Unrealistic Expectations and Misguided Learning*

Paul Heidhues, ESMT

Botond Kőszegi, Central European University

Philipp Strack, UC Berkeley

January 11, 2016

Abstract

We explore the learning process and behavior of an individual with unrealistically high expectations about ability ("overconfidence") when outcomes also depend on an external fundamental that affects the optimal action. Moving beyond existing results in the literature, we show that the agent's belief regarding the fundamental converges under weak conditions. Furthermore, we identify a broad class of situations in which "learning" about the fundamental is self-defeating: it leads the individual systematically away from the correct belief and toward lower performance. Due to her overconfidence, the agent—even if initially correct—becomes too pessimistic about the fundamental. As she adjusts her behavior in response, she lowers outcomes and hence becomes even more pessimistic about the fundamental, perpetuating the misdirected learning. The greater is the loss from choosing a suboptimal action, the *further* the agent's action ends up from optimal. We argue that the decision situations in question are common in economic settings, including delegation, organizational, public-policy, and labor-leisure choices. We partially characterize environments in which self-defeating learning occurs, and show that the decisionmaker is indistinguishable from a rational agent if, and in some cases only if, a specific *non*-identifiability condition is satisfied. In contrast to an overconfident agent, an underconfident agent's misdirected learning is self-limiting and therefore not very harmful.

*We are grateful to Daniel Benjamin, Volker Nocke, Demian Pouzo, as well as seminar and conference audiences for insightful comments and feedback. Heidhues and Kőszegi thank the European Research Council for financial support under Starting Grant #313341. Part of this research was carried out while Heidhues and Kőszegi visited the WZB, which they thank for its hospitality. Kinga Marczell provided excellent research assistance.

1 Introduction

Large literatures in psychology and economics suggest that in many situations, individuals have unrealistically positive beliefs about their traits or prospects, and researchers have began to investigate the nature of this "overconfidence" and study its implications for economic interactions.[1] One important question concerning individuals with overconfident or otherwise biased beliefs is how they update these beliefs when relevant information comes in. Indeed, classical results identify conditions under which learning leads to correct beliefs (e.g., Savage, 1954, Chapter 3), and more recent research explores ways in which a biased learning process can lead to overconfident beliefs (e.g., Gervais and Odean, 2001, Chiang, Hirshleifer, Qian and Sherman, 2011).

With overconfident individuals, however, interesting and potentially important learning questions arise not only regarding the overconfident beliefs themselves, but also regarding other variables—simply because overconfidence affects inferences about other variables. For example, if a team member is overly full of herself and is hence disappointed by her team's performance, she might conclude that her teammates are less talented or lazier than she thought—and this can affect her perceived optimal behavior, such as how much work to delegate to others.

In this paper, we investigate an individual's learning process and behavior when outcomes depend both on a factor that she is overconfident about and on an additional fundamental that affects the optimal action. Moving beyond existing results in the literature, we show that the agent's belief regarding the fundamental converges under weak conditions. Furthermore, we identify a broad and economically important class of situations in which the agent's inferences regarding the fundamental are self-defeating: they lead her systematically away from the correct belief and toward lower performance. Continuing with the team-production example, as the overconfident team member becomes unduly pessimistic about her teammates, she increases her control of the team, lowering team performance. She misinterprets this low performance as reflecting even more negatively on her teammates, perpetuating the misdirected learning further. Perversely, the greater is the loss from choosing a suboptimal action, the *further* the agent's action ends up from optimal. We partially characterize environments in which self-defeating learning occurs, and show that the

[1] We use the term "overconfidence" to mean broadly any unrealistic beliefs, especially about ability or other important personal characteristics, that lead a person to expect good outcomes. The same expression is often used more specifically to denote overly narrow confidence intervals (Moore and Healy, 2008). We review evidence for and theoretical work on overconfidence in Section 5.

decisionmaker is indistinguishable from a rational agent if, and in some cases only if, a specific *non*-identifiability condition is satisfied. In contrast to an overconfident agent, an underconfident agent's misdirected learning is self-limiting and therefore not very harmful.

We present our framework in Section 2.1. In each period $t \in \{1, 2, 3, \dots\}$, the agent produces observable output $q_t = Q(e_t, a, b_t)$, which depends on her action e_t, her ability or other output-relevant parameter a, and external conditions b_t beyond her control. We assume that Q is increasing in a and b_t, and that the optimal action is increasing in b_t.[2] The states b_t are independent normally distributed random variables with mean equal to a fixed fundamental Φ and precision equal to h_ϵ. The agent does not directly observe Φ or b_t, but attempts to learn about them from her observations and to adjust her action optimally in response. Crucially, the agent is overconfident: while her true ability is A, she believes with certainty that it is $\tilde{a} > A$.

In Section 2.2, we argue that beyond team production, this reduced-form model captures a number of economically important situations in which individuals may be overconfident. A person may not know how much of a certain activity, such as work or training, maximizes well-being or performance. A principal may not know the optimal level of control or explicit incentives for her employees. And a policymaker may not know what scale of intervention, such as regulations in an industry or restrictions on drug use, leads to the best outcomes.

In Section 3, we study the limit version of our model in which noise vanishes ($h_\epsilon \to \infty$), beginning in Sections 3.1 and 3.2 with identifying our central mechanism, summarizing a variety of casual evidence for it, and discussing key economic implications. To illustrate in a specific case, suppose for a moment that the optimal action does not depend on ability, and that the agent starts off with the correct mean belief about the fundamental, taking the optimal action. Yet because she is overconfident, the output she obtains is to her unexpectedly low. To explain this observation, she concludes that the fundamental is lower than she thought, and adjusts her action accordingly. But because she now takes a suboptimal action, her output—instead of achieving the optimum she was expecting to reach by the adjustment—falls. As a result, she becomes even more pessimistic about the fundamental, continuing the vicious circle of self-defeating learning.

More generally, we establish that for any prior, the agent's belief regarding the fundamental changes monotonically over time and converges to a—not necessarily unique—limiting belief that

[2] Note that so long as these effects are monotonic, the above assumptions are just normalizations.

is below the true fundamental. Similarly to Esponda and Pouzo (forthcoming), the limiting belief has the intuitive consistency property that—given her false beliefs and her action induced by these beliefs—the level of output the agent expects coincides with the level of output she produces. Furthermore, the agent's limiting belief satisfies a surprising and perverse comparative static: the more important it is for her to take the right action—that is, the greater is the loss from a suboptimal action—the *further* her belief and behavior end up from optimal. Intuitively, when choosing a suboptimal action is more harmful, the agent hurts herself more through her misguided learning. To develop a consistent theory of her observations, therefore, she must become more pessimistic about the world.

We also consider what happens if—such as when a dissatisfied team member can replace her teammate—the agent can choose between the above task and an outside option. Consistent with the literature on overconfidence, the agent might be too prone to enter into and initially persist in the task. In contrast to received wisdom, however, our model predicts that the agent's growing pessimism about the fundamental may induce her to *exit* the task too easily, and by implication to jump too much between tasks. This prediction is consistent with the observation that many documented effects of overconfidence in economic settings, such as the pursuit of mergers and innovations by overconfident CEOs or the creation of new businesses by overconfident entrepreneurs, pertain to *new* directions.

In Section 3.3, we partially characterize the conditions on Q under which self-defeating misguided learning—self-defeating in the sense that the ability to update her action based on what she has learned makes an agent with an approximately correct prior about the fundamental worse off—does versus does not occur. Roughly, self-defeating learning occurs if the optimal action depends on the fundamental, and either depends sufficiently less on ability, or does so in the opposite way. As a conceptually interesting case, we show that the agent's behavior is always indistinguishable from someone's with a correct belief about ability, and therefore optimal, if and only if output has the form $Q(e_t, B(a, b_t))$—that is, ability and the fundamental are *not* separately identifiable. This conclusion contrasts with the lesson from classical learning settings that non-identifiability hinders efficient learning. Intuitively, because ability and the fundamental do not have independent effects on output, the agent's misinference about the fundamental can fully compensate her overconfidence, and hence overconfidence does not adversely affect her. If her team's output depends on her effort and the team's total ability, for instance, she is able to correctly infer total ability, and chooses effort optimally.

In Section 3.4, we compare underconfident agents (for whom $\tilde{a} < A$) to overconfident agents when output has the simple form $Q(e_t, a, b_t) = a + b_t - L(|e_t - b_t|)$ for a loss function L. We identify an interesting asymmetry: while an overconfident agent's utility loss from misguided learning can be an arbitrarily large multiple of her overconfidence $\tilde{a} - A$, an underconfident agent's limiting utility loss is bounded by her underconfidence $A - \tilde{a}$. To understand the intuition, consider again an agent who starts off with a correct initial belief about the fundamental. Upon observing a better performance than she expected, the underconfident agent concludes that the fundamental is better than she thought, and revises her action. The resulting utility loss, however, leads her to reassess her first optimistic revision of her belief, bringing her belief back toward the true fundamental. In this sense, her misinference regarding the fundamental—which with overconfidence was self-reinforcing—is now self-correcting.

In Section 4, we consider the version of our model with non-trivial noise. We use tools from stochastic approximation theory to show that under weak conditions on Q, the agent's belief converges with probability one to a point where the external conditions $\tilde{b}_t$ that the agent thinks she has observed on average equal to her belief about the fundamental. Within our setting, this convergence result substantially supersedes previous ones in the literature, including those of Berk (1966)—where the agent does not make decisions based on her beliefs—and Esponda and Pouzo (forthcoming)—where convergence to stable beliefs is established only for nearby priors and in expectation approximately optimal actions. Even with convergence guaranteed, however, with noise the agent's limiting belief depends in a complicated way on the shape of Q. While our results are therefore weaker than in the noiseless case, we confirm the main mechanisms above also in this version of our model.

In Section 5, we relate our paper to the two big literatures our paper connects, that on overconfidence and that on learning with misspecified models. From a methodological perspective, our model is a special case of Esponda and Pouzo's (forthcoming) framework for games when players have misspecified models. Because we have an individual-decisionmaking problem with a specific structure, we can derive novel and economically important results that are not possible in the general framework. In particular, we are unaware of previous research studying the implications of overconfidence for inferences about decision-relevant other variables, and how these inferences interact with behavior. In addition, to our knowledge our paper is the only one to analyze an individual-decisionmaking problem that features systematic self-defeating learning.

In Section 6, we conclude by discussing some potential applications of our framework for multi-person situations, and by pointing out limitations of our approach. Most importantly, our assumption that the agent has a degenerate overconfident belief about ability prevents us from studying whether and how the forces we identify interact with the forces that generate widespread overconfident beliefs in the population.

2 Learning Environment

In this section, we introduce our basic framework, outline possible economic applications of it, and perform a few preliminary steps of analysis.

2.1 Setup

In each period $t \in \{1, 2, 3, \dots\}$, the agent produces observable output $q_t \in \mathbb{R}$ according to the twice differentiable output function $Q(e_t, a, b_t)$, which depends on her action $e_t \in (\underline{e}, \overline{e})$, her ability $a \in \mathbb{R}$, and an unobservable external state $b_t \in \mathbb{R}$ beyond her control.[3] We assume that $b_t = \Phi + \epsilon_t$, where Φ is an underlying fixed fundamental and the ϵ_t are independent normally distributed random variables with mean zero and precision h_ϵ. The agent's prior is that Φ is distributed normally with mean $\tilde{\phi}_0$ and precision h_0.

Throughout the paper, we impose the following largely technical assumptions on Q:

Assumption 1 (Technical Assumptions)**.** (i) $Q_{ee} < 0$ and is bounded, and $\lim_{e \to \underline{e}} Q_e(e, a, b_t) > 0 > \lim_{e \to \overline{e}} Q_e(e, a, b_t)$ for all a, b_t; (ii) Q_a and Q_b are bounded from above and below by positive numbers; (iii) $Q_{eb} > 0$; (iv) $Q_{eb}/|Q_{ee}|$ is bounded from above and Q_{ea} is bounded from below.

Part (i) guarantees that there is always a unique myopically optimal action. Part (ii) implies that output is increasing in the state and ability, and Part (iii) implies that the optimal action is increasing in the state. Note that so long as the effects implied by Parts (ii) and (iii) are monotonic,

[3] In Footnote 10 and Proposition 8, we use output functions with kinks. In the latter case, it is not immediately obvious that the methods we develop for the differentiable case apply, so we prove convergence of the agent's belief directly.

our directional assumptions on them are just normalizations. We use the bounds on the derivatives of Q_a and Q_b for our proof that the agent's belief converges, and the positive lower bound on Q_b also guarantees that the agent can always find an explanation for her observations. Finally, the bounds in Part (iv) ensure that the agent's action and inference are Lipschitz continuous in her belief.

While the above assumptions are technical, the following one is substantial: it is a sufficient condition for self-defeating misguided learning to occur.

Assumption 2 (Sufficient Condition for Self-Defeating Learning)**.** $Q_{ea} \leq 0$.

Assumption 2 implies that an increase in ability either has no effect on the optimal action, or its effect is opposite to that of the state. We explore the implications of misguided learning under this assumption, and also study what happens when the assumption is not satisfied.

Crucially, we assume that the agent is overoptimistic about her ability: while her true ability is A, she believes with certainty that it is $\tilde{a} > A$. Let $\Delta = |A - \tilde{a}|$ denote the degree of the agent's overconfidence. Given her inflated self-assessment, the agent updates her belief about the fundamental in a Bayesian way, and chooses her action in each period to maximize perceived discounted expected output.

We specify the agent's belief about ability as degenerate for two main reasons. Technically, the assumption is invaluable for tractability. If the decisionmaker was uncertain about ability in addition to the fundamental, we would have a multi-dimensional updating problem that also features an experimentation motive, and we do not know how to analyze this complicated problem. Nevertheless, continuity suggests that even if the agent was somewhat uncertain about ability and her true ability was in the support of her prior, the mechanism of our model would remain operational, albeit only for a while.

Perhaps more importantly, we view an assumption of overconfident beliefs that are not updated downwards as broadly realistic. Quite directly, such an assumption is consistent with the view of many psychologists that individuals are extremely reluctant to revise self-views downwards (e.g., Baumeister, Smart and Boden, 1996). More generally, this kind of assumption can be thought of as a stand-in for forces explored in the psychology and economics literatures (but not explicitly modeled here) that lead individuals to maintain unrealistically positive beliefs. Viewed from this perspective, any model in which the agent eventually learns her true ability contradicts observed widespread overconfidence among individuals who have had plenty of opportunity to learn about themselves. Of course, as we emphasize in the conclusion, it would be interesting to study how the forces maintaining overconfidence interact with the forces we explore.

Consistent with the above arguments, in describing and discussing our model and results we interpret a as ability and Φ as an external feature of the world. But more broadly, a could stand for any variable—possibly an external one—that leads the agent to be steadfastly and unrealistically optimistic about her prospects, and Φ could stand for any variable—possibly an internal one—about which the agent is willing to draw inferences. For example, she may have overly positive views about her country or organization; and she may be willing to draw inferences about internal skills that are not central to her self-esteem (e.g., mathematical talent for an athlete).

2.2 Applications

In this section, we argue that several economically important settings fit our reduced-form model of Section 2.1. In each of these settings, it seems plausible to assume that the agent has unrealistic expectations regarding what can be achieved, and (consistent with Assumption 2) the optimal action is either much more sensitive to the state than to ability, or depends on ability in the opposite way than on the state.

Delegation. The decisionmaker is working in a team with another agent, and must decide how much of the work to delegate. The output of the team is $Q(e_t, a, b_t) = af(e_t) + b_t g(e_t)$, where $e_t \in (0,1)$ is the proportion of the job the decisionmaker delegates, b_t is the teammate's ability, and $f(e), g(e) > 0, f'(e), f''(e), g''(e) < 0, g'(e) > 0$. Then, the higher is the decisionmaker's ability and the lower is the teammate's ability, the lower is the optimal amount of delegation.

Control in Organizations. A principal is deciding on the incentive system to use for an agent who chooses two kinds of effort, overt effort x_t^o (e.g., writing reports) and discretionary effort x_t^d (e.g., helping others in the organization). The principal can incentivize overt effort, for instance through monitoring (e.g., requiring and reading reports) or explicit incentives written on an objective signal of overt effort. For simplicity, we assume that the principal chooses x_t^o directly. Consistent with the literature on multitasking starting with Holmström and Milgrom (1991), we also assume that discretionary effort is a decreasing function of overt effort. In addition, we suppose that discretionary effort is an increasing function of the agent's intrinsic motivation b_t. Writing discretionary effort as $x_t^d = x^d(x_t^o, b_t)$, the principal's profit is $R(a, x_t^o, x^d(x_t^o, b_t))$, where a is the quality of the organization, the principal's ability, or other factor affecting overall productivity. Supposing that the optimal overt effort is decreasing in intrinsic motivation,[4] this model reduces to our setting with $Q(e_t, a, b_t) = R(a, -e_t, x_t^d(-e_t, b_t))$.

An alternative interpretation of the same framework is that x_t^o is the agent's "mechanical" input into the organization, x_t^d is her "creative" input, and b_t is her ability. In this interpretation, creative input depends on creative effort—which is a substitute to overt effort—and ability.

Public policy. A policymaker is choosing a policy e_t to maximize the performance of some aspect of the economy, $Q(e_t, a, b_t) = a + b_t - L(e_t - b_t)$, where b_t is the optimal policy, $L(e_t - b_t)$ is the loss from a potentially suboptimal choice, and a is the policymaker's ability or her party's or country's potential. Several specific examples broadly fit these features.

First, suppose that q_t is the health outcomes of the population, and b_t is a measure of the prevailing quality of the healthcare market, which depends on factors such as adverse selection, externalities, and private decisionmakers' conflicts of interest. The policymaker's goal is to align the amount of deregulation, e_t, with the optimal level. More widespread problems with the market mean greater optimal regulation, so that e_t is increasing in b_t.

