Games and Economic Behavior 142 (2023) 315–338

Contents lists available at ScienceDirect

Games and Economic Behavior

journal homepage: www.elsevier.com/locate/geb

A multi-agent model of misspecified learning with overconfidence

Cuimin Ba [a,*], Alice Gindin [b]

[a] University of Pittsburgh, Posvar Hall, Pittsburgh, 15260, PA, USA
[b] Middlebury College, Warner Hall, Middlebury, 05753, VT, USA

A R T I C L E   I N F O

Article history:
Received 28 January 2021
Available online 18 August 2023

Keywords:
Overconfidence
Misspecified learning
Multiple agents
Informational externalities

A B S T R A C T

This paper studies the long-term interaction between two overconfident agents who choose how much effort to exert while learning about their environment. Overconfidence causes agents to underestimate either a common fundamental, such as the underlying quality of their project, or their counterpart's ability, to justify their worse-than-expected performance. We show that in many settings, agents create informational externalities for each other. When informational externalities are positive, the agents' learning processes are mutually-reinforcing: one agent best responding to his own overconfidence causes the other agent to reach a more distorted belief and take more extreme actions, generating a positive feedback loop. The opposite pattern, mutually-limiting learning, arises when informational externalities are negative. We also show that in our multi-agent environment, overconfidence can lead to Pareto improvement in welfare. Finally, we prove that under certain conditions, agents' beliefs and effort choices converge to a steady state that is a Berk-Nash equilibrium.

© 2023 Elsevier Inc. All rights reserved.

1. Introduction

Overconfidence is a widely documented psychological bias. Experimental work demonstrates that individuals often remain overconfident even when confronted with evidence of their bias (Langer and Roth, 1975) by attributing successes to themselves and failures to others or the environment (Miller and Ross, 1975; Ross and Sicoly, 1979; Campbell and Sedikides, 1999).

Both economists and psychologists have explored what happens when a single overconfident agent interacts with their environment. For example, Camerer and Lovallo (1999) find that the overconfidence of entrepreneurs can lead to excessive business entry and losses. Heidhues et al. (2018) discuss how a single overconfident agent learns about and ends up underestimating how talented his team is at performing joint tasks. As a result, the agent exerts sub-optimal effort, leading to a welfare loss. However, when working on a task within a team, all members of the team learn about the environment and adjust their effort simultaneously. In this paper, we consider what happens when *multiple* persistently overconfident agents interact with each other while learning and exerting effort. We show that this can change the learning dynamics as well as the welfare impact of overconfidence. In particular, we find that the direction of this change could depend on whether the object agents learn about is a common fundamental or their teammate's ability.

* Corresponding author.
E-mail addresses: cub8@pitt.edu (C. Ba), agindin@middlebury.edu (A. Gindin).
URLs: https://cuiminba.com (C. Ba), https://alicegindin.com (A. Gindin).

https://doi.org/10.1016/j.geb.2023.08.007
0899-8256/© 2023 Elsevier Inc. All rights reserved.

To ground this idea, consider two engineers who work together on a joint project. Both engineers are overconfident in their research skills. Suppose for now that neither engineer knows the underlying quality of the project assigned by their supervisor and thus both learn about this quality over time by working on the project and receiving evaluations for their periodic joint progress. Both share knowledge and experience gained from reading articles or testing out different methods, so the joint output depends on both engineers' efforts and abilities as well as the project's underlying quality. Our model predicts that the two engineers will both attribute more of the research output to their own ability than is actually warranted, thus underestimating the project's quality. Each engineer's misperception of the project's quality will distort his own choice of effort. Depending on whether the return to effort decreases or increases in the project's quality, an engineer either shirks (because the return to effort on a worse idea is lower) or exerts more effort (to compensate for the project's low quality and churn out a product nevertheless).

Fixing the second engineer's effort, suppose the first engineer's optimal effort increases as his belief about quality becomes lower to compensate for the low quality. We show that the first engineer converges to a low belief about the project's quality, in turn earning lower utility from his excessive effort. If we allow the second engineer to adjust his effort, the first engineer becomes considerably more disappointed by the new output which corresponds to the higher total effort. The extra disappointment exacerbates the drop in his inference about the project quality and encourages him to exert even more effort. If we also assume that efforts are complementary, then this leads to a feedback loop which causes effort to increase and inferences to decrease more than they would if only one engineer adjusted their effort. When the presence of a second engineer who simultaneously adjusts his effort makes the first engineer's beliefs more extreme and vice versa, we call this channel *mutually-reinforcing* learning. However, unlike the single-engineer case, it is now possible that the extra effort leads to higher payoffs for both engineers due to the common-good nature of joint research efforts.

Now let's entertain the other possibility that the underlying quality of the project is common knowledge but the engineers are unsure of their coworker's ability. As in the previous case, suppose that as the engineers' belief in their counterpart's ability drops, their optimal effort increases to compensate. Again, fixing the second engineer's effort, the first engineer exerts excessive effort and has a low belief in the ability of the second engineer. In contrast with the previous case, now if the second engineer can adjust his effort, the first engineer becomes less disappointed by the new output. This makes the first engineer partially correct his underestimation of the second engineer's ability. This is because the higher effort from the second engineer in turn decreases the marginal return to the first engineer's research skill, thus lowering his unrealistically high expectations caused by overconfidence. If in addition efforts are substitutes, this creates a negative feedback loop which we describe as *mutually-limiting* learning.

We formalize this intuition in an infinite-horizon environment, where two agents, i and j, choose how much effort to exert in each period. Each agent's payoff in a given period is the joint output minus their individual cost of effort. The output is determined by both agents' efforts and abilities, as well as some common fundamental (such as the quality of the research idea in the above example). We assume that each agent has a degenerate belief about the value of their own ability and study the case where that point belief is higher than the true value of ability. Moreover, each agent starts with a non-degenerate prior about either the fundamental or the other's ability and updates this belief over time. In each period, each agent chooses a level of effort to myopically maximize his payoff given the effort of the other agent and his belief.

As illustrated in the example, our two-agent model generates two key insights that are not present in the single-agent case. First, we find that agent j's effort provides *an informational externality* by affecting agent i's inference problem over the fundamental through two channels.[1] The first channel is a direct one, a change in agent j's effort changes the signal structure for agent i whenever the marginal product of agent i's unknown variable or his ability is changed by agent j's effort. The second channel takes effect when payoffs exhibit complementarity or substitutability between the two agents' efforts. Complementarity or substitutability of efforts implies that a change in agent j's effort causes a change in agent i's effort, further altering agent i's payoff distributions. Depending on whether the direct and indirect effects are (co)monotonic, informational externalities could be categorized as positive, negative, or ambiguous. We show that the agents' learning processes are mutually-reinforcing when informational externalities are *positive* and mutually-limiting when informational externalities are *negative*. Interestingly, under natural assumptions, informational externalities can only be positive when agents learn about a common fundamental and can only be negative when agents learn about each other's ability.

Second, in a single-agent model, the agent faces an individual decision-making problem and thus misspecification can only result in distorted inferences and suboptimal effort, generating worse payoffs. By contrast, we show that the effect of misspecification is not always negative when multiple agents interact. The idea is very simple: since individual optimization fails to be socially efficient due to payoff externalities, Pareto improvement can be obtained if the efforts are distorted slightly upwards by overconfidence.

We demonstrate the insights above by analyzing agents' long-run beliefs and effort choices which form a Berk-Nash equilibrium (Esponda and Pouzo, 2016). We show that under certain conditions, agents converge to this equilibrium. Our proof augments the contraction argument in Heidhues et al. (2018) to accommodate the additional agent. For convergence to hold, we need agents' informational externalities to be unambiguously positive or negative so that one agent's optimization does not impede the other agent's belief updating and leads to oscillation.

[1] Murooka and Yamamoto (2021) also find that strategic interaction could give rise to informational externalities in misspecified learning problems. We became aware of their paper as we were completing this paper.

Finally, we discuss how our insights extend to settings with underconfident agents. Due to an asymmetry in how the agents draw inferences, an opposite pattern emerges—positive informational externalities lead to mutually-limiting learning while negative informational externalities lead to mutually-reinforcing learning.

1.1. Related literature

This paper builds on the single-agent learning setting in Heidhues et al. (2018). They find that overconfidence leads to distorted beliefs and a reduction in welfare, which are both exacerbated as the agent re-optimizes his effort causing a self-defeating learning pattern. Augmenting their setting, we explore how multiple overconfident agents influence each others' learning process. The presence of multiple agents gives rise to informational externalities and payoff gains relative to the single-agent environment as well as the correctly specified environment. In recent work, Murooka and Yamamoto (2021) identify similar informational externalities in a multi-agent setting. The key difference between the papers is that agents in Murooka and Yamamoto (2021) are overconfident about the same joint ability and only learn about a common fundamental, while in this paper, agents are overconfident in their own personal ability, allowing us to consider cases where agents learn specifically about one another. In their setting, mutually-limiting learning occurs only when the output function maps efforts asymmetrically into outputs; in contrast, mutually-limiting learning arises naturally in our setting when agents learn about each other's ability, even if the output function is symmetric.

There is a growing literature that explores the implications of model misspecification on learning.[2] Esponda and Pouzo (2016) propose the solution concept, Berk-Nash equilibrium, for such games with misspecification. In recent years, there has been substantial progress in showing the convergence of beliefs to the Berk-Nash equilibrium in general environments. Bohren and Hauser (2021) characterize conditions under which correct learning, incorrect learning, or cyclical learning arise in a binary-state learning environment. Heidhues et al. (2021) consider a class of single-agent learning problems where the agent's posterior admits a one-dimensional summary statistic and find conditions under which the agent's belief converges to a point mass with probability 1. Esponda et al. (2021) study a single-agent problem with finite actions, focusing on the dynamics of the frequency of actions and characterizing asymptotic outcomes as the solutions of a differential inclusion; Fudenberg et al. (2021) study a similar setting with finite actions, but obtain characterization based on a stronger assumption of uniform optimality of an action to any long-term beliefs. Frick et al. (2023) propose stronger conditions than Kullback-Leibler divergence dominance that establish convergence. They abstract from specifying actions and preferences and develop a more general setup only in terms of signals and states. This general formulation nests social learning problems as well as the single-agent active learning problem, but does not allow signals to depend endogenously on the subjective belief of a second agent through his action. Therefore, none of their techniques are directly applicable to our multi-agent model that assumes continuous actions and states. The contraction argument used in this paper and Heidhues et al. (2018) rely on structural properties of the payoff functions.

In line with our findings, the literature on overconfidence suggests that this bias can be helpful or detrimental depending on the context. Camerer and Lovallo (1999) find that overconfidence leads to excessive business entry and high rates of new business failure. On the other hand, Gervais and Goldstein (2007) show that overconfidence can improve every team member's welfare in a single-period compensation contract problem.

Finally, this paper relies on the assumption that agents tend to be persistently overconfident, which is well supported by the psychology literature on overconfidence.[3] One illustration of this is the "better-than-average effect", where a large majority of individuals believe themselves to be better than average at a specific skill. For example, Langer and Roth (1975) find that when individuals guess a coin flip outcome, they attribute correct guesses to skill and incorrect guesses to bad luck, and thus believe that they are skilled at predicting coin flips in spite of a 50% success rate. In a job performance setting, Svenson (1981) finds that the vast majority of truck drivers he surveyed believed they were more skilled and safer at driving than the average driver surveyed. Benoît et al. (2015) find that people overplace themselves in quiz performance in a manner that cannot be explained by a model of rational expected utility maximization.

Several papers provide justifications for persistent overconfidence. Anderson et al. (2012) find that overconfident individuals are perceived as more competent which in turn leads to higher social status. This can reinforce feelings of overconfidence, despite contrary evidence. Though this channel provides insight into how individuals might initially develop overconfidence, we do not introduce it into our model. Instead, we assume agents are already overconfident to focus on misperception's effect on the learning process. Ba (2023) shows that overconfidence is persistent even when the decision-maker is aware of potential model misspecification and considers competing models with varying levels of self-perception. Heifetz et al. (2007) demonstrate that overconfidence can arise in an evolutionary environment where agents in a game receive higher payoffs when they maximize objective functions predicated on higher ability than the agent possesses. This

[2] Papers that model the consequences of other types of misspecification include overestimating the informativeness of actions of other agents (Eyster and Rabin, 2010; Bohren, 2016; Gagnon-Bartsch and Rabin, 2017), taste projection (Gagnon-Bartsch, 2017), confirmation bias (Rabin and Schrag, 1999), gambler's fallacy (He, 2022), misspecified beliefs about the type distribution (Frick et al., 2020), and misspecified prior beliefs (Nyarko, 1991; Fudenberg et al., 2017). All such models lead to inefficient long-run actions.

[3] We focus on overconfidence in abilities. There are also studies finding that agents are overconfident in the precision of their beliefs (Moore and Healy, 2008; Moore et al., 2015).

mechanism is closely tied with our welfare results in Section 4.4, where agents' overconfidence can be welfare improving because overconfidence mitigates a public good problem.

The remainder of this paper proceeds as follows. Section 2 describes the model and Section 3 defines the steady state of our learning dynamics—a Berk-Nash Equilibrium adapted to our non-stationary environment. Section 4 contains the main result of the paper, in which we explore the patterns of mutually-reinforcing and mutually-limiting learning in the equilibrium, and analyze the welfare implications. Section 5 shows that in the presence of unambiguously positive or negative informational externalities, the two-agent learning process will converge to the Berk-Nash equilibrium. Section 6 provides extensions including allowing for underconfidence. Section 7 concludes.

2. Multi-agent learning environment

Environment There are two agents, indexed by $i \in I := \{1, 2\}$ and $j \neq i$. In each period $t \in \{1, 2, ...\}$, each agent i simultaneously chooses an effort level e_t^i from a compact set $[\underline{e}, \overline{e}] \subset \mathbb{R}$. Each agent obtains a common output q_t that is determined by their efforts, e_t^i and e_t^j, their individual abilities, a^i and a^j, a common fundamental, ϕ, and random noise. We write this output as $q_t = Q\,(e_t^i, e_t^j, a^i, a^j, \phi) + \epsilon_t$, where ϵ_t is a zero-mean i.i.d. random variable drawn from some continuous distribution with a positive and log-concave density f with full support over $\mathbb{R}$. The output function Q is deterministic and twice continuously differentiable, with its derivatives having polynomial growth in a^i, a^j and ϕ.[4] Additionally, each agent i incurs an individual cost $c(e_t^i) > 0$ that is strictly increasing in his own effort. All past outputs and efforts are publicly observable.

Learning with misspecification We denote the true values of the agents' ability levels by $A^i, A^j \in (\underline{a}, \overline{a})$ and the true common fundamental by $\Phi \in (\underline{\Phi}, \overline{\Phi})$; they are deterministic variables that remain unchanged throughout. Agents are overconfident in their own ability. In particular, agent i believes that his true ability is actually given by $\tilde{a}^i \in (\underline{a}, \overline{a})$ and $\tilde{a}^i > A^i$. Agents' self-perceptions $\tilde{a}^i$ and $\tilde{a}^j$ are common knowledge. Agents realize that their counterpart may be subject to overconfidence but fail to recognize their own overconfidence.[5]

We simultaneously consider two different learning problems in which the agents are either learning about the common fundamental or learning about each other's ability.[6] The psychology literature documents self-serving attribution biases in both environments (Miller and Ross, 1975). For a unified analysis, let ψ^i denote the object that agent i tries to learn about from outputs and Ψ^i denote its true value.

Case 1. Learning about a common fundamental. The agents know each other's true ability A^1, A^2 but are uncertain about the value of the fundamental Φ. We capture this assumption by specifying $\psi^i = \phi$ for all i.

Case 2. Learning about each other. The agents know the true common fundamental but are unsure about the other agent's true ability. Here, agents learn about different objects, $\psi^i = a^j$ for all i and $j \neq i$.

In both cases, the agents are aware of and accept the fact that their self-perception may be different from their counterpart's assessment of them. Due to overconfidence, the agents use a misspecified model to learn about the object ψ^i. Let Π_t^i and π_t^i denote the c.d.f. and p.d.f. of agent i's posterior about the unknown ψ^i at the end of period t. We assume that the prior Π_0^i has finite moments and bounded strictly positive continuous density π_0^i with potentially unbounded support $(\underline{\psi}, \overline{\psi}) \subset \mathbb{R}$, which corresponds to $(\underline{\phi}, \overline{\phi})$ in Case 1 and $(\underline{a}, \overline{a})$ in Case 2. Note that this full-support assumption ensures that any mislearning is a result of overconfidence rather than misspecified priors. We make the following assumptions about the output function.

Assumption 1. *For all i and $j \neq i$: (i) $Q_{a^i} := \partial Q / \partial a^i$ and $Q_\phi := \partial Q / \partial \phi$ are strictly bounded and positive; (ii) the signs of $Q_{e^i a^i} := \partial^2 Q / \partial e^i \partial a^i$ and $Q_{e^i \psi^i} := \partial^2 Q^i / \partial e^i \partial \psi^i$ are different, $Q^i_{e^i \psi^i} \neq 0$, and the signs do not vary with i; (iii) $\forall e^i, e^j$, there always exists $\phi^i \in \left(\underline{\phi}, \overline{\phi}\right)$ and $a^j \in (\underline{a}, \overline{a})$ such that $Q\,\left(e^i, e^j, \tilde{a}^i, A^j, \phi^i\right) = Q\,\left(e^i, e^j, A^i, A^j, \Phi\right) = Q\,\left(e^i, e^j, \tilde{a}^i, a^j, \Phi\right)$.*

The first condition says that a higher ability and a larger fundamental positively influence the common output. The second condition guarantees both agents are optimizing and making inferences in a predictable direction. For example, consider the engineer who underestimates the quality of the project idea due to his overconfidence in his own ability (Case 1). Note that $\psi^i = \phi$. Next, suppose $Q_{e^i \phi} > 0$ and $Q_{e^i a^i} \leq 0$. Then this agent should lower his effort in response to his lowered belief in the fundamental. If instead, both cross derivatives were positive, then more structure would be needed to determine how the agent should best respond.[7] We also assume $Q_{e^i \psi^i} \neq 0$ to rule out the uninteresting case where agents always exert the same effort. The assumption that the signs of cross derivatives do not change with i is without

[4] Log-concavity of f and polynomial growth of Q ensure beliefs and expected output have well defined moments following arbitrary histories.

