) = 11.78$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 1,089.68$
Panel B: Mixed feedback			
11	Demand and supply-side, low price elasticity of supply (PES)		N = 2,250; Cohort size = 9; 5 cohorts; T = 50; $E(\lvert\epsilon^i\rvert) = 17.01$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 178.12$
12	Demand and supply side, high PES		N = 2,700; Cohort size = 9; 6 cohorts; T = 50; $E(\lvert\epsilon^i\rvert) = 3.39$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 20.24$
Panel C: Negative feedback			
13	Fundamental value (FV) = 56	Bao et al. (2012), JEDC	N = 960; Cohort size = 6; 8 cohorts; T = 20; $E(\lvert\epsilon^i\rvert) = 2.31$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 20.54$
14	FV = 41		N = 1,104; Cohort size = 6; 8 cohorts; T = 23; $E(\lvert\epsilon^i\rvert) = 3.43$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 52.67$
15	FV = 62		N = 1,056; Cohort size = 6; 8 cohorts; T = 22; $E(\lvert\epsilon^i\rvert) = 3.59$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 91.61$
16	Price prediction	Bao et al. (2013), EER	N = 1,200; Cohort size = 6; 4 cohorts; T = 50; $E(\lvert\epsilon^i\rvert) = 2.46$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 18.51$
17	Price prediction and quantity decision		N = 2,100; Cohort size = 6; 7 cohorts; T = 50; $E(\lvert\epsilon^i\rvert) = 4.46$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 44.50$
18	Price prediction, paired with another participant who makes quantity decision		N = 1,800; Cohort size = 6; 6 cohorts; T = 50; $E(\lvert\epsilon^i\rvert) = 3.52$; $E[\text{Var}(\lvert\epsilon^i\rvert)] = 26.30$

Table A.2. Pearce-Hall in positive feedback experiments. Logit estimation for panel data with random intercepts (Panel A) and M-estimator with fixed effects (Panel B). Standard errors in parentheses, robust for clustering at the participant level, and Bonferroni-corrected (for 18 tests); p-values in brackets.

Experiment	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Dependent variable = Y										
λ	0.21	0.14	-0.21	0.03	-0.00	-0.00	0.01	0.13	-0.23	0.52
	(0.13)	(0.16)	(0.17)	(0.07)	(0.08)	(0.07)	(0.07)	(0.10)	(0.11)	(0.16)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[0.846]	[0.030]
Constant	0.14	0.18	0.49	0.32	0.42	0.42	0.37	0.08	0.37	0.05
	(0.08)	(0.09)	(0.11)	(0.04)	(0.04)	(0.04)	(0.03)	(0.06)	(0.07)	(0.10)
	[1.000]	[0.817]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[1.000]	[0.000]	[1.000]
Observations	852	1,053	978	4,558	4,576	4,572	4,551	2,246	2,269	1,138
Number of Participants	48	48	48	96	96	96	96	48	48	24
Panel B: Dependent variable = Δ										
λ	0.13	1.11	0.51	0.20	0.05	-0.02	0.33	0.16	-0.59	0.66
	(0.30)	(0.68)	(0.87)	(0.29)	(0.42)	(0.38)	(0.34)	(0.12)	(0.26)	(0.43)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[0.501]	[1.000]
Observations	852	1,053	978	4,558	4,576	4,572	4,551	2,246	2,269	1,138
Number of Participants	48	48	48	96	96	96	96	48	48	24

Table A.3. Pearce-Hall in mixed feedback experiments. Logit estimation for panel data with random intercepts (Panel A) and M-Estimation with fixed effects (Panel B). Standard errors in parentheses, robust for clustering at the participant level, and Bonferroni-corrected (for 18 tests); p-values in brackets.

Experiment	(11)	(12)
Panel A: Dependent variable = Y		
λ	0.39	0.46
	(0.10)	(0.10)
	[0.001]	[0.000]
Constant	0.24	0.16
	(0.07)	(0.06)
	[0.005]	[0.217]
Observations	2,142	2,513
Number of Participants	45	54
Panel B: Dependent variable = Δ		
λ	0.34	0.33
	(0.12)	(0.17)
	[0.136]	[0.947]
Observations	2,142	2,513
Number of Participants	45	54

Table A.4. Pearce-Hall in negative feedback experiments. Logit estimation for panel data with random intercepts (Panel A) and M-Estimator with fixed effects (Panel B). Standard errors in parentheses, robust for clustering at the participant level, and Bonferroni-corrected (for 18 tests); p-values in brackets.