A second relevant example is drug policy. Suppose that q_t is the well-being of the population in relation to drug-related crime, b_t is the underlying condition of the population with respect to drug use, and e_t is the degree of drug liberalization. The extent of restrictions must be optimally aligned with underlying conditions to minimize drug-related crime.

Military policy aimed at minimizing attacks on the army in an occupied enemy region provides a third example. Suppose q_t measures the well-being of the army, which is a decreasing function of the level of attacks, b_t is the friendliness of the local population, and e_t is the level of restraint in engagement. The friendlier is the local population, the higher is the optimal level of restraint in engagement.

Work and Social Status. A person's social status is given by $a + f(e_t, b_t)$, where a is how likable she is, e_t is the amount of leisure she takes, and b_t captures the social norm for work or consumption, with higher b_t corresponding to a lower norm. We assume that f is increasing in b_t and decreasing in e_t, and that $f_e(e_t, b_t)$ is increasing in b_t. This means that the more the agent

[4] One simple functional form capturing our discussion is $R(a, x_t^o, x_t^d(x_t^o, b_t)) = a + x_t^o + x_t^d(x_t^o, b_t)$, where the discretionary effort function satisfies $\frac{\partial x_t^d(x_t^o, b_t)}{\partial x_t^o} < 0$, $\frac{\partial x_t^d(x_t^o, b_t)}{\partial b_t} > 0$, and $\frac{\partial^2 x_t^d(x_t^o, b_t)}{\partial x_t^o \partial b_t} < 0$.

works or consumes, the greater is her social status, but the higher is the norm, the lower is her social status. Furthermore, the higher is the norm for work, the greater is the marginal social-status return of work. The agent's overall happiness or life satisfaction is $Q(e_t, a, b_t) = a + f(e_t, b_t) + b(e_t)$, where $b(e_t)$ is her utility from leisure.[5]

2.3 Preliminaries

We begin the analysis of our model by noting a few basic properties. After observing output q_t generated by the realized state b_t, the agent believes that the realized state was a $\tilde{b}_t$ satisfying

$$Q(e_t, \tilde{a}, \tilde{b}_t) = q_t = Q(e_t, A, b_t)\,. \tag{1}$$

By Assumption 1, there is a unique $\tilde{b}_t$ satisfying Equation 1. Since b_t does not depend on e_t, therefore, the agent believes that whatever action she chooses, she will infer the same signal $\tilde{b}_t$ about Φ. This implies that she chooses her action in each period to maximize that period's perceived expected output. In addition, since the agent believes that $\tilde{b}_s \sim \mathcal{N}(\Phi, h_\epsilon^{-1})$, at the end of period $t \geq 1$ she believes that Φ is distributed with mean

$$\tilde{\phi}_t = \sum_{s=1}^{t} \frac{h_\epsilon \tilde{b}_s + h_0 \tilde{\phi}_0}{t h_\epsilon + h_0} \tag{2}$$

and precision $h_0 + t h_\epsilon$. Hence, at the beginning of period t, the agent believes that $\tilde{b}_t$ is normally distributed with mean $\tilde{\phi}_{t-1}$ and variance $h_\epsilon^{-1} + [h_0 + (t-1)h_\epsilon]^{-1}$. Let $\tilde{\mathbb{E}}_{t-1,\tilde{\phi}_{t-1}}$ be her subjective expectation with respect to this distribution.

Lemma 1 (Agent Chooses Myopic Decision Rule)**.** *Suppose Assumption 1 holds. In each period t, there is a unique optimal action given by*

$$e^*(t, \tilde{\phi}_{t-1}) = \arg\max_e \tilde{\mathbb{E}}_{t-1,\tilde{\phi}_{t-1}} \left[Q(e, \tilde{a}, \tilde{b}) \right]. \tag{3}$$

Furthermore, $e^*(t, \tilde{\phi}_{t-1})$ is differentiable and increasing in $\tilde{\phi}_{t-1}$.

The proof of Lemma 1 also implies that there is a unique confident action

$$e^*(\tilde{\phi}) = \arg\max_e \tilde{\mathbb{E}}_{\infty,\tilde{\phi}} \left[Q(e, \tilde{a}, \tilde{b}) \right]$$

that the agent perceives as optimal if she is confident that the fundamental is $\tilde{\phi}$. Furthermore, this confident action $e^*(\tilde{\phi})$ is differentiable and increasing in $\tilde{\phi}$.

[5] Reinterpreting e_t as the amount of rest and b_t as the state determining the optimal training schedule, our example describes training decisions by athletes as well.

3 Noiseless Limit

In this section, we investigate properties of the agent's beliefs and behavior in the limiting case of our model when noise vanishes ($h_\epsilon \to \infty$). While somewhat artificial, this version allows us to study not only the agent's limiting belief, but also her full path of beliefs, as well as to pose a question on the observational equivalence between rational and overconfident agents. In the next section, we show that all of our qualitative results survive in a weaker form with non-trivial noise.

Formally, we assume that in every period, the agent observes $Q(e_t, A, \Phi)$ with probability one, and to update her belief she uses the limit of the updating rule (2) as $h_\epsilon \to \infty$, so that

$$\tilde{\phi}_t = \frac{\sum_{s=1}^t \tilde{b}_s}{t} = \frac{t-1}{t}\tilde{\phi}_{t-1} + \frac{1}{t}\tilde{b}_t.$$

To avoid some clumsy statements regarding the first period, we assume that in the noiseless limit the precision of the prior also approaches infinity, but at a slower speed than the signals ($h_0 \to \infty$ and $h_\epsilon/h_0 \to \infty$). This does not affect any of our points.

3.1 Example and Economic Implications

We begin by presenting the main insights and economic implications of our model using a simple example. We consider an output function of the form

$$Q(e_t, a, b_t) = a + b_t - L(e_t - b_t), \tag{4}$$

where L is a symmetric loss function with $|L'(x)| < k < 1$ for all x. To make our points most starkly, we also suppose that the agent's mean belief is initially correct: $\tilde{\phi}_0 = \Phi$.

Fixed Action. As a benchmark case, we suppose for a moment that e_t is exogenously given and constant over time at level $e \geq \Phi$. To explain output $q_1 = A + \Phi - L(e - \Phi)$ given her overconfidence, the agent solves $\tilde{a} + b_1 - L(e - \Phi) = q_1 = A + \Phi - L(e - \Phi)$. This yields $\tilde{\phi}_1 = b_1 < \Phi$. Intuitively, the agent is surprised by the low output she observes, and explains it by concluding that the fundamental is worse than she thought. This tendency to attribute failures to external factors provides a formalization of part of the self-serving attributional bias documented by Miller and Ross (1975) and the literature following it. Indeed, one account of the bias is that individuals have high expectations for outcomes, and update when outcomes fall short (Tetlock and Levi, 1982, Campbell and Sedikides, 1999).

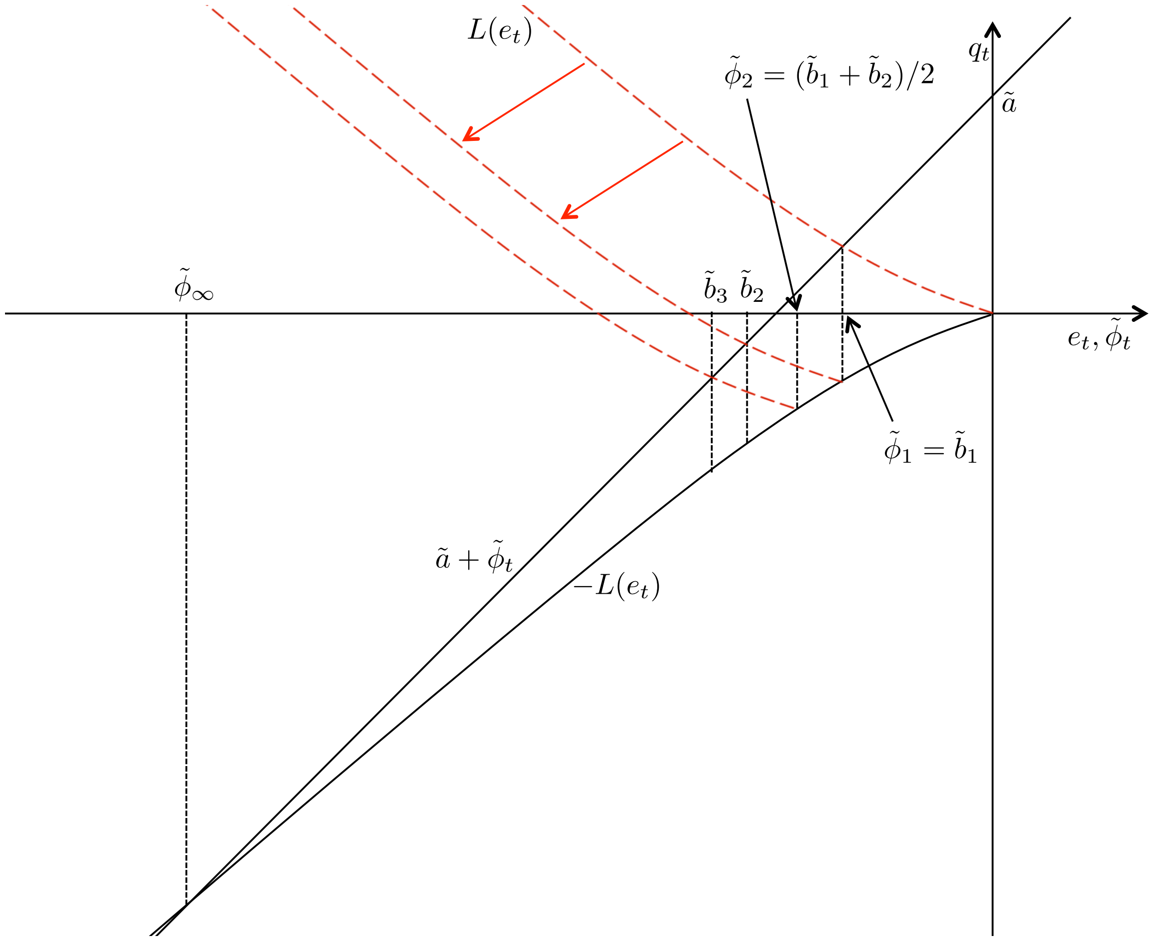

Figure 1: Learning Dynamics

Since the agent takes the same action every period and hence observes the same output, after period 1 she does not update her belief further ($\tilde{\phi}_t = \tilde{\phi}_1$ for all $t > 1$). And while her inference about the fundamental is misguided—it takes her away from her correct prior—in the current setup with a fixed action it is harmless or potentially even beneficial. For instance, because the agent now correctly predicts her output, she makes the correct choice when deciding whether to choose this task over an outside option with a given level of utility.[6]

Endogenous Action. We next turn to illustrating the main theoretical point of our paper: that when the agent can adjust her behavior to what she thinks she has learned about the fundamental,

[6] As a simple implication, if her high belief about ability is due to ego utility as in Kőszegi (2006), then misdirected learning allows her to have her cake and eat it too: she can maintain the pleasure from believing that her ability is high, while not suffering any losses associated with incorrect beliefs.

she is often drawn into extremely suboptimal behavior over time.

Suppose that in each period the agent chooses her leisure e_t optimally given her beliefs. We illustrate the agent's learning dynamics in Figure 1, normalizing Φ and A to zero. The agent's belief about the fundamental ($\tilde{\phi}_t$) as well as her action (e_t) are on the horizontal axis, and output is on the vertical axis. The "output possibilities curve" $-L(e_t)$ represents the agent's level of output as a function of her action, and the "perceived achievable output line" $\tilde{a}+\tilde{\phi}_t$ represents the output level the agent believes is reachable as a function of the fundamental $\tilde{\phi}_t$.

In period 1, the agent chooses $e_1 = \tilde{\phi}_0 = 0$, and ends up with an output of zero. What perceived state $\tilde{b}_1$ is consistent with this observation? The highest red dashed curve is the loss function starting at $(0,0)$. Her perceived state $\tilde{b}_1$ and hence also her mean belief $\tilde{\phi}_1$ is where this curve intersects the perceived achievable output line: at this point, the difference between the achievable output and actual output is exactly the loss from choosing the in her mind suboptimal action of zero.

In period 2, then, the agent chooses $e_2 = \tilde{\phi}_1 = \tilde{b}_1$ and ends up with an output of $-L(e_2)$. To identify the $\tilde{b}_2$ consistent with this observation, we draw a copy of the loss function starting from her current effort-output location, and find the intersection of this curve with the perceived achievable output line. This is the perceived state in period 2, and because $\tilde{b}_2 < \tilde{\phi}_1$, we have $\tilde{\phi}_2 = (\tilde{b}_1 + \tilde{b}_2)/2 < \tilde{\phi}_1$. Continuing with this logic identifies the agent's learning dynamics.

The intersection of the output possibilities curve and the perceived achievable output line gives the agent's limiting belief $\tilde{\phi}_\infty < \Phi - \Delta$. To see this, note that if $\tilde{\phi}_t > \tilde{\phi}_\infty$, then $\tilde{\phi}_\infty < \tilde{b}_t < \tilde{\phi}_t$, so the agent's belief keeps drifting downwards toward $\tilde{\phi}_\infty$. In contrast, if $\tilde{\phi}_t = \tilde{\phi}_\infty$, then $\tilde{b}_t = \tilde{\phi}_t$, so this belief is stable. Intuitively, a mean belief of $\tilde{\phi}_\infty$ provides a coherent theory of what is going on, so the agent feels no need to update her belief: given her false belief about the fundamental, she chooses an action that leads her to produce exactly as much as she expects. Unfortunately for her, her utility loss relative to the optimal action can be an arbitrarily large multiple of her overconfidence Δ.

The agent's limiting belief has a puzzling feature: it is a terrible explanation for her early observations. In particular, in the limit the agent believes that if she takes the action e_1 or e_2, her expected output is far lower than what she observed in periods 1 and 2. Given her early observations, how does the agent end up with such a wrong theory? By reacting to what she thinks she has learned, she generates misleading observations that move her further from the correct theory. In addition, she endogenously generates so many of these misleading observations that they overwhelm the initial less misleading observations she made.

The self-defeating learning our model predicts is plausible in all the applications in Section 2.2. In the context of personal decisions, our model provides one possible mechanism for self-destructive behavior in normal individuals observed by psychologists (for a review, see for instance Baumeister and Scher, 1988). For instance, Kasser (2003) argues that materialism (the pursuit of material possessions and pleasure) can lead to a vicious cycle whereby a person pursues consumption to become happier, she becomes less happy as a result, and as she attempts to compensate with even more consumption, she exacerbates the problem.[7]

In the context of manager-subordinate relationships, Manzoni and Barsoux (1998, 2002) describe a common problem—the "set-up-to-fail syndrome"—that is due to a dynamic similar to our mechanism. When the subordinate performs below expectations, the manager reacts by increasing control and assigning less inspiring tasks. This undermines the employee's motivation, lowering performance and eliciting further negative reactions from the manager. Similarly, the idea that organizations can get caught in dysfunctional vicious circles has been described in sociology (March and Simon, 1958, Crozier, 1964, Masuch, 1985). Management is dissatisfied with the performance of the organization, and attempts to improve things with increased monitoring, control, or other bureaucratic measures. These measures backfire, for instance because they alienate employees. Management reacts by tightening bureaucratic measures further.[8]

Although we have not found discussions that are reminiscent of our mechanism, self-defeating interpretations of worse-than-expected outcomes also seem plausible in the context of policymaking. In the war on drugs, for instance, a policymaker may interpret drug problems as indicating that she should crack down on drugs, only to make the problem worse and the reaction harsher.

Note that one channel for the agent to develop overconfidence is through a drop in her ability—for instance due to aging—that is not accompanied by a change in her belief about ability. In this

[7] Similarly, overtraining by performance athletes is a widely discussed problem in academic, professional, and popular writings (Budgett, 1998, for example). Looking for better performance, an athlete begins to train harder than previously. Because the body cannot cope with such a demanding training schedule, performance suffers. The athlete reacts by training even harder, exacerbating the problem.

[8] Masuch (1985) also argues that one reason for such vicious cycles is that people have unrealistically high expectations regarding how well things will work. The cycle ends when management runs out of control tools, or, as described for instance by Argyris (1982), when management settles into a suboptimal set of practices.

case, relatively small drops in ability can trigger our mechanism and lead to a drastic drop in her belief about the world—as when an aging person concludes that things are much worse than they used to be.

Comparative Statics with Respect to the Loss Function. Because it provides a surprising and potentially testable prediction of our model, we consider the comparative statics of the agent's beliefs with respect to the loss function. As is clear from Figure 1, an increase in the loss function leads the agent to initially adjust her belief by less (i.e., to increase $\tilde{\phi}_1$). Intuitively, with a higher loss function a smaller deviation from the perceived-optimal action is necessary to explain the agent's surprise in period 1, and hence a smaller adjustment in the state is necessary to explain her observation.

But the opposite comparative static holds for the agent's limiting belief: the more important it is for her action to be aligned with the fundamental, the *further* from the fundamental she ends up. If the loss function shifts up, then in Figure 1 the curve $-L(e_t)$ shifts down, and the limiting belief $\tilde{\phi}_\infty$ therefore moves to the left. Intuitively, a steeper loss function means that the agent hurts herself more through her misinferences. To develop a consistent theory of her observations, therefore, she must become more pessimistic about the world.

Outside Options. In the above model, we have assumed that the agent participates in the task in every period regardless of her beliefs. It is natural to consider an environment in which she has an outside option, such as another task she could perform. In the manager-employee setting, for instance, a manager can keep working with the current employee, or replace her. Let the perceived utility of the outside option be $\tilde{u}$. Interesting new issues arise if $\tilde{a} > \tilde{u} > A - L(A - \tilde{\phi}_\infty) = -L(\tilde{\phi}_\infty)$—i.e., $\tilde{u}$ is in the intermediate range—which implies that the agent starts off with the task but eventually quits.