[5] For justification, consider experiments from Kennedy et al. (2013) where overconfident subjects maintained nearly consistent levels of overconfidence after learning others felt they overestimated their ability. They also remained overconfident after recognizing their peers were overconfident.

[6] We do not consider the case where agents learn about both the common fundamental and each other's ability simultaneously, since the agents' subjective models would be under-identified.

[7] Heidhues et al. (2018) also make this assumption. We discuss the role of this assumption in more detail at the end of Section 4.3.

loss of generality.[8] If Q is symmetric for agents i and j, then this assumption is automatically satisfied. The third condition guarantees that for any fixed action profile, agent i can always perfectly justify the distribution of the outputs by attaching probability 1 to an incorrect fundamental value or teammate ability level.

Actions The agents are myopic and maximize their payoff in the current period. Since the agents' self-perceptions as well as the history of payoffs and efforts are public information, their posteriors $\left\{\pi^1_{t-1}, \pi^2_{t-1}\right\}$ are common knowledge. Therefore, the agents can use iterated deletion of dominated strategies to determine their play. The following regularity assumption ensures that in each period, the induced game is dominance solvable (See Lemma 4).

Assumption 2. *For all i and $j \neq i$: (i) the return to effort is diminishing, $Q_{e^ie^i} - c''(e^i) < 0$ for all e^i; (ii) $Q_{e^i}\left(\underline{e}, e^j, \tilde{a}^i, a^j, \phi\right) - c'(\underline{e}) > 0 > Q_{e^i}\left(\overline{e}, e^j, \tilde{a}^i, a^j, \phi\right) - c'(\overline{e})$ for all e^j, a^j, ϕ; (iii) the diminishing return dominates any complementarity or substitutability between efforts, $|Q_{e^ie^i} - c''(e^i)| > |Q_{e^ie^j}|$, with $Q_{e^ie^j} \geq 0$ for all values or $Q_{e^ie^j} \leq 0$ for all values.*

In each period, agent i must form some belief over what action player j is going to play so he can maximize his own stage payoff. With dominance solvability it is clear how agent i forms his conjecture about agent j's action; agent i employs iterated deletion of dominated strategies until he arrives at the uniquely rationalizable action profile and uses that to inform his play. All this requires is Assumption 2 and the common knowledge that both agents use their subjective models to make decisions. Moreover, this is equivalent to assuming that agents play a Nash equilibrium each period, which boils down to the following restriction: agents choose efforts $\left\{e^1_t, e^2_t\right\}$, in which e^i_t is myopically optimal against e^j_t given belief π^i_{t-1}. However, if we were to simply impose that the agents play the stage game Nash Equilibrium in each period, it would be unclear how each player formed the correct conjecture about what action the other player was going to take.

Contrasting with the assumption from Esponda and Pouzo (2016), where players assume they are in a stationary environment, we model players to be a little more sophisticated so that they understand the underlying distribution of outputs depends on their counterpart's actions and thus varies over time. The set of Berk-Nash equilibria we identify in Section 3, nevertheless, is the same as those identified by Esponda and Pouzo (2016) if players start with conjectures on each other's actions that are correct in equilibrium. This is because the Berk-Nash equilibrium is a steady-state concept. By assuming common knowledge of non-stationarity, we have a more natural interpretation and a clearer picture of how agents form beliefs dynamically—it is hard to isolate how inferences are affected by overconfidence over time if agents are also misspecified about the game structure.

2.1. Examples

We present a few parametric examples that satisfy the assumptions in the paper. We will revisit the first two to illustrate our results in later sections.

Example 1. Consider two engineers who work on a joint project for a large firm. Each team member's payoff depends on a common fundamental representing the quality of the project idea. They are both overconfident in their research ability and periodically split a bonus reliant on the project's profitability to the firm. Each engineer also experiences a convex cost to exerting effort. For a concrete functional form, let $Q\left(e^i, e^j, a^i, a^j, \phi\right) = \phi\left(e^i + e^j + e^ie^j + a^i + a^j\right)$ and $c^i(e^i) = \frac{1}{2}\kappa(e^i)^2$, where $\kappa \geq \overline{\phi}$ and $\phi > 0$. The engineers' expected bonus, Q, is increasing in their effort and ability as well as the quality of the project. In this example, the agents learn about the fundamental, ϕ, which governs the productivity of the project ($\psi^i = \phi$). Notice that engineer i's effort and the fundamental are complements—a higher belief in the fundamental motivates a greater input of effort.[9] The efforts of the engineers are also complements; as they share knowledge and experience gained from reading articles or testing out different methods their marginal productivity improves.

Example 2. Consider a modified version of the teamwork setting in the previous example, but now the engineers learn about each other ($\psi^i = a^j$). The engineers still experience the same convex effort cost. In each period, the engineers split a bonus and each gets $Q(e^i, e^j, a^i, a^j, \phi) = \log(e^ia^i + e^ja^j + \phi)$. In this example, since the returns for the project are concave in the sum of the agent's effort and abilities, agent i's effort and his coworker's unknown ability are substitutes—a higher belief in his coworker induces lower effort from agent i. In addition, the efforts of the agents are substitutes.

Example 3. The legislature passes a law that must be implemented by two federal agencies who work together to create a series of rules that enforce different aspects of the law.[10] The two agencies learn about the underlying quality of the law,

[8] If $sgn(Q_{e^ia^i}) \neq sgn(Q_{e^ja^j})$, we can change the orientation of e^j and then our framework still applies.

[9] This differs from the motivating example in the introduction where we assumed the common fundamental and effort were substitutes. Assuming complementarity here allows us to use a simpler functional form.

[10] For instance consider two US agencies: the SEC and CFTC. Both agencies are tasked with regulating financial products. In the case of regulating financial swaps, the SEC writes rules pertaining to specific securities-based swaps while the CFTC writes the rules for all other types. The agencies have similar policy goals and often share information in order to create better and more consistent rules (Bils, 2020).

ϕ, while dedicating effort e^i towards writing each rule. Each agency is overconfident in its ability, a^i, to write good rules. In each period, the output is given by $Q(e^i, e^j, a^i, a^j, \phi) = a^i + a^j + \phi - L(\phi - e^i) - L(\phi - e^j) + \lambda e^i e^j$ and $c^i(e^i) = \frac{1}{2}\kappa\left(e^i\right)^2$, where $\lambda > 0$, $\kappa > \lambda$, and L is a positive, convex loss function with $|L'| < 1$. The agencies would like to match the time and resources they put towards writing rules to the underlying quality of the law, which is captured by the loss function L. The agency will pass better rules if they have a higher capacity to write quality rules (higher a^i) as well as if the underlying legislation is of high quality (higher ϕ), and will put more effort into writing rules (higher e^i) as the other agency works harder (higher e^j).

3. Steady state

Following the definition developed in Esponda and Pouzo (2016), we now define the Berk-Nash equilibrium for both Case 1 and Case 2. As will be shown in Section 5, the action process described earlier almost surely converges to a steady state that constitutes such an equilibrium. An equilibrium consists of strategies that are optimal given equilibrium beliefs which minimize the Kullback-Leibler (henceforth KL) divergence.[11]

Definition 1. A strategy profile $\boldsymbol{e} \in [\underline{e}, \overline{e}]^2$ is a pure-strategy Berk-Nash equilibrium if there exists a probability distribution $\pi^i \in \Delta\left(\underline{\psi}, \overline{\psi}\right)$ for each i such that:

(i) Effort e^i is optimal given π^i and e^j. That is, in Case 1,

$$e^i \in \arg\max_{e^{i\prime} \in [\underline{e}, \overline{e}]} \mathbb{E}_{\pi^i}\left[Q\left(e^{i\prime}, e^j, \tilde{a}^i, A^j, \psi^i\right)\right] - c(e^{i\prime}), \tag{1}$$

and in Case 2,

$$e^i \in \arg\max_{e^{i\prime} \in [\underline{e}, \overline{e}]} \mathbb{E}_{\pi^i}\left[Q\left(e^{i\prime}, e^j, \tilde{a}^i, \psi^i, \Phi\right)\right] - c(e^{i\prime}), \tag{2}$$

(ii) For all $\psi^i \in \operatorname{supp} \pi^i$, we have $\psi^i \in \arg\min_{\psi^{i\prime} \in \left(\underline{\psi}, \overline{\psi}\right)} K^i\left(\boldsymbol{e}, \psi^{i\prime}\right)$, where in Case 1,

$$K^i\left(\boldsymbol{e}, \psi^{i\prime}\right) := \mathbb{E}\left[\log \frac{f\left(q - Q\left(\boldsymbol{e}, A^i, A^j, \Phi\right)\right)}{f\left(q - Q\left(\boldsymbol{e}, \tilde{a}^i, A^j, \psi^{i\prime}\right)\right)}\right], \tag{3}$$

and in Case 2,

$$K^i\left(\boldsymbol{e}, \psi^{i\prime}\right) := \mathbb{E}\left[\log \frac{f\left(q - Q\left(\boldsymbol{e}, A^i, A^j, \Phi\right)\right)}{f\left(q - Q\left(\boldsymbol{e}, \tilde{a}^i, \psi^{i\prime}, \Phi\right)\right)}\right]. \tag{4}$$

The expectation is taken with respect to the true distribution of $q = Q\left(\boldsymbol{e}, A^i, A^j, \Phi\right) + \epsilon$.

By strict concavity of the payoff function, we rule out mixed strategy equilibria. It is straightforward to see that if the learning process ever converges, the steady state must be a pure Berk-Nash equilibrium: intuitively, if efforts converge, then in the long run, an agent must converge to the beliefs that best fit the data. These will be the beliefs that minimize the KL divergence. Meanwhile, in the limit, the efforts must best respond to the limit beliefs as well as the opponent's action. To characterize the equilibrium, let $\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right) := \left(e^{*i}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right), e^{*j}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right)\right)$ denote the myopically optimal action profile where each agent i assigns probability 1 to ψ^i. This is the Nash equilibrium of a one-shot game when we fix the beliefs in the unknown variable to a Dirac measure at $\boldsymbol{\psi}$. It is straightforward to show its existence and uniqueness.

Lemma 1. *Under Assumption 2, a unique action profile $\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right)$ exists, $\forall \tilde{\boldsymbol{a}}, \boldsymbol{\psi}$.*

Next, we define the *gap* function for each player to capture the discrepancy between the actual average output and agent i's expected average output. For both $i \in I$, define $g^i : [\underline{e}, \overline{e}]^2 \times (\underline{\psi}, \overline{\psi}) \to \mathbb{R}$ such that in Case 1,

$$g^i\left(\boldsymbol{e}, \psi^i\right) := Q\left(\boldsymbol{e}, A^i, A^j, \Phi\right) - Q\left(\boldsymbol{e}, \tilde{a}^i, A^j, \psi^i\right), \tag{5}$$

and in Case 2,

[11] Kullback-Leibler divergence, also known as relative entropy, is a common measure of distance between two distributions. By Gibb's inequality, the Kullback-Leibler divergence is weakly positive and equal to zero if and only if the two distributions being compared coincide almost everywhere.

$$g^i\left(\boldsymbol{e},\psi^i\right) := Q\left(\boldsymbol{e},A^i,A^j,\Phi\right) - Q\left(\boldsymbol{e},\tilde{a}^i,\psi^i,\Phi\right). \tag{6}$$

We will call $\boldsymbol{g}(\boldsymbol{e},\boldsymbol{\psi}) := \left(g^i\left(\boldsymbol{e},\psi^i\right), g^j\left(\boldsymbol{e},\psi^j\right)\right) = 0$ the *no-gap condition*. Intuitively, fixing efforts $\boldsymbol{e}$, the solution to the no-gap condition, $\boldsymbol{\psi}$, is the value of the unknown variable that agent i finds most likely since it perfectly matches the distribution of outputs. The no-gap condition exactly characterizes the point where the weighted Kullback-Leibler divergence is minimized to 0 for both agents. We thus obtain the following lemma.

Lemma 2. *Under Assumptions 1 and 2, there exists at least one Berk-Nash equilibrium. Moreover, each equilibrium $\boldsymbol{e}_\infty$ is associated with a supporting belief that is a Dirac measure at $\boldsymbol{\psi}_\infty$, which satisfies the following:*

(i) Optimality: $\boldsymbol{e}_\infty = \boldsymbol{e}^*\left(\tilde{\boldsymbol{a}},\boldsymbol{\psi}_\infty\right)$.
(ii) Consistency: $\boldsymbol{g}\left(\boldsymbol{e}_\infty,\boldsymbol{\psi}_\infty\right) = 0$.

To streamline exposition, we sometimes denote the equilibrium beliefs and efforts as functions of the self-perception levels, such as $\boldsymbol{\psi}_\infty\left(\tilde{\boldsymbol{a}}\right)$ and $\boldsymbol{e}_\infty\left(\tilde{\boldsymbol{a}}\right)$.

In order to establish global convergence, we follow Heidhues et al. (2018) who assume there is a unique Berk-Nash equilibrium. Note that the uniqueness of $\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}},\boldsymbol{\psi}\right)$ is insufficient since there could be multiple equilibrium beliefs, supporting different optimal action profiles.

Assumption 3. *There exists a unique Berk-Nash equilibrium.*

Lemma 3 provides sufficient conditions on the fundamentals of the stage game for Assumption 3. Specifically, uniqueness is guaranteed if agents are not too misspecified. Section 6 considers how our insights extend to scenarios where Assumption 3 fails.

Lemma 3. *Suppose Assumptions 1 and 2 hold. There exist $\Delta^1, \Delta^2 > 0$, such that whenever $\left|\tilde{a}^i - A^i\right| < \Delta^i, \forall i = 1,2$, there is a unique Berk-Nash equilibrium.*

4. Main results

In this section, we explore the properties of the steady state, in particular how the discrepancy between $\boldsymbol{\psi}_\infty$ and the true value of the unknown variable Ψ varies in settings with or without strategic interaction between agents. We first define the concept of informational externalities, then demonstrate how they can cause different learning patterns in Cases 1 and 2. Finally, we examine the welfare implications.

4.1. Informational externality

Payoff externalities are a well-known concept which describes the direct influence of agent j' actions on agent i's utility, often discussed in the context of a common good problem. We find that agent j's action may also have an impact on agent i's beliefs, formalized below as *informational externalities*. Notably, the notion of informational externality is distinct from informative actions in social learning environments. The agents do not obtain any additional information from each other's effort choices; instead, the agents face a signal structure that varies as efforts change.

Definition 2. We say agent j's action creates an *informational externality* for agent i when the solution ψ^i to the no-gap condition $g^i(\boldsymbol{e},\psi^i) = 0$ depends on e^j, or equivalently, at least one of $Q_{e^j a^i}$, $Q_{e^j \psi^i}$, or $Q_{e^i e^j}$ is nonzero.

Agent j's action affects agent i's inference about ψ^i when the conditions in Definition 2 are met. The informational externality works both directly and indirectly. To understand the direct channel, first notice that a different e^j changes the underlying distribution of the outputs, consequently distorting agent i's belief updating process. This may push agent i's belief upwards or downwards, which critically depends on the signs of $Q_{e^j a^i}$ and $Q_{e^j \psi^i}$. Meanwhile, the indirect effect operates through agent i's optimization process. When $Q_{e^i e^j} \neq 0$, a different e^j changes the marginal product of e^i and thus the optimal choice of the latter. This feeds back to the direct channel by once again changing the underlying output distribution. Informational externalities, just like payoff externalities, can be categorized as positive or negative based on the signs of the aforementioned cross derivatives.

Definition 3. The informational externality of agent j's action over i is *positive* if $Q_{e^i e^j} \geq 0$, $\mathrm{sgn}(Q_{e^i \psi^i}) = \mathrm{sgn}(Q_{e^j \psi^i})$, and $\mathrm{sgn}(Q_{e^i a^i}) = \mathrm{sgn}(Q_{e^j a^i})$; it is *negative* if $Q_{e^i e^j} \leq 0$, $\mathrm{sgn}(Q_{e^i \psi^i}) \neq \mathrm{sgn}(Q_{e^j \psi^i})$, and $\mathrm{sgn}(Q_{e^i a^i}) \neq \mathrm{sgn}(Q_{e^j a^i})$; otherwise, it is *neither positive or negative.*

The direction of an informational externality captures whether e^j and e^i exert the same effect over i's inference about ψ^i. It is easier to understand this technical definition in terms of complementarity or substitutability between efforts and through the no-gap condition. When agent j's action creates a positive informational externality, the agents' efforts are complements. Furthermore, the parameter value ψ^i that solves $g^i\left(\boldsymbol{e},\psi^i\right)=0$ is either both increasing or both decreasing in e^i and e^j. In sum, these observations imply that the efforts affect agent i's belief in the same direction and are mutually-reinforcing. By contrast, the negative informational externality of agent j amounts to substitutable efforts with opposite influences on agent i's belief.

Definition 3 distinguishes the informational externality imposed by agent j's action over i from the one imposed by agent i's action over j. Nevertheless, with the assumption that both agents react to overconfidence in the same direction (Assumption 1), the informational externalities between the two agents must point in the same direction. We now show that efforts create positive informational externalities in Example 1 and negative informational externalities in Example 2.