Experiment	(13)	(14)	(15)	(16)	(17)	(18)
Panel A: Dependent variable = Y						
λ	-0.39	0.04	0.17	-0.02	-0.03	-0.04
	(0.18)	(0.16)	(0.16)	(0.13)	(0.11)	(0.11)
	[0.617]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Constant	-0.06	-0.08	-0.22	-0.00	0.00	-0.02
	(0.09)	(0.09)	(0.09)	(0.07)	(0.06)	(0.06)
	[1.000]	[1.000]	[0.204]	[1.000]	[1.000]	[1.000]
Observations	791	918	826	1,087	1,846	1,537
Number of Participants	48	48	48	24	42	36
Panel B: Dependent variable = Δ						
λ	-0.56	-0.15	0.14	-0.04	-0.08	-0.03
	(0.21)	(0.13)	(0.13)	(0.11)	(0.10)	(0.09)
	[0.171]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Observations	791	918	826	1,087	1,846	1,537
Number of Participants	48	48	48	24	42	36

Table A.5. Ordered-logit modeling of direction of change in learning rate α, Robust Expectations Adaptation (REA). Listed are estimates of the parameters of the equation displayed in (11), per Experiment. Robust standard errors clustered at cohort level in parentheses. p-values in bracket; those for γ, δ, β are Bonferroni-corrected for 54 tests.

Panel A: Positive feedback

Experiment	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
γ	1.71	1.04	1.63	2.09	1.93	1.92	1.90	1.58	1.78	2.02
	(0.16)	(0.16)	(0.21)	(0.14)	(0.09)	(0.12)	(0.12)	(0.15)	(0.18)	(0.20)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
δ	-1.27	-0.37	-1.08	-1.12	-0.80	-0.65	-0.86	-0.75	-1.24	-0.98
	(0.15)	(0.26)	(0.21)	(0.20)	(0.14)	(0.15)	(0.16)	(0.24)	(0.35)	(0.08)
	[0.000]	[1.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.086]	[0.021]	[0.000]
β	0.67	-0.18	0.40	0.53	0.26	0.28	0.38	0.22	0.71	0.56
	(0.14)	(0.29)	(0.22)	(0.17)	(0.12)	(0.12)	(0.12)	(0.20)	(0.15)	(0.03)
	[0.000]	[1.000]	[1.000]	[0.081]	[1.000]	[1.000]	[0.113]	[1.000]	[0.000]	[0.000]
$\bar{\mu}_1$	0.68	0.15	0.47	0.87	0.63	0.66	0.69	0.57	0.70	1.05
	(0.13)	(0.18)	(0.18)	(0.11)	(0.06)	(0.08)	(0.09)	(0.13)	(0.09)	(0.21)
	[0.000]	[0.397]	[0.009]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$\bar{\mu}_2$	0.89	0.25	0.54	0.91	0.65	0.69	0.73	0.65	0.77	1.11
	(0.13)	(0.21)	(0.15)	(0.11)	(0.06)	(0.08)	(0.09)	(0.11)	(0.10)	(0.20)
	[0.000]	[0.230]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	894	1,077	994	4,598	4,590	4,594	4,590	2,283	2,302	1,152
Number of Participants	48	48	48	96	96	96	96	48	48	24

Panel B: Mixed and negative feedback

Experiment	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
γ	4.04	4.02	2.16	1.82	2.03	1.30	1.15	1.29
	(0.39)	(0.16)	(0.16)	(0.25)	(0.16)	(0.32)	(0.20)	(0.24)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.003]	[0.000]	[0.000]
δ	-1.84	-2.00	-1.85	-1.51	-1.52	-1.50	-1.35	-1.17
	(0.41)	(0.26)	(0.24)	(0.33)	(0.32)	(0.14)	(0.32)	(0.43)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.380]
β	1.13	0.97	1.23	0.60	0.54	0.77	0.65	0.57
	(0.43)	(0.27)	(0.20)	(0.15)	(0.19)	(0.10)	(0.18)	(0.25)
	[0.473]	[0.019]	[0.000]	[0.004]	[0.265]	[0.000]	[0.014]	[1.000]
$\bar{\mu}_1$	2.64	2.21	1.05	0.43	0.49	0.49	0.37	0.43
	(0.36)	(0.19)	(0.13)	(0.16)	(0.21)	(0.21)	(0.11)	(0.11)
	[0.000]	[0.000]	[0.000]	[0.009]	[0.018]	[0.019]	[0.000]	[0.000]
$\bar{\mu}_2$	2.69	2.39	1.66	1.12	1.27	0.72	0.71	0.89
	(0.35)	(0.23)	(0.12)	(0.15)	(0.08)	(0.09)	(0.13)	(0.24)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	2,160	2,586	910	1,088	998	1,150	2,012	1,726
Number of Participants	45	54	48	48	48	24	42	36