We consider two cases. If $\tilde{u} > A$, then the agent should not be in this task in the first place, so both her entry into it and her persistence in it are suboptimal. The prediction that overconfident individuals—overestimating their ability to perform well—are more likely to enter into and persist in ability-sensitive enterprises is consistent with common intuition and evidence from both psychology and economics. Overconfidence is often invoked as one explanation for why many new businesses fail in the first few years, and indeed Landier and Thesmar (2009) find that entrepreneurs of small startups have overconfident expectations about future growth. Similarly, several studies suggest that CEO overconfidence is associated with a higher likelihood of pursuing risky actions, such as making acquisitions (Malmendier and Tate, 2008) and undertaking innovation (Galasso and Simcoe, 2011, Hirshleifer, Low and Teoh, 2012). And persisting in tasks (often for too long) is regarded as one of the main characteristics of overconfidence (McFarlin, Baumeister and Blascovich, 1984, for example).

In contrast, if $\tilde{u} < A$, then the agent should be in the task, so overconfidence generates suboptimal exit: ironically, the agent stops performing the task *because* she overestimates her ability to do well in it. Intuitively, she is more prone to exit than a realistic agent because her repeated negative inference about the world eventually negates the effect of her overconfidence. Because the agent does not update her belief about a, this reaction can generate interesting dynamics when there are multiple types of alternative tasks for her to choose. The logic of our model suggests that she will first seek out another ability-sensitive task in which she believes a different fundamental determines outcomes, and then successively jump from one such task to the next. And once she runs out of these tasks, she chooses a less ability-sensitive task and sticks with it.

The prediction that overconfidence leads individuals to eventually quit superior tasks, to jump too much between tasks, and to eventually prefer less ability-sensitive tasks, contrasts with the typical view on the implications of overconfidence. While we have not found direct evidence for this prediction, it is consistent with the observation that many documented effects of overconfidence in economic settings, such as the tendency of overconfident CEO's to undertake mergers or innovations, pertain to *new* directions. In particular, Landier and Thesmar (2009) find that serial entrepreneurs—those who start a new business following another one—have more unrealistic expectations than first-time entrepreneurs. This observation is consistent with the idea that overoptimistic entrepreneurs do not simply persist in a single project, but also jump to others.[9]

3.2 Self-Defeating Learning: General Results

We now derive the results on self-defeating learning—illustrated in a specific setting above—for general production functions Q satisfying Assumption 2. In the next subsection, we discuss what happens when Assumption 2 does not hold.

[9] It is important to note, however, that alternative interpretations are also possible. For instance, Landier and Thesmar suggest that previous experience increases overoptimism because entrepreneurs attribute failures to bad luck, and successes to ability.

To characterize the agent's learning process, we define the "surprise function" $\Gamma(\tilde{\phi})$ as the difference between actual output and expected output:

$$\Gamma(\tilde{\phi}) = Q(e^*(\tilde{\phi}), A, \Phi) - Q(e^*(\tilde{\phi}), \tilde{a}, \tilde{\phi})\,. \tag{5}$$

It is easy to see that the surprise function governs the agent's updating process. To start, if $\Gamma(\tilde{\phi}_t) = 0$—the agent gets exactly the output she expects—then $\tilde{b}_t = \tilde{\phi}_t$, so that the agent sees no reason to update her belief about the fundamental. Hence, the zeros of $\Gamma(\cdot)$ constitute consistent theories of the world. We define this set as C:

$$C = \{\tilde{\phi} \,|\, \Gamma(\tilde{\phi}) = 0\}.$$

In our simple example of Section 3.1, C is a singleton given by the intersection of the output possibilities curve and the perceived achievable output line.

If, on the other hand, the agent is surprised about output ($\Gamma(\tilde{\phi}_t) \neq 0$), then she must update her belief about the fundamental. In particular, if $\Gamma(\tilde{\phi}_t) < 0$—she gets lower output than she expects—then $\tilde{b}_t < \tilde{\phi}_t$, so she updates her belief downwards; and if $\Gamma(\tilde{\phi}_t) > 0$—she gets higher output than she expects—then $\tilde{b}_t > \tilde{\phi}_t$, so she updates her belief upwards. It turns out that updating is monotonic:

Proposition 1. *Suppose Assumptions 1 and 2 hold. If* $\Gamma(\tilde{\phi}_0) < 0$*, then* $\tilde{\phi}_t$ *is strictly decreasing in* t *and converges to* $\max\{\phi \in C \colon \phi < \tilde{\phi}_0\}$*. If* $\Gamma(\tilde{\phi}_0) > 0$*, then* $\tilde{\phi}_t$ *is strictly increasing in* t *and converges to* $\min\{\phi \in C \colon \phi > \tilde{\phi}_0\}$*. If* $\Gamma(\tilde{\phi}_0) = 0$*, then* $\tilde{\phi}_t$ *is constant.*

These implications are illustrated in Figure 2, where there are three possible consistent beliefs. If the prior is above $\tilde{\phi}^1_\infty$, then the agent's belief decreases monotonically over time and approaches $\tilde{\phi}^1_\infty$. If her prior is between $\tilde{\phi}^2_\infty$ and $\tilde{\phi}^1_\infty$, then her belief increases monotonically over time, and again converges to $\tilde{\phi}^1_\infty$. If the agent's prior is below $\tilde{\phi}^2_\infty$, her belief converges to $\tilde{\phi}^3_\infty$ from the side from which it starts off.

To understand why the agent reacts to a change in her belief in a way that reinforces the change—thus generating monotonically evolving beliefs over time—note that because $Q_{ea} \leq 0$ and $Q_{eb} > 0$, the fact that the agent overestimates her ability and underestimates the state implies that she underestimates Q_e. This means that an increase in her action generates a positive surprise—spurring further increases in the action—and a decrease in the action generates a negative surprise—spurring further decreases in the action.

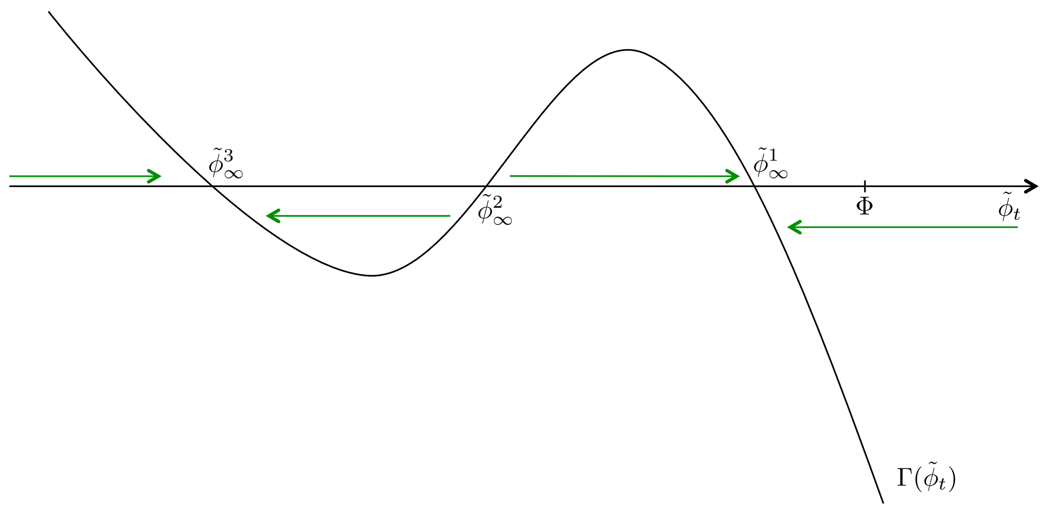

Figure 2: Convergence and Dependence of Limiting Belief on Prior

An interesting possibility is that—as illustrated in Figure 2—there are multiple consistent beliefs, and which one the agent converges to depends on her prior. Multiple consistent theories can arise if (unlike in the loss-function specification) the agent's overconfidence has different effects on expected output for different fundamentals. In particular, if ability has a large effect on output for a lower fundamental and a small effect for a higher fundamental, then both the lower and the higher fundamental can constitute consistent theories. If the agent believes the lower fundamental, then—since the true fundamental is much higher—output is (for a realist) inexplicably high, but since she is overconfident and output is very sensitive to ability, in the end output is consistent with her theory. And if the agent believes in the higher fundamental, then output is less inexplicably high, which is also consistent with her overconfidence since here output is less sensitive to ability.[10]

[10] We demonstrate through an example that a production function with multiple consistent beliefs is indeed possible. For simplicity, we use an output function with kinks and linear parts; our output function can obviously be approximated arbitrarily closely by a differentiable output function satisfying Assumptions 1 and 2. Let $Q(e_t, a, b_t) = T(a, b_t - (b_t - e_t)_+/2) - e_t/4$. We specify T such that (i) it is optimal to choose $e_t = \tilde{\phi}_{t-1}$ for any t; (ii) $T_{ax}(a, x) \leq 0$ for all a, x; and (iii) $T(a, x)$ to be increasing in a and x with derivatives bounded from above and below.

Assuming (i)—which we will check later—the belief $\tilde{\phi}$ is consistent if and only if $T(A, \tilde{\phi} + (\Phi - \tilde{\phi})/2) = T(\tilde{a}, \tilde{\phi})$. We construct $T(A, x)$ and $T(\tilde{a}, x)$ so that for $\Phi = 10$, both $\tilde{\phi} = 0$ and $\tilde{\phi} = 6$ are consistent theories. Let $T(\tilde{a}, x) = 10 + x$. Then, for the two theories to be consistent, we need $T(A, 5) = T(\tilde{a}, 0) = 10$ and $T(A, 8) = T(\tilde{a}, 6) = 16$. We let $T(A, x)$ be three-piece linear in x with a slope of one below 5 and above 8, and a slope of 2 between 5 and 8.

To extend the above to a full function $T(a, x)$, we let $T(a, x) = T(A, x) - (A - a)$ for $a < A$, $T(a, x) = T(\tilde{a}, x) + (a - \tilde{a})$ for $a > \tilde{a}$, and we linearly interpolate between A and $\tilde{a}$. It is now straightforward to check properties (i)-(iii): (i) follows as $T_x \geq 1$; (ii) as $T_x(\tilde{a}, \cdot) = 1 \leq T_x(A, \cdot)$ the property holds by the piecewise linear construction; and (iii) as $T(A, \cdot) \leq T(\tilde{a}, \cdot)$, $T_x \leq 2$, and $\min\{1, \frac{2}{\tilde{a} - A}\} \leq T_a \leq \max\{1, \frac{5}{\tilde{a} - A}\}$.

A simple but important corollary of Proposition 1 is that the agent ends up being too pessimistic, and if she has an approximately correct or overly optimistic initial belief about the fundamental, then she perpetually revises her belief downwards:

Corollary 1. *Suppose Assumptions 1 and 2 hold. From period 1 on, the agent is overly pessimistic about the fundamental (i.e., for any* $\tilde{\phi}_0$*,* $\tilde{\phi}_t < \Phi$ *for* $t \geq 1$*). For any initial belief* $\tilde{\phi}_0 \geq \Phi$*, the agent permanently revises her belief downwards (i.e.,* $\tilde{\phi}_t$ *is strictly decreasing).*

Given Proposition 1, the corollary follows from the fact that—as illustrated in Figure 2—the surprise function is negative for any $\tilde{\phi} \geq \Phi$. And this fact is intuitive: given her overconfidence about ability, if the agent is not too pessimistic about the fundamental, she will be negatively surprised about output.

We now generalize the insight from the example that the agent ends up further from truth when it is more important to be close. For this purpose, we define

$$\begin{aligned} R(a,\phi) &= Q(e^*(a,\phi),a,\phi) \\ L(e,a,\phi) &= Q(e^*(a,\phi),a,\phi) - Q(e,a,\phi), \end{aligned} \tag{6}$$

and observe that $Q(e,a,\phi) = R(a,\phi) - L(e,a,\phi)$. Intuitively, $R(a,\phi)$ is the achievable level of output given a and ϕ, and $L(e,a,\phi)$ is the loss relative to the achievable output due to choosing a suboptimal action. We compare two technologies Q_1, Q_2 with corresponding R_1, R_2, L_1, L_2:

Proposition 2. *Suppose Assumptions 1 and 2 hold,* $e_1^*(a,\phi) = e_2^*(a,\phi)$ *and* $R_1(a,\phi) = R_2(a,\phi)$ *for all* a,ϕ*, and* $L_1(e,a,\phi) < L_2(e,a,\phi)$ *for all* $a,\phi,e \neq e^*(a,\phi)$*. Then, for any initial belief* $\tilde{\phi}_0$*, the agent's limiting belief is lower under technology* Q_2 *than under technology* Q_1*.*

The key step in the proof of Proposition 2 is that an increase in the loss function decreases the surprise function. As in Section 3, an increase in the loss function means that the agent hurts herself more through her misinference-induced suboptimal behavior, and this must mean that she is more negatively surprised by her output. And examining Figure 2 makes it intuitively clear that a downward shift in the surprise function lowers the agent's limiting belief for any initial belief.

3.3 When Learning Is Not Detrimental

To sharpen our intuitions regarding when self-defeating learning occurs, in this section we identify some conditions in which it does not. To begin, we note a mathematically obvious, but useful fact:

that if the optimal action is not sensitive to the state, then self-defeating learning does not occur.

Proposition 3. *Suppose Assumption 1(i) holds and output takes the form $Q(S(e_t, a), b_t)$. Then, the agent never changes her action.*

If the optimal action depends on ability, the agent takes the wrong action, and if she is surprised by low output, she makes an incorrect inference regarding the fundamental. But if the optimal action does not depend on the fundamental, these reactions are not self-reinforcing. Since the agent does not change her action in response to her change in belief, she is not surprised by her output, and stops changing her action or updating her belief.

Beyond supposing that the optimal action depends on the state, in Section 3.2 we have assumed that $Q_{ea} \leq 0$. We now discuss the implications of the possibility that $Q_{ae} > 0$. To understand the key issue, suppose that the agent starts off with the correct belief about the fundamental. Then, because she is overconfident and her ability and action are complements, her initial action is too high. As in the rest of our analysis, her observation of the surprisingly low output in period 1 leads her to revise her belief about the fundamental downwards—and, as a consequence, to choose a lower action next time. Because her initial action was too high, this adjustment can increase output. That is, the agent's misdirected learning can counteract the effect of the agent's overconfidence. It is possible that in the limit misdirected learning increases output. It is, however, also possible that misdirected learning lowers the action below optimal, at which point the logic by which further learning occurs is analogous to that in Section 3.2.

To say more about when each of the above two possibilities obtains, we next identify circumstances under which an overconfident agent behaves in the same way as an agent who is correct about her ability—who in our noiseless model identifies the fundamental correctly in period 1, and behaves optimally from period 2 on.

Proposition 4 (Non-Identifiability and Optimal Behavior)**.** *Suppose Assumption 1 holds. The following are equivalent:*

I. *For any A, $\tilde{a}$, and Φ, there is a unique belief $\tilde{\phi}(A, \tilde{a}, \Phi)$ such that for any $\tilde{\phi}_0$, $\tilde{\phi}_t = \tilde{\phi}(A, \tilde{a}, \Phi)$ for any $t \geq 1$ and the agent chooses the optimal action from period 2 on.*

II. *The agent's behavior is identical to that with an output function of the form $Q(e_t, B(a, b_t))$, where $Q_B, B_a, B_b > 0$, and for any $a, b, \tilde{a}$, there is some $\tilde{b}$ such that $B(a, b) = B(\tilde{a}, \tilde{b})$.*

Proposition 4 says that agents with wrong beliefs about ability behave optimally if and only if behavior can be described by an output function that depends only on a summary statistic of ability and the state, and not independently on the two variables. If this is the case, then the agent can correctly deduce the level of the summary statistic from output in period 1, so that she correctly predicts how changes in her action affect output. As a result, she both chooses the optimal action and gets the output she expects—and hence she does not adjust her action any further. Our proof establishes that this kind of production function is not only sufficient, but also necessary for learning to be optimal in steady state.

An interesting aspect of Proposition 4 is that the agent is able to find the optimal action exactly when the problem is *not* identifiable—when her observations of output do not allow her to separately learn A and Φ. This beneficial role of non-identifiability is in direct contrast with what one might expect based on the statistical learning literature, where non-identifiability is *defined* as a property of the environment that hinders learning. Yet it is exactly the non-identifiability of the problem that allows the overconfident agent to choose her action well: because ability and the state do not have independent effects on output, the agent's misinference about the fundamental can fully compensate her overconfidence regarding ability, and hence overconfidence does not adversely affect her.

In the team production setting, for instance, suppose that the agent—instead of making delegation decisions—chooses her effort level, and output depends on effort and the total ability of the team. Then, although the agent still underestimates her teammates, she is able to deduce the team's total ability from output. As a result, she chooses the optimal action.

Notice that combined with our assumptions that $Q_{eb}, Q_a, Q_b > 0$, statement II of Proposition 4 implies that $Q_{ea} > 0$. Furthermore, since the optimal action depends only on a summary statistic of ability and the state, ability and the state affect the optimal action to the same extent (controlling for their effect on output). Combining this insight with that of Proposition 3 indicates that changes in the agent's belief about the fundamental eventually induces actions that are significantly lower than optimal—and hence self-defeating learning occurs—if the optimal action is sufficiently more sensitive to the state than to ability. And adding our insights from Section 3.2, we conclude that self-defeating learning occurs when the optimal action depends on the state of the world, and either does not depend so much on ability, or it does so in an opposite way than on the state.