Example 1 *(cont.).* Recall that engineers are learning about the productivity of the project ($\psi^i=\phi$), with an output function given by $Q\left(e^i,e^j,a^i,a^j,\phi\right)=\phi(e^i+e^j+e^ie^j+a^i+a^j)$. Simple calculations establish that $Q_{e^ie^j}>0$, $Q_{e^i\psi^i}=Q_{e^i\phi}>0$, $Q_{e^j\psi^i}=Q_{e^j\phi}>0$, and $Q_{e^ia^i}=Q_{e^ja^i}=0$. Therefore, the engineers exert positive informational externalities over one another.

Example 2 *(cont.).* In contrast, engineers in this example learn about one another's ability ($\psi^i=a^j$), with an output function given by $Q\left(e^i,e^j,a^i,a^j,\phi\right)=\log(e^ia^i+e^ja^j+\phi)$. Simple calculations tell us that $Q_{e^ie^j}<0$, $Q_{e^i\psi^i}=Q_{e^ia^j}<0$, $Q_{e^j\psi^i}=Q_{e^ja^j}>0$, $Q_{e^ia^i}>0$, $Q_{e^ja^i}<0$. Therefore, the engineers exert negative informational externalities over one another.

More generally, in Case 1, where agents learn about a common fundamental, they impose positive informational externalities *if and only if* efforts are complements; in addition, the agents *cannot* exhibit negative informational externalities on one another. Since overconfidence distorts agent i's and j's optimal effort choices in the same direction, the sign of $Q_{e^i\phi}$ does not vary with i. It then follows from $\psi^i=\phi$ that the sign of $Q_{e^i\psi^i}$ must be aligned with $Q_{e^j\psi^i}$, ruling out negative informational externalities.

In Case 2, however, we observe the opposite. When agents learn about one another, agents impose negative informational externalities *if and only if* efforts are substitutes; in addition, agents *cannot* exhibit positive informational externalities. Since we require overconfidence and the resulting underestimation of the unknown to affect agent j's optimal effort in the same direction for both agents, $Q_{e^ja^j}$ and $Q_{e^ia^i}$ have the same sign. By Assumption 1, the sign of $Q_{e^ja^j}$ must disagree with the sign of $Q_{e^j\psi^j}$, or equivalently, the sign of $Q_{e^ja^i}$.[12] Therefore, $sgn(Q_{e^ia^i})\neq sgn(Q_{e^ja^i})$, which rules out positive informational externalities.

4.2. Mutually-reinforcing learning

We consider the following question: how do the agents mutually influence their beliefs in the unknown variable through interactive learning and optimization? Heidhues et al. (2018) show that a single agent's learning is self-defeating in the sense that allowing an overconfident agent to adjust his own actions results in a more extreme belief about the unknown variable. This encourages more extreme actions and leads to even lower outputs. We first show that the presence of a second actively-optimizing agent reinforces this pattern when there are positive informational externalities.

Consider a special learning environment in which we fix agent j's action at $e_S^j:=e_\infty^j(\boldsymbol{A})$, i.e. the effort choice that best responds to correct self-perceptions, but allow agent i to optimize his action in each period. Maintaining the assumption that there exists a unique steady-state, denote the steady-state inferences as $\boldsymbol{\psi}_{\boldsymbol{S}}=\left(\psi_S^i,\psi_S^j\right)$ and actions as $\boldsymbol{e}_{\boldsymbol{S}}=\left(e_S^i,e_S^j\right)$.[13] We now compare this steady state with $\left(\boldsymbol{e}_\infty,\boldsymbol{\psi}_\infty\right)$, i.e. the steady state when we allow both agents to adjust actions. The following proposition shows that when agents create positive informational externalities, the steady-state underestimation always becomes more severe as more agents actively participate in action optimization. We describe such learning processes as *mutually-reinforcing*.

Proposition 1. *Suppose Assumptions 1 to 3 hold and agent j's action has a positive informational externality over agent i. Then both agents' underestimation of their unknown variables ψ^i gets **reinforced** when agent j is free to optimize compared to when agent j's action is fixed at the level e_S^j, i.e. $\boldsymbol{\psi}_\infty\left(\tilde{\boldsymbol{a}}\right)<\boldsymbol{\psi}_{\boldsymbol{S}}<\boldsymbol{\Phi}$.*

The message conveyed by Proposition 1 is twofold. First, a self-defeating pattern emerges. Agent j underestimates the unknown variable more when he is allowed to optimize, i.e. $\psi_\infty^j\left(\tilde{\boldsymbol{a}}\right)<\psi_S^j<\Psi^j$. More importantly, a mutually-reinforcing pattern can be observed by noting that $\psi_\infty^i\left(\tilde{\boldsymbol{a}}\right)<\psi_S^i<\Psi^i$, which means agent i's inference also becomes more extreme

[12] In Section 4.3, we discuss the role of this assumption and whether relaxing it could lead to different predictions.

[13] We can ensure the uniqueness of the steady state with similar sufficient conditions to those in Lemma 3.

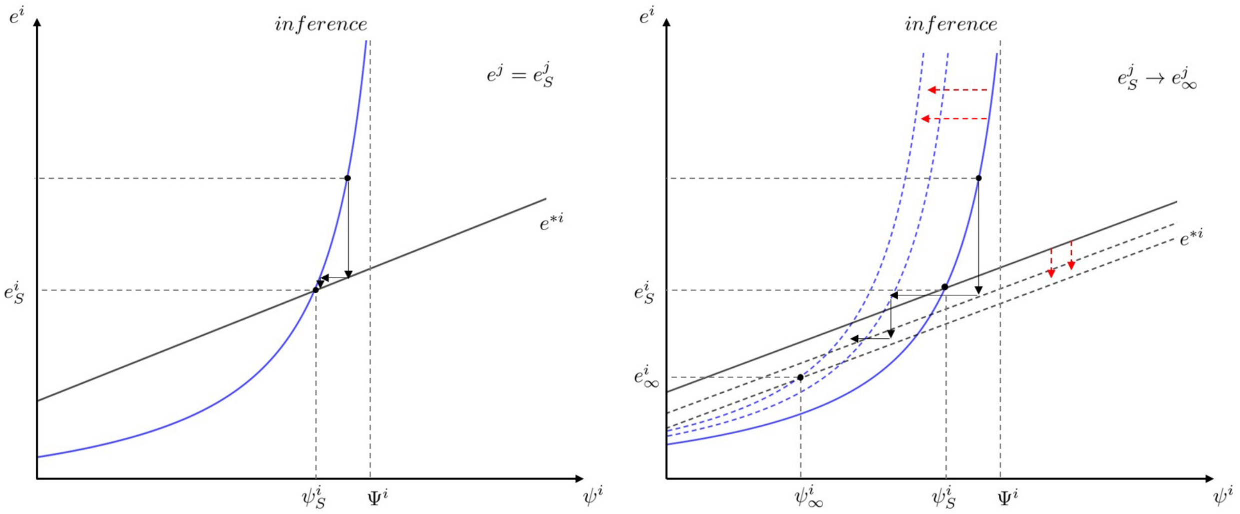

Fig. 1. Mutually-reinforcing learning. The left panel shows how allowing agent i to change her action induces a lower inference for agent i, with e^j fixed at e^j_S, and the right panel shows how agent i gets an even lower inference when agent j is also allowed to revise actions. The black straight line depicts the optimal action given a belief ψ^i, while the blue curve describes the belief ψ^i derived from the no-gap condition $g^i = 0$ with e^i and e^j given. In the right panel, the two curves shift in the directions of the red arrows, capturing the effect of a changing e^j. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

when agent j can freely change actions. Hence, even if only one agent was to fix his action at the level that best responds to correct self-perceptions, *both* agents would understand the environment better.[14]

To illustrate the mechanism, we describe the learning dynamics heuristically for Case 1 with complementary efforts. For this purpose, we assume $Q_{e^i\psi^i} > 0$ (that is, $Q_{e^i\phi} > 0$) and $Q_{e^i a^i} \leq 0$ for both i in the illustration. Note that Example 1 satisfies this assumption. When agent i holds a degenerate belief at ψ^i and his coworker's effort is fixed at e^j, agent i optimally chooses an optimal level of effort e^{*i} which increases in both ψ^i and e^j. We can plot this *optimal action* function in the $e^i - \psi^i$ domain (see Fig. 1). In addition, solving the no-gap equation $g^i(\boldsymbol{e}, \psi^i) = 0$ yields the *inference formation* function, which is the increasing blue curve we plot in the figure.

Fixing agent j's effort at e^j_S, as agent i starts to make inferences and optimize from the point plotted in the left panel of Fig. 1, agent i scales down his effort because he underestimates the common fundamental ϕ. This decrease in effort results in lower output in the following period, resulting in an even lower belief from agent i. Eventually, agent i reaches the belief ψ^i_S, lower than the belief he started with. Suppose now agent j is also given the chance to optimize; the dynamics change dramatically. Since both agents have the tendency to scale down their effort and the output function admits complementarity between efforts. Agent i now exerts lower effort than he did when his coworker was constrained to play a fixed higher action. In the right panel of Fig. 1, this is captured by a downward shift of the optimal action curve. Next, the decrease in agent j's effort leads to a larger negative gap between the true and expected outputs. To see this, note that positive informational externalities and our assumptions imply

$$\frac{\partial g^i(\boldsymbol{e}, \psi^i)}{\partial e^j} = Q_{e^j}(\boldsymbol{e}, A^i, A^j, \Phi) - Q_{e^j}(\boldsymbol{e}, \tilde{a}^i, A^j, \psi^i) > 0. \tag{7}$$

Hence, the evaluation of the common fundamental by agent i experiences a larger decline, i.e. the belief formation curve for agent i shifts to the left. To understand this, note that the common fundamental ϕ and the teammate's effort e^j are complements, so the marginal return on ϕ decreases in response to a lower e^j (similarly the marginal return on his own ability, a^i, weakly increases). Hence, agent i believes the fundamental has to be much worse to justify his own underperformance. The same process repeats until both agents reach the steady state with action e^i_∞ and belief ψ^i_∞. These steady-state values may be much more extreme than the initial action and belief values of e^i_S and ψ^i_S.

We now examine mutually-reinforcing learning from another perspective. Given that the presence of an actively-optimizing second agent reinforces one agent's mislearning, it is reasonable to expect that mislearning becomes more severe when the second agent is more biased. Proposition 2 confirms this intuition. As one or both agents become more overcon-

[14] We choose $e^j_S = e^j_\infty(\boldsymbol{A})$ for ease of interpretation, but it is worth noting that Proposition 1 holds for more general choices of e^j_S. Specifically, if $Q_{e^i\psi^i} > 0$, then it holds for all $e^j_S > e^j_\infty(\tilde{\boldsymbol{a}})$; if instead $Q_{e^i\psi^i} < 0$, then it holds for all $e^j_S < e^j_\infty(\tilde{\boldsymbol{a}})$. In either scenario, the condition ensures that e^j_S moves away from $e^j_\infty(\tilde{\boldsymbol{a}})$ in the direction of $e^j_\infty(\boldsymbol{A})$.

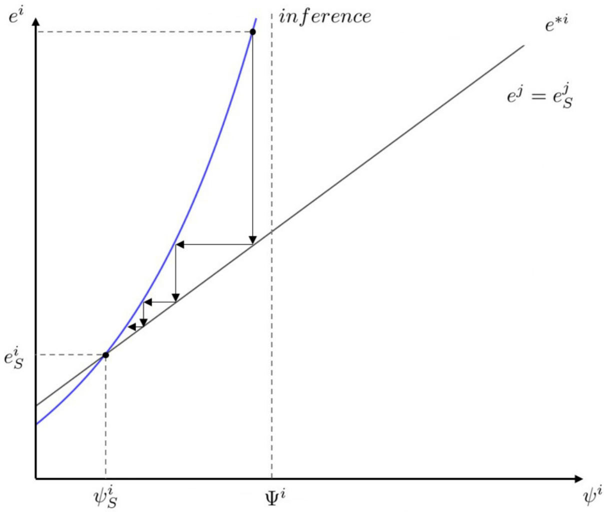

Fig. 2. Mutually-limiting learning. The left panel shows the steady state with e^j fixed at e^j_S, and the right panel shows how agent i gets a relatively higher inference when agent j is also allowed to revise actions.

fident, i.e. increase their self-perceptions, positive informational externalities imply that each agent has a worse evaluation of the unknown variable.

Proposition 2. *Suppose Assumptions 1 to 3 hold and there exist positive informational externalities between the agents. When any of the agents is more overconfident, both agents' underestimation of their unknown variable ψ^i is more severe. That is, let $\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{a}}' \in (\underline{a}, \overline{a})^2$ and $\tilde{\boldsymbol{a}} > \tilde{\boldsymbol{a}}' > \boldsymbol{A}$, then $\boldsymbol{\psi}_\infty(\tilde{\boldsymbol{a}}) < \boldsymbol{\psi}_\infty(\tilde{\boldsymbol{a}}') < \boldsymbol{\Psi}$.*

The pattern of mutually-reinforcing learning generates novel policy implications. First, in many interesting economic problems where there are often multiple agents interacting with each other, even a slight bias of overconfidence can be magnified to induce a nontrivial discrepancy between agents' beliefs and the truth, thereby driving agents' actions far away from optimal. Second, effective intervention can involve restricting the action choices of certain agents. Specifically, mutually-reinforcing learning implies that even intervention that only targets a subset of agents can have effects on every agent involved.

4.3. Mutually-limiting learning

When informational externalities are negative, agents' learning processes become *mutually-limiting*; allowing another agent to freely optimize will make the original agent's belief distortion less severe. Similarly, an increase in a second agent's overconfidence will cause the first agent's inferences to be closer to the true value of the unknown. In addition, we know from Section 4.1 that mutually-limiting learning arises when agents learn about each other (Case 2) and their efforts form substitutes.

Proposition 3. *Suppose agent j's action has a negative informational externality over agent i. The following are true:*

(i) Agent i's underestimation of the unknown variable is less severe when agent j is free to optimize than when agent j's action is fixed at e^j_S, i.e. $\psi^i_S < \psi^i_\infty(\tilde{\boldsymbol{a}}) < \Psi$.
(ii) As agent j becomes more overconfident, agent i's underestimation is smaller. That is, for any $\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{a}}' > \boldsymbol{A}$ such that $\tilde{a}^j > \tilde{a}'^j$ and $\tilde{a}^i = \tilde{a}'^i$, it is true that $\psi^i_\infty(\tilde{\boldsymbol{a}}') < \psi^i_\infty(\tilde{\boldsymbol{a}}) < \Psi$.

We illustrate mutually-limiting learning in the context of Case 2 with substitutable efforts. To better compare with mutually-reinforcing learning, we continue to assume $Q_{e^i\psi^i} > 0$ (that is, $Q_{e^ia^j} > 0$) for both i. As before, both the optimal action curve and the inference formation curve are upward-sloping (see Fig. 2). However, when we allow agent j to freely optimize, the decrease in his effort moves the optimal effort curve upwards due to substitutability. Meanwhile, we now have a larger positive gap between the true and expected outputs since negative informational externalities and our assumptions imply

$$\frac{\partial g^i(\boldsymbol{e}, \psi^i)}{\partial e^j} = Q_{e^j}(\boldsymbol{e}, A^i, A^j, \Phi) - Q_{e^j}(\boldsymbol{e}, \tilde{a}^i, A^j, \psi^i) < 0. \tag{8}$$

Therefore, for each fixed e^i, agent i now has a more favorable view of his teammate's ability, i.e. the belief formation curve shifts to the right. This process eventually leads to action e^i_∞ and belief ψ^i_∞, which are less extreme than their single-agent steady state counterparts.

In sum, our analysis shows that mutually-reinforcing learning may occur when agents learn about a common fundamental, but not when agents learn about each other's ability. Conversely, mutually-limiting learning may arise in the latter case, but not in the former. How much does this observation hinge on the parametric conditions in Assumption 1, namely, $Q_{e^i a^i}$ has the same sign as $Q_{e^j a^j}$ but the opposite sign as $Q_{e^i \psi^i}$? The former assumption is a mere normalization, but the latter is critical to our analysis. To see this, consider an alternative environment where $sgn(Q_{e^i a^i}) = sgn(Q_{e^i \psi^i}) > 0$. In this case, the optimal effort and inference formation curves could be non-monotone, and thus the effect of e^j on i's inference could also be non-monotone. Without imposing further structure over the functional form of Q, we cannot directly infer whether positive or negative informational externalities lead to mutually-reinforcing or mutually-limiting learning.

4.4. Welfare analysis

In this subsection, we analyze the welfare implications of overconfidence. We first discuss different channels impacting welfare and briefly analyze the single-agent case where overconfidence almost always leads to utility loss. We then move on to the two-agent case, where we derive two main insights. First, overconfidence is not always bad—it sometimes makes everyone better off, but only in a multi-agent environment; second, we can characterize the direction of the welfare change and how mutual learning reinforces or limits it by verifying a few conditions when the bias is small.

With self-perceptions $\tilde{\boldsymbol{a}}$, agent i's objective average payoff could be written as

$$Q\left(\boldsymbol{e}_\infty(\tilde{\boldsymbol{a}}), A^i, A^j, \Phi\right) - c(e^i_\infty(\tilde{\boldsymbol{a}})), \tag{9}$$

where $e^i_\infty(\tilde{\boldsymbol{a}}) = e^{*i}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}_\infty(\tilde{\boldsymbol{a}})\right)$ depend on $\tilde{a}^j$ and $\psi^j_\infty(\tilde{\boldsymbol{a}})$ only indirectly through agent j's effort. It is then clear that any impact on agent i's payoff comes from four different sources: (i) the distortion of e^i due to *overconfidence* (the difference between $\tilde{a}^i$ and A^i); (ii) the distortion of e^i due to *false inference*, (the discrepancy between ψ^i and Ψ^i); (iii) the distortion of e^i due to the distortion of e^j *(complementarity/substitutability)*; (iv) the direct effect of e^j on Q (*payoff externality*).