Table A.6. Logit modeling of direction of change in learning rate α, Robust Expectations Adaptation (REA), in positive feedback experiments. Listed are estimates of the parameters of the equation displayed in (12) and (13), per Experiment. Robust standard errors clustered at participant level in parentheses. p-values in bracket; those for γ, δ, β are Bonferroni-corrected for 54 tests.

Experiment	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Discrete explanatory variable										
γ	1.76	1.05	1.63	2.10	1.94	1.93	1.95	1.60	1.79	2.04
	(0.21)	(0.15)	(0.19)	(0.26)	(0.12)	(0.11)	(0.10)	(0.11)	(0.15)	(0.26)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
δ	-1.27	-0.36	-1.05	-1.12	-0.80	-0.66	-0.91	-0.76	-1.23	-0.97
	(0.28)	(0.25)	(0.26)	(0.17)	(0.14)	(0.14)	(0.12)	(0.18)	(0.21)	(0.26)
	[0.000]	[1.000]	[0.002]	[0.000]	[0.000]	[0.000]	[0.000]	[0.002]	[0.000]	[0.011]
β	0.66	-0.18	0.41	0.53	0.26	0.28	0.38	0.21	0.70	0.56
	(0.21)	(0.19)	(0.20)	(0.12)	(0.10)	(0.11)	(0.09)	(0.15)	(0.14)	(0.21)
	[0.118]	[1.000]	[1.000]	[0.001]	[0.451]	[0.521]	[0.001]	[1.000]	[0.000]	[0.365]
$\bar{\mu}$	-0.77	-0.20	-0.52	-0.88	-0.64	-0.67	-0.70	-0.61	-0.73	-1.11
	(0.16)	(0.12)	(0.15)	(0.14)	(0.07)	(0.07)	(0.06)	(0.09)	(0.10)	(0.20)
	[0.000]	[0.090]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	852	1,053	978	4,558	4,576	4,572	4,551	2,246	2,269	1,138
Number of Participants	48	48	48	96	96	96	96	48	48	24
Panel B: Continuous explanatory variable										
γ	0.39	0.55	0.55	1.34	1.38	1.45	1.29	0.64	0.72	1.10
	(0.18)	(0.11)	(0.18)	(0.10)	(0.08)	(0.07)	(0.07)	(0.15)	(0.17)	(0.18)
	[1.000]	[0.000]	[0.152]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.001]	[0.000]
δ	1.47	0.69	0.88	0.01	0.01	0.01	0.01	0.54	0.15	0.04
	(0.36)	(0.18)	(0.18)	(0.00)	(0.00)	(0.00)	(0.00)	(0.11)	(0.04)	(0.01)
	[0.002]	[0.005]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.013]
β	-0.88	-0.26	-0.56	-0.01	-0.00	-0.01	-0.00	-0.26	-0.14	-0.01
	(0.30)	(0.08)	(0.17)	(0.00)	(0.00)	(0.00)	(0.00)	(0.08)	(0.03)	(0.01)
	[1.000]	[0.001]	[1.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.275]	[1.000]	[0.000]
$\bar{\mu}$	-0.03	-0.16	0.02	-0.52	-0.44	-0.46	-0.40	-0.24	0.04	-0.67
	(0.10)	(0.04)	(0.10)	(0.06)	(0.05)	(0.04)	(0.04)	(0.08)	(0.09)	(0.14)
	[1.000]	[0.000]	[0.810]	[0.000]	[0.000]	[0.000]	[0.000]	[0.005]	[0.663]	[0.000]
Observations	852	1,053	978	4,558	4,576	4,572	4,551	2,246	2,269	1,138
Number of Participants	48	48	48	96	96	96	96	48	48	24

Table A.7. Logit modeling of direction of change in learning rate α, Robust Expectations Adaptation (REA), in mixed feedback experiments. Listed are estimates of the parameters of the equation displayed in (12) and (13), per Experiment. Robust standard errors clustered at participant level in parentheses. p-values in bracket; those for γ, δ, β are Bonferroni-corrected for 54 tests.