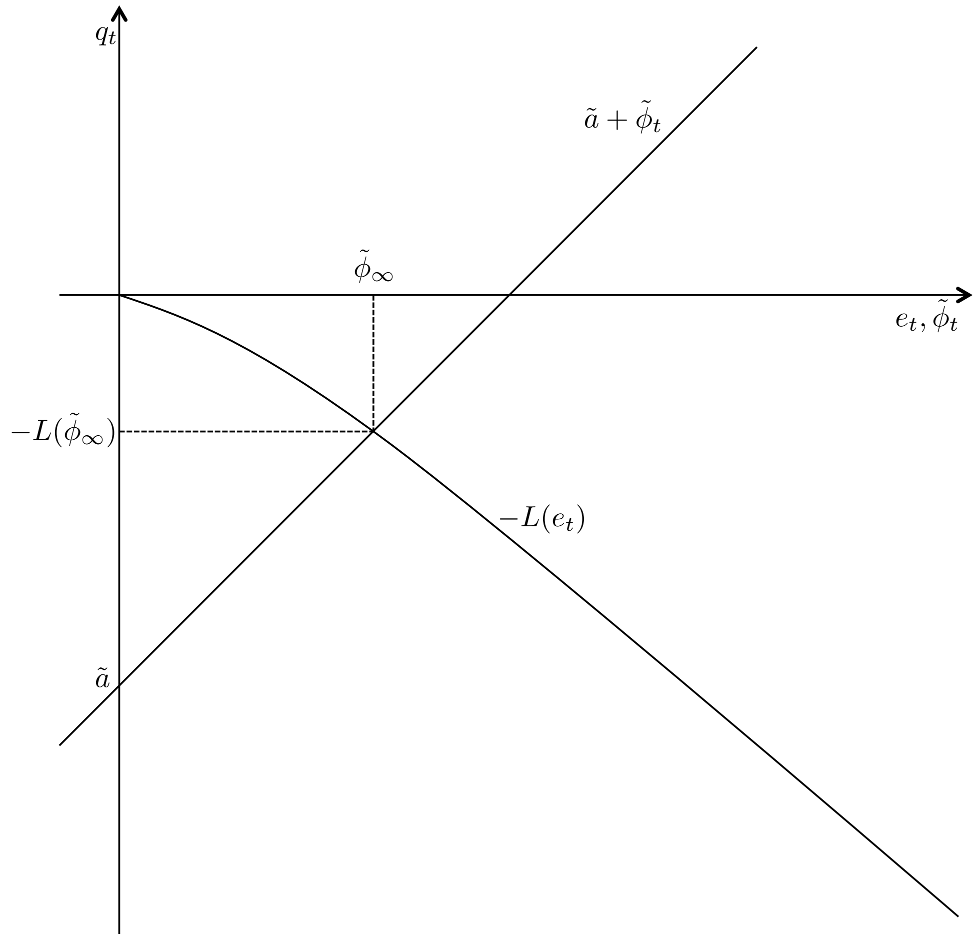

Figure 3: Limiting Belief and Loss with Underconfidence

3.4 Underconfidence

While available evidence indicates that people are on average overconfident, it is clear that some are instead "underconfident"—having unrealistically *low* beliefs about their ability and other traits. We briefly discuss how this modifies our results, restricting attention to the loss-function specification of Section 3.1.

The key insight regarding the agent's limiting belief is easiest to see graphically. In Figure 3, we redraw the relevant parts of Figure 1 for underconfident beliefs, again normalizing A and Φ to 0. As for overconfident agents, any possible limiting belief is given by the intersection of the output possibilities curve $-L(e_t)$ and the perceived achievable output line $\tilde{a} + \tilde{\phi}_t$. If $\tilde{a} < A$, the two curves intersect to the right of the true fundamental: since the agent is pessimistic about her own ability, she becomes overly optimistic about the fundamental. Furthermore, it is apparent from the figure that in the limit the agent's loss from underconfidence is bounded by Δ:

Proposition 5 (Limiting Belief and Loss with Underconfidence)**.** *Suppose* $Q(e_t, a, b_t) = a + b_t - L(e_t - b_t)$ *for a symmetric convex loss function* L *with* $L'(x) < k < 1$ *for all* x*,* $\tilde{a} < A$*, and there is a vanishingly small amount of noise. Then, for any* $\tilde{\phi}_0$ *the agent's belief converges to the same limiting belief* $\tilde{\phi}_\infty$*, and* $0 < \tilde{\phi}_\infty - \Phi < \Delta$ *and* $L(\tilde{\phi}_\infty - \Phi) < \Delta$*.*

These results contrast sharply with those in the overconfident case, where the limiting belief is always more than Δ away from the true fundamental ($\Phi - \tilde{\phi}_\infty > \Delta$), and the associated loss can be an arbitrarily large multiple of Δ. Underconfidence therefore has a very different effect on the agent's misguided learning than an identical degree of overconfidence. Consistent with such an asymmetry in our managerial example, while Manzoni and Barsoux emphasize that the set-up-to-fail syndrome is a severe consequence of a manager's unrealistically high expectations regarding a subordinate, they do not identify similar consequences from unrealistically low expectations.

To understand the intuition, consider again an agent who starts off with the correct initial belief about the fundamental ($\tilde{\phi}_0 = \Phi$). Upon observing a better performance than she expected, she concludes that the fundamental is better than she thought, and revises her action. The resulting utility loss, however, leads her to reassess her first optimistic revision of her belief, bringing her belief back toward the true fundamental. In this sense, the agent's misinference regarding the fundamental is self-correcting—in contrast to the logic in the case of overconfidence, where the misinference is self-reinforcing. Moreover, because a utility loss of Δ or more cannot be explained by a combination of underconfidence in the amount of Δ and an unrealistically positive belief about the fundamental (which increases expected output), any consistent belief must generate a utility loss of less than Δ.

In addition to explaining that the limiting loss from underconfidence can be bounded, the above logic also makes clear that—unlike with overconfidence and vanishingly little noise—the agent's path to this limit is often non-monotonic.

4 Noisy Model

In this section, we analyze our model when noise is non-trivial. First, going beyond previous results in the literature, we show that the agent's belief converges under general conditions on Q. Second, we establish versions of all of our results above also for the noisy model.

4.1 Convergence of the Agent's Belief

Establishing convergence is technically challenging because of the endogeneity of the agent's action: as the agent updates her belief, she changes her action, thereby changing the objective distribution of the perceived signal she observes and uses to update her belief. This means that we cannot apply results from the statistical learning literature, such as that of Berk (1966), where the observer does not choose actions based on her belief. Within our more specific domain, our convergence result also substantially supersedes that of Esponda and Pouzo (forthcoming, Theorem 3), which applies only for priors close to the limiting belief and for actions that are only asymptotically optimal.

We first identify key features of how the agent interprets output. For this purpose, we define $\tilde{b}(b_t, e_t)$ as the signal $\tilde{b}_t$ (given by the solution to Equation (1)) that the agent perceives as a function of the true signal b_t and her action e_t.

Proposition 6. *Suppose Assumption 1 holds. Then, the function $\tilde{b}$ has the following properties: (1) $0 \leq b_t - \tilde{b}(b_t, e_t) \leq \Delta\overline{Q}_a/\underline{Q}_b$ for all b_t, e_t; (2) $\tilde{b}$ is differentiable in e_t; (3) if Assumption 2 holds in addition, then $\tilde{b}$ is strictly increasing in e_t.*

The agent's inferences satisfy three important properties. First, Assumption 1 implies that inferences are downward biased—the agent perceives worse news in each period than is warranted. Because the agent believes she is more able than she actually is, any given level of output conveys worse news to her about the world than is realistic. Second, by the differentiability properties of Q, the agent's perceived signal is differentiable in her action. Third, if Assumption 2 holds, then the agent's inferences are more biased if she takes a lower action. Since $\tilde{a} > A$ and $\tilde{b}_t < b_t$, the assumptions that $Q_{ea} \leq 0$ and $Q_{eb} > 0$ imply that the agent underestimates Q_e. To explain the change in output resulting from an increase in her action, therefore, she must become more optimistic about the fundamental.

Using Equation (2), the dynamics of the agent's belief can be written as

$$\tilde{\phi}_{t+1} = \tilde{\phi}_t + \gamma_t \left[\tilde{b}_{t+1} - \tilde{\phi}_t\right], \text{ where } \gamma_t = \frac{h_\epsilon}{(t+1)h_\epsilon + h_0}. \tag{7}$$

We define the function $g : \mathbb{N} \times \mathbb{R} \to \mathbb{R}$ as the expectation an outside observer holds about $\tilde{b}_{t+1} - \tilde{\phi}_t$:

$$g(t, \tilde{\phi}_t) = \mathbb{E}_{b_{t+1}} \left[\tilde{b}(b_{t+1}, e^*(t+1, \tilde{\phi}_t))|\Phi\right] - \tilde{\phi}_t. \tag{8}$$

The above expectation is the one of an outside observer who knows the true fundamental Φ and the agent's subjective belief $\tilde{\phi}_t$, so that—being able to deduce the action $e^*(t+1, \tilde{\phi}_t)$—she knows the distribution of $\tilde{b}_{t+1}$. The function g can be thought of as the agent's mean surprise regarding the fundamental, and in this sense it is a noisy variant of the surprise function Γ we have defined for output in Section 3.2.

As we are interested in how the agent's subjective belief is updated in the limit, we define $g(\tilde{\phi}) = \lim_{t\to\infty} g(t, \tilde{\phi})$. Generalizing our definition for the noiseless limit, we denote by C the set of points where g intersects zero, i.e., the agent's mean surprise equals zero:

$$C = \{\tilde{\phi} \colon g(\tilde{\phi}) = 0\}. \tag{9}$$

Intuitively, the agent cannot have a long-run belief $\tilde{\phi}$ outside C as her limiting action $e^*(\tilde{\phi})$ would generate subjective signals that push her systematically away from $\tilde{\phi}$. Specifically, when $g(\tilde{\phi}_t) > 0$, then the agent eventually generates signals that are on average above $\tilde{\phi}_t$, so that her belief drifts upwards; and if $g(\tilde{\phi}_t) < 0$, then the agent's belief eventually drifts downwards. Hence, any limiting belief must be in C. Furthermore, not every belief in $\tilde{\phi} \in C$ is stable: whenever g is negative on the left of $\tilde{\phi}$ and positive on the right of $\tilde{\phi}$, the agent's perceived signals would still in expectation push her away from $\tilde{\phi}$. We thus have the following definition:

Definition 1 (Stability)**.** A point $\tilde{\phi} \in C$ is stable if there exists a $\delta > 0$ such that $g(\phi) < 0$ for $\phi \in (\tilde{\phi}, \tilde{\phi} + \delta)$ and $g(\phi) > 0$ for $\phi \in (\tilde{\phi} - \delta, \tilde{\phi})$.

We denote by H the set of stable points in C. Our main result in this subsection shows that H completely describes the set of possible long-run beliefs.

Theorem 1. *Suppose Assumption 1 holds, and that C is the union of a countable number of points. Then, almost surely $\tilde{\phi}_\infty = \lim_{t\to\infty} \tilde{\phi}_t$ exists and lies in H. Furthermore, $-\Delta\, \overline{Q}_a / \underline{Q}_b \leq \tilde{\phi}_\infty - \Phi \leq -\Delta\, \underline{Q}_a / \overline{Q}_b$.*

The mathematical intuition behind Theorem 1 is the following. Because the agent's posterior can be fully described by its mean and variance and the latter only depends on time t, we study dynamics of the agent's mean belief $\tilde{\phi}_t$. Although—as we have explained above—the agent's subjective signals are not independently and identically distributed over time, we use a result from stochastic approximation that in our context requires the perceived signals to be approximately independent over time conditional on the current subjective belief. To use this result, for $s \geq t$ we approximate the agent's time-dependent action $e^*(s, \tilde{\phi}_s)$ by the time-independent confident action $e^*(\tilde{\phi}_s)$—that is, the action the agent would choose if she was fully confident of the fundamental. We show that the mistake we make in approximating $e^*(s, \tilde{\phi}_s)$ in this way is of order $1/s$, and since the perceived signal $\tilde{b}_s$ is Lipschitz continuous in e_s, the mistake we make in approximating $\tilde{b}_s$ is also of order $1/s$. Since when updating the agent's newest signal gets a weight of order $1/s$, this means that the error in approximating the change in her beliefs is of order $1/s^2$, and hence the total error in approximating beliefs from period t onwards is of order $\sum_{s=t}^{\infty} 1/s^2$, i.e., finite. Furthermore, as $t \to \infty$, the approximation error from period t onwards goes to zero.

Given the above considerations, on the tail we can think of the dynamics as driven by those that would prevail if the agent chose $e^*(\tilde{\phi}_t)$ in period $t+1$. The expected change in the mean belief in period $t+1$ is therefore a function of $g(\tilde{\phi}_t)$ defined above. Note, however, that the expected change is a time-dependent function of $g(\tilde{\phi}_t)$—as the agent accumulates observations, she puts less and less weight on her new observation. Stochastic approximation theory therefore defines a new time scale τ_t and a new process $z(\tau)$ that "speeds up" $\tilde{\phi}_t$ to keep its expected steps constant over time, also making $z(\tau)$ a continuous-time process by interpolating between points. It then follows that on the tail the realization of z can be approximated by the ordinary differential equation

$$z'(\tau) = g(z(\tau)), \tag{10}$$

which is a deterministic equation. Intuitively, since $z(\tau)$ is a sped-up version of $\tilde{\phi}_t$, as τ increases any given length of time in the process $z(\tau)$ corresponds to more and more of the agent's observations. Applying a version of the law of large numbers to these many observations, the change in the process is close to deterministic. In this sense, on the tail the noisy model reduces to a noiseless model. Now because the solution to Equation (10) converges, the agent's belief converges.

To show this convergence, the bounds we imposed on the derivatives of Q play a key role: they ensure that the agent's belief does not diverge to minus infinity and imply the bounds on H. Since the agent overestimates her ability, she eventually develops pessimistic views about the world. But since the signal has at least some influence on output and the influence of ability on output is bounded, the agent cannot make arbitrarily large misinferences about the signal as a consequence of her overconfidence.[11]

A simple corollary of Theorem 1 is that if C is a singleton, then the agent's belief converges to a unique limiting belief:

Corollary 2. *Suppose Assumption 1 holds. If there exists a unique ϕ^* with $g(\phi^*) = 0$, then $\tilde{\phi}_t$ converges to ϕ^* almost surely.*

Unfortunately, the single-crossing condition in Corollary 2 is not a condition on the primitives of the model. In a specific setting, it may be easy to check the condition numerically. Furthermore, we can derive sufficient conditions on the primitives for the condition to hold.

First, C is a singleton if overconfidence is sufficiently small ($\tilde{a}$ is close to A). If $\tilde{a} = A$, then $\tilde{b}_t = b_t$, so $g(\tilde{\phi}_t) = \Phi - \tilde{\phi}_t$, and g has a unique root. By continuity of g in $\tilde{a}$, g has a unique root if $\tilde{a}$ is sufficiently close to A.

Second, as in the noiseless case, C is a singleton if Q has the loss-function specification introduced in Section 3:

Lemma 2. *If $Q(a, b, e) = a + b - L(e - b)$ for a symmetric loss function L satisfying $|L'(x)| < k < 1$ for all x, then C is a singleton.*

4.2 Properties of the Limiting Beliefs

We now study the properties of the agent's learning and limiting belief. Even given that her belief converges, the agent's learning process is complicated because the effect of noise on output depends on the shape of Q, so that the agent's inferences depend on the comparison of the shape of Q at the perceived and true fundamentals. Due to this complication, our results are weaker than for the noiseless limit above.

Self-Defeating Learning. We first generalize our insight that the agent's learning has a self-defeating nature in that—so long as she starts off not too far from the correct action—the possibility

[11] Our model can produce divergence of the agent's belief to minus infinity if we do not assume that the influence of the signal on output is bounded from below ($Q_b > \underline{Q}_b$) or that the influence of the action is bounded from above ($Q_a < \overline{Q}_a$).

to update her action leads her further away from the true fundamental Φ. Denote by $\hat{\phi}_t^e$ the subjective belief the agent holds when she has taken the fixed action e in all periods. Then:

Proposition 7. *Suppose Assumptions 1 and 2 hold, and that C is the union of countably many points. The sequence $\hat{\phi}_t^e$ converges almost surely to a unique point. If $e \geq e^*(\Phi)$, then*

$$\lim_{t\to\infty} \tilde{\phi}_t < \lim_{t\to\infty} \hat{\phi}_t^e < \Phi.$$

The intuition is the same as in the noiseless limit. Because she is overconfident, the agent's belief—in the noiseless case immediately, in the noisy case eventually—becomes too pessimistic. This leads her to choose lower actions and to generate further surprisingly low output observations, perpetuating the misguided learning.

The Greater Is the Loss from Incorrect Beliefs, the More Incorrect Is the Limiting Belief. We now establish a version of our insight—formalized for the noiseless limit of our model in Proposition 2—that an increase in the loss function moves the agent's limiting belief further from the true fundamental. Since a general analysis of this comparative static has proved intractable, we focus on our loss-function specification with a linear loss function: $Q(e_t, a, b_t) = a + b_t - \eta|e_t - b_t|$, where $\eta < 1$. We have:

Proposition 8 (Limiting Belief with Linear Loss Function.)**.** *For any η, there is a $\tilde{\phi}_\infty(\eta)$ such that for any $\tilde{\phi}_0$, $\tilde{\phi}_t$ converges to $\tilde{\phi}_\infty(\eta)$ almost surely. Furthermore, the limiting belief $\tilde{\phi}_\infty(\eta)$ is strictly decreasing in η.*

The intuition for Proposition 8 is the same as for Proposition 2: the higher is the loss function, the more the agent suffers from her misinference regarding the fundamental, and hence—to explain her suffering—the more pessimistic about the fundamental she becomes. The complication that arises in the noisy case is that an increase in the loss function lowers not only actual output, but also the output the agent expects, so that the net effect on the agent's mean surprise is unclear. But since the agent's action is too low while she thinks it is optimal, for the majority of realizations of the state the increase in the loss she actually suffers due to the increase in the loss function is greater than the increase in the loss she thinks she suffers. This means that she interprets the typical output realization more negatively, so that the logic from the noiseless case extends. Although this intuition suggests that Proposition 8 should hold far more generally than for linear loss functions, we have not been able to prove a broader statement.