An easy observation is that the sum of effect (i) and effect (ii) can never be positive: misconceptions always impair the agent's ability to choose the correct actions. With a single actively-optimizing agent, or in a case where output Q does not depend on e^j, effects (iii) and (iv) are eliminated. As a result, in a single-agent setting, the agent always enjoys lower or equal utility. We summarize these observations below.

Claim 1. If the output function has the following form

$$Q\left(e^i, e^j, a^i, a^j, \phi\right) = \overline{Q}\left(e^i, a^i, a^j, \phi\right), \tag{10}$$

then an incorrect belief about a^i makes the agent i weakly worse off in the steady state.

When there are multiple agents, sources (iii) and (iv) start to take effect. The presence of a second agent introduces a public-good problem since agents maximize their individual payoff without internalizing their positive payoff externality over one another. Therefore, we may see increases in both agents' expected payoffs if the misspecification turns out to incentivize more (but not too much) effort. The extent to which overconfidence harms or benefits the agents depends on the properties of Q and how much $\tilde{\boldsymbol{a}}$ deviates from $\boldsymbol{A}$.

Proposition 4 partially characterizes the welfare impact of slight overconfidence. Since e^i has been optimized, the change in agent i's payoff will be dictated by the change of e^j and the derivative of agent i's payoff with respect to e^j, while the effect of a distorted e^i is secondary compared to agent j's payoff externality. That is, effect (iv) dominates effects (i) to (iii), and the impact of overconfidence on welfare is mainly determined by changes in steady-state efforts. Notably, the welfare impact highly depends on the direction of informational externalities. Positive externalities allow for an easy determination of the welfare change from the primitives, while negative externalities make the welfare impact sensitive to differences in agents' self-perceptions.

Proposition 4. *Suppose Assumptions 1 to 3 hold and $Q_{e^i \psi^i} > 0$ for both $i = 1, 2$.*

1. *If agents impose positive informational externalities over each other, then there exists $\delta > 0$ such that with self-perceptions $\tilde{\boldsymbol{a}} \in B^+_\delta(\boldsymbol{A})$, both agents are worse off than when they are correctly specified.*[15]
2. *If agents impose negative informational externalities and efforts e^i, e^j are strict substitutes, then there exists $\delta > 0$ such that with $\tilde{a}^i = A^i, \tilde{a}^j \in B^+_\delta(A^j)$, agent i is worse off while agent j is better off. If in addition Q is symmetric with respect to i and j, then there exists $\delta > 0$ such that with $\tilde{a}^i = \tilde{a}^j \in B^+_\delta(A^j)$, both agents are worse off.*

[15] $B^+_\delta(x) = \{y : y > x, \|y - x\| < \delta\}$ is defined as the upper right-hand area inside an x-centered circle with radius δ.

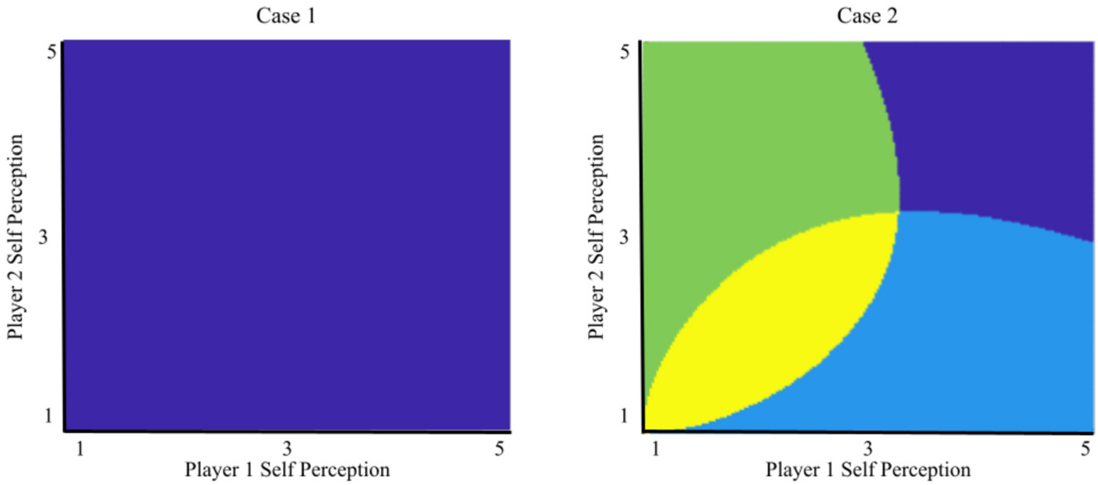

Fig. 3. Welfare changes for Examples 1 and 2. Both panels display the welfare change for both agents under different levels of confidence, given that they have true ability $A^1 = A^2 = 1$. Purple areas represent confidence levels where both agents experience a decrease in welfare compared to the case with correctly specified beliefs. In green areas, player 1 is better off and player 2 is worse off, while in blue areas player 1 is worse off and player 2 is better off. Yellow areas denote cases where both players are better off. The payoff functions used are symmetric, and for the first panel, the functional form is $Q(e^1, e^2, a^1, a^2, \phi) = \phi(e^1 + e^2 + e^1e^2 + a^1 + a^2)$, while for the second panel, it is $Q(e^1, e^2, a^1, a^2, \phi) = log(e^1a^1 + e^2a^2 + \phi)$. The cost function is $c_i(e^i) = 5(e^i)^2$ for both panels, and $\Phi = 3$ for both panels. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

If $Q_{e^i\psi^i} < 0$ for both $i = 1, 2$ then all payoffs change in opposite directions.

In a simple setting without informational externalities, agents' effort choices depend only on their incorrect self-perceptions and the resulting underestimation of the unknown variable. Mild overconfidence increases both agent's efforts and corrects inefficiency if $Q_{e^ia^i} \geq 0$ and $Q_{e^i\psi^i} < 0$, while the opposite pattern arises if $Q_{e^ia^i} \leq 0$ and $Q_{e^i\psi^i} > 0$. Under positive informational externalities, the welfare change is larger in the same direction due to the mutually reinforced mis-learning, leading to even more distorted effort. However, with negative informational externalities, if $Q_{e^i\psi^i} > 0$, the decrease in agent j's effort creates upward pressure on agent i's effort, counteracting the effect of agent i's overconfidence. When the agents are equally overconfident and Q is symmetric across agents, the agents still exert lower effort—negative informational externalities only mitigate the welfare loss. If only agent j is overconfident, agent j works less but agent i works harder in response to a lower e^j, which makes agent j better off and agent i worse off. Continuity ensures that this continues to hold when agent i is slightly overconfident. To illustrate, we compute the steady states and plot the direction of the welfare changes for Examples 1 and 2 in Fig. 3.

5. Convergence

In this section, we prove the convergence of the multi-agent learning processes to the unique Berk-Nash equilibrium under positive or negative informational externalities. We make use of a simple and intuitive lemma from Heidhues et al. (2018) stating that the support of any agent i's long-term belief cannot contain an element ψ^i if its implied distribution of outputs exhibits systematic mismatch with true distribution, i.e. the subjective expectation of outputs is consistently lower or higher than the objective expectation. We then use a contraction argument similar to theirs, exploiting the structural properties of the output function to eliminate a subset of actions given all possible beliefs, which in turn further rules out a subset of beliefs.

The added player brings non-trivial complications to the proof. To prevent one agent's converging learning process from being disrupted by the other agent's changing optimization, we have to impose additional structure to control for the agents' mutual influence.

Assumption 4. *The informational externalities are either both positive or both negative.*

Now we are ready to state the theorem for the convergence of beliefs and actions for the two-agent environment.

Theorem 1. *Suppose Assumptions 1 to 4 hold, then the agents' actions almost surely converge to the Berk-Nash equilibrium actions, $\boldsymbol{e}_\infty$, and their beliefs almost surely converge in distribution to the Dirac measure at $\boldsymbol{\psi}_\infty$.*

Fig. 1 offers some key insights to understand the convergence mechanism. In the left panel, our assumption dictates that the belief formation curve intersects with the optimal action curve at only one point, and the former must be steeper than

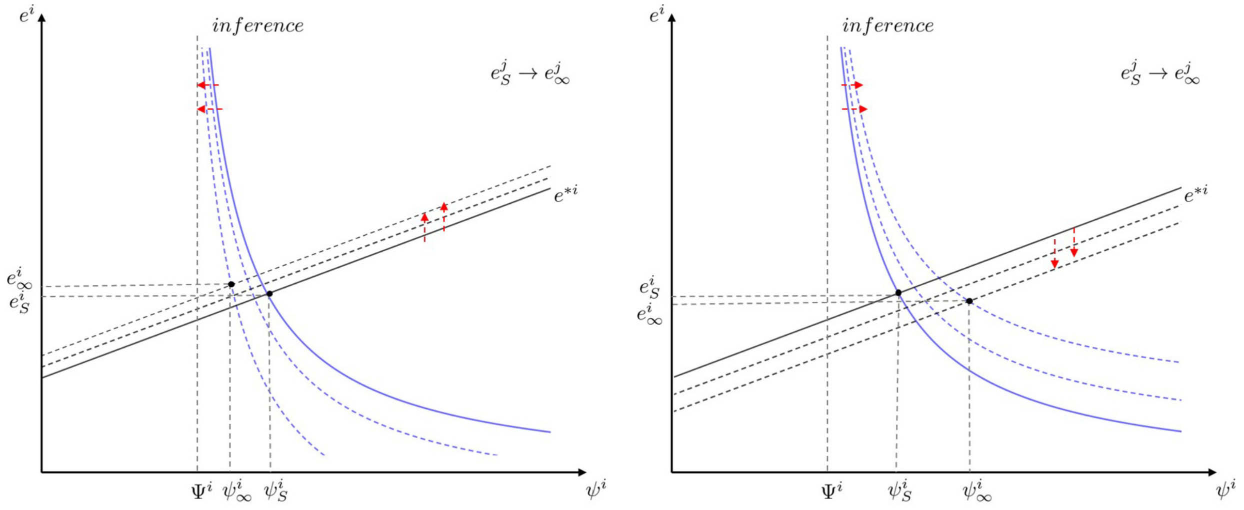

Fig. 4. Mutual learning with underconfidence. The left panel shows the heuristic learning dynamics of an underconfident agent when informational externalities are positive. Steady-state inferences are lower than when agent j's effort choice is fixed. The right panel shows a reverse pattern when informational externalities are negative.

the latter at this point. In a single-agent environment, iterated elimination of dominated actions and infeasible beliefs eliminates all but the crossing point—the Berk-Nash equilibrium profile. However, in the two-agent environment, agents' mutual influence must be considered—the iterated elimination has to be run simultaneously for both agents. When informational externalities are neither positive nor negative, agent i's inference, computed from the no-gap condition, and his optimal action can be non-monotone functions of e^j. This possibly creates cycles in which agents have jointly oscillating actions and beliefs and fail to converge.

6. Extensions

6.1. Underconfidence

In this section, we explore the implications of underconfidence. The assumption that $\tilde{\boldsymbol{a}} > \boldsymbol{A}$ is important to the direction of mutual learning. As pointed out by Heidhues et al. (2018), with underconfidence, the single-agent learning process is self-limiting. Similarly, in this model, with positive informational externalities the two-agent learning processes are mutually-limiting. In contrast to the overconfidence case, assuming $Q_{e^i\psi^i} > 0$, the belief formation curve is *downward* sloping: as the agents exert more effort, the marginal return of the unknown variable increases, causing them to overestimate the unknown to a lesser extent. Therefore, allowing agent j to freely optimize induces a lower steady-state belief that is closer to the truth. Conversely, with negative informational externalities, allowing agent j to freely optimize causes agent i's overestimation more severe. However, the reinforcement is not as severe as in the overconfidence case, because agent j's effort change is constrained through self-limiting learning. Nevertheless, depending on the output function, the resulting belief of agent i could still be arbitrarily far away from the truth. This contrasts the single-agent underconfidence case, where misinference is self-correcting and limited.

These phenomena are illustrated by Fig. 4. Notably, since the belief formation curve and the optimal effort curve are not comonotone, the effect of adding another actively-optimizing agent on the first agent's effort choice becomes ambiguous.

Additionally, there is no guarantee that the learning processes will converge. Assuming $Q_{e^i\psi^i} > 0$ and $Q_{e^ia^i} \leq 0$, as agents reach a lower belief in $\boldsymbol{\psi}$, they react by exerting lower efforts, pushing up beliefs again. Nevertheless, the following theorem shows that underconfident agents also converge to the unique Berk-Nash equilibrium, as long as the agents are only mildly underconfident. This assumption ensures that the belief formation curve is steeper than the optimal action curve over the relevant region, enabling the use of the contraction argument.[16] Note that Heidhues et al. (2018), working on a similar but single-agent setting, does not offer a convergence result for the underconfidence case.

Theorem 2. *Suppose Assumptions 1 to 4 hold. When $A^i - \tilde{a}^i$ is positive and sufficiently small for both i, the agents' actions almost surely converge to the Berk-Nash equilibrium $\boldsymbol{e}_\infty$, and their beliefs almost surely converge in distribution to the Dirac measure at $\boldsymbol{\psi}_\infty$.*

[16] The mild underconfidence assumption is crucial for the use of the contraction argument, but it may not be a necessary condition for convergence. However, no existing results from the literature can be directly applied here.

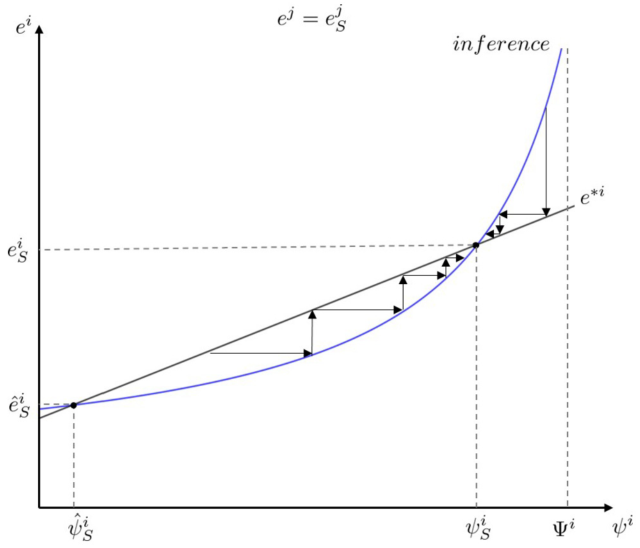

Fig. 5. Multiple Berk-Nash equilibria.

6.2. Multiple equilibria

The existence of multiple Berk-Nash equilibria does not affect our key message of mutually-reinforcing and mutually-limiting learning, but it does pose difficulties in proving the convergence of the learning processes. In Fig. 5, we illustrate single-agent heuristic learning dynamics where two equilibria, $\left(e^i_S, \psi^i_S\right)$ and $\left(\hat{e}^i_S, \hat{\psi}^i_S\right)$, exist. The contraction argument fails because there can be no further elimination of actions when the lower and upper bounds of actions are given by $\hat{e}^i_S$ and e^i_S. Thus, this paper does not provide a proof of convergence for such settings. Our analysis of mutual learning patterns remains valid for the former equilibrium $\left(e^i_S, \psi^i_S\right)$ but not for the latter equilibrium $\left(\hat{e}^i_S, \hat{\psi}^i_S\right)$. Moreover, note that $\left(e^i_S, \psi^i_S\right)$ is a more plausible steady state since the agent's belief tends to drift towards ψ^i_S as he optimizes and updates while drifting away from $\hat{\psi}^i_S$. In fact, it is possible to use the stochastic approximation tools from Esponda et al. (2021) to show convergence to $\left(e^i_S, \psi^i_S\right)$ in a finite-action environment—but unfortunately, their techniques do not directly apply to continuous-action settings like ours.[17]

6.3. Multiple agents

While our paper focuses on the two-player case, the results could be extended to an arbitrary number of agents. Mutually-reinforcing learning is exacerbated with more overconfident agents exerting positive informational externalities over each other. Moreover, the Pareto improvement when agents are slightly overconfident remains possible and potentially gets stronger. However, the parametric assumptions become increasingly complex with more agents, as more cross derivatives are involved.

7. Conclusion

In this paper, we develop a two-agent learning model with overconfidence and introduce a new notion of informational externalities to capture how one agent's actions affect the other agent's inference. When positive informational externalities are present, we find a mutually-reinforcing learning pattern, where strategic interaction exacerbates the underestimation of the common fundamental and leads to more extreme actions. In contrast, negative informational externalities give rise to a mutually-limiting learning pattern. These patterns are absent in Heidhues et al. (2018), where only one agent actively learns and optimizes. Furthermore, our welfare implications starkly contrast with those of single-agent models as there can be Pareto improvements resulting from overconfidence, and mutually-reinforcing learning can amplify this welfare gain.

One possible future direction is to consider strategic manipulation of informational externalities. For example, in a setting where agents are non-myopic and consider the impact of their actions over the other agent's beliefs, agents may be

[17] The more recent paper Murooka and Yamamoto (2021) generalizes this characterization based on stochastic approximation to allow for continuous actions and also study a multi-agent learning problem under overconfidence, but like ours their convergence argument only applies when there is a unique steady state.

incentivized to play actions that are suboptimal in the stage game in order to profitably distort the other player's long-term beliefs.

Declaration of competing interest

None

Data availability

No data was used for the research described in the article.

Acknowledgments

We are deeply indebted to Aislinn Bohren, George Mailath, Annie Liang, and Kevin He for their guidance and support. We would like to thank Alvaro Sandroni, Andrew Postlewaite, Takuo Sugaya, Steven Matthews, participants of multiple conferences and seminars at the University of Pennsylvania, and the anonymous referees for helpful comments and suggestions.