Experiment	(11)	(12)
Panel A: Discrete explanatory variable		
γ	4.03	3.97
	(0.35)	(0.25)
	[0.000]	[0.000]
δ	-1.78	-1.84
	(0.37)	(0.27)
	[0.000]	[0.000]
β	1.06	0.89
	(0.36)	(0.26)
	[0.174]	[0.039]
$\bar{\mu}$	-2.65	-2.26
	(0.33)	(0.23)
	[0.000]	[0.000]
Observations	2,142	2,513
Number of Participants	45	54
Panel B: Continuous explanatory variable		
γ	1.27	2.07
	(0.22)	(0.20)
	[0.000]	[0.000]
δ	0.16	0.33
	(0.02)	(0.07)
	[0.000]	[0.000]
β	-0.12	-0.13
	(0.02)	(0.07)
	[0.000]	[0.000]
$\bar{\mu}$	-0.94	-1.41
	(0.18)	(0.16)
	[0.000]	[0.000]
Observations	2,142	2,513
Number of Participants	45	54

Table A.8. Logit modeling of direction of change in learning rate α, Robust Expectations Adaptation (REA), in negative feedback experiments. Listed are estimates of the parameters of the equation displayed in (12) and (13), per Experiment. Robust standard errors clustered at participant level in parentheses. p-values in bracket; those for γ, δ, β are Bonferroni-corrected for 54 tests.

Experiment	(13)	(14)	(15)	(16)	(17)	(18)
Panel A: Discrete explanatory variable						
γ	2.22	1.82	2.09	1.37	1.17	1.35
	(0.20)	(0.19)	(0.23)	(0.22)	(0.16)	(0.21)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
δ	-1.85	-1.40	-1.41	-1.58	-1.40	-1.20
	(0.29)	(0.28)	(0.27)	(0.30)	(0.20)	(0.29)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.002]
β	1.12	0.56	0.50	0.80	0.65	0.55
	(0.22)	(0.19)	(0.21)	(0.19)	(0.11)	(0.13)
	[0.000]	[0.134]	[1.000]	[0.001]	[0.000]	[0.001]
$\bar{\mu}$	-1.28	-0.79	-0.98	-0.62	-0.53	-0.65
	(0.13)	(0.12)	(0.15)	(0.13)	(0.08)	(0.11)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	791	918	826	1,087	1,846	1,537
Number of Participants	48	48	48	24	42	36
Panel B: Continuous explanatory variable						
γ	0.69	1.10	1.32	-0.23	-0.05	0.17
	(0.20)	(0.24)	(0.19)	(0.25)	(0.15)	(0.19)
	[0.031]	[0.000]	[0.000]	[1.000]	[1.000]	[1.000]
δ	0.38	0.03	0.03	0.47	0.14	0.20
	(0.10)	(0.02)	(0.02)	(0.11)	(0.03)	(0.05)
	[0.005]	[1.000]	[1.000]	[0.001]	[0.000]	[0.001]
β	-0.16	0.01	-0.01	-0.12	-0.04	-0.06
	(0.05)	(0.01)	(0.01)	(0.03)	(0.01)	(0.02)
	[0.024]	[0.001]	[0.000]	[1.000]	[1.000]	[1.000]
$\bar{\mu}$	-0.42	-0.59	-0.70	0.05	-0.06	-0.18
	(0.12)	(0.14)	(0.10)	(0.12)	(0.09)	(0.11)
	[0.000]	[0.000]	[0.000]	[0.686]	[0.488]	[0.105]
Observations	791	918	826	1,087	1,846	1,537
Number of Participants	48	48	48	24	42	36

Table A.9. Linear modeling of direction of change in learning rate α, Robust Expectations Adaptation (REA), in positive feedback experiments, allowing for heterogeneity in intercept or/and slope. Listed are estimates of the parameters of the equation displayed in (14), per Experiment. Standard errors robust to clustering at participant level in parentheses. p-values in brackets; those for γ, δ, β are Bonferroni-corrected for 54 tests.