When Learning Is Not Detrimental. To conclude our analysis of the noisy model, we reconsider the questions in Section 3.3 regarding situations in which self-defeating learning does not occur. To start, Proposition 3, which states that when the optimal action does not depend on the state, learning has no effect on behavior, trivially extends to the noisy case. Next, we consider conditions under which the agent behaves optimally. Although for the noisy model we have not been able to formulate an "if-and-only-if" result of the type in Proposition 4—nor even to find an appropriate definition of what it means for the agent to be observationally equivalent to a realistic agent—we state an "if" result that again connects optimal behavior to non-identifiability:

Proposition 9. *If output has the form $Q(e_t, g(a) + b_t)$, then for any prior $\tilde{\phi}_0$ and any true fundamental Φ, the agent's action converges to the optimal one almost surely.*

If output depends only on $g(a) + b_t$, then the agent is able to use repeated observations to correctly infer $g(A)+\Phi$ and thereby learn to take the optimal action. Furthermore, this separability condition on the output function is a non-identifiability condition: it implies that learning A and Φ separately is impossible. This means that non-identifiability is again beneficial for the agent.

5 Related Literature

Our theory connects two literatures, that on overconfidence and that on learning with misspecified models. While we discuss other more specific differences below, to our knowledge our paper is the only one to analyze an individual-decisionmaking problem in which the agent's own behavior in response to a misinterpretation of information generates data that reinforces the misinterpretation. We are also unaware of previous research studying the implications of overconfidence for inferences about other, decision-relevant exogenous variables.

5.1 Overconfidence

Our paper studies the implications of unrealistic expectations regarding a variable for learning about another variable. In many applications, the most plausible source of such unrealistic expectations is overconfidence—the topic of an extensive literature in economics and psychology.

A plethora of classical evidence in psychology as well as economics suggests that on average people have unrealistically positive views of their traits and prospects (e.g., Weinstein, 1980, Svenson, 1981, Camerer and Lovallo, 1999). Recently, Benoît and Dubra (2011) have argued that much of this evidence is also consistent with Bayesian updating and correct priors, and thus does not conclusively demonstrate overconfidence. In response, a series of careful experimental tests have documented overconfidence in the laboratory in a way that is immune to the Benoît-Dubra critique (Burks, Carpenter, Goette and Rustichini, 2013, Charness, Rustichini and van de Ven, 2014, Benoît, Dubra and Moore, 2015). In addition, there is empirical evidence that consumers are overoptimistic regarding future self-control (Shui and Ausubel, 2004, DellaVigna and Malmendier, 2006, for instance), that truck drivers persistently overestimate future productivity (Hoffman and Burks, 2013), that genetically predisposed individuals underestimate their likelihood of having Huntington's disease (Oster, Shoulson and Dorsey, 2013), that unemployed individuals overestimate their likelihood of finding a job (Spinnewijn, 2012), and that some CEOs are overoptimistic regarding the future performance of their firms (Malmendier and Tate, 2005). Moreover, in all of these domains the expressed or measured overconfidence predicts individual choice behavior. For example, CEOs' overconfidence predicts the likelihood of acquiring other firms (Malmendier and Tate, 2008), of using internal rather than external financing (Malmendier and Tate, 2005), of using short-term debt (Graham, Harvey and Puri, 2013), of engaging in financial misreporting (Schrand and Zechman, 2012), and of engaging in innovative activity (Hirshleifer et al., 2012). While all of these papers look at the relationship between overconfidence and behavior, they do not theoretically investigate the implications of overconfidence for (misdirected) learning about other variables.

A number of theoretical papers explain why agents become (or seem to become) overconfident. In one class of papers, the agent's learning process is tilted in favor of moving toward or stopping at confident beliefs (Gervais and Odean, 2001, Zábojnik, 2004, Bénabou and Tirole, 2002, Kőszegi, 2006). In other papers, non-common priors or criteria lead agents to take actions that lead the average agent to expect better outcomes than others (Van den Steen, 2004, Santos-Pinto and Sobel, 2005). Finally, some papers assume that an agent simply chooses unrealistically positive beliefs because she derives direct utility from such beliefs (Brunnermeier and Parker, 2005, Oster et al., 2013). While these papers provide foundations for overconfident beliefs and some feature learning, they do not analyze how overconfidence affects learning about other variables.

Many researchers take the view that overconfidence can be individually and socially beneficial even beyond providing direct utility.[12] Our theory is not in contradiction with this view, but it does predict circumstances under which overconfidence can be extremely harmful.

5.2 Learning with Misspecified Models

On a basic level, an overconfident agent has an incorrect view of the world, and hence our paper is related to the literature on learning with misspecified models. Most closely related to our theory, Esponda and Pouzo (forthcoming) develop a general framework for studying repeated games in which players have misspecified models. Our model is a special case of theirs in which there is one player. Building on Berk (1966), Esponda and Pouzo establish that if actions converge, beliefs converge to a limit at which a player's predicted distribution of outcomes is closest to the actual distribution. Our limiting beliefs have a similar property. Because of our specific setting, we derive stronger results on convergence of beliefs and establish many other properties of the learning process and the limiting beliefs.

Taking the interpretation that at most one prior can be correct, multi-agent models with non-common priors can also be viewed as analyzing learning with misspecified models. In this literature, papers ask how different agents' beliefs change relative to each other, but do not study the interaction with behavior. Dixit and Weibull (2007) construct examples in which individuals with different priors interpret signals differently, so that the same signal can push their beliefs further from each other. Similarly, Acemoglu, Chernozhukov and Yildiz (forthcoming) consider Bayesian agents with different prior beliefs regarding the conditional distribution of signals given (what we call) the fundamental, and show that the agents' beliefs regarding the fundamental do not necessarily converge.[13]

Misdirected learning also occurs in social learning settings when individuals have misspecified models of the world. In Eyster and Rabin (2010) and in many cases also in Bohren (2013), agents do not sufficiently account for redundant information in previous actions. With more and more redundant actions accumulating, this mistake is amplified, preventing learning even in the long

[12] See, e.g., Taylor and Brown (1988) for a review of the relevant psychology literature, and Bénabou and Tirole (2002) and de la Rosa (2011) for economic examples.

[13] Andreoni and Mylovanov (2012) and Kondor (2012) develop closely related models within the common-prior paradigm. In their models, there are two sources of information about a one-dimensional fundamental. Agents receive a series of public signals, and private signals on how to interpret the public signals. Although beliefs regarding the public information converge, beliefs regarding the fundamental do not, as agents keep interpreting the public information differently.

run. We are interested in self-defeating learning in individual-decisionmaking rather than in social environments, and explore the implications of a very different mistake.

The interaction between incorrect inferences and behavior we study is somewhat reminiscent of Ellison's (2002) model of academic publishing, in which researchers who are biased about the quality of their own work overestimate publishing standards, making them tougher referees and thereby indeed toughening standards. In contrast to our work, updating is ad-hoc, and the model relies on "feedback" from others on the evolution of the publishing standard.

There is also a considerable literature on the learning implications of various mistakes in interpreting information (see for instance Rabin and Schrag, 1999, Rabin, 2002, Eyster and Rabin, 2010, Madarász, 2009, Rabin and Vayanos, 2010, Benjamin, Rabin and Raymond, 2012). Overconfidence is a different type of mistake—in particular, it is not directly an inferential mistake—so our results have no close parallels in this literature.

Methodologically, our theory confirms Fudenberg's (2006) point that it is often insufficient to do behavioral economics by modifying one assumption of a classical model, as one modeling change often justifies other modeling changes as well. In our setting, the agent's false belief about her ability leads her to draw incorrect inferences regarding the fundamental, so assuming that an overconfident agent is otherwise classical may be misleading.

6 Conclusion

While our paper focuses exclusively on individual decisionmaking, the possibility of self-defeating learning likely has important implications for multi-agent situations. For example, it has been recognized in the literature that managerial overconfidence can benefit a firm both because it leads the manager to overvalue bonus contracts and because it can lead her to exert greater effort (de la Rosa, 2011, for example). Yet for tasks with the properties we have identified, misguided learning can also induce a manager to make highly suboptimal decisions. Hence, our analysis may have implications on the optimal allocation of decisionmaking authority for a manager.

A completely different issue is that different individuals may have different interpretations as to what explains unexpectedly low performance. For instance, a Democrat may interpret poor health outcomes as indicating problems with the private market, whereas a Republican may think that the culprit is government intervention. In the formalism of our model, one side believes that output is increasing in b_t, while the other side believes that output is decreasing in b_t. Similarly to Dixit and Weibull's (2007) model of political polarization, decisionmakers with such opposing theories may prefer to adjust policies in different directions. Our model highlights that unrealistically high expectations regarding outcomes can play an important role in political polarization. Furthermore, our model makes predictions on how the two sides interpret each other's actions. It may be the case, for instance, that if a Republican has the power to make decisions, she engages in self-defeating learning as our model predicts, with a Democrat looking on in dismay and thinking that—if only she had the power—a small adjustment in the opposite direction would have been sufficient. If the Democrat gets power, she adjusts a little in the opposite direction, initially improving performance, but then she engages in self-defeating learning of her own. Frequent changes in power therefore help keep self-defeating learning in check. Meanwhile, an independent who also has unrealistic expectations but is unsure about which side's theory of the world is correct always tends to move her theory toward the theory of the party in opposition, as that party's theory is better at explaining current observations.

The assumption in our model that the agent has a degenerate overconfident belief about ability—while invaluable for tractability—prevents us from studying a number of relevant questions. As a case in point, it would be important to understand whether and how the forces we identify interact with the forces that generate widespread overconfident beliefs in the population. For instance, our results suggest one possible reason for the persistence—though not for the emergence—of overconfidence. Since (as discussed in Section 3.1) the agent often abandons tasks and prefers tasks for which previous learning does not apply, she can keep updating mostly about external circumstances, slowing down learning about her ability.

References

Acemoglu, Daron, Victor Chernozhukov, and Muhamet Yildiz, "Fragility of Asymptotic Agreement under Bayesian Learning," *Theoretical Economics*, forthcoming.

Andreoni, James and Tymofiy Mylovanov, "Diverging Opinions," *American Economic Journal: Microconomics*, 2012, *4* (1), 209–232.

Argyris, Chris, *Reasoning, Learning and Action*, San Francisco, CA: Jossey-Bass, 1982.

Baumeister, Roy F. and Steven J. Scher, "Self-Defeating Behavior Patterns Among Normal Individuals: Review and AnalysisofCommon Self-Destructive Tendencies," *Psychological Bulletin*, 1988, *104*, 3–22.

—, **Laura Smart, and Joseph M. Boden**, "Relation of Threatened Egotism to Violence and Aggression: The Dark Side of High Self-Esteem," *Psychological Review*, 1996, *103* (1), 5–33.

Bénabou, Roland and Jean Tirole, "Self-Confidence and Personal Motivation," *Quarterly Journal of Economics*, 2002, *117* (3), 871–915.

Benjamin, Daniel J., Matthew Rabin, and Collin Raymond, "A Model of Non-Belief in the Law of Large Numbers," 2012. Working Paper, Cornell University.

Benoît, JeanPierre and Juan Dubra, "Apparent Overconfidence," *Econometrica*, 2011, *79* (5), 1591–1625.

—, —, **and Don A. Moore**, "Does the Better-Than-Average Effect Show That People Are Overconfident? Two Experiments.," *Journal of the European Economic Association*, 2015, *13* (2), 293–329.

Berk, Richard H., "Limiting Behavior of Posterior Distributions when the Model Is Incorrect," *Annals of Mathematical Statistics*, 1966, *37* (1), 51–58.

Bohren, Aislinn, "Informational herding with Model Mispecification," June 2013. Working Paper.

Brunnermeier, Markus and Jonathan Parker, "Optimal Expectations," *American Economic Review*, 2005, *95* (4), 1092–1118.

Budgett, Richard, "Fatigue and Underperformance in Athletes: The Overtraining Syndrome.," *British Journal of Sports Medicine*, 1998, *32* (2), 107–110.

Burks, Stephen V., Jeffrey P. Carpenter, Lorenz Goette, and Aldo Rustichini, "Overconfidence and Social Signalling," *The Review of Economic Studies*, 2013, *80* (3), 949–983.

Camerer, Colin and Dan Lovallo, "Overconfidence and Excess Entry: An Experimental Approach," *American Economic Review*, 1999, *89* (1), 306–318.

Campbell, W. Keith and Constantine Sedikides, "Self-Threat Magnifies the Self-Serving Bias: A Meta-Analytic Integration," *Review of General Psychology*, 1999, *3* (1), 23–43.

Charness, Gary, Aldo Rustichini, and Jeroen van de Ven, "Self-Confidence and Strategic Behavior," 2014. Working Paper.

Chiang, Yao-Min, David Hirshleifer, Yiming Qian, and Ann E. Sherman, "Do Investors Learn from Experience? Evidence from Frequent IPO Investors," *Review of Financial Studies*, 2011, *24* (5), 1560–1589.

Crozier, Michel, *The Bureaucratic Phenomenon*, Chicago: University of Chicago Press, 1964.

de la Rosa, Leonidas Enrique, "Overconfidence and Moral Hazard," *Games and Economic Behavior*, 2011, *73* (2), 429–451.

DellaVigna, Stefano and Ulrike Malmendier, "Paying Not to Go to the Gym," *American Economic Review*, 2006, *96* (3), 694–719.

den Steen, Eric Van, "Rational Overoptimism (and Other Biases)," *American Economic Review*, 2004, *94* (4), 1141–1151.

Dixit, Avinash K. and Jörgen Weibull, "Political Polarization," *Proceedings of the National Academy of Sciences*, 2007, *104* (18), 7351–7356.

Ellison, Glenn, "Evolving Standards for Academic Publishing: A q-r Theory," *Journal of Political Economy*, October 2002, *110* (5), 994–1034.

Esponda, Ignacio and Demian Pouzo, "Berk-Nash Equilibrium: A Framework for Modeling Agents with Misspecified Models," *Econometrica*, forthcoming.

Eyster, Erik and Matthew Rabin, "Naïve Herding in Rich-Information Settings," *American Economic Journal: Microeconomics*, 2010, *2* (4), 221–243.

Fudenberg, Drew, "Advancing Beyond Advances in Behavioral Economics," *Journal of Economic Literature*, 2006, *44* (3), 694–711.

Galasso, Alberto and Timothy S. Simcoe, "CEO Overconfidence and Innovation," *Management Science*, 2011, *57* (8), 1469–1484.

Gervais, Simon and Terrance Odean, "Learning To Be Overconfident," *Review of Financial Studies*, 2001, *14* (1), 1–27.

Graham, John R., Campbell R. Harvey, and Manju Puri, "Managerial Attitudes and Corporate Actions," *Journal of Financial Economics*, 2013, *109* (1), 103 – 121.

Hirshleifer, David, Angie Low, and Siew Hong Teoh, "Are Overconfident CEOs Better Innovators?," *Journal of Finance*, 08 2012, *67* (4), 1457–1498.

Hoffman, Mitch and Stephen Burks, "Training Contracts, Worker Overconfidence, and the Provision of Firm-Sponsored General Training," 2013. Working Paper.

Holmström, Bengt and Paul Milgrom, "Multi-Task Principal-Agent Analyzes: Incentive Contracts, Asset Ownership,andJob Design," *Journal of Law, Economics and Organization*, 1991, *7*, 24–52. special issue.

Kasser, Tim, *The High Price of Materialism*, Bradford, 2003.

Kőszegi, Botond, "Ego Utility, Overconfidence, and Task Choice," *Journal of the European Economic Association*, 2006, *4* (4), 673–707.

Kondor, Péter, "The More We Know about the Fundamental, the Less We Agree on the Price," *Review of Economic Studies*, 2012, *79* (3), 1175–1207.

Kushner, Harold J and George Yin, *Stochastic approximation and recursive algorithms and applications*, Vol. 35, Springer, 2003.

Landier, Augustin and David Thesmar, "Financial Contracting with Optimistic Entrepreneurs," *Review of Financial Studies*, 2009, *22* (1), 117–150.

Madarász, Kristóf, "Information Projection: Model and Applications," 2009. SSRN Paper #1407190.

Malmendier, Ulrike and Geoffrey Tate, "CEO Overconfidence and Corporate Investment," *Journal of Finance*, 2005, *60* (6), 2661–2700.

_ **and** _ , "Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction," *Journal of Financial Economics*, 2008, *89* (1), 20–43.

Manzoni, Jean-François and Jean-Louis Barsoux, "The Set-Up-to-Fail Syndrome," *Harvard Business Review*, March-April 1998, pp. 101–113.

_ **and** _ , *The Set-Up-to-Fail Syndrome*, Harvard Business School Press, 2002.

March, James G. and Herbert A. Simon, *Organizations*, New York: Wiley, 1958.

Masuch, Michael, "Vicious circles in organizations," *Administrative Science Quarterly*, 1985, pp. 14–33.

McFarlin, Dean B., Roy F. Baumeister, and Jim Blascovich, "On Knowing When to Quit: Task Failure, Self-Esteem, Advice, and Nonproductive Persistence," *Journal of Personality*, 1984, *52* (2), 138–155.

Miller, Dale T. and Michael Ross, "Self-serving Biases in the Attribution of Causality: Fact or Fiction?," *Psychological Bulletin*, 1975, *82* (2), 213–225.

Moore, Don A. and Paul J. Healy, "The Trouble with Overconfidence," *Psychological Review*, 2008, *115* (2), 502–517.

Oster, Emily, Ira Shoulson, and E. Ray Dorsey, "Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease," *American Economic Review*, 2013, *103* (2), 804–30.

Rabin, Matthew, "Inference by Believers in the Law of Small Numbers," *Quarterly Journal of Economics*, 2002, *117* (3), 775–816.

_ **and Dimitri Vayanos**, "The Gambler's and Hot-Hand Fallacies: Theory and Applications," *Review of Economic Studies*, 2010, *77* (2), 730–778.

_ **and Joel Schrag**, "First Impressions Matter: A Model of Confirmatory Bias," *Quarterly Journal of Economics*, 1999, *114* (1), 37–82.

Santos-Pinto, Luis and Joel Sobel, "A Model of Positive Self-Image in Subjective Assessments," *American Economic Review*, 2005, *95* (5), 1386–1402.