Appendix A. Preliminary lemmas

We first define notation that allows us to present a unified analysis for both Case 1 and Case 2. For $i \in I$, define $Q^i : [\underline{e}, \overline{e}]^2 \times (\underline{a}, \overline{a}) \times (\underline{\psi}, \overline{\psi})$ such that

$$Q^i(e^i, e^j, a^i, \psi^i) = \begin{cases} Q(e^i, e^j, a^i, A^j, \psi^i) - c(e^i) & \text{in Case 1,} \\ Q(e^i, e^j, a^i, \psi^i, \Phi) - c(e^i) & \text{in Case 2.} \end{cases}$$

Assumption 2 implies that Q^i is strictly concave in e^i. Then the no-gap condition introduced by Eqs. (5) and (6) could be written as the same condition: for all $i \in I$,

$$g^i(\boldsymbol{e}, \psi^i) = Q^i(\boldsymbol{e}, A^i, \Psi^i) - Q^i(\boldsymbol{e}, \tilde{a}^i, \psi^i) = 0.$$

Define $G^i(\boldsymbol{\psi}) := g^i\left(\boldsymbol{e^*}(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}), \psi^i\right)$ and denote $\boldsymbol{G}(\boldsymbol{\psi}) = \left(G^i(\boldsymbol{\psi}), G^j(\boldsymbol{\psi})\right)$. This is the gap function when every agent actively optimizes according to a degenerate belief at ψ^i. The Berk-Nash equilibrium belief satisfies $G^i(\boldsymbol{\psi}_\infty) = 0, \forall i$. Let $\overline{\kappa}_a \geq \max\left\{Q_{a^1}, Q_{a^2}\right\}$ denote the upper bound on Q_{a^i} and $0 < \underline{\kappa}_\psi \leq \min\left\{Q_{\psi^1}, Q_{\psi^2}\right\}$ denote the lower bound on Q_{ψ^i} throughout.

Proof of Lemma 1. Suppose $e^i \in \arg\max_{e \in [\underline{e}, \overline{e}]} Q^i\left(e, e^j, \tilde{a}^i, \psi^i\right)$, then strict concavity implies that e^i is unique for a fixed e^j. Since Q^i is twice continuously differentiable, Brouwer's fixed-point theorem implies that a fixed point exists. Suppose there are two different fixed points $\left(e^i, e^j\right)$ and $\left(\hat{e}^i, \hat{e}^j\right)$, and without loss of generality $e^i > \hat{e}^i$, then

$$Q^i_{e^i}\left(e^i, e^j, \tilde{a}^i, \psi^i\right) = 0, \forall i,$$
$$Q^i_{e^i}\left(\hat{e}^i, \hat{e}^j, \tilde{a}^i, \psi^i\right) = 0, \forall i.$$

Assumption 2 then implies $|e^j - \hat{e}^j| > |e^i - \hat{e}^i|, \forall i, j \neq i$. Since it cannot hold for every i, we obtain a contradiction. □

Proof of Lemma 2. By Gibb's inequality, the KL divergence is weakly positive and equates 0 if and only if the two distributions coincide almost everywhere. Therefore,

$$\mathbb{E}\left[\log \frac{f\left(q - Q^i\left(\boldsymbol{e}, A^i, \Psi^i\right)\right)}{f\left(q - Q^i\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)\right)}\right] \geq 0,$$

where equality is obtained if and only if $Q^i\left(\boldsymbol{e}, A^i, \Psi^i\right) = Q^i\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)$. Hence, the agent i's equilibrium belief is a Dirac measure at such ψ^i. Since the equilibrium must be optimal given the belief, it follows that $\left(\boldsymbol{e^*}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}_\infty\right), \boldsymbol{\psi}_\infty\right)$ is a pure-strategy Berk-Nash equilibrium if and only if the no-gap conditions hold.

Existence: Since $Q^i_{e^i}\left(\boldsymbol{e^*}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right), \tilde{a}^i, \psi^i\right) = 0, \forall i$ and Q^i is twice continuously differentiable, $e^{*i}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right)$ and $e^{*j}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right)$ are continuous in $\boldsymbol{\psi}$ and $\tilde{\boldsymbol{a}}$. Moreover, for all $\boldsymbol{e}$,

$$
\begin{aligned}
& Q^i\left(\boldsymbol{e}, \tilde{a}^i, \Psi-\frac{\overline{\kappa}_a}{\underline{\kappa}_\psi}\left(\tilde{a}^i-A^i\right)\right) \\
\leq & Q^i\left(\boldsymbol{e}, A^i, \Psi^i\right)+\overline{\kappa}_a\left(\tilde{a}^i-A^i\right)-\underline{\kappa}_\psi \frac{\overline{\kappa}_a}{\underline{\kappa}_\psi}\left(\tilde{a}^i-A^i\right) \\
= & Q^i\left(\boldsymbol{e}, A^i, \Psi^i\right) .
\end{aligned}
$$

It follows that $G^i\left(\Psi-\frac{\overline{\kappa}_a}{\underline{\kappa}_\psi}\left(\tilde{a}^i-A^i\right), \psi^j\right) \geq 0, \forall \psi^j$. Since $G^i\left(\Psi, \psi^j\right)<0, \forall \psi^j$, by the Brouwer's fixed-point theorem, there exists at least one root of $\boldsymbol{G}$ over the domain of $\mathbb{R}^2$. By Assumption 1, the root is inside the support of the prior belief. □

Proof of Lemma 3. By the implicit function theorem, $\partial e_i^*(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}) / \partial \psi^k$ is a continuous function of $\boldsymbol{\psi}$ and $\tilde{\boldsymbol{a}}$, $\forall i, k$. Thus,

$$
\begin{aligned}
\frac{\partial G^i(\boldsymbol{\psi})}{\partial \psi^k}= & Q_{e^i}^i\left(\boldsymbol{e}^{*}(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}), A^i, \Psi^i\right) \frac{\partial e^{* i}(\tilde{\boldsymbol{a}}, \boldsymbol{\psi})}{\partial \psi^k}+Q_{e^j}^i\left(\boldsymbol{e}^{*}(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}), A^i, \Psi^i\right) \frac{\partial e^{* j}(\tilde{\boldsymbol{a}}, \boldsymbol{\psi})}{\partial \psi^k} \\
& -Q_{e^i}^i\left(\boldsymbol{e}^{*}(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}), \tilde{a}^i, \psi^i\right) \frac{\partial e^{* i}(\tilde{\boldsymbol{a}}, \boldsymbol{\psi})}{\partial \psi^k}-Q_{e^j}^i\left(\boldsymbol{e}^{*}(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}), \tilde{a}^i, \psi^i\right) \frac{\partial e^{* j}(\tilde{\boldsymbol{a}}, \boldsymbol{\psi})}{\partial \psi^k} \\
& -\mathbf{1}_i(k) \cdot Q_{\psi^i}^i\left(\boldsymbol{e}^{*}(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}), \tilde{a}^i, \psi^i\right)
\end{aligned}
$$

is a continuous function of ψ and $\tilde{a}$, $\forall i$, where $\mathbf{1}_i(k)=1$ if $i=k$ and 0 otherwise. When the derivatives are evaluated at $\tilde{a}^i=A^i$ and $\psi^i=\Psi^i$,

$$
\begin{aligned}
& \frac{\partial G^i(\boldsymbol{\psi})}{\partial \psi^i}|_{(\tilde{a}^i, \psi^i)=(A^i, \Psi^i)}=-Q_{\psi^i}^i\left(e^*(\boldsymbol{\psi}), A^i, \Psi^i\right)<0, \\
& \frac{\partial G^i(\boldsymbol{\psi})}{\partial \psi^j}|_{(\tilde{a}^i, \psi^i)=(A^i, \Psi^i)}=0 .
\end{aligned}
$$

Continuity then implies that there exist $\Delta^i>0, i=1,2$, such that for any $a^i \in\left(A^i, A^i+\Delta^i\right)$ and for any $\psi^i \in \left[\Psi^i-\frac{\overline{\kappa}_a}{\underline{\kappa}_\psi}\left|\tilde{a}^i-A^i\right|, \Psi^i\right] \subset(\underline{\psi}, \overline{\psi})$, the following are true:

$$
\frac{\partial G^i(\boldsymbol{\psi})}{\partial \psi^i}<0,\left|\frac{\partial G^i(\boldsymbol{\psi})}{\partial \psi^j}\right|<\left|\frac{\partial G^i(\boldsymbol{\psi})}{\partial \psi^i}\right|, \forall i . \tag{A.1}
$$

Suppose there are two different roots, $\tilde{\boldsymbol{\psi}}$ and $\hat{\boldsymbol{\psi}}$, and assume without loss of generality $\tilde{\psi}^i<\hat{\psi}^i$. By the second inequality in Eq. (A.1), if $\boldsymbol{G}(\tilde{\boldsymbol{\psi}})=\boldsymbol{G}(\hat{\boldsymbol{\psi}})=0$, then it must be that $|\tilde{\psi}^j-\hat{\psi}^j|>|\tilde{\psi}^i-\hat{\psi}^i|, \forall i, j \neq i$. The statement contradicts itself. □

Lemma 4. *The stage game at time t is dominance solvable. That is, given the priors π_{t-1}^i and π_{t-1}^j, there exists a unique rationalizable action profile $\left(e_t^i, e_t^j\right)$.*

Proof. First, consider the case of complementary efforts, $Q_{e^i e^j} \geq 0$. Let $\left[\underline{e}_0^j, \overline{e}_0^j\right]=\left[\underline{e}_0^i, \overline{e}_0^i\right]=[\underline{e}, \overline{e}]$, and recursively define for all i and $\tau \geq 1$,

$$
\begin{aligned}
& \underline{b}_\tau^i=\arg \max _{e^i} \mathbb{E}_{\pi_{\tau-1}^i}\left[Q^i\left(e^i, \underline{e}_{\tau-1}^j, \tilde{a}^i, \psi^i\right)\right]:=\arg \max _{e^i} h_\tau^i\left(e^i, \underline{e}_{\tau-1}^j, \tilde{a}^i\right), \\
& \overline{b}_\tau^i=\arg \max _{e^i} \mathbb{E}_{\pi_{\tau-1}^i}\left[Q^i\left(e^i, \overline{e}_{\tau-1}^j, \tilde{a}^i, \psi^i\right)\right]:=\arg \max _{e^i} h_\tau^i\left(e^i, \overline{e}_{\tau-1}^j, \tilde{a}^i\right) .
\end{aligned}
$$

The existence of such $\underline{b}_\tau^i$ and $\overline{b}_\tau^i$ follow from Assumption 2 and the continuity of Q^i. Since Q^i is twice continuously differentiable, h_τ^i is also twice continuously differentiable. Note that complementarity between efforts implies $\frac{\partial h_\tau^i}{\partial e^i \partial e^j} \geq 0$, so $\underline{b}_\tau^i \leq \overline{b}_\tau^i$. Let $\left[\underline{e}_\tau^i, \overline{e}_\tau^i\right]:=\left[\underline{e}_{\tau-1}^i, \overline{e}_{\tau-1}^i\right] \cap\left[\underline{b}_\tau^i, \overline{b}_\tau^i\right]$. By Assumption 2, $\left[\underline{e}_1^i, \overline{e}_1^i\right] \subsetneq\left[\underline{e}_0^i, \overline{e}_0^i\right]=[\underline{e}, \overline{e}], \forall i$. Using $\frac{\partial h_\tau^i}{\partial e^i \partial e^j} \geq 0$ again, we know $[\underline{e}_\tau^j, \overline{e}_\tau^j] \subset[\underline{e}_{\tau-1}^j, \overline{e}_{\tau-1}^j]$ implies that $\left[\underline{e}_{\tau+1}^i, \overline{e}_{\tau+1}^i\right] \subset\left[\underline{e}_\tau^i, \overline{e}_\tau^i\right], \forall i, j \neq i$ and $\forall \tau>1$. By the Nested Intervals Theorem, we know that each agent's set of rationalizable actions $\left[\underline{e}_\tau^i, \overline{e}_\tau^i\right]$ converges to the an interval with boundary points which are fixed points of mutual optimization. By Lemma 1, there is only one such fixed point. Therefore, there is a unique rationalizable action profile, $\left(e_t^i, e_t^j\right)$, and it satisfies

$$e_t^i = \arg\max_{e^i} \mathbb{E}_{\pi_{t-1}^i}\left[Q^i\left(e^i, e_t^i, \tilde{a}^i, \psi^i\right)\right], \forall i.$$

Next, consider the case of substitute efforts, $Q_{e^ie^j} \le 0$. Analogously define $\left[\underline{e}_0^i, \overline{e}_0^i\right] = [\underline{e}, \overline{e}]$ and recursively define for all i and $\tau \ge 1$,

$$\underline{b}_\tau^i = \arg\max_{e^i} \mathbb{E}_{\pi_{t-1}^i}\left[Q^i\left(e^i, \overline{e}_{\tau-1}^j, \tilde{a}^i, \psi^i\right)\right] := \arg\max_{e^i} h_\tau^i\left(e^i, \overline{e}_{\tau-1}^j, \tilde{a}^i\right),$$
$$\overline{b}_\tau^i = \arg\max_{e^i} \mathbb{E}_{\pi_{t-1}^i}\left[Q^i\left(e^i, \underline{e}_{\tau-1}^j, \tilde{a}^i, \psi^i\right)\right] := \arg\max_{e^i} h_\tau^i\left(e^i, \underline{e}_{\tau-1}^j, \tilde{a}^i\right).$$

Substitute efforts implies $\frac{\partial h_\tau^i}{\partial e^i \partial e^j} \le 0$, so similarly we have $\underline{b}_\tau^i \le \overline{b}_\tau^i$. Let $\left[\underline{e}_\tau^i, \overline{e}_\tau^i\right] := \left[\underline{e}_{\tau-1}^i, \overline{e}_{\tau-1}^i\right] \cap \left[\underline{b}_\tau^i, \overline{b}_\tau^i\right]$. The rest of the steps are identical to the previous case. □

Appendix B. Proofs for Section 4

Proof of Proposition 1. In this proof, we show a more general version of Proposition 1. Specifically, we show mutual-reinforced learning for any $e_S^j \in [\underline{e}, \overline{e}]$ such that if $Q_{e^i\psi^i} > 0$, then we have $e_S^j > e_\infty^j(\tilde{\boldsymbol{a}})$, and if instead $Q_{e^i\psi^i} < 0$, then $e_S^j < e_\infty^j(\tilde{\boldsymbol{a}})$. Note that $e_S^j = q_\infty^j(\boldsymbol{A})$ always satisfies the condition. We assume $Q_{e^i\psi^i} > 0$ for both i for this proof. Note that we can accommodate $Q_{e^i\psi^i} < 0, \forall i$ by replacing e^i, e^j with $-e^i, -e^j$ and substituting the constraint with $e_S^j < e_\infty^j(\tilde{\boldsymbol{a}})$.

We start by showing $e_S^i > e_\infty^i$ and $\psi_S^i > \psi_\infty^i$. Fixing agent j's effort at some level e^j, consider the following two equations,

$$g^i(\boldsymbol{e}, \psi^i) = Q^i\left(\boldsymbol{e}, A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right) = 0, \tag{B.1}$$
$$Q_{e^i}^i\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right) = 0, \tag{B.2}$$

where e^i and ψ^i are the unknown variables. Notice that (e_S^i, ψ_S^i) and $(e_\infty^i, \psi_\infty^i)$ are solutions to Eqs. (B.1) and (B.2) when agent j's effort is taken to be e_S^j and $e_\infty^j(\tilde{\boldsymbol{a}})$, respectively.

We denote the set of possible actions and beliefs under a fixed e^j by $D_e^i(e^j)$ and $D_\psi^i(e^j)$ respectively, i.e. $(e^i, \psi^i) \in D_e^i(e^j) \times D_\psi^i(e^j)$ iff it satisfies Eqs. (B.1) and (B.2). By Assumptions 1 and 2, both $D_e^i(e^j)$ and $D_\psi^i(e^j)$ are nonempty and compact for any e^j. Further, let $\hat{\psi}^i(e^j)$ represent the largest element in $D_\psi^i(e^j)$, and $\hat{e}^i(e^j)$ represent the corresponding effort in $D_e^i(e^j)$. Since $(\hat{e}^i(e^j), \hat{e}^j(e^i))$ are continuous over a compact convex set, Brouwer's fixed point theorem implies that there exists a fixed point which is a Berk-Nash equilibrium. We know that $\boldsymbol{e}_\infty$ is a fixed point of the correspondence $\left(D_e^i(e^j), D_e^j(e^i)\right)$ and that it is unique by assumption, so $\boldsymbol{e}_\infty$ must also be the fixed point of $(\hat{e}^i(e^j), \hat{e}^j(e^i))$, which implies $e_\infty^i = \hat{e}^i(e_\infty^j)$ and $\psi_\infty^i = \hat{\psi}^i(e_\infty^j)$. By our assumption of a unique steady state at e_S^j, we also have $e_S^i = \hat{e}^i(e_S^j)$ and $\psi_S^i = \hat{\psi}^i(e_S^j)$.