Experiment	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Heterogeneity in intercept only										
γ	0.42	0.27	0.42	0.48	0.44	0.45	0.45	0.38	0.43	0.48
	(0.05)	(0.04)	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.05)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
δ	-0.29	-0.09	-0.25	-0.24	-0.16	-0.14	-0.20	-0.17	-0.29	-0.22
	(0.07)	(0.06)	(0.06)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.05)	(0.05)
	[0.005]	[1.000]	[0.011]	[0.000]	[0.000]	[0.002]	[0.000]	[0.015]	[0.000]	[0.018]
β	0.15	-0.04	0.11	0.12	0.06	0.07	0.09	0.05	0.17	0.13
	(0.05)	(0.05)	(0.05)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.04)
	[0.317]	[1.000]	[1.000]	[0.002]	[0.748]	[0.599]	[0.001]	[1.000]	[0.000]	[0.273]
μ^i	0.32	0.44	0.34	0.29	0.35	0.33	0.33	0.35	0.32	0.24
	(0.04)	(0.03)	(0.03)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.04)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	852	1,053	978	4,558	4,576	4,572	4,551	2,246	2,269	1,138
Number of Participants	48	48	48	96	96	96	96	48	48	24
Panel B: Heterogeneity in intercept and slope										
γ^i	0.43	0.27	0.42	0.49	0.45	0.45	0.46	0.38	0.43	0.50
	(0.06)	(0.03)	(0.05)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
δ	-0.30	-0.09	-0.25	-0.24	-0.17	-0.14	-0.20	-0.17	-0.28	-0.23
	(0.07)	(0.06)	(0.06)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.05)	(0.05)
	[0.002]	[1.000]	[0.016]	[0.000]	[0.000]	[0.003]	[0.000]	[0.020]	[0.000]	[0.009]
β	0.17	-0.05	0.12	0.12	0.07	0.07	0.09	0.04	0.17	0.14
	(0.05)	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)	(0.02)	(0.04)	(0.03)	(0.04)
	[0.114]	[1.000]	[1.000]	[0.002]	[0.415]	[0.627]	[0.001]	[1.000]	[0.000]	[0.161]
μ^i	0.30	0.44	0.34	0.28	0.33	0.33	0.32	0.36	0.32	0.22
	(0.03)	(0.02)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	852	1,053	978	4,558	4,576	4,572	4,551	2,246	2,269	1,138
Number of Participants	48	48	48	96	96	96	96	48	48	24

Table A.10. Linear modeling of direction of change in learning rate α, Robust Expectations Adaptation (REA), in mixed feedback experiments, allowing for heterogeneity in intercept or/and slope. Listed are estimates of the parameters of the equation displayed in (14), per Experiment. Standard errors robust to clustering at participant level in parentheses. p-values in brackets; those for γ, δ, β are Bonferroni-corrected for 54 tests.

Experiment	(11)	(12)
Panel A: Heterogeneity in intercept only		
γ	0.76	0.75
	(0.03)	(0.02)
	[0.000]	[0.000]
δ	-0.27	-0.26
	(0.04)	(0.04)
	[0.000]	[0.000]
β	0.13	0.10
	(0.03)	(0.03)
	[0.006]	[0.073]
μ^i	0.04	0.10
	(0.02)	(0.02)
	[0.000]	[0.000]
Observations	2,142	2,513
Number of Participants	45	54
Panel B: Heterogeneity in intercept and slope		
γ^i	0.74	0.76
	(0.03)	(0.02)
	[0.000]	[0.000]
δ	-0.25	-0.27
	(0.04)	(0.03)
	[0.000]	[0.000]
β	0.11	0.11
	(0.03)	(0.03)
	[0.102]	[0.022]
μ^i	0.06	0.09
	(0.02)	(0.02)
	[0.001]	[0.000]
Observations	2,142	2,513
Number of Participants	45	54

Table A.11. Linear modeling of direction of change in learning rate α, Robust Expectations Adaptation (REA), in negative feedback experiments, allowing for heterogeneity in intercept or/and slope. Listed are estimates of the parameters of the equation displayed in (14), per Experiment. Standard errors robust to clustering at participant level in parentheses. p-values in brackets; those for γ, δ, β are Bonferroni-corrected for 54 tests.