Savage, Leonard J., *The Foundations of Statistics*, New York, Wiley, 1954.

Schrand, Catherine M. and Sarah L.C. Zechman, "Executive overconfidence and the slippery slope to financial misreporting," *Journal of Accounting and Economics*, 2012, *53* (1–2), 311 – 329.

Shui, Haiyan and Lawrence M. Ausubel, "Time Inconsistency in the Credit Card Market," 2004. Available at http://ssrn.com/abstract=586622.

Spinnewijn, Johannes, "Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs," 2012. Working Paper.

Svenson, Ola, "Are We Less Risky and More Skillful Than Our Fellow Drivers?," *Acta Psychologica*, 1981, *18*, 473–474.

Taylor, Shelley E. and Jonathan D. Brown, "Illusion and Well-Being: A Social Psychological Perspective on Mental Health," *Psychological Bulletin*, 1988, *103* (2), 193–210.

Tetlock, Philip E. and Ariel Levi, "Attribution Bias: On the Inconclusiveness of the Cognition-Motivation Debate," *Journal of Experimental Social Psychology*, 1982, *18* (1), 68–88.

Weinstein, Neil D., "Unrealistic Optimism About Future Life Events," *Journal of Personality and Social Psychology*, 1980, *39*, 806–820.

Zábojnik, Ján, "A Model of Rational Bias in Self-Assessments," *Economic Theory*, 2004, *23* (2), 259–282.

A Notation

Throughout the proofs in the Appendix, we use the following additional notation for the bounds on the derivatives of Q we imposed in Assumption 1. We let $\underline{Q}_{ee} > -\infty$ be a lower bound on Q_{ee}. Let $\underline{Q}_b, \overline{Q}_b > 0$ be bounds on Q_b such that $\underline{Q}_b \leq Q_b \leq \overline{Q}_b$; similarly, let $\underline{Q}_a, \overline{Q}_a > 0$ be bounds satisfying $\underline{Q}_a \leq Q_a \leq \overline{Q}_a$. Let $\underline{Q}_{ea} > -\infty$ and $0 < \kappa < \infty$ be bounds such that $Q_{ea} \geq \underline{Q}_{ea}$ and $Q_{eb}(e_t, a, b_t)/|Q_{ee}(e_t, a, b_t)| \leq \kappa$ for all e_t, a, b_t. Note that since $|\underline{Q}_{ee}| < \infty$, the latter bound implies also that Q_{eb} is bounded from above. Let $\overline{Q}_{eb} < \infty$ be a bound such that $Q_{eb} \leq \overline{Q}_{eb}$.

B Proofs

Proof of Lemma 1. We established in the text that the agent believes at the end of period $t-1$ that b_t is normally distributed with mean $\tilde{\phi}_{t-1}$ and variance $h_\epsilon^{-1} + [h_0 + (t-1)h_\epsilon]^{-1}$. Now because Q is strictly concave in e_t, so is the expectation $\tilde{\mathbb{E}}_{t-1,\tilde{\phi}_{t-1}}\left[Q(e, \tilde{a}, \tilde{b})\right]$ and thus if an optimal action exists it is unique and characterized by the first order condition

$$0 = \tilde{\mathbb{E}}_{t-1,\tilde{\phi}_{t-1}}\left[Q_e(e^*(t, \tilde{\phi}_{t-1}), \tilde{a}, \tilde{b})\right]. \tag{11}$$

Since for all a, b_t, $\lim_{e \to \underline{e}} Q_e(e, a, b_t) < 0$ and $\lim_{e \to \overline{e}} Q_e(e, a, b_t) > 0$, an optimal action exists. As $|Q_{ee}|$ and $|Q_{eb}|$ are bounded we can apply the dominated convergence theorem to show that the right-hand-side of Equation 11 is differentiable in e and ϕ. We can thus apply the implicit function Theorem to get

$$\frac{\partial e^*(t, \tilde{\phi}_{t-1})}{\partial \tilde{\phi}_{t-1}} = -\frac{\tilde{\mathbb{E}}_{t,\tilde{\phi}_{t-1}}\left[Q_{eb}(e^*(t, \tilde{\phi}_{t-1}), \tilde{a}, \tilde{b})\right]}{\tilde{\mathbb{E}}_{t,\tilde{\phi}_{t-1}}\left[Q_{ee}(e^*(t, \tilde{\phi}_{t-1}), \tilde{a}, \tilde{b})\right]} > 0\,,$$

where the inequality follows from the facts that that $Q_{eb} > 0$ and $Q_{ee} < 0$ by Assumption 1. □

Proof of Proposition 1. We begin with the following lemma:

Lemma 3. *The belief $\tilde{\phi}_t$ is strictly increasing in e_t.*

Proof of Lemma 3. The agent's perceived signal $\tilde{b}_t$ at the end of period t is implicitly defined through

$$Q(e_t, \tilde{a}, \tilde{b}_t) = Q(e_t, A, b_t)\,. \tag{12}$$

Because Q is strictly increasing in the state of the world, there is a unique belief $\tilde{b}_t$ satisfying the above equation. We begin by establishing that since $Q_{ae} \leq 0$, the perceived signal $\tilde{b}_t$ is strictly increasing in the effort chosen in period t. To see this formally, subtract $Q(e_t, \tilde{a}, b_t)$ on both sides of Equation 12

$$Q(e_t, \tilde{a}, \tilde{b}_t) - Q(e_t, \tilde{a}, b_t) = Q(e_t, A, b_t) - Q(e_t, \tilde{a}, b_t),$$

and rewrite it as

$$\int_{\tilde{b}_t}^{b_t} Q_b(e_t, \tilde{a}, z)\mathrm{d}z = \int_0^{\tilde{a}} Q_a(e_t, z, b_t)\mathrm{d}z.$$

Since $Q_{ae} \leq 0$, the right hand side of the above equation weakly decreases in the chosen effort e_t, and since $Q_{be} > 0$ the left hand side strictly increases unless $\tilde{b}_t$ strictly increases. Since $\tilde{\phi}_t$ is a weighted average between $\tilde{b}_t$ and $\tilde{\phi}_{t-1}$, the fact that $\tilde{b}_t$ strictly increases in e completes the proof. □

We now prove the proposition.

Step 1. We now establish that the belief strictly decreases in period t, i.e. $\tilde{\phi}_t < \tilde{\phi}_{t-1}$, if and only if $\Gamma(\tilde{\phi}_{t-1}) < 0$. Using that $e_t = e^*(\tilde{\phi}_{t-1})$, that $b_t = \Phi$ in our noiseless limit, and subtracting $Q(e^*(\tilde{\phi}_{t-1}), \tilde{a}, \tilde{\phi}_{t-1})$ from both sides of Equation 12 yields

$$Q(e^*(\tilde{\phi}_{t-1}), \tilde{a}, \tilde{b}_t) - Q(e^*(\tilde{\phi}_{t-1}), \tilde{a}, \tilde{\phi}_{t-1}) = \Gamma(\tilde{\phi}_{t-1})\,.$$

Since $Q_b > 0$, it follows that $\tilde{b}_t < \tilde{\phi}_{t-1}$, and because $\tilde{\phi}_t$ is a weighted average of $\tilde{b}_t$ and $\tilde{\phi}_{t-1}$, $\tilde{\phi}_t < \tilde{\phi}_{t-1}$.

Step 2. We next establish that the sequence of beliefs is monotone. Observe that the belief at the end of period t, $\tilde{\phi}_t$, depends only on e_t and $\tilde{\phi}_{t-1}$, and by Lemma 3 $\tilde{\phi}_t$ is strictly increasing in e_t and it is also obviously increasing in $\tilde{\phi}_{t-1}$. Furthermore, by Lemma 1, $e^*(\tilde{\phi}_{t-1})$ itself is strictly increasing in $\tilde{\phi}_{t-1}$. Hence, $\tilde{\phi}_t$ is strictly increasing in $\tilde{\phi}_{t-1}$.

We conclude from Steps 1 and 2 that the sequence $\tilde{\phi}_t$ is strictly decreasing if $\Gamma(\tilde{\phi}_0) < 0$ and strictly increasing if $\Gamma(\tilde{\phi}_0) > 0$.

Step 3. Consider first the case in which $\Gamma(\tilde{\phi}_0) < 0$. Let

$$\hat{\phi} = \max\{\phi \in C \colon \phi \leq \tilde{\phi}_0\} .$$

We proof by induction that $\tilde{\phi}_t > \hat{\phi}$ for all t. Suppose $\tilde{\phi}_{t-1} > \hat{\phi}$. The optimal effort e_t is increasing in $\tilde{\phi}_{t-1}$, so that $e_t = e^*(\tilde{\phi}_{t-1}) > e^*(\hat{\phi})$. Since the belief $\tilde{\phi}_t$ is a strictly increasing function of e_t and increasing in $\tilde{\phi}_{t-1}$, and since $e_t > e^*(\hat{\phi})$ and $\phi_{t-1} > \hat{\phi}$, we have

$$\tilde{\phi}_t = \tilde{\phi}(e^*(\tilde{\phi}_{t-1}), \tilde{\phi}_{t-1}) > \tilde{\phi}(e^*(\hat{\phi}), \hat{\phi}) .$$

Because $\hat{\phi} \in C$, we also have $\tilde{\phi}(e^*(\hat{\phi}), \hat{\phi}) = \hat{\phi}$. Hence, $\tilde{\phi}_t > \hat{\phi}$. By induction, thus, $\tilde{\phi}_t > \hat{\phi}$ for all t.

In case $\Gamma(\tilde{\phi}_0) > 0$, an analogous argument shows that for $\hat{\phi}' = \min\{\phi \in C \colon \phi \geq \tilde{\phi}_0\}$, $\tilde{\phi}_t < \hat{\phi}'$ for all t.

Step 4. Again, we consider the case $\Gamma(\tilde{\phi}_0) < 0$ first. Since the belief is monotone decreasing for $\Gamma(\tilde{\phi}_0) < 0$, and it is bounded from below by $\hat{\phi}$, it thus converges. We next show that it converges to $\hat{\phi}$ by showing that any limit point must satisfy $\Gamma(\phi) = 0$. As in the noiseless limit $b_t = \Phi$ one has

$$\Gamma(\tilde{\phi}_t) = Q(e^*(\tilde{\phi}_t), A, \Phi) - Q(e^*(\tilde{\phi}_t), \tilde{a}, \tilde{\phi}_t) = Q(e^*(\tilde{\phi}_t), \tilde{a}, \tilde{b}_t) - Q(e^*(\tilde{\phi}_t), \tilde{a}, \tilde{\phi}_t) \geq \underline{Q}_b(\tilde{b}_t - \tilde{\phi}_t),$$

where the first line follows from Equation 12 and the second from the facts that $Q_b \geq \underline{Q}_b$ and $\tilde{b}_t < \tilde{\phi}_t$. Let $\tilde{\phi}_\infty = \lim_{t\to\infty} \tilde{\phi}_t$. We have that

$$\phi_{t+1} - \phi_t = \frac{1}{t}(\tilde{b}_t - \tilde{\phi}_t) \leq \frac{1}{\underline{Q}_b\, t} \inf_{\phi \in [\tilde{\phi}_\infty, \phi_0]} \Gamma(\phi) .$$

Thus, if we define $d := \inf_{\phi \in [\tilde{\phi}_\infty, \phi_0]} \Gamma(\phi) < 0$, we have

$$\tilde{\phi}_t = \tilde{\phi}_0 - \sum_{s=1}^{t-1}(\tilde{\phi}_{t+1} - \tilde{\phi}_t) \leq \tilde{\phi}_0 - \frac{|d|}{\underline{Q}_b} \sum_{s=1}^{t} \frac{1}{s} .$$

Consequently, $\tilde{\phi}$ diverges to minus infinity whenever $d \neq 0$ and thus by definition of d we have $\Gamma(\tilde{\phi}_\infty) = 0$.

In case $\Gamma(\tilde{\phi}_0) > 0$, a very similar argument establishes that beliefs converge to $\hat{\phi}' = \min\{\phi \in C\colon \phi \geq \tilde{\phi}_0\}$. □

Proof of Corollary 1. It follows from the updating rule in the noiseless limit case that $\tilde{\phi}_1 = \tilde{b}_1$; $\tilde{b}_1$ is implicitly defined through

$$Q(e^*(\tilde{\phi}_0, \tilde{a}, \tilde{b}_1) = Q(e^*(\tilde{\phi}_0), A, \Phi).$$

Since $Q_a, Q_b > 0$, $\tilde{b}_1 < \Phi$ as otherwise the left-hand side above is greater than the right-hand side. Hence, $\tilde{\phi}_1 < \Phi$. As beliefs are initially decreasing, $\Gamma(\tilde{\phi}_0) < 0$, and the remainder of the corollary follows immediately from Proposition 1. □

Proof of Proposition 2. Consider the set of points where the agent revises his belief upwards

$$D_i = \{\phi\colon \Gamma_i(\phi) \geq 0\}, \quad \text{for } i = 1, 2.$$

Note that $\Gamma_i(\phi) = R_i(A, \Phi) - R_i(\tilde{a}, \phi) - L_i(e^*(\tilde{a}, \phi), A, \Phi)$. Thus, $\Gamma_1 > \Gamma_2$ pointwise and $D_1 \supset D_2$.

First, let $\tilde{\phi}_0 \in D_2$. By Proposition 1 the limit belief is given by

$$\tilde{\phi}_{\infty,i} = \inf\{\phi \notin D_i\colon \phi \geq \tilde{\phi}_0\} = \inf\left(D_i^C \cap [\tilde{\phi}_0, \infty)\right) \text{ for } i = 1, 2.$$

As $D_1^C \subset D_2^C$ the infimum strictly increases and $\tilde{\phi}_{\infty,2} < \tilde{\phi}_{\infty,1}$. Second, consider the case where $\tilde{\phi}_0 \notin D_2$ and $\tilde{\phi}_0 \in D_1$. By Proposition 1, with technology Q_1 beliefs are updated upwards and with Q_2 strictly downwards, so that $\tilde{\phi}_{\infty,2} < \tilde{\phi}_0 \leq \tilde{\phi}_{\infty,1}$. Third, let $\tilde{\phi}_0 \notin D_2$ and $\tilde{\phi}_0 \notin D_1$. Proposition 1 implies that the limit belief is given by

$$\tilde{\phi}_{\infty,i} = \sup\{\phi \in D_i\colon \phi \leq \tilde{\phi}_0\} = \sup\left(D_i \cap (-\infty, \tilde{\phi}_0]\right) \text{ for } i = 1, 2.$$

As $D_1 \supset D_2$ the supremum strictly decreases and $\tilde{\phi}_{\infty,2} < \tilde{\phi}_{\infty,1}$. Fourth, observe that $\tilde{\phi}_0 \in D_2$ and $\tilde{\phi}_0 \notin D_1$ contradicts $d_2 \subset D_1$. Finally, since, $\Gamma_1 > \Gamma_2$ it follows that $C_1 \cap C_2 = \emptyset$. Hence $\tilde{\phi}_{\infty,2} < \tilde{\phi}_{\infty,1}$ in all cases above. □

Proof of Proposition 3. By Assumption 1(i) the agent has a unique optimal action, and since Q is twice differentiable, the agent's optimal effort in period t must satisfy the following first-order condition:

$$Q_S(S(e_t, a), b_t)\ S_{e_t}(e_t, a) = 0.$$

Since the optimal action is unique, $Q_S(S(e_t,a),b_t) \neq 0$ at the optimum, and hence at the optimum $S_{e_t}(e_t,a) = 0$. Therefore, the optimal action is independent of the state of the world, and hence the agent's beliefs. Thus, the agent never changes her action. $\square$

Proof of Proposition 4. $I \Rightarrow II$. Denote by $\tilde{\phi}(A,\tilde{a},\phi)$ the unique steady-state beliefs associated with $A,\tilde{a},\phi$, and note that lead the agent to take the optimal action. That the action is optimal when the agent holds beliefs $\tilde{\phi}(A,\tilde{a},\phi)$ is equivalent to

$$e^*(A,\phi) = \arg\max_e Q(e,A,\phi) = \arg\max_e Q(e,\tilde{a},\tilde{\phi}(A,\tilde{a},\phi)). \tag{13}$$

Furthermore, the fact that these are steady-state beliefs means that the agent gets no surprise:

$$Q(e^*(A,\phi),A,\phi) = Q(e^*(A,\phi),\tilde{a},\tilde{\phi}(A,\tilde{a},\phi)). \tag{14}$$

We establish some properties of $e^*(a,\phi)$. First, we know that $e^*_\phi(a,\phi) > 0$. Second, we show that $e^*_a(a,\phi) > 0$. Take any $a' > a$. By Equation (13), we have $e^*(a,\phi) = e^*(\tilde{a},\tilde{\phi}(a,\tilde{a},\phi))$, so that

$$e^*(a',\phi) = e^*(a,\tilde{\phi}(a',a,\phi)).$$

Now by Equation (14), we know that $\tilde{\phi}(a,\tilde{a},\phi)$ is strictly decreasing in $\tilde{a}$. Furthermore, since $\tilde{\phi}(a',a',\phi) = \phi$, we can conclude that $\tilde{\phi}(a',a,\phi) > \phi$. But then $e^*(a,\tilde{\phi}(a',a,\phi)) > e^*(a,\phi)$, establishing our claim that $e^*_a(a,\phi) > 0$.

Given these properties, we can define $B(a,\phi) = e^*(a,\phi)$, and this function satisfies the properties in the proposition. Furthermore, using again that $e^*(a,\phi) = e^*(\tilde{a},\tilde{\phi}(a,\tilde{a},\phi))$, note that for any $a,\tilde{a},\phi$, the unique ϕ' satisfying $B(a,\phi) = B(\tilde{a},\phi')$ is $\phi' = \tilde{\phi}(a,\tilde{a},\phi)$. Now we define

$$\hat{Q}(e,a,\phi) = B(a,\phi) - L(|e - B(a,\phi)|),$$

where $L(\cdot)$ is any strictly increasing function satisfying $L'(x) < 1$ everywhere. By construction, for any initial beliefs the agent chooses the same action with the output function $\hat{Q}$ as she does with the output function Q. In addition, upon observing q_1, she can infer the value of B, so that she forms new beliefs equal to $\tilde{\phi}(a,\tilde{a},\phi)$. Again by construction, from here she does not change beliefs, and behaves the same way as she would with output function Q.