Differentiate Eqs. (B.1) and (B.2) with respect to e^j,

$$\left[Q_{e^i}^i\left(\boldsymbol{e}, A^i, \Psi^i\right) - Q_{e^i}^i\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)\right]\frac{\partial e^i}{\partial e^j} + \left[Q_{e^j}^i\left(\boldsymbol{e}, A^i, \Psi^i\right) - Q_{e^j}^i\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)\right] = Q_{\psi^i}^i\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)\frac{\partial \psi^i}{\partial e^j},$$
$$Q_{e^ie^i}^i\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)\frac{\partial e^i}{\partial e^j} + Q_{e^i\psi^i}^i\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)\frac{\partial \psi^i}{\partial e^j} + Q_{e^ie^j}^i\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right) = 0.$$

Simplify and then obtain

$$\frac{\partial e^i}{\partial e^j} = \frac{-Q_{\psi^i}^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)\, Q_{e^ie^j}^i(\boldsymbol{e}, \tilde{a}^i, \psi^i) - \left[Q_{e^j}^i(\boldsymbol{e}, A^i, \Psi^i) - Q_{e^j}^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)\right] Q_{e^i\psi^i}^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)}{Q_{e^i\psi^i}^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)\left(Q_{e^i}^i(\boldsymbol{e}, A^i, \Psi^i) + Q_{\psi^i}^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)\frac{Q_{e^ie^i}^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)}{Q_{e^i\psi^i}^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)}\right)},$$
$$\frac{\partial \psi^i}{\partial e^j} = \frac{Q_{e^ie^i}^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)\left[Q_{e^j}^i(\boldsymbol{e}, A^i, \Psi^i) - Q_{e^j}^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)\right] - Q_{e^i}^i(\boldsymbol{e}, A^i, \Psi^i)\, Q_{e^ie^j}^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)}{Q_{e^i\psi^i}^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)\left(Q_{e^i}^i(\boldsymbol{e}, A^i, \Psi^i) + Q_{\psi^i}^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)\frac{Q_{e^ie^i}^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)}{Q_{e^i\psi^i}^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)}\right)}.$$

When e^i and ψ^i are taken to be $\hat{e}^i(e^j)$ and $\hat{\psi}^i(e^j)$, Lemma 5 and positive informational externalities imply that both derivatives are positive. Since $e_S^j > e_\infty^j$, we infer that $e_S^i > e_\infty^i$ and $\psi_S^i > \psi_\infty^i$.

Next we prove $\psi_S^j > \psi_\infty^j$. Observe that ψ_S^j is given by

$$Q^j\left(\boldsymbol{e}_S, A^j, \Psi^j\right) - Q^j\left(\boldsymbol{e}_S, \tilde{a}^j, \psi_S^j\right) = 0.$$

Since $Q^i_{e^k \psi} \geq 0, \forall k$, the function $Q^i\left(\boldsymbol{e}, A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)$ is increasing in e^i and e^j. Hence,

$$\begin{aligned} 0 &= Q^j\left(\boldsymbol{e_S}, A^j, \Psi^j\right) - Q^j\left(\boldsymbol{e_S}, \tilde{a}^j, \psi^j_S\right) \\ &= Q^j\left(\boldsymbol{e}_\infty, A^j, \Psi^j\right) - Q^j\left(\boldsymbol{e}_\infty, \tilde{a}^j, \psi^j_\infty\right) \\ &< Q^j\left(\boldsymbol{e_S}, A^j, \Psi^j\right) - Q^j\left(\boldsymbol{e_S}, \tilde{a}^j, \psi^j_\infty\right). \end{aligned}$$

The inequality implies $\psi^j_S > \psi^j_\infty$. Since the agents are overconfident, their equilibrium beliefs are always below the true levels $\boldsymbol{\Psi}$. Therefore, $\boldsymbol{\psi}_\infty\left(\tilde{\boldsymbol{a}}\right) < \boldsymbol{\psi_S} < \boldsymbol{\Psi}$. □

Lemma 5. *Suppose $Q_{e^i\psi^i} > 0$ for both i. Given e^j, when $e^i = \hat{e}^i(e^j)$ and $\psi^i = \hat{\psi}^i(e^j)$, we have*

$$Q^i_{e^i}\left(\boldsymbol{e}, A^i, \Psi^i\right) + Q^i_{\psi^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right) \frac{Q^i_{e^ie^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)}{Q^i_{e^i\psi^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)} < 0, \forall i.$$

Proof. Note that $\hat{e}^i(e^j)$ and $\hat{\psi}^i(e^j)$ satisfy

$$Q^i\left(\hat{e}^i(e^j), e^j, A^i, \Psi^i\right) - Q^i\left(\hat{e}^i(e^j), e^j, \tilde{a}^i, \hat{\psi}^i(e^j)\right) = 0.$$

Let $e^{\dagger i}(\psi^i)$ denote the solution to $Q^i_{e^i}\left(e^i, e^j, \tilde{a}^i, \psi^i\right) = 0$ for a given ψ^i. Then the above equation could be written as

$$Q^i\left(e^{\dagger i}(\hat{\psi}^i(e^j)), e^j, A^i, \Psi^i\right) - Q^i\left(e^{\dagger i}(\hat{\psi}^i(e^j)), e^j, \tilde{a}^i, \hat{\psi}^i(e^j)\right) = 0.$$

Since $Q^i\left(e^i, e^j, A^i, \Psi^i\right) - Q^i\left(e^i, e^j, \tilde{a}^i, \Psi^i\right) < 0$ for all e^i, e^j, we infer that for any $\psi'^i > \hat{\psi}^i\left(e^j\right)$,

$$Q^i\left(e^{\dagger i}(\psi'^i), e^j, A^i, \Psi^i\right) - Q^i\left(e^{\dagger i}(\psi'^i), e^j, \tilde{a}^i, \psi'^i\right) < 0,$$

otherwise there exists an element in $D^i_\psi\left(e^j\right)$ larger than $\hat{\psi}^i(e^j)$, contradicting the definition of $\hat{\psi}^i(\cdot)$. This implies that when $\psi^i = \hat{\psi}^i\left(e^j\right)$,

$$\begin{aligned} &\frac{\partial\left[Q^i\left(e^{\dagger i}(\psi^i), e^j, A^i, \Psi^i\right) - Q^i\left(e^{\dagger i}(\psi^i), e^j, \tilde{a}^i, \psi^i\right)\right]}{\partial \psi^i} < 0, \\ &\Rightarrow Q^i_{e^i}\left(e^{\dagger i}(\psi^i), e^j, A^i, \Psi^i\right)\frac{\partial e^{\dagger i}}{\partial \psi^i} - Q^i_{\psi^i}\left(e^{\dagger i}(\psi^i), e^j, \tilde{a}^i, \psi^i\right) < 0. \end{aligned}$$

Since $Q^i_{e^i}\left(e^{\dagger i}(\psi^i), e^j, \tilde{a}^i, \psi^i\right) = 0$, we know that

$$Q^i_{e^ie^i}\left(e^{\dagger i}(\psi^i), e^j, \tilde{a}^i, \psi^i\right)\frac{\partial e^{\dagger i}}{\partial \psi^i} = -Q^i_{e^i\psi^i}\left(e^{\dagger i}(\psi^i), e^j, \tilde{a}^i, \psi^i\right).$$

Plugging this back into the previous inequality, we obtain that when $\psi^i = \hat{\psi}^i\left(e^j\right)$ and $e^i = e^{\dagger i}(\psi^i)$,

$$Q^i_{e^i}\left(\boldsymbol{e}, A^i, \Psi^i\right) + Q^i_{\psi^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right) \frac{Q^i_{e^ie^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)}{Q^i_{e^i\psi^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)} < 0. \quad \square$$

Lemma 6. *Let $\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{a}}' \in \left(\underline{a}, \overline{a}\right)^2$ and $\tilde{\boldsymbol{a}}' < \tilde{\boldsymbol{a}}$. Suppose agents create positive informational externalities for each other. Then we have:*

(i) If $Q^i_{e^i\psi^i} > 0, \forall i$, then $\boldsymbol{e}_\infty\left(\tilde{\boldsymbol{a}}'\right) > \boldsymbol{e}_\infty\left(\tilde{\boldsymbol{a}}\right)$;
(ii) If $Q^i_{e^i\psi^i} < 0, \forall i$, then $\boldsymbol{e}_\infty\left(\tilde{\boldsymbol{a}}'\right) < \boldsymbol{e}_\infty\left(\tilde{\boldsymbol{a}}\right)$.

Now suppose agents create negative informational externalities for each other, then:

(iii) If $Q^i_{e^i\psi^i} > 0, \forall i$, then $e^i_\infty(\boldsymbol{a})$ decreases in a^i but increases in a^j;
(iv) If $Q^i_{e^i\psi^i} < 0, \forall i$, then $e^i_\infty(\boldsymbol{a})$ increases in a^i but decreases in a^j.

Proof. We only show parts (i) and (iii) since parts (ii) and (iv) follow from analogous arguments. Differentiating the following equations that determine the steady state with respect to a^i and a^j,

$$Q^i\left(\boldsymbol{e}_\infty(\boldsymbol{a}), A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e}_\infty(\boldsymbol{a}), a^i, \psi^i_\infty(\boldsymbol{a})\right) = 0,$$
$$Q^i_{e^i}\left(\boldsymbol{e}_\infty(\boldsymbol{a}), a^i, \psi^i_\infty(\boldsymbol{a})\right) = 0,$$

we obtain,

$$\frac{\partial e^i_\infty(\boldsymbol{a})}{\partial a^k} = \frac{I(i,k)\omega^{ik}}{\omega^{ii}\omega^{jj} + \omega^{ij}\omega^{ji}}\left(Q^k_{e^k a} - Q^k_{e^k\psi^k}\frac{Q^k_a}{Q^k_{\psi^k}}\right),$$

where $I(i,k) = -1$ if $i = k$ and $I(i,k) = 1$ if $i \neq k$, and $\omega^{ik} = Q^{-k}_{e^{-k}e^{-i}} + Q^{-k}_{e^{-k}\psi^{-k}}\frac{\left(Q^{-k,A}_{e^{-i}} - Q^{-k}_{e^{-i}}\right)}{Q^{-i}_{\psi^{-i}}} = \frac{Q^{-k}_{e^{-k}\psi^{-k}}}{Q^{-i}_{\psi^{-i}}}\left(Q^{-k,A}_{e^{-i}} - Q^{-k}_{e^{-i}} + Q^{-i}_{\psi^{-i}}\frac{Q^{-k}_{e^{-k}e^{-i}}}{Q^{-k}_{e^{-k}\psi^{-k}}}\right)$. All derivatives are evaluated at $\boldsymbol{e}_\infty(\boldsymbol{a}), \boldsymbol{a}, \boldsymbol{\psi}_\infty(\boldsymbol{a})$, except $Q^{-k,A}_{e^{-k}}$ which is evaluated at $\boldsymbol{e}_\infty(\boldsymbol{a}), \boldsymbol{A}, \boldsymbol{\Psi}$. From Lemma 5, $\omega^{ii}, \omega^{jj} < 0$. Positive informational externalities imply that $\omega^{ij}, \omega^{ji} > 0$, whereas negative informational externalities imply that $\omega^{ij}, \omega^{ji} < 0$. Therefore, when $Q^i_{e^i\psi^i} > 0, \forall i$, we have $\frac{\partial e^i_\infty(\boldsymbol{a})}{\partial a^i} < 0$; in addition, $\frac{\partial e^i_\infty(\boldsymbol{a})}{\partial a^j} < 0$ if the informational externalities are positive, whereas $\frac{\partial e^i_\infty(\boldsymbol{a})}{\partial a^j} > 0$ when they are negative. This yields the desired results in (i) and (iii). □

Proof of Proposition 2. We only consider the case where $Q^i_{e^i\psi^i} > 0, \forall i$ since analogous arguments could be applied to the other case where $Q^i_{e^i\psi^i} < 0, \forall i$. It follows from this assumption and positive informational externalities that the function $Q^i(\boldsymbol{e}, A^i, \Psi^i) - Q^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)$ is increasing in e^i and e^j, so Lemma 6 implies that

$$\begin{aligned}
0 =& Q^i\left(\boldsymbol{e}_\infty(\tilde{\boldsymbol{a}}), A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e}_\infty(\tilde{\boldsymbol{a}}), \tilde{a}^i, \psi^i_\infty(\tilde{\boldsymbol{a}})\right)\\
<& Q^i\left(\boldsymbol{e}_\infty(\tilde{\boldsymbol{a}}'), A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e}_\infty(\tilde{\boldsymbol{a}}'), \tilde{a}^i, \psi^i_\infty(\tilde{\boldsymbol{a}})\right)\\
\leq& Q^i\left(\boldsymbol{e}_\infty(\tilde{\boldsymbol{a}}'), A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e}_\infty(\tilde{\boldsymbol{a}}'), \tilde{a}'^i, \psi^i_\infty(\tilde{\boldsymbol{a}})\right).
\end{aligned}$$

Since $Q^i(\boldsymbol{e}_\infty(\tilde{\boldsymbol{a}}'), A^i, \Psi^i) - Q^i(\boldsymbol{e}_\infty(\tilde{\boldsymbol{a}}'), \tilde{a}'^i, \psi^i_\infty(\tilde{\boldsymbol{a}}')) = 0$, it follows that $\psi^i_\infty(\tilde{\boldsymbol{a}}') > \psi^i_\infty(\tilde{\boldsymbol{a}})$ for all i. □

Proof of Proposition 3. Consider the case where $Q^i_{e^i\psi^i} > 0$ for all i. Then the negative informational externality of j's action implies that $Q^i_{e^j\psi^i} \leq 0$.

Part (i). As in the proof of Proposition 1, we have

$$\frac{\partial e^i}{\partial e^j} = \frac{-Q^i_{\psi^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)Q^i_{e^ie^j}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right) - \left[Q^i_{e^j}\left(\boldsymbol{e}, A^i, \Psi^i\right) - Q^i_{e^j}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)\right]Q^i_{e^i\psi^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)}{Q^i_{e^i\psi^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)\left(Q^i_{e^i}\left(\boldsymbol{e}, A^i, \Psi^i\right) + Q^i_{\psi^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)\frac{Q^i_{e^ie^i}(\boldsymbol{e}, \tilde{a}^i, \psi^i)}{Q^i_{e^i\psi^i}(\boldsymbol{e}, \tilde{a}^i, \psi^i)}\right)},$$

$$\frac{\partial \psi^i}{\partial e^j} = \frac{Q^i_{e^ie^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)\left[Q^i_{e^j}\left(\boldsymbol{e}, A^i, \Psi^i\right) - Q^i_{e^j}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)\right] - Q^i_{e^i}\left(\boldsymbol{e}, A^i, \Psi^i\right)Q^i_{e^ie^j}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)}{Q^i_{e^i\psi^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)\left(Q^i_{e^i}\left(\boldsymbol{e}, A^i, \Psi^i\right) + Q^i_{\psi^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)\frac{Q^i_{e^ie^i}(\boldsymbol{e}, \tilde{a}^i, \psi^i)}{Q^i_{e^i\psi^i}(\boldsymbol{e}, \tilde{a}^i, \psi^i)}\right)}.$$

Note that the negative informational externality implies that both $Q^i_{e^ie^j}$ and $Q^i_{e^j}(\boldsymbol{e}, A^i, \Psi^i) - Q^i_{e^j}(\boldsymbol{e}, \tilde{a}^i, \psi^i)$ are non-positive and at least one of them has to be strictly negative. It follows that $\frac{\partial e^i}{\partial e^j}, \frac{\partial \psi^i}{\partial e^j} < 0$. Therefore, $e^j_S = e^j_\infty(\boldsymbol{A}) > e^j_\infty(\tilde{\boldsymbol{a}})$ implies $\psi^i_S < \psi^i_\infty(\tilde{\boldsymbol{a}}) < \Psi^i$.

Part (ii). By Lemma 6, $e^i_\infty(\boldsymbol{a})$ is increasing in a^j and $e^j_\infty(\boldsymbol{a})$ is decreasing in a^j. So $e^i_\infty(\tilde{\boldsymbol{a}}) > e^i_\infty(\tilde{\boldsymbol{a}}')$ and $e^j_\infty(\tilde{\boldsymbol{a}}) < e^j_\infty(\tilde{\boldsymbol{a}}')$. It follows that

$$\begin{aligned}
0 =& Q^i\left(\boldsymbol{e}_\infty(\tilde{\boldsymbol{a}}), A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e}_\infty(\tilde{\boldsymbol{a}}), \tilde{a}^i, \psi^i_\infty(\tilde{\boldsymbol{a}})\right)\\
>& Q^i\left(\boldsymbol{e}_\infty(\tilde{\boldsymbol{a}}'), A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e}_\infty(\tilde{\boldsymbol{a}}'), \tilde{a}^i, \psi^i_\infty(\tilde{\boldsymbol{a}})\right)
\end{aligned}$$

$$=Q^i\left(\boldsymbol{e}_\infty\left(\tilde{\boldsymbol{a}}'\right), A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e}_\infty\left(\tilde{\boldsymbol{a}}'\right), \tilde{a}'^i, \psi^i_\infty\left(\tilde{\boldsymbol{a}}\right)\right).$$

Since $Q^i\left(\boldsymbol{e}_\infty\left(\tilde{\boldsymbol{a}}'\right), A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e}_\infty\left(\tilde{\boldsymbol{a}}'\right), \tilde{a}'^i, \psi^i_\infty\left(\tilde{\boldsymbol{a}}'\right)\right) = 0$, this implies that $\psi^i_\infty\left(\tilde{\boldsymbol{a}}'\right) < \psi^i_\infty\left(\tilde{\boldsymbol{a}}\right)$. The proof is analogous when $Q^i_{e^i\psi^i} < 0$ for both i. □

Proof of Proposition 4. Differentiate agent i's true average payoff $Q\left(\boldsymbol{e}_\infty(\boldsymbol{a}), A^i, A^j, \Phi\right) - c(e^i_\infty(\boldsymbol{a}))$ with respect to a^i and a^j and evaluate the derivative at $\boldsymbol{A}$, then we obtain

$$\frac{\partial[Q\left(\boldsymbol{e}_\infty(\boldsymbol{a}), A^i, A^j, \Phi\right) - c(e^i_\infty(\boldsymbol{a}))]}{\partial a^i} = Q_{e^j}\left(\boldsymbol{e}_\infty(\boldsymbol{a}), A^i, A^j, \Phi\right)\frac{\partial e^j_\infty(\boldsymbol{a})}{\partial a^i},$$
$$\frac{\partial[Q\left(\boldsymbol{e}_\infty(\boldsymbol{a}), A^i, A^j, \Phi\right) - c(e^i_\infty(\boldsymbol{a}))]}{\partial a^j} = Q_{e^j}\left(\boldsymbol{e}_\infty(\boldsymbol{a}), A^i, A^j, \Phi\right)\frac{\partial e^j_\infty(\boldsymbol{a})}{\partial a^i}.$$

When $\boldsymbol{a} = \boldsymbol{A}$, the optimality of $e^j_\infty(\boldsymbol{A})$ implies that $Q_{e^j}\left(\boldsymbol{e}_\infty(\boldsymbol{a}), A^i, A^j, \Phi\right) = c'(e^j_\infty(A)) > 0$.