Experiment	(13)	(14)	(15)	(16)	(17)	(18)
Panel A: Heterogeneity in intercept only						
γ	0.53	0.46	0.50	0.34	0.29	0.33
	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.05)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
δ	-0.42	-0.36	-0.32	-0.39	-0.34	-0.29
	(0.07)	(0.07)	(0.06)	(0.07)	(0.05)	(0.07)
	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.010]
β	0.25	0.13	0.12	0.20	0.16	0.13
	(0.05)	(0.04)	(0.05)	(0.04)	(0.03)	(0.03)
	[0.000]	[0.212]	[0.814]	[0.010]	[0.000]	[0.008]
μ^i	0.21	0.30	0.26	0.35	0.37	0.34
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	791	918	826	1,087	1,846	1,537
Number of Participants	48	48	48	24	42	36
Panel B: Heterogeneity in intercept and slope						
γ^i	0.54	0.46	0.49	0.34	0.28	0.33
	(0.06)	(0.07)	(0.07)	(0.03)	(0.03)	(0.03)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
δ	-0.42	-0.35	-0.33	-0.39	-0.35	-0.29
	(0.07)	(0.07)	(0.06)	(0.07)	(0.05)	(0.07)
	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.011]
β	0.25	0.13	0.13	0.19	0.16	0.13
	(0.05)	(0.04)	(0.05)	(0.04)	(0.03)	(0.03)
	[0.000]	[0.184]	[0.394]	[0.012]	[0.000]	[0.013]
μ^i	0.21	0.30	0.26	0.35	0.37	0.34
	(0.03)	(0.04)	(0.04)	(0.02)	(0.02)	(0.02)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	791	918	826	1,087	1,846	1,537
Number of Participants	48	48	48	24	42	36

Table A.12. M-Estimator modeling of magnitude of change in learning rate α, Robust Expectations Adaptation (REA), in positive feedback experiments. Listed are estimates of the parameters of the equation displayed in (15) (Panel A) and (16) (Panel B), per Experiment. Standard error robust to clustering at participant level in parentheses. p-values in brackets; those for γ, δ, β are Bonferroni-corrected for 54 tests.

Experiment	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Discrete measurement of surprise										
γ	2.03	1.40	5.29	4.06	5.77	5.23	4.14	1.07	2.15	4.70
	(0.31)	(0.57)	(0.83)	(0.36)	(0.46)	(0.39)	(0.31)	(0.11)	(0.24)	(0.61)
	[0.000]	[0.998]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
δ	-0.94	2.18	-1.84	0.18	0.75	1.02	0.62	0.12	-0.51	-1.28
	(0.49)	(1.22)	(1.51)	(0.70)	(0.90)	(0.83)	(0.62)	(0.19)	(0.39)	(0.75)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
β	0.92	-1.74	2.25	-0.22	-0.67	-0.65	-0.61	-0.13	0.85	1.08
	(0.46)	(1.06)	(1.32)	(0.64)	(0.79)	(0.68)	(0.53)	(0.15)	(0.27)	(0.74)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Observations	894	1,077	994	4,598	4,590	4,594	4,590	2,283	2,302	1,152
Number of Participants	48	48	48	96	96	96	96	48	48	24
Panel B: Continuous measurement of surprise										
γ	1.71	2.10	3.93	4.10	6.00	5.31	4.37	1.15	0.88	3.83
	(0.28)	(0.50)	(0.73)	(0.37)	(0.46)	(0.40)	(0.31)	(0.13)	(0.25)	(0.76)
	[0.000]	[0.006]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.055]	[0.002]
δ	-0.34	0.94	0.64	0.45	0.63	1.03	0.19	-0.04	1.78	0.40
	(0.39)	(1.02)	(1.09)	(0.23)	(0.42)	(0.43)	(0.26)	(0.14)	(0.29)	(0.91)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[0.000]	[1.000]
β	-0.41	0.94	0.26	0.03	0.03	0.03	0.00	-0.02	0.34	0.04
	(0.36)	(1.80)	(1.06)	(0.01)	(0.02)	(0.02)	(0.00)	(0.07)	(0.00)	(0.06)
	[1.000]	[1.000]	[1.000]	[0.320]	[1.000]	[1.000]	[1.000]	[1.000]	[0.000]	[1.000]
Observations	894	1,077	994	4,598	4,590	4,594	4,590	2,283	2,302	1,152
Number of Participants	48	48	48	96	96	96	96	48	48	24

Table A.13. M-Estimator modeling of magnitude of change in learning rate α, Robust Expectations Adaptation (REA), in mixed feedback experiments. Listed are estimates of the parameters of the equation displayed in (15) (Panel A) and (16) (Panel B), per Experiment. Standard error robust to clustering at participant level in parentheses. p-values in brackets; those for γ, δ, β are Bonferroni-corrected for 54 tests.