$II \Rightarrow I$. Because $Q_B, B_a, B_b > 0$, for any action e there is a unique $\tilde{b}$ such that $Q(e,B(A,\Phi)) = Q(e,B(\tilde{a},\tilde{b}))$. Hence, $\tilde{\phi}_t = \tilde{b}$ for all $t \geq 1$ and $\tilde{\phi}(a,\tilde{a},\Phi)$ in the proposition thus equals $\tilde{b}$. For all $t \geq 2$,

e_t is implicitly defined through $Q_e(e_t, B(\tilde{a}, \tilde{\phi}_{t-1})) = 0$. Since $B(\tilde{a}, \tilde{\phi}_{t-1}) = B(\tilde{a}, \tilde{b}) = B(A, \Phi))$, the agent thus chooses the optimal action for which $Q_e(e_t, B(A, \Phi)) = 0$ from period 2 on. $\square$

Proof of Proposition 5. First, note that as $Q(e, a, b) = a + b - L(e - b)$ is increasing in b, a, the agent is under-confident $\tilde{a} < 0$ and the perceived signal $\tilde{b}_t$ satisfies

$$Q(e_t, \tilde{a}, \tilde{b}_t) = Q(e_t, A, b_t) \tag{15}$$

we have that $\tilde{b}_t > b_t = \Phi$ for all t. Thus, the limit mean belief is bounded from below by the fundamental Φ

$$\tilde{\phi}_\infty = \lim_{t\to\infty} \frac{1}{t} \sum_{s\le t} \tilde{b}_s \ge \lim_{t\to\infty} \frac{1}{t} \sum_{s\le t} b_s = \Phi\,.$$

By the symmetry of the loss function, the agent takes the action that equals her posterior belief in the previous period $\tilde{\phi}_{t-1} = e_t$. Beliefs converge to $\tilde{\phi}_\infty$ if and only if $\lim_{t\to\infty} \tilde{b}_t = \tilde{\phi}_\infty$. Thus, if the beliefs converge, the expected loss $L(e_t - \tilde{b}_t)$ from the agent's perspective vanishes in the limit for which $\lim_{t\to\infty} e_t = \lim_{t\to\infty} \tilde{\phi}_{t-1} = \lim_{t\to\infty} \tilde{\phi}_t = \lim_{t\to\infty} \tilde{b}_t$. As the agent expects to suffer no loss in the limit, the limit mean belief satisfies

$$\tilde{a} + \tilde{\phi}_\infty = A + \Phi - L(\tilde{\phi}_\infty - \Phi)\,.$$

Rearranging for $|\Delta| = -(\tilde{a} - A)$ and using the fact that L is strictly positive whenever $\tilde{\phi}_\infty \neq \Phi$ yields

$$|\Delta| = \tilde{\phi}_\infty - \Phi + L(\tilde{\phi}_\infty - \Phi) > \tilde{\phi}_\infty - \Phi\,.$$

Since we already established that $\tilde{\phi}_\infty > \Phi$, we conclude that $0 < \tilde{\phi}_\infty - \Phi < |\Delta|$. Since the slope of the loss function is less than one everywhere, this also implies that $L(\tilde{\phi}_\infty - \Phi) < |\Delta|$.

Finally, we argue that mean beliefs converge. Rewriting Equation 15 yields

$$|\Delta| + \Phi - L(e_t - \Phi) = \tilde{b}_t - L(e_t - \tilde{b}_t).$$

Since $\tilde{b}_t > \Phi$, we know that for any period $\tau > 1$, $e_\tau = \phi_{\tau-1} = \sum_{t=1}^{\tau-1} \tilde{b}_t > \Phi$. Since $L(\cdot)$ is weakly convex, $\tilde{b}_\tau$ is weakly decreasing in e_τ. Defining $\tilde{b}^{max}$ implicitly through $|\Delta| + \Phi = \tilde{b}^{max} - L(\tilde{b}^{max} - \Phi)$, thus, implies that for all $\tau > 1$, $\tilde{b}_\tau \in [\Phi, \tilde{b}^{max}]$.

Since $\tilde{b}_t$ is weakly decreasing in the chosen action $e_t = \tilde{\phi}_{t-1}$ and there is a unique consistent belief $\hat{\phi}$, for all $\tilde{\phi}_t > \hat{\phi}$ one has $\tilde{b}_t \le \hat{\phi}$ and for all $\tilde{\phi}_t < \hat{\phi}$ one has $\tilde{b}_t \ge \hat{\phi}$. For any $\tilde{\phi} \neq \hat{\phi}$ let

$\delta = |\tilde{\phi} - \hat{\phi}|/2$. If for any period $t \geq \tau$ the agent chooses an action $e_t \in (\hat{\phi} - \delta, \hat{\phi} + \delta)$, the sum $\sum_{t=\tau}^{\infty}(1/t)|\tilde{b}_t - \hat{\phi}| \geq \sum_{t=\tau}^{\infty}(1/t)\delta = \infty$, and hence beliefs cannot converge to some limit belief other than $\hat{\phi}$. By definition $\Gamma(\hat{\phi}) = 0$.

To establish that beliefs converge, we next show that for every $\delta > 0$, there exists a period $\hat{t}(\delta)$ such that if beliefs $\tilde{\phi}_t$ enter the interval $(\hat{\phi} - \delta, \hat{\phi} + \delta)$ in some period $t \geq \hat{t}(\delta)$, then they never leave this interval. Let $t(\delta)$ be the smallest t for which $[\tilde{b}^{max} - \Phi] < \hat{t}(\delta)\delta$. Then, for any $t \geq \hat{t}(\delta)$, $|\tilde{\phi}_{t+1} - \tilde{\phi}_t| = (1/(t+1))[\tilde{b}^{max} - \Phi] < \delta$, and since for $\tilde{\phi}_t < \hat{\phi}$ beliefs increase, one has $\hat{\phi} - \delta < \tilde{\phi}_t < \tilde{\phi}_t + \delta < \hat{\phi} + \delta$. Similarly, since for $\tilde{\phi}_t > \hat{\phi}$ beliefs decrease, one has $\hat{\phi} + \delta > \tilde{\phi}_t > \tilde{\phi}_t - \delta > \hat{\phi} - \delta$, which establishes that for every $\delta > 0$, there exists a period $\hat{t}(\delta)$ such that if beliefs $\tilde{\phi}_t$ enter the interval $(\hat{\phi} - \delta, \hat{\phi} + \delta)$ in some period $t \geq \hat{t}(\delta)$, then they never leave this interval.

To complete the proof that beliefs converge, we argue that the sequence of beliefs $\tilde{\phi}_t$ must eventually enter the interval $(\hat{\phi} - \delta, \hat{\phi} + \delta)$. Suppose otherwise. Since the belief sequence is strictly increasing below and strictly decreasing above this interval, and cannot converge to any point outside of this interval, it must therefore jump over $\hat{\phi}$ infinitely often. Hence it must cross $\hat{\phi}$ after period $t(\delta)$ without entering the interval $(\hat{\phi} - \delta, \hat{\phi} + \delta)$. To do so, $|\tilde{\phi}_{t+1} - \tilde{\phi}_t| > 2\delta$, contradicting the fact established above that for all $t \geq \hat{t}(\delta)$, $|\tilde{\phi}_{t+1} - \tilde{\phi}_t| < \delta$.

Hence for any $\delta > 0$, there exists some period $\bar{t}$ such that for all $t > \bar{t}$, $|\tilde{\phi}_t - \tilde{\phi}_\infty| < \delta$. Thus, mean beliefs converge to $\hat{\phi}$. □

Proof of Proposition 6. (1) To establish the bounds in the proposition, observe that Equation 1 implies that $\tilde{b}_t < b_t$ since output increases in ability and the state. We are left to show that $b_t - \tilde{b}_t$ is bounded from above. Since $Q_a < \overline{Q}_a$ and $Q_b > \underline{Q}_b$,

$$\overline{Q}_a \times (\tilde{a} - A) \geq Q(e_t, \tilde{a}, b_t) - Q(e_t, A, b_t) = Q(e_t, \tilde{a}, b_t) - Q(e_t, \tilde{a}, \tilde{b}_t) \geq \underline{Q}_b \times (b_t - \tilde{b}_t)\,.$$

Thus, $b_t - \tilde{b}_t \leq \overline{Q}_a/\underline{Q}_b \times \Delta$.

(2) We apply the Implicit Function Theorem to Equation 1 to show that $\tilde{b}_t$ is differentiable in e_t with derivative

$$\frac{\partial \tilde{b}_t}{\partial e_t} = \frac{Q_e(e_t, A, b_t) - Q_e(e_t, \tilde{a}, \tilde{b}_t)}{Q_{\tilde{b}}(e_t, \tilde{a}, \tilde{b}_t)} = \frac{[Q_e(e_t, A, b_t) - Q_e(e_t, A, \tilde{b}_t)] + [Q_e(e_t, A, \tilde{b}_t) - Q_e(e_t, \tilde{a}, \tilde{b}_t)]}{Q_b(e_t, \tilde{a}, \tilde{b}_t)}\,.$$

(3) As $Q_b > 0$, $\tilde{b}_t < b_t$ and $Q_{eb} > 0$, $\tilde{a} > A$ and $Q_{ea} \leq 0$ (by Assumption 2), it follows that $\frac{\partial \tilde{b}_t}{\partial e_t} > 0$. □

In the proof of Theorem 1, we make use of the following fact.

Lemma 4 (Convergence to the Limiting Action.)**.** *Suppose Assumption 1 holds. There is a $d > 0$ such that for any t and any $\tilde{\phi}_{t-1}$, we have $|e^*(t, \tilde{\phi}_{t-1}) - e^*(\tilde{\phi})| \leq (1/t)d$.*

Proof of Lemma 4. Consider the optimal action as a function of the posterior variance $\sigma_t = \frac{1}{h_\epsilon} + \frac{1}{h_0+(t-1)h_\epsilon}$. As the optimal action e^* is interior and the output Q is differentiable with respect to the action, we have that

$$0 = \mathbb{E}\left[Q_e(e^*(\sigma, \tilde{\phi}), \tilde{a}, \tilde{\phi} + \sigma\eta) \mid \eta \sim \mathcal{N}(0,1)\right].$$

As $|Q_{ee}|$ and $|Q_{eb}|$ are bounded we can apply the dominated convergence theorem to show that the right-hand-side of Equation 11 is differentiable in e and η. We can thus apply the implicit function Theorem to get

$$\frac{\partial e^*(\sigma, \tilde{\phi})}{\partial \sigma} = -\frac{\mathbb{E}\left[Q_{eb}(e^*(\sigma, \tilde{\phi}), \tilde{a}, \tilde{\phi} + \sigma\eta)\eta \mid \eta \sim \mathcal{N}(0,1)\right]}{\mathbb{E}\left[Q_{ee}(e^*(\sigma, \tilde{\phi}), \tilde{a}, \tilde{\phi} + \sigma\eta) \mid \eta \sim \mathcal{N}(0,1)\right]}.$$

Using that $Q_{eb}(e_t, a, b_t)/|Q_{ee}(e_t, a, b_t)| \leq \kappa$ for all e_t, a, b_t, thus,

$$\begin{aligned}
\left|\frac{\partial e^*(\sigma, \tilde{\phi})}{\partial \sigma}\right| &\leq \frac{\mathbb{E}\left[Q_{eb}(e^*(\sigma, \tilde{\phi}), \tilde{a}, \tilde{\phi} + \sigma\eta)|\eta| \mid \eta \sim \mathcal{N}(0,1)\right]}{\mathbb{E}\left[|Q_{ee}(e^*(\sigma, \tilde{\phi}), \tilde{a}, \tilde{\phi} + \sigma\eta)| \mid \eta \sim \mathcal{N}(0,1)\right]} \\
&\leq \frac{\mathbb{E}\left[\kappa|Q_{ee}(e^*(\sigma, \tilde{\phi}), \tilde{a}, \tilde{\phi} + \sigma\eta)||\eta| \mid \eta \sim \mathcal{N}(0,1)\right]}{\mathbb{E}\left[|Q_{ee}(e^*(\sigma, \tilde{\phi}), \tilde{a}, \tilde{\phi} + \sigma\eta)| \mid \eta \sim \mathcal{N}(0,1)\right]} = \kappa\mathbb{E}\left[|\eta| \mid \eta \sim \mathcal{N}(0,1)\right] =: d\,.
\end{aligned}$$

We thus have that

$$\begin{aligned}
|e^*(\sigma_t, \phi) - e^*(\sigma_\infty, \phi)| &= \left|\int_{\sigma_\infty}^{\sigma_t} \frac{\partial e^*(z, \tilde{\phi})}{\partial z}\mathrm{d}z\right| \leq \int_{\sigma_\infty}^{\sigma_t} \left|\frac{\partial e^*(z, \tilde{\phi})}{\partial z}\right| \mathrm{d}z \\
&\leq d(\sigma_t - \sigma_\infty) = d\frac{1}{h_0 + (t-1)h_\epsilon}\,.
\end{aligned}$$

□

Proof of Theorem 1. The result is Theorem 2.1 in Kushner and Yin (2003), page 127 applied to dynamics given in Equation (7)

$$\tilde{\phi}_{t+1} = \tilde{\phi}_t + \gamma_t\left[\tilde{b}_{t+1} - \tilde{\phi}_t\right],$$

which (when constrained to a compact set) is a special case of the dynamics of Equation (1.2) on page 120 in Kushner and Yin. We will first establish that the sequence of mean beliefs is bounded with probability one, and then apply the theorem for the case in which the dynamics are bounded with probability one and there is no constraint set. Second, we verify that the conditions for the theorem apply.

First, by Proposition 6, $b_t - \tilde{b}_t \in [0, \overline{Q}_a/\underline{Q}_b\Delta]$. Since $\tilde{\phi}_0 = \phi_0$ and $\phi_t - \tilde{\phi}_t = (1-\gamma_t)(\phi_{t-1} - \tilde{\phi}_{t-1}) + \gamma_t(b_t - \tilde{b}_t)$ for $t \geq 1$, $\phi_t - \tilde{\phi}_t \in [0, \overline{Q}_a/\underline{Q}_b\Delta]$ is bounded. Hence, because $(\phi_t)_t$ converges to Φ almost surely, $\limsup_{t\to\infty} \tilde{\phi}_t \leq \Phi$ and $\liminf_{t\to\infty} \tilde{\phi}_t \geq \Phi - \overline{Q}_a/\underline{Q}_b\Delta$ are almost surely bounded.

Second, we now check the conditions that are necessary for the Theorem to be applicable. Below we follow the enumeration of assumptions used in Kushner and Yin:

(1.1) The theorem requires $\sum_{t=1}^{\infty} \gamma_t = \infty$ and $\lim_{t\to\infty} \gamma_t = 0$. Immediate from the definition of $\gamma_t = \frac{h_\epsilon}{(t+1)h_\epsilon + h_0}$.

(A2.1) The theorem requires $\sup_t \mathbb{E}\left[|\tilde{b}_{t+1} - \tilde{\phi}_t|^2 \mid \Phi\right] < \infty$, where the expectation condition on the true state and is taken at time 0. We have that

$$\tilde{b}_{t+1} - \tilde{\phi}_t = (\tilde{b}_{t+1} - b_{t+1}) + (b_{t+1} - \phi_t) + (\phi_t - \tilde{\phi}_t)\,.$$

By the norm inequality for the L_2 norm we have that

$$\begin{aligned}\sqrt{\mathbb{E}\left[|\tilde{b}_{t+1} - \tilde{\phi}_t|^2 \mid \Phi\right]} &\leq \sqrt{\mathbb{E}\left[|\tilde{b}_{t+1} - b_{t+1}|^2 \mid \Phi\right]} + \sqrt{\mathbb{E}\left[|b_{t+1} - \phi_t|^2 \mid \Phi\right]} + \sqrt{\mathbb{E}\left[|\phi_t - \tilde{\phi}_t|^2 \mid \Phi\right]} \\ &\leq 2\Delta\overline{Q}_a/\underline{Q}_b + \sqrt{\mathbb{E}\left[|b_{t+1} - \Phi|^2 \mid \Phi\right]} + \sqrt{\mathbb{E}\left[|\phi_t - \Phi|^2 \mid \Phi\right]} \\ &= 2\Delta\overline{Q}_a/\underline{Q}_b + h_\epsilon^{-1/2} + \sqrt{\mathbb{E}\left[|\phi_t - \Phi|^2 \mid \Phi\right]}\,.\end{aligned}$$

Where the second to last line follows as $|\phi_t - \tilde{\phi}_t| \leq \Delta\overline{Q}_a/\underline{Q}_b$ and $|b_t - \tilde{b}_t| \leq \Delta\overline{Q}_a/\underline{Q}_b$. The result follows as the expected L_2 distance of the correct belief of an outsider ϕ_t from the state Φ is monotone decreasing.

(A2.2) The theorem requires the existence of a function g and a sequence of random variables $(\beta_t)_t$ such that $\mathbb{E}\left[\tilde{b}_{t+1} - \tilde{\phi}_t \mid \Phi, \tilde{\phi}_t\right] = g(\tilde{\phi}_t) + \beta_t$. Hence, we define

$$\beta_t = \mathbb{E}\left[\tilde{b}_{t+1} - \tilde{\phi}_t \mid \Phi, \tilde{\phi}_t\right] - g(\tilde{\phi}_t) = g(t, \tilde{\phi}_t) - g(\tilde{\phi}_t)\,,$$

where g is defined in Equation 8.