When there are positive informational externalities, by Lemma 6, we know that if $Q_{e^i\psi^i} > 0$ for both i, then $\frac{\partial e^j_\infty(\boldsymbol{a})}{\partial a^i}, \frac{\partial e^j_\infty(\boldsymbol{a})}{\partial a^j} < 0$ and thus both agents are worse off with slight overconfidence; if instead $Q_{e^i\psi^i} < 0$ for both i, then $\frac{\partial e^j_\infty(\boldsymbol{a})}{\partial a^i}, \frac{\partial e^j_\infty(\boldsymbol{a})}{\partial a^j} > 0$ and thus both agents are better off with slight overconfidence.

On the other hand, when there are negative informational externalities, by Lemma 6, we know that if $Q_{e^i\psi^i} > 0$ for both i, then $\frac{\partial e^j_\infty(\boldsymbol{a})}{\partial a^j} < 0$ but $\frac{\partial e^j_\infty(\boldsymbol{a})}{\partial a^i} > 0$. Therefore, if agent j is slightly overconfident and agent i is correctly specified, then agent i is better off but agent j becomes weakly worse off (strictly worse off if e^i and e^j are strict substitutes). The opposite pattern arises when $Q_{e^i\psi^i} < 0$ for both i. Finally, when Q is symmetric with respect to i and j and $a^i = a^j$, the calculations in the proof of Lemma 6 imply that $\frac{\partial e^j_\infty(\boldsymbol{a})}{\partial a^j} / \frac{\partial e^j_\infty(\boldsymbol{a})}{\partial a^i} = -Q^i_{e^ie^i}/Q^i_{e^ie^j} < -1$, where the inequality follows from Assumption 2. Hence,

$$\frac{\partial\left[Q\left(\boldsymbol{e}_\infty(a,a), A^i, A^j, \Phi\right) - c(e^i_\infty(a,a))\right]}{\partial a} < 0,$$

which implies that the agents are worse off when they are equally slightly overconfident. □

Appendix C. Proofs for Section 5

Following Heidhues et al. (2018) (HKS), for each agent i, we define $m^i_t\left(\psi^i\right)$ to keep track of the actual gap in the average output at time t when the unknown variable is believed to take the value of ψ^i,

$$m^i_t\left(\psi^i\right) = Q^i\left(\boldsymbol{e_t}, A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e_t}, \tilde{a}^i, \psi^i\right), \forall i.$$

Define the lowest upper bound and the highest lower bound for any agent i' long-run beliefs as follows,

$$\underline{\psi}^i_\infty := \sup\left\{\psi^i : \lim_{t\to\infty} \Pi^i_t\left(\psi^i\right) = 0 \text{ almost surely}\right\},$$
$$\overline{\psi}^i_\infty := \inf\left\{\psi^i : \lim_{t\to\infty} \Pi^i_t\left(\psi^i\right) = 1 \text{ almost surely}\right\}.$$

Note that all probabilistic statements are with respect to the objective probability measure, unless indicated otherwise. Let $\tilde{\mathbb{P}}^i_t$ denote agent i's subjective probability measure conditional on the history up to time t. Write the vectors of bounds as $\underline{\boldsymbol{\psi}}_\infty = \left(\underline{\psi}^i_\infty \underline{\psi}^j_\infty\right)$ and $\overline{\boldsymbol{\psi}}_\infty = \left(\overline{\psi}^i_\infty, \overline{\psi}^j_\infty\right)$. We show that $\underline{\boldsymbol{\psi}}_\infty$ and $\overline{\boldsymbol{\psi}}_\infty$ are bounded in Lemma 8.

Next, we proceed by stating an important lemma established in HKS that could be easily reformulated in a two-agent environment. Lemma 7 shows that if some fundamental level ψ^i is in the support of the long-run beliefs, the average output associated with ψ^i should not be consistently higher or lower than the average output associated with Ψ^i. We replicate the proof to show that this lemma extends to our two-agent environment.

Lemma 7 *(Heidhues et al. (2018), Lemma 13). The following are true:*

(a) For all i, if $\liminf_{t\to\infty} m^i_t\left(\psi^i\right) \geq \underline{m} > 0$ for all $\psi^i \in (l,h) \subset \left(\underline{\psi}, \overline{\psi}\right)$, then

$$\lim_{t\to\infty} \tilde{\mathbb{P}}^i_t\left[\Psi^i \in [l,h)\right] = 0.$$

(b) For all i, if $\limsup_{t\to\infty} m_t^i\left(\psi^i\right) \le \overline{m} < 0$ *for all* $\psi^i \in (l,h) \subset \left(\underline{\psi}, \overline{\psi}\right)$*, then*

$$\lim_{t\to\infty} \tilde{\mathbb{P}}_t^i\left[\Psi^i \in (l,h]\right] = 0.$$

Proof. We only show (a) since the proof of (b) is analogous. Let $l_0^i(\psi^i)$ denote the subjective prior log-likelihood assigned to ψ^i by the agent. By the Bayes rule, $l_t^i(\psi^i) := \sum_{s=1}^t \log f\left(m_t^i(\psi^i) + \epsilon_s\right) + l_0^i(\psi^i)$ denote the log-likelihood assigned to ψ^i by agent i and $\mathcal{L}(x) := \mathbb{E}\log f(x+\epsilon)$. By Lemma 12 in Heidhues et al. (2018), there exists $r > 0$ such that

$$\liminf_{t\to\infty} \inf_{\psi^i\in(l,h)} \frac{l_t^{i\prime}(\psi^i)}{t} \ge \liminf_{t\to\infty} \frac{1}{t}\sum_{s=1}^{t} \mathcal{L}'(\underline{m})\left[-Q_{\psi^i}^i(\boldsymbol{e_t}, \tilde{a}^i, \psi^i)\right] \ge r,$$

where the first inequality is an implication of a generalized law of large numbers and the second inequality follows from the log-concavity of f and the observation that $\mathcal{L}$ is strictly concave and reaches its peak at 0. Next, notice that when t is large,

$$\begin{aligned}
\tilde{\mathbb{P}}_t^i\left[\Psi^i \in [(l+h)/2, h]\right] &= \int_{(l+h)/2}^{h} \pi_t^i(\psi^i)d\psi^i = \int_{l}^{(l+h)/2} \pi_t^i(\psi^i)\frac{\pi_t^i(\psi^i + (h-l)/2)}{\pi_t^i(\psi^i)}d\psi^i \\
&= \int_{l}^{(l+h)/2} \pi_t^i(\psi^i)\exp\left(\int_{\psi^i}^{\psi^i+(h-l)/2} l_t^{i\prime}(y)dy\right)d\psi^i \\
&\ge \tilde{\mathbb{P}}_t^i\left[\Psi^i \in [l, (l+h)/2]\right]\exp\frac{rt(h-l)}{2},
\end{aligned}$$

which implies that as t grows to infinity, $\tilde{\mathbb{P}}_t^i\left[\Psi^i \in [l,(l+h)/2]\right]$ goes to 0. Applying this argument iteratively yields the desired result. □

Lemma 8. *For all i, we have that* $\Psi^i - \frac{\overline{\kappa}_a}{\underline{\kappa}_\psi}\left(\tilde{a}^i - A^i\right) \le \underline{\psi}_\infty^i$ *and* $\overline{\psi}_\infty^i \le \Psi^i$.

Proof. Suppose $\Psi^i - \frac{\overline{\kappa}_a}{\underline{\kappa}_\psi}\left(\tilde{a}^i - A^i\right) > \underline{\psi}_\infty^i$, then

$$\begin{aligned}
&\liminf_{t\to\infty} m_t^i\left(\underline{\psi}_\infty^i\right) \\
=&\liminf_{t\to\infty}\left[Q^i\left(\boldsymbol{e_t}, A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e_t}, \tilde{a}^i, \underline{\psi}_\infty^i\right)\right] \\
>&\liminf_{t\to\infty}\left[Q^i\left(\boldsymbol{e_t}, A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e_t}, \tilde{a}^i, \Psi^i - \frac{\overline{\kappa}_a}{\underline{\kappa}_\psi}\left(\tilde{a}^i - A^i\right)\right)\right] \\
>&-\overline{\kappa}_a\left(\tilde{a}^i - A^i\right) + \underline{\kappa}_\psi\frac{\overline{\kappa}_a}{\underline{\kappa}_\psi}\left(\tilde{a}^i - A^i\right) = 0.
\end{aligned}$$

It then follows from Lemma 7 that a small neighborhood of $\underline{\psi}_\infty^i$ will be assigned zero probability by the agent almost surely in the long run, which is a contradiction to the definition of $\underline{\psi}_\infty^i$. Hence, $\Phi - \frac{\overline{\kappa}_a}{\underline{\kappa}_\psi}\left(\tilde{a}^i - A^i\right) \le \underline{\psi}_\infty^i$. Analogously we can prove $\overline{\psi}_\infty^i \le \Psi^i$. □

Define

$$\begin{aligned}
\underline{e}_\infty^i &:= \sup\left\{e^i : e^i \le \liminf_{t\to\infty} e_t^i \text{ almost surely}\right\}, \\
\overline{e}_\infty^i &:= \inf\left\{e^i : e^i \ge \limsup_{t\to\infty} e_t^i \text{ almost surely}\right\}.
\end{aligned}$$

Define $E_\infty := \left[\underline{e}_\infty^1, \overline{e}_\infty^1\right] \times \left[\underline{e}_\infty^2, \overline{e}_\infty^2\right]$, which is the set of efforts that may be taken by agents in the long run. In addition, define $E_\infty^D := \left\{\boldsymbol{e} : \exists\boldsymbol{\psi} \in \left[\underline{\psi}_\infty^1, \overline{\psi}_\infty^1\right] \times \left[\underline{\psi}_\infty^2, \overline{\psi}_\infty^2\right], \forall i, \text{s.t. } \boldsymbol{e^*}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right) = \boldsymbol{e}\right\}$, which is the set of action profiles that constitutes a Nash equilibrium when both agents hold degenerate beliefs in the unknown variables that are in the support of the long-run subjective distribution. The next lemma shows that E_∞ is a subset of E_∞^D.

Lemma 9. $E_\infty \subseteq E_\infty^D$.

Proof. By definition, $\boldsymbol{e_t}$ satisfies

$$\tilde{\mathbb{E}}_{\pi_{t-1}^i}\left(Q_{e^i}^i\left(\boldsymbol{e_t}, \tilde{a}^i, \psi^i\right)\right) = 0, \forall i.$$

Continuity of $Q_{e^i}^i$ implies that there exists $\hat{\boldsymbol{\psi}}_{\boldsymbol{t}} \in \operatorname{supp}(\Pi_t^i)$ such that $\forall i$,

$$Q_{e^i}^i\left(\boldsymbol{e_t}, \tilde{a}^i, \hat{\psi}_t^i\right) = 0.$$

We know that the support of Π_t^i is contained in $\left[\underline{\psi}_\infty^i, \overline{\psi}_\infty^i\right]$ when t is large enough almost surely. So $\hat{\psi}_t^i$ lies inside the support of Π_t^i when t is large. Therefore, when t is large, $\underline{\psi}_\infty^i \leq \psi_t^i \leq \overline{\psi}_\infty^i, \forall i$ almost surely, implying that $\boldsymbol{e_t} \in E_\infty^D$ almost surely. Hence, $E_\infty \subseteq E_\infty^D$. □

Lemma 10. *$\frac{\partial e^{*i}(\tilde{\boldsymbol{a}}, \boldsymbol{\psi})}{\partial \psi^i}$ has the same sign as $Q_{e^i\psi^i}^i$, while $\frac{\partial e^{*j}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}^{\boldsymbol{i}}\right)}{\partial \psi^i}$ has the same sign as $Q_{e^i\psi^i}^i Q_{e^i e^j}^j$.*

Proof. Note that

$$Q_{e^i}^i\left(\boldsymbol{e^*}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right), \tilde{a}^i, \psi^i\right) = 0, \forall i.$$

Therefore,

$$\frac{\partial e^{*i}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right)}{\partial \psi^i} = \frac{-Q_{e^i\psi^i}^i Q_{e^j e^j}^j}{Q_{e^i e^i}^i Q_{e^j e^j}^j - Q_{e^i e^j}^i Q_{e^i e^j}^j}, \quad \frac{\partial e^{*j}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right)}{\partial \psi^i} = \frac{Q_{e^i\psi^i}^i Q_{e^i e^j}^j}{Q_{e^i e^i}^i Q_{e^j e^j}^j - Q_{e^i e^j}^i Q_{e^i e^j}^j}.$$

It follows that $\frac{\partial e^{*i}(\tilde{\boldsymbol{a}}, \boldsymbol{\psi})}{\partial \psi^i}$ has the same sign as $Q_{e^i\psi^i}^i$ and $\frac{\partial e^{*j}(\tilde{\boldsymbol{a}}, \boldsymbol{\psi})}{\partial \psi^i}$ has the same sign as $Q_{e^i\psi^i}^i Q_{e^i e^j}^j$. □

Proof of Theorem 1. By Lemma 7 and the continuity of Q^i, it must be true that $\liminf_{t\to\infty} m_t^i\left(\underline{\psi}_\infty^i\right) \leq 0$ almost surely, because otherwise Lemma 7 implies that there exists a small neighborhood around $\underline{\psi}_\infty^i$ to which agent i almost surely assigns probability 0 in the limit, which contradicts the definition of $\underline{\psi}_\infty^i$. Analogously, it must be that $\limsup_{t\to\infty} m_t^i\left(\overline{\psi}_\infty^i\right) \geq 0$.

Case (i): Assume for now that both agents are overconfident and create positive externalities. It is sufficient to show convergence under the assumption that $Q_{e^i\psi^i}^i > 0$ and $Q_{e^i a^i}^i \leq 0$. Differentiating g with respect to e^j, we obtain

$$\frac{\partial g^i\left(\boldsymbol{e}, \psi^i\right)}{\partial e^j} = Q_{e^j}^i\left(\boldsymbol{e}, A^i, \Psi^i\right) - Q_{e^j}^i\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right). \tag{C.1}$$

Since $Q_{e^k\psi^i}^i > 0, Q_{e^k a^i}^i \leq 0, \forall i, k$, we know that $\frac{\partial g^i(\boldsymbol{e}, \psi^i)}{\partial e^k} > 0$, $\forall i, k$. By Lemma 10, since $E_\infty \subseteq E_\infty^D$, we have $\boldsymbol{e^*}\left(\tilde{\boldsymbol{a}}, \underline{\boldsymbol{\psi}}_\infty\right) \leq \boldsymbol{e_t} \leq \boldsymbol{e^*}\left(\tilde{\boldsymbol{a}}, \overline{\boldsymbol{\psi}}_\infty\right)$ almost surely when t is large enough. Hence,

$$\begin{aligned}
&0 \geq \liminf_{t\to\infty} m_t^i\left(\underline{\psi}_\infty^i\right) = \liminf_{t\to\infty} g^i\left(\boldsymbol{e}_t, \underline{\psi}_\infty^i\right) \geq G^i\left(\underline{\boldsymbol{\psi}}_\infty\right), \forall i, \\
&0 \leq \limsup_{t\to\infty} m_t^i\left(\overline{\psi}_\infty^i\right) = \limsup_{t\to\infty} g^i\left(\boldsymbol{e}_t, \overline{\psi}_\infty^i\right) \leq G^i\left(\overline{\boldsymbol{\psi}}_\infty\right), \forall i.
\end{aligned} \tag{C.2}$$

Since $\boldsymbol{e^*}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right)$ is increasing in ψ^j and $\frac{\partial g^i(\boldsymbol{e}, \psi^i)}{\partial e^k} \geq 0$, $\forall i, k$, we know that $G^i\left(\boldsymbol{\psi}\right) = g^i\left(\boldsymbol{e^*}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right), \psi^i\right)$ is increasing in ψ^j. Therefore, $G^i\left(\underline{\psi}_\infty^i, \psi^j\right) \leq 0$ for any $\psi^j \leq \underline{\psi}_\infty^j$ and $G^i\left(\overline{\psi}_\infty^i, \psi^j\right) \geq 0$ for any $\psi^j \geq \overline{\psi}_\infty^j$. Moreover, by Lemma 8, for all i, ψ^j, we have $G^i\left(\Psi^i, \psi^j\right) < 0$ and $G^i\left(\Psi^i - \frac{\overline{\kappa}_a}{\underline{\kappa}_\psi}\left(\tilde{a}^i - A^i\right), \psi^j\right) > 0$. For convenience, let $\underline{\psi}^i = \Psi^i - \frac{\overline{\kappa}_a}{\underline{\kappa}_\psi}\left(\tilde{a}^i - A^i\right)$ and $\overline{\psi}^i = \Psi^i$ for all i. The above results can be summarized by:

$$\begin{aligned}
&\boldsymbol{G}\left(\underline{\boldsymbol{\psi}}_\infty\right) \leq 0, \boldsymbol{G}\left(\underline{\boldsymbol{\psi}}\right) > 0, \boldsymbol{G}\left(\overline{\boldsymbol{\psi}}_\infty\right) \geq 0, \boldsymbol{G}\left(\overline{\boldsymbol{\psi}}\right) < 0, \\
&G^i\left(\underline{\psi}^i, \underline{\psi}_\infty^j\right) > 0, G^j\left(\underline{\psi}^i, \underline{\psi}_\infty^j\right) \leq 0, \\
&G^i\left(\underline{\psi}_\infty^i, \underline{\psi}^j\right) \leq 0, G^j\left(\underline{\psi}_\infty^i, \underline{\psi}^j\right) > 0,
\end{aligned} \tag{C.3}$$

$$G^i\left(\overline{\psi}^i, \overline{\psi}^j_\infty\right) < 0, G^j\left(\overline{\psi}^i, \overline{\psi}^j_\infty\right) \geq 0,$$
$$G^i\left(\overline{\psi}^i_\infty, \overline{\psi}^j\right) \geq 0, G^j\left(\overline{\psi}^i_\infty, \overline{\psi}^j\right) < 0.$$

By Brouwer's fixed point theorem, there exist $\hat{\boldsymbol{\psi}}, \tilde{\boldsymbol{\psi}}$ such that $\tilde{\boldsymbol{\psi}} \in \left[\boldsymbol{\Psi} - \frac{\underline{\kappa}_a}{\underline{\kappa}_\psi}\left(\tilde{\boldsymbol{a}} - \boldsymbol{A}\right), \underline{\boldsymbol{\psi}}_\infty\right]$, $\hat{\boldsymbol{\psi}} \in \left[\overline{\boldsymbol{\psi}}_\infty, \boldsymbol{\Psi}\right]$, and $\boldsymbol{G}\left(\hat{\boldsymbol{\psi}}\right) = \boldsymbol{G}\left(\tilde{\boldsymbol{\psi}}\right) = 0$. Because the root of $\boldsymbol{G}(\boldsymbol{\psi}) = 0$ is unique by assumption, it must be that $\hat{\boldsymbol{\psi}} = \tilde{\boldsymbol{\psi}} = \underline{\boldsymbol{\psi}}_\infty = \overline{\boldsymbol{\psi}}_\infty = \boldsymbol{\psi}_\infty$ and thus $E_\infty = E^D_\infty = \{\boldsymbol{e}_\infty\}$.