Experiment	(11)	(12)
Panel A: Discrete measurement of surprise		
γ	3.17	3.67
	(0.25)	(0.18)
	[0.000]	[0.000]
δ	0.14	-0.30
	(0.32)	(0.24)
	[1.000]	[1.000]
β	0.02	0.23
	(0.33)	(0.22)
	[1.000]	[1.000]
Observations	2,160	2,586
Number of Participants	45	54
Panel B: Continuous measurement of surprise		
γ	3.19	3.60
	(0.20)	(0.16)
	[0.000]	[0.000]
δ	0.12	-0.16
	(0.09)	(0.10)
	[1.000]	[1.000]
β	-0.00	-0.02
	(0.00)	(0.01)
	[1.000]	[1.000]
Observations	2,160	2,586
Number of Participants	45	54

Table A.14. M-Estimator modeling of magnitude of change in learning rate α, Robust Expectations Adaptation (REA), in negative feedback experiments. Listed are estimates of the parameters of the equation displayed in (15) (Panel A) and (16) (Panel B), per Experiment. Standard error robust to clustering at participant level in parentheses. p-values in brackets; those for γ, δ, β are Bonferroni-corrected for 54 tests.

Experiment	(13)	(14)	(15)	(16)	(17)	(18)
Panel A: Discrete measurement of surprise						
γ	1.31	0.82	0.97	0.72	0.62	0.74
	(0.16)	(0.12)	(0.11)	(0.17)	(0.10)	(0.11)
	[0.000]	[0.000]	[0.000]	[0.013]	[0.000]	[0.000]
δ	-0.82	-0.48	-0.64	-0.95	-0.79	-0.61
	(0.22)	(0.16)	(0.10)	(0.25)	(0.15)	(0.17)
	[0.033]	[0.238]	[0.000]	[0.041]	[0.000]	[0.044]
β	0.70	0.22	0.14	0.42	0.44	0.31
	(0.15)	(0.10)	(0.08)	(0.14)	(0.08)	(0.08)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Observations	910	1,088	998	1,150	2,012	1,726
Number of Participants	48	48	48	24	42	36
Panel B: Continuous measurement of surprise						
γ	0.94	0.72	0.98	0.54	0.43	0.60
	(0.13)	(0.11)	(0.13)	(0.15)	(0.10)	(0.11)
	[0.000]	[0.000]	[0.000]	[0.062]	[0.004]	[0.000]
δ	-0.16	-0.23	-0.60	-0.59	-0.36	-0.33
	(0.17)	(0.10)	(0.09)	(0.18)	(0.10)	(0.13)
	[1.000]	[1.000]	[0.000]	[0.192]	[0.049]	[0.797]
β	-0.03	0.00	-0.01	-0.02	-0.00	-0.01
	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
	[1.000]	[1.000]	[0.153]	[1.000]	[1.000]	[1.000]
Observations	910	1,088	998	1,150	2,012	1,726
Number of Participants	48	48	48	24	42	36

B Experimental Instruction for a Typical LtFE

The following instructions are taken from Bao et al. (2024).

Welcome to our experiment! You are participating in an experiment with a real monetary reward. There are 6 participants in each market. In other words, your payoff from the experiment depends on your decision and the decisions of the other 5 participants in your market. The experiment consists of 50 periods. When the experiment ends, we will pay you according to the total number of points you earned. The exchange rate is 80 points = 1 RMB.

The experiment is anonymous. You do not know the identity of the other participants, nor do they know yours. You are not allowed to communicate with others during the experiment, and please place your cell phone in the place assigned by the experimenter.

The following is a detailed description of the experimental setup. Please read it carefully and listen to the explanation by the experimenter. If you have any questions, please feel free to ask.

Your role in the experiment is a financial advisor to an investment fund that wants to optimally invest a large amount of money. The fund is a major participant in the market of some risky assets. The experiment consists of 50 periods. Before the beginning of each period, you must predict the asset price of the risky asset for the investment fund. Based on your prediction, the fund will decide the unit of the risky asset to purchase or sell. The investment fund has two investment options for the limited fund: a risk-free investment (e.g., government bond), with an interest rate of 5%; and a risky investment (e.g., stock), where the value of the dividend of the stock is 3.3 points. According to finance theory, the fundamental value of the risky asset is positively correlated with its dividend and negatively correlated with the interest rate of the risk-free asset.