(A2.3) The theorem requires $g(\tilde{\phi})$ to be continuous in $\tilde{\phi}$. By Lemma 1, $e^*(\tilde{\phi})$ is differentiable and thus continuous in $\tilde{\phi}$. By Proposition 6, $\tilde{b}$ is differentiable in e and thus continuous e. Hence, $g(\tilde{\phi}) = \mathbb{E}\left[\tilde{b}(e^*(\tilde{\phi}), b) - \tilde{\phi} \mid \Phi\right]$ is continuous in $\tilde{\phi}$.

(A2.4) The theorem requires $\sum_{t=1}^{\infty}(\gamma_t)^2 < \infty$. Immediate from the definition of $\gamma_t = \frac{h_\epsilon}{(t+1)h_\epsilon + h_0}$.

(A2.5) The theorem requires $\sum_{t=1}^{\infty}\gamma_t|\beta_t| < \infty$ *w.p.*1. Observe that

$$\beta_t = \mathbb{E}\left[\tilde{b}(e^*(t,\tilde{\phi}),b) - \tilde{b}(e^*(\tilde{\phi}),b) \mid \Phi\right] .$$

To bound β_t we first bound the derivative of the perceived signal with respect to the action $\frac{\partial \tilde{b}}{\partial e}$. The equality below follows from the proof of Proposition 6 and the inequality from Part (v) of Assumption 1

$$\begin{aligned}\frac{\partial \tilde{b}_t}{\partial e_t} &= \frac{[Q_e(e_t,A,b_t) - Q_e(e_t,A,\tilde{b}_t)] + [Q_e(e_t,A,\tilde{b}_t) - Q_e(e_t,\tilde{a},\tilde{b}_t)]}{Q_b(e_t,\tilde{a},\tilde{b}_t)}\\ &\le \frac{\overline{Q}_{eb}|b_t - \tilde{b}_t| + |\underline{Q}_{ea}| \times |A - \tilde{a}|}{\underline{Q}_b} =: d_1 .\end{aligned}$$

Hence,

$$|\beta_t| \le \mathbb{E}\left[|\tilde{b}(e^*(t,\tilde{\phi}),b) - \tilde{b}(e^*(\tilde{\phi}),b)| \mid \Phi\right] \le d_1\,\mathbb{E}\left[|e^*(t,\tilde{\phi}) - e^*(\tilde{\phi})|\right] .$$

By Lemma 4 $|e^*(t,\tilde{\phi}) - e^*(\tilde{\phi})| < (1/t)d_2$ for some $d_2 > 0$. Consequently,

$$\sum_{t=1}^{\infty}\gamma_t|\beta_t| \le d_1 d_2 \sum_{t=1}^{\infty}\frac{h_\epsilon}{(t+1)h_\epsilon + h_0} \times \frac{1}{t} < \infty.$$

(A2.6) The theorem requires the existence of a real valued function f that satisfies $f_{\tilde{\phi}}(\tilde{\phi}) = g(\tilde{\phi})$ and that f is constant on each connected subset of C. We thus define f by

$$f(\tilde{\phi}) = -\int_x^{\tilde{\phi}} g(z)\mathrm{d}z\,,$$

for some x. As in our case C is the union of countably many points, f trivially satisfies all conditions of the theorem. □

Proof of Proposition 7. If the agent uses the action e in every period her long-run mean belief is given by

$$\hat{\phi} = \mathbb{E}_b\left[\tilde{b}(b,e)|\Phi\right].$$

When the agent can change her actions, by Theorem 1 the mean beliefs converge to some limit belief $\tilde{\phi}_\infty \le \Phi - \Delta\,\overline{Q}_b/\underline{Q}_a$. Since by Lemma 3 the action is monotone in the state, $e^*(\tilde{\phi}_\infty) \le$

$e^*(\Phi - \Delta\, \overline{Q}_b/\underline{Q}_a) < e^*(\Phi) \leq e$. By Proposition 6 the perceived signal $\tilde{b}(b, e)$ is strictly increasing in the action e for every realization of b. Thus $\tilde{b}(b, e) < \tilde{b}(b, e^*(\tilde{\phi}_\infty))$, and hence one has

$$\tilde{\phi} = \mathbb{E}_b\ [\tilde{b}(b, e^*(\tilde{\phi}_\infty))|\Phi] < \mathbb{E}_b\ [\tilde{b}(b, e)|\Phi] = \hat{\phi}.$$

Furthermore, by Proposition 6, $\tilde{b}(e, b) < b$ and therefore $\hat{\phi} = \mathbb{E}_b\ [\tilde{b}(b, e)|\Phi] < \Phi = \mathbb{E}[b]$. □

Proof of Proposition 8. To prove the proposition, we first solve for the perceived signal. Second, we show that there is a unique limiting belief. Since the linear-loss specification does not satisfy Assumption 1, we then—in a third setp—use the fact that there is a unique limiting belief to establish directly that mean beliefs converge. Finally, we establish the comparative static in η.

Lemma 5. *In the linear loss case, the perceived signal satisfies*

$$\tilde{b}_t(b_t, e_t) = \begin{cases} b_t - \frac{\Delta}{1+\eta} & \textit{if}\ \ b_t \leq e_t \\ b_t - \frac{\Delta}{1+\eta} - \frac{2\eta}{1+\eta}(b_t - e_t) & \textit{if}\ \ e_t < b_t < e_t + \frac{\Delta}{1-\eta} \\ b_t - \frac{\Delta}{1-\eta} & \textit{if}\ \ b_t \geq e_t + \frac{\Delta}{1-\eta}. \end{cases} \tag{16}$$

Proof of Lemma 5. We distinguish two cases: $e_t < b_t$ and $e_t > b_t$. We have that

$$e_t > b_t \Rightarrow q_t = A + (1+\eta)b_t - \eta\, e_t < A + e_t\,,$$

and

$$e_t < b_t \Rightarrow q_t = A + (1-\eta)b_t + \eta\, e_t > A + e_t\,.$$

Hence, the perceived signal $\tilde{b}_t$ the agent infers from observing the output q_t in period t is given by

$$\tilde{b}_t = \begin{cases} \frac{q_t + \eta e_t - \tilde{a}}{1+\eta} & \text{for } q_t < \tilde{a} + e_t \\ \frac{q_t - \eta e_t - \tilde{a}}{1-\eta} & \text{for } q_t > \tilde{a} + e_t \end{cases}.$$

Because $\tilde{a} > A$, $q_t < A + e_t$ implies $q_t < \tilde{a} + e_t$. Since $q_t = a + b_t - \eta|e_t - b_t|$, we have that in this case

$$\tilde{b}_t = \frac{(A + (1+\eta)b_t - \eta\, e_t) + \eta\, e_t - \tilde{a}}{1+\eta} = b_t - \frac{\tilde{a} - A}{1+\eta}\,.$$

Similar, as $q_t > \tilde{a} + e_t$ implies $q_t > A + e_t$ it follows that

$$\tilde{b}_t = \frac{(A + (1-\eta)b_t + \eta\, e_t) - \eta\, e_t - \tilde{a}}{1-\eta} = b_t - \frac{\tilde{a} - A}{1-\eta}\,.$$

In case $A + e_t < q_t < \tilde{a} + e_t$ we have that

$$\tilde{b}_t = \frac{(A + (1-\eta)b_t + \eta\, e_t) + \eta\, e_t - \tilde{a}}{1+\eta} = b_t - \frac{\tilde{a} - A}{1+\eta} - \frac{2\,\eta}{1+\eta}\,(b_t - e_t)\,. \qquad \square$$

With a linear loss function, it is optimal for the agent to choose the action equal to the expected state of the world, so that $e^*(t, \tilde{\phi}) = \tilde{\phi}$. Hence, in this example

$$g(t, \tilde{\phi}) = \mathbb{E}_b\; [\tilde{b}(b, \tilde{\phi})|\Phi] - \tilde{\phi} = g(\tilde{\phi}).$$

We now show

Lemma 6. *In the linear loss case there is a unique limiting belief satisfying $\tilde{\phi}_\infty < \Phi - \Delta$.*

Proof of Lemma 6. Substituting the expected signal $\tilde{b}_t$ as a function of the mean belief $\tilde{\phi}$ and Φ into the definition of g gives

$$g(\tilde{\phi}) = \mathbb{E}\left[b_t - \frac{\Delta}{1+\eta} - \frac{2\eta}{1+\eta}\min\left\{(b_t - \tilde{\phi})^+, \frac{\Delta}{1-\eta}\right\}\right] - \tilde{\phi}\,,$$

where $b_t \sim \mathcal{N}(\Phi, h_\epsilon^{-1})$. Let us first show that there exists a unique a unique belief $\hat{\phi} \in (\Phi - \frac{\Delta}{1-\eta}, \Phi - \frac{\Delta}{1+\eta})$ such that $g(\hat{\phi}) = 0$. Substituting establishes that $g\left(\Phi - \frac{\Delta}{1-\eta}\right) > 0$ and $g\left(\Phi - \frac{\Delta}{1+\eta}\right) < 0$. The derivative of g is given by

$$g_{\tilde{\phi}}(\tilde{\phi}) = \frac{2\eta}{1+\eta}\mathbb{E}\left[\mathbf{1}_{\{\tilde{\phi} < b_t < \tilde{\phi} + \frac{\Delta}{1-\eta}\}}\right] - 1 < 0\,,$$

where we use that $\eta < 1$ and that the probability that b_t lies in a given interval is less than one. As g is continuous and strictly decreasing, there exists a unique $\hat{\phi} \in (\Phi - \frac{\Delta}{1-\eta}, \Phi - \frac{\Delta}{1+\eta})$.

It remains to show that $g(\Phi - \Delta) < 0$. Evaluating g at $\Phi - \Delta$ gives

$$g(\Phi - \Delta) = \Delta\frac{\eta}{1+\eta}\left\{1 - 2\mathbb{E}\left[\min\{(\frac{b - \Phi}{\Delta} + 1)^+, \frac{1}{1-\eta}\}\right]\right\}.$$

Hence it suffices to establish the expectation in the equation above is greater than 1/2, which holds since

$$\mathbb{E}\left[\min\{(\frac{b-\Phi}{\Delta} + 1)^+, \frac{1}{1-\eta}\}\right] \geq \mathbb{E}\left[\min\{(\frac{b-\Phi}{\Delta} + 1)^+, 1\}\right] > \mathbb{P}\left[\frac{b - \Phi}{\Delta} \geq 0\right] = \frac{1}{2}\,. \qquad \square$$

Lemma 7. *In the linear loss case beliefs converge to the unique limiting belief ϕ_∞.*

Proof of Lemma 7. We beginng by establishing that for almost every sequence of ϵ's there exists an interval I containing $\tilde{\phi}_\infty$ and a t_I such that for all $t > t_I$ the mean belief is in that interval $\tilde{\phi}_t \in I$. First, Lemma 5 shows that for every true signal b_t the perceived signal $\tilde{b}_t$ is increasing in e_t. Second, Lemma 5 implies that

$$b_t - \frac{\Delta}{1-\eta} \le \tilde{b}_t \le b_t - \frac{\Delta}{1+\eta} .$$

Hence, we have that the mean belief of the agent is bounded by the true mean belief minus $\frac{\Delta}{1+\eta}$ from above for all $t \ge 1$;

$$\begin{aligned} \tilde{\phi}_t &= \sum_{s=0}^{t-1} \frac{h_\epsilon \tilde{b}_s + h_0 \phi_0}{h_\epsilon t + h_0} \le \sum_{s=0}^{t-1} \frac{h_\epsilon b_s + h_0 \phi_0}{h_\epsilon t + h_0} - \frac{h_\epsilon t}{h_\epsilon t + h_0} \frac{\Delta}{1+\eta} = \phi_t - \frac{h_\epsilon t}{h_\epsilon t + h_0} \frac{\Delta}{1+\eta} \\ &\le \phi_t - \frac{h_\epsilon}{h_\epsilon + h_0} \frac{\Delta}{1+\eta} , \end{aligned}$$

and by the true mean belief minus $\frac{\Delta}{1-\eta}$ from below

$$\tilde{\phi}_t \ge \phi_t - \frac{h_\epsilon t}{h_\epsilon t + h_0} \frac{\Delta}{1-\eta} = \phi_t - \frac{\Delta}{1-\eta} .$$

By the strong law of large numbers the mean belief of an outside observer converges to the fundamental almost surely $\mathbb{P}[\lim_{t\to\infty} \phi_t = \Phi] = 1$. By convergence of the mean beliefs of an outside observer, it follows that for almost every sequence of $\epsilon's$ the subjective mean belief $\tilde{\phi}$ stays in an interval, i.e. for every $\delta > 0$ there exists a t_δ such that for all $t \ge t_\delta$

$$\tilde{\phi}_t \in \left[\phi - \frac{\Delta}{1-\eta} - \delta, \phi - \frac{h_\epsilon}{h_\epsilon + h_0} \frac{\Delta}{1+\eta} + \delta\right] = I_\delta .$$

Finally, we show that the mean belief $(\tilde{\phi}_t)_{t=0,1,\ldots}$ converges to $\hat{\phi}$ for almost every sequence of ϵ's. Fix a sequence of $\epsilon's$. Let, $[u,v]$ be the smallest interval such that the beliefs $\tilde{\phi}_t$ never leave it. As the mean belief of the agent never leave the interval $[u,v]$ and the actions satisfy $e_t = \tilde{\phi}_t$, her actions never leave the interval $[u,v]$. As $\tilde{b}_t$ is increasing in e_t, a lower bound on the perceived signals is given by

$$\tilde{b}_t \ge b_t - \frac{\Delta}{1+\eta} - \frac{2\eta}{1+\eta} \min\left\{(b_t - u)^+, \frac{\Delta}{1-\eta}\right\} . \tag{17}$$

Note, that the right-hand side is a sequence of identically, independently distributed random variables whose average converges for almost every sequence of $\epsilon's$ to $g(u) + u$. This implies that the belief $\tilde{\phi}$ is bounded from below by $g(u) + u - \delta$ in the long-run with probability one for every $\delta > 0$. Suppose $u < \hat{\phi}$, by definition of $\hat{\phi}$ and the single crossing property of g we have $g(u) > 0$ which contradicts the fact that $[u, v]$ is the minimal interval that the beliefs do not leave. Similar, the perceived signals are upper bounded by

$$\tilde{b}_t \leq b_t - \frac{\Delta}{1+\eta} - \frac{2\eta}{1+\eta} \min\left\{(b_t - v)^+, \frac{\Delta}{1-\eta}\right\} ,$$

and hence the long-run belief is upper bounded by $g(v)+v+\delta$ with probability one for every $\delta > 0$. Suppose $v > \hat{\phi}$. By definition of $\hat{\phi}$ this implies $g(v) < 0$ which contradicts the fact that $[u, v]$ is the minimal interval that the beliefs do not leave. Hence, we have that $u \geq \hat{\phi}$ and $v \leq \hat{\phi}$, which together with $u \leq v$ implies that the beliefs converge to $\hat{\phi}$ for almost every path.

To establish the proposition, we are left to show that $\tilde{\phi}_\infty$ decreases in η. We do so, by showing that g is decreasing in η at $\tilde{\phi}_\infty$. The long run belief $\tilde{\phi}_\infty$ solves the equation

$$0 = g(\tilde{\phi}_\infty) = \Phi - \tilde{\phi}_\infty - \frac{1}{1+\eta} - \frac{2\eta}{1+\eta}\mathbb{E}\left[\min\{(b - \tilde{\phi}_\infty)^+, \frac{\Delta}{1-\eta}\}\right] . \tag{18}$$

Taking the derivative with respect to η yields

$$g_\eta(\tilde{\phi}_\infty) = \frac{\Delta}{(1+\eta^2)} - \frac{2}{(1+\eta)^2}\mathbb{E}\left[\min\{(b - \tilde{\phi}_\infty)^+, \frac{\Delta}{1-\eta}\}\right] - \frac{2\eta}{1+\eta}\frac{\Delta}{(1-\eta)^2}\mathbb{P}\left[b \geq \tilde{\phi}_\infty + \frac{\Delta}{1-\eta}\right] .$$

Solving for the expectation using the Equation 18 and rewritting we have

$$g_\eta(\tilde{\phi}_\infty) = (\Phi - \tilde{\phi}_\infty)\frac{-1}{(1+\eta)\eta} + \frac{\Delta}{(1+\eta)\eta} - \frac{2\eta}{1+\eta}\frac{\Delta}{(1-\eta)^2}\mathbb{P}\left[b \geq \tilde{\phi}_\infty + \frac{\Delta}{1-\eta}\right] .$$

As $(\Phi - \tilde{\phi}_\infty) > \Delta$, we have that $g_\eta(\tilde{\phi}_\infty) < 0$. In the proof of Lemma 6, we established that $g_{\tilde{\phi}}(\tilde{\phi}_\infty < 0)$. Using the implicit function theorem and these facts, it follows that

$$\frac{d\tilde{\phi}_\infty}{d\eta} = -\frac{g_\eta(\tilde{\phi}_\infty)}{g_{\tilde{\phi}}(\tilde{\phi}_\infty)} < 0 . \qquad \square$$

Proof of Proposition 9. We first establish that there is a unique limiting mean belief $\tilde{\phi}_\infty$. The perceived signal $\tilde{b}_t$ satisfies

$$Q(e_t, g(\tilde{a}) + \tilde{b}_t) = Q(e_t, g(A) + b_t).$$

By Assumption 1, Q is strictly increasing in $g(a) + b_t$, and therefore $g(\tilde{a}) + \tilde{b}_t = g(A) + b_t$. Thus, for any sequence of action each $\tilde{b}_t$ is normally distributed with mean $g(A) - g(\tilde{a}) + b_t$ and variance h_ϵ^{-1}. Hence, the mean beliefs converge with probability one to $\tilde{\phi}_\infty = g(A) - g(\tilde{a}) + \Phi$.

Since the aggregator $g(A) + b_t \sim \mathcal{N}(g(A) + \Phi, h_\epsilon^{-1})$ and the agent expects the aggregator $g(\tilde{a}) + \tilde{\phi}_\infty$ to have the same distribution in the limit, for each action e the expected distribution of output is the same, and hence the agent's action converges to the unique optimal limit action. $\square$