Case (ii): Assume instead that both agents are overconfident create negative informational externalities. Again assume that $Q^i_{e^i\psi^i} > 0$ and $Q^i_{e^ia^i} \leq 0$. Analogous to Case (i), we will derive a contradiction if $\underline{\boldsymbol{\psi}}_\infty \neq \overline{\boldsymbol{\psi}}_\infty$. Since informational externalities are negative, the signs of Eq. (C.1) are different: $\frac{\partial g^i(\boldsymbol{e},\psi^i)}{\partial e^i} > 0$ and $\frac{\partial g^i(\boldsymbol{e},\psi^i)}{\partial e^j} < 0$. By Lemma 10, when t is large enough, $e^{*i}\left(\tilde{\boldsymbol{a}}, \left(\underline{\psi}^i_\infty, \overline{\psi}^j_\infty\right)\right) \leq e^i_t \leq e^{*i}\left(\tilde{\boldsymbol{a}}, \left(\overline{\psi}^i_\infty, \underline{\psi}^j_\infty\right)\right), \forall i$. Hence,

$$0 \geq \liminf_{t\to\infty} g^i\left(\boldsymbol{e}_t, \underline{\psi}^i_\infty\right) = \liminf_{t\to\infty} m^i_t\left(\underline{\psi}^i_\infty\right) \geq G^i\left(\underline{\psi}^i_\infty, \overline{\psi}^j_\infty\right), \forall i,$$
$$0 \leq \limsup g^i\left(\boldsymbol{e}_t, \overline{\psi}^i_\infty\right) = \limsup_{t\to\infty} m^i_t\left(\overline{\psi}^i_\infty\right) \leq G^i\left(\overline{\psi}^i_\infty, \underline{\psi}^j_\infty\right), \forall i.$$

In addition, $G^i(\boldsymbol{\psi}) = g^i\left(\boldsymbol{e}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right), \psi^i\right)$ is decreasing in ψ^j. Therefore, we obtain some inequalities different from those in Eq. (C.3):

$$G^i\left(\underline{\psi}^i_\infty, \overline{\psi}^j_\infty\right) \leq 0, G^j\left(\underline{\psi}^i_\infty, \overline{\psi}^j_\infty\right) \geq 0,$$
$$G^i\left(\underline{\psi}^i, \overline{\psi}^j\right) \geq 0, G^j\left(\underline{\psi}^i, \overline{\psi}^j\right) \leq 0,$$
$$G^i\left(\underline{\psi}^i, \overline{\psi}^j_\infty\right) \geq 0, G^j\left(\underline{\psi}^i, \overline{\psi}^j_\infty\right) \geq 0,$$
$$G^i\left(\underline{\psi}^i_\infty, \overline{\psi}^j\right) \leq 0, G^j\left(\underline{\psi}^i_\infty, \overline{\psi}^j\right) \leq 0,$$
$$G^i\left(\overline{\psi}^i_\infty, \underline{\psi}^j_\infty\right) \geq 0, G^j\left(\overline{\psi}^i_\infty, \underline{\psi}^j_\infty\right) \leq 0,$$
$$G^i\left(\overline{\psi}^i, \underline{\psi}^j\right) \leq 0, G^j\left(\overline{\psi}^i, \underline{\psi}^j\right) \geq 0,$$
$$G^i\left(\overline{\psi}^i, \underline{\psi}^j_\infty\right) \leq 0, G^j\left(\overline{\psi}^i, \underline{\psi}^j_\infty\right) \leq 0,$$
$$G^i\left(\overline{\psi}^i_\infty, \underline{\psi}^j\right) \geq 0, G^j\left(\overline{\psi}^i_\infty, \underline{\psi}^j\right) \geq 0.$$

Again, there exist two different roots to $\boldsymbol{G}(\boldsymbol{\psi}) = 0$ if $\underline{\boldsymbol{\psi}}_\infty \neq \overline{\boldsymbol{\psi}}_\infty$, which are located in $[\underline{\psi}^i, \underline{\psi}^i_\infty] \times [\overline{\psi}^j_\infty, \overline{\psi}^j]$ and $[\overline{\psi}^i_\infty, \overline{\psi}^i] \times [\underline{\psi}^j, \underline{\psi}^j_\infty]$ respectively. This contradicts the assumption of a unique steady state. □

Appendix D. Proofs for Section 6

Proof of Theorem 2. We prove the result when there are positive externalities, $Q^i_{e^i\psi^i} > 0$ and $Q^i_{e^ia^i} \leq 0$; the proof for other cases is analogous.

Similarly define $\overline{\psi}^i = \Psi^i - \frac{\underline{\kappa}_a}{\underline{\kappa}_\psi}\left(\tilde{a}^i - A^i\right)$ and $\underline{\psi}^i = \Psi^i$ for all i. Since the agents are underconfident, we now have $\frac{\partial g^i(\boldsymbol{e},\psi^i)}{\partial e^k} < 0, \forall i, k$. With positive informational externalities we again have $\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \underline{\boldsymbol{\psi}}_\infty\right) \leq \boldsymbol{e_t} \leq \boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \overline{\boldsymbol{\psi}}_\infty\right)$ when t is large enough. Therefore,

$$0 \geq \liminf_{t\to\infty} m^i_t\left(\underline{\psi}^i_\infty\right) \geq Q^i\left(\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \overline{\boldsymbol{\psi}}_\infty\right), A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \overline{\boldsymbol{\psi}}_\infty\right), \tilde{a}^i, \underline{\psi}^i_\infty\right), \forall i,$$
$$0 \leq \limsup_{t\to\infty} m^i_t\left(\overline{\psi}^i_\infty\right) \leq Q^i\left(\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \underline{\boldsymbol{\psi}}_\infty\right), A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \underline{\boldsymbol{\psi}}_\infty\right), \tilde{a}^i, \overline{\psi}^i_\infty\right), \forall i.$$

Therefore, $g^i\left(\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \overline{\boldsymbol{\psi}}_\infty\right), \underline{\psi}^i_\infty\right) \leq 0 \leq g^i\left(\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \underline{\boldsymbol{\psi}}_\infty\right), \overline{\psi}^i_\infty\right), \forall i$. For each i, define $h^i(\boldsymbol{\zeta}, \boldsymbol{\psi}) := g^i\left(\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \boldsymbol{\zeta}\right), \psi^i\right)$. Differentiate h^i with respect to ζ^k and ψ^k, we obtain

$$\frac{\partial h^i(\boldsymbol{\zeta},\boldsymbol{\psi})}{\partial \zeta^k} = \left(Q^{i,A}_{e^i} - Q^i_{e^i}\right)\frac{\partial e^{*i}(\tilde{\boldsymbol{a}},\boldsymbol{\zeta})}{\partial \zeta^k} + \left(Q^{i,A}_{e^j} - Q^i_{e^j}\right)\frac{\partial e^{*j}(\tilde{\boldsymbol{a}},\boldsymbol{\zeta})}{\partial \zeta^k}$$
$$\frac{\partial h^i(\boldsymbol{\zeta},\boldsymbol{\psi})}{\partial \psi^i} = -Q^i_{\psi^i}, \quad \frac{\partial h^i(\boldsymbol{\zeta},\boldsymbol{\psi})}{\partial \psi^j} = 0,$$

where $Q^{i,A}_{e^k}$ which denotes the derivative of Q^i w.r.t. e^k evaluated at $\boldsymbol{e^*}(\tilde{\boldsymbol{a}},\boldsymbol{\zeta}), \boldsymbol{A}, \boldsymbol{\psi}$. Hence, when $(\tilde{\boldsymbol{a}},\boldsymbol{\zeta},\boldsymbol{\psi}) = (\boldsymbol{A},\boldsymbol{\Psi},\boldsymbol{\Psi})$,

$$\frac{\partial h^i(\boldsymbol{\zeta},\boldsymbol{\psi})}{\partial \zeta^i} = 0, \frac{\partial h^i(\boldsymbol{\zeta},\boldsymbol{\psi})}{\partial \zeta^j} = 0, \frac{\partial h^i(\boldsymbol{\zeta},\boldsymbol{\psi})}{\partial \psi^i} = -Q^i_{\psi^i}, \frac{\partial h^i(\boldsymbol{\zeta},\boldsymbol{\psi})}{\partial \psi^j} = 0.$$

There thus exists δ such that when $\tilde{\boldsymbol{a}} \in B_\delta(\boldsymbol{A})$: (i) the beliefs are also restricted to a small neighborhood, i.e. $\overline{\boldsymbol{\psi}}_\infty, \underline{\boldsymbol{\psi}}_\infty$ are close to $\boldsymbol{\Psi}$; (ii) for all i and $j \neq i$, $\frac{\partial h^i(\zeta,\psi)}{\partial \psi^i} < -\frac{1}{2}\underline{\kappa}_\psi < 0$, and $\left|\frac{\partial h^i(\zeta,\psi)}{\partial \zeta^i}\right|, \left|\frac{\partial h^i(\zeta,\psi)}{\partial \zeta^j}\right| < \frac{1}{4}\underline{\kappa}_\psi, \left|\frac{\partial h^i(\zeta,\psi)}{\partial \psi^j}\right| = 0$. Therefore,

$$0 \ge g^i\left(\boldsymbol{e^*}\left(\tilde{\boldsymbol{a}},\overline{\boldsymbol{\psi}}_\infty\right),\underline{\psi}^i_\infty\right) - g^i\left(\boldsymbol{e^*}\left(\tilde{\boldsymbol{a}},\underline{\boldsymbol{\psi}}_\infty\right),\overline{\psi}^i_\infty\right) \ge \frac{1}{4}\underline{\kappa}_\psi\left(\overline{\psi}^i_\infty - \underline{\psi}^i_\infty - \overline{\psi}^j_\infty + \underline{\psi}^j_\infty\right),$$
$$0 \ge g^j\left(\boldsymbol{e^*}\left(\tilde{\boldsymbol{a}},\overline{\boldsymbol{\psi}}_\infty\right),\underline{\psi}^j_\infty\right) - g^j\left(\boldsymbol{e^*}\left(\tilde{\boldsymbol{a}},\underline{\boldsymbol{\psi}}_\infty\right),\overline{\psi}^j_\infty\right) \ge \frac{1}{4}\underline{\kappa}_\psi\left(\overline{\psi}^j_\infty - \underline{\psi}^j_\infty - \overline{\psi}^i_\infty + \underline{\psi}^i_\infty\right),$$

which hold at the same time if and only if $\overline{\boldsymbol{\psi}}_\infty = \underline{\boldsymbol{\psi}}_\infty = \boldsymbol{\psi}_\infty$. $\square$

References

Anderson, C., Brion, S., Moore, D.A., Kennedy, J.A., 2012. A status-enhancement account of overconfidence. J. Pers. Soc. Psychol. 103 (4), 718.
Ba, C., 2023. Robust Misspecified Models and Paradigm Shifts. Working paper.
Benoît, J.-P., Dubra, J., Moore, D.A., 2015. Does the better-than-average effect show that people are overconfident?: two experiments. J. Eur. Econ. Assoc. 13 (2), 293–329.
Bils, P., 2020. Policymaking with multiple agencies. Am. J. Polit. Sci. 64 (3), 634–648.
Bohren, J.A., 2016. Informational herding with model misspecification. J. Econ. Theory 163, 222–247.
Bohren, J.A., Hauser, D.N., 2021. Learning with heterogeneous misspecified models: characterization and robustness. Econometrica 89 (6), 3025–3077.
Camerer, C., Lovallo, D., 1999. Overconfidence and excess entry: an experimental approach. Am. Econ. Rev. 89 (1), 306–318.
Campbell, W.K., Sedikides, C., 1999. Self-threat magnifies the self-serving bias: a meta-analytic integration. Rev. Gen. Psychol. 3 (1), 23–43.
Esponda, I., Pouzo, D., 2016. Berk-Nash equilibrium: a framework for modeling agents with misspecified models. Econometrica 84 (3), 1093–1130.
Esponda, I., Pouzo, D., Yamamoto, Y., 2021. Asymptotic behavior of Bayesian learners with misspecified models. J. Econ. Theory 195, 105260.
Eyster, E., Rabin, M., 2010. Naive herding in rich-information settings. Am. Econ. J. Microecon. 2 (4), 221–243.
Frick, M., Iijima, R., Ishii, Y., 2020. Misinterpreting others and the fragility of social learning. Econometrica 88 (6), 2281–2328.
Frick, M., Iijima, R., Ishii, Y., 2023. Belief convergence under misspecified learning: a martingale approach. Rev. Econ. Stud. 90 (2), 781–814.
Fudenberg, D., Lanzani, G., Strack, P., 2021. Limit points of endogenous misspecified learning. Econometrica 89 (3), 1065–1098.
Fudenberg, D., Romanyuk, G., Strack, P., 2017. Active learning with a misspecified prior. Theor. Econ. 12 (3), 1155–1189.
Gagnon-Bartsch, T., 2017. Taste Projection in Models of Social Learning. Working paper.
Gagnon-Bartsch, T., Rabin, M., 2017. Naive Social Learning, Mislearning, and Unlearning. Working paper.
Gervais, S., Goldstein, I., 2007. The positive effects of biased self-perceptions in firms. Rev. Finance 11 (3), 453–496.
He, K., 2022. Mislearning from censored data: the Gambler's fallacy and other correlational mistakes in optimal-stopping problems. Theor. Econ. 17 (3), 1269–1312.
Heidhues, P., Kőszegi, B., Strack, P., 2018. Unrealistic expectations and misguided learning. Econometrica 86 (4), 1159–1214.
Heidhues, P., Kőszegi, B., Strack, P., 2021. Convergence in models of misspecified learning. Theor. Econ. 16 (1), 73–99.
Heifetz, A., Shannon, C., Spiegel, Y., 2007. What to maximize if you must. J. Econ. Theory 133 (1), 31–57.
Kennedy, J.A., Anderson, C., Moore, D.A., 2013. When overconfidence is revealed to others: testing the status-enhancement theory of overconfidence. Organ. Behav. Hum. Decis. Process. 122 (2), 266–279.
Langer, E.J., Roth, J., 1975. Heads I win, tails it's chance: the illusion of control as a function of the sequence of outcomes in a purely chance task. J. Pers. Soc. Psychol. 32 (6), 951.
Miller, D.T., Ross, M., 1975. Self-serving biases in the attribution of causality: fact or fiction? Psychol. Bull. 82 (2), 213.
Moore, D.A., Healy, P.J., 2008. The trouble with overconfidence. Psychol. Rev. 115 (2), 502.
Moore, D.A., Tenney, E.R., Haran, U., 2015. Overprecision in judgment. In: The Wiley Blackwell Handbook of Judgment and Decision Making, vol. 2, pp. 182–209.
Murooka, T., Yamamoto, Y., 2021. Multi-Player Bayesian Learning with Misspecified Models. OSIPP Discussion Paper 21E001. Osaka School of International Public Policy, Osaka University.
Nyarko, Y., 1991. Learning in mis-specified models and the possibility of cycles. J. Econ. Theory 55 (2), 416–427.
Rabin, M., Schrag, J.L., 1999. First impressions matter: a model of confirmatory bias. Q. J. Econ. 114 (1), 37–82.
Ross, M., Sicoly, F., 1979. Egocentric biases in availability and attribution. J. Pers. Soc. Psychol. 37 (3), 322.
Svenson, O., 1981. Are we all less risky and more skillful than our fellow drivers? Acta Psychol. 47 (2), 143–148.