Note that your prediction is the only determinant of the fund's purchasing/selling decision. The more accurate you predict, the more money the fund will earn. Accordingly, your earnings in each period in the experiment solely depend on your forecasting accuracy in each period. At the end of the experiment, we will pay you according to the total points you earn in 50 periods.

Participants need to complete a test before the beginning of the experiment. Please answer the questions carefully.

1. The determination of the asset price

The stock price is determined by the following mechanism: when the total demand for risky asset in the market is larger than the total supply, or when the total assets firms want to purchase is larger than the total assets firms want to sell, the price will increase. Conversely, if the total demand for risky asset in the market is smaller than the total supply, the price will decrease. This rule is generally consistent with the reality.

There are some large investment funds in the market, where each of the investment funds is advised by a financial advisor played by a participant in the experiment (like you). Generally speaking, the funds will buy more of the risky asset if the financial advisor forecasts that the price of the risky asset will increase, whereas they will sell more assets if the financial advisor forecasts that the price of the risky asset will decrease. The

total demand and total supply of the asset are determined by the total purchasing/selling decisions of these large investment funds in the market.

2. Your task in the experiment

Your only task in the experiment is to forecast the market price in each period. At the beginning of the experiment, you need to submit your forecast for the price in the first period, where the forecast should range between 0 to 100. After all participants have submitted their forecasts for the first period, the investment fund played by the computer will make the decisions on purchasing/selling quantities based on each participant's predictions. After that, the experimental program will determine the asset price in the current period using the total purchasing and selling quantities and reveal it to everyone. Based on your forecasting error, your earnings (in points) for period 1 will be calculated.

Next, you need to submit your forecasts for the price in the second period. After all participants have submitted their predictions for the second period, the market price in the second period will be calculated based on all the predictions and their corresponding trading decisions. This process continues for 50 periods. In each period, the available information comprises the previous market prices, your previous predictions, and your previous earnings. **Specifically, the experimental procedure in each period is as follows:**

In general, at period t ($t \geq 2$), participants need to **predict the asset price in the current period t at the beginning of period t.** When forecasting the price, the following information will be disclosed on the user interface: **the interest rate of risk-free asset γ, expected dividend of risky asset y, participant's previous predictions up to $t-1$, previous prices up to period $t-1$, previous total earnings up to period $t-1$.**

Participants only need to fill up **the price forecast for the current period** in the corresponding experimental program. After collecting the price forecasts from all participants, the program will disclose the market price in period t, P_t, and the earnings in point that is calculated based on the predicting error between the price forecasts $P^e_{h,t}$ and market price P_t, **at the end of period t.** In other words, **the actual asset price that was predicted at the beginning of each period will be disclosed at the end of each period.**

Simply put, when $t \geq 2$ such as at the beginning of period 5, participants will need to predict the market price in period 5. After all participants submit their forecasts on the price for period 5, the actual market price for the asset in period 5 will be disclosed at the end of period 5. Next is to predict the price in period 6 at the beginning of period 6, where the actual market price for period 6 will be derived using the forecasting price and its corresponding demand/supply equation. And so on.

Specifically, when $t = 1$, participants only need to submit the price forecasts, where no market price and earnings (in points) will be disclosed. The price forecasts for period 1 need to range between 0 and 100.

Note that 60 seconds is given to you for forecasting in each period. Please submit your price forecasts before the end of the countdown. Except for the first period, where the price forecasts need to range between 0 to 100, all price forecasts from period 2 could range between 0 to 1000. All predictions could have an accuracy up to 2 decimal points.

3. Your payoff

Earnings in each period will depend only on the forecasting accuracy in the corresponding period. The more accurate you predict the asset price in each period, the higher your aggregate earnings will be. In other words, as your prediction error increases, or as the difference between the actual stock price and your price forecasts increases, your payment decreases. When your forecast equals the stock price, you get 100 points. When your prediction error is greater than 7, you get 0 points. Hence, your earnings in each period are:

$$\text{earning} = \max\left\{100 - \frac{100}{49} \times (\text{prediction error})^2, 0\right\}$$

The earnings with regards to the prediction error is plotted as follows: