A Multi-Agent Model of Misspecified Learning with Overconfidence*

Cuimin Ba [†] Alice Gindin [‡]

First version: Aug 27, 2019
This version: May 23, 2022

Abstract

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.

*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, and participants of multiple conferences and seminars at the University of Pennsylvania for helpful comments and suggestions.

[†]University of Pennsylvania. Email: cuiminba@sas.upenn.edu.
[‡]University of Pennsylvania. Email: agindin@sas.upenn.edu.

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 outside 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 overconfidence of entrepreneurs can lead to excessive business entry and losses. Heidhues, Kőszegi, and Strack (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 suboptimal effort, leading to a welfare loss. However, when working on a task within a team, all members of the team learn about their decision 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.

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 of the engineers knows the underlying quality of the project assigned by their supervisor and thus they learn about this quality over time by working on the project and being evaluated 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 will find himself considerably more disappointed by the new output that 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 additionally 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. We call the mechanism where the presence of a second engineer simultaneously adjusting his effort causes the first engineer's beliefs to become more extreme *mutually-reinforcing* learning. We use this terminology because the second engineer's effort is reinforcing the distortions that overconfidence creates for the first engineer and vice versa. However, unlike the single-engineer case, it is now possible that the extra efforts lead 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. Similarly to 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 due to overconfidence. If in addition efforts are substitutes, then 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, we assume that agents either learn about the fundamental or the other agent's ability. In particular, each agent starts with a non-degenerate prior about either the fundamental or the other's ability and updates this belief over time. 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 single-agent case. First, we find that agent j's effort not only generates a positive payoff externality for agent i, it also 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. When informational externalities are *positive*, the agents' learning processes are *mutually-reinforcing* in the following sense: as more agents are permitted to adjust their effort according to myopic optimality, or as any agent becomes more overconfident, the inferences of all agents become more extreme—underestimation gets more severe. By contrast, if informational externalities are *negative*, the learning processes are *mutually-limiting* and the impact of overconfidence is alleviated. Interestingly, 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.

The second insight which the two-agent model highlights is how the presence of a second agent impacts the long-run welfare of the first agent. 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. We then show that under certain conditions, agents

[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 project.

converge to this Berk-Nash equilibrium, meaning that agents choose the optimal amount of effort with respect to a belief that best fits their observations. Our proof augments the contraction argument in Heidhues, Kőszegi, and Strack (2018) to accommodate the additional agent. For convergence to hold, agents' informational externalities must be either positive or negative so that one agent's optimization does not impede the other agent's belief updating and lead to oscillation.

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 information externalities lead to mutually-limiting learning while negative information externalities lead to mutually-reinforcing learning.

Related Literature

This paper builds on the single-agent learning setting in Heidhues, Kőszegi, and Strack (2018). They find that overconfidence leads to distorted beliefs and 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) also identify similar informational externalities in a multi-agent setting. The key difference between the papers is that the 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 equi-

[2]The consequences of misspecified models have been investigated in various settings other than overconfidence. For example, 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, 2018), misspecified beliefs about the type distribution (Frick, Iijima, and Ishii, 2019), misspecified prior beliefs (Nyarko, 1991; Fudenberg, Romanyuk, and Strack,

librium, 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 (2019) characterize conditions under which correct learning, incorrect learning, or cyclical learning arise in a binary-state learning environment. Heidhues, Kőszegi, and Strack (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, Pouzo, and Yamamoto (2019) 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, Lanzani, and Strack (2020) 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, Iijima, and Ishii (2021) 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, Kőszegi, and Strack (2018) rely on structural properties of the payoff functions.

In line with our findings, the literature on overconfidence suggests that overconfidence can be helpful or detrimental depending on the context. For example, Camerer and Lovallo (1999) use experiments to simulate entrepreneurs deciding whether or not to start a new business. They find that overconfidence leads to excessive entry followed by large rates of new business failure, consistent with the high rates of new business failure that Dunne, Roberts, and Samuelson (1988) find using US plant-level data. On the other hand, Gervais and Goldstein (2007) show that overconfidence can improve the welfare of all team members in a compensation contract problem. They focus on a one-period problem and do not allow agents to learn about other fundamentals.

Finally, this paper relies on the assumption that agents tend to be persistently overconfident about their abilities, which is well supported by the psychology literature on overconfidence.[3] One illustration of this is the "better-than-average effect", which shows that most

2017) have all been shown to lead to inefficient actions in the long run.

[3]There are also studies finding that agents are overconfident in the precision of their beliefs (Moore and

individuals in the population believe themselves to be better than the population average at some skill. For example, Langer and Roth (1975) find that when individuals correctly guess the outcome of a coin flip, they attribute it to skill while attributing incorrect guesses to bad luck. Many individuals thus believed they were particularly skilled at predicting the coin flip despite mounting evidence that they were only correct 50% of the time. Svenson (1981) finds that when a group of truck drivers were asked to compare their driving to the group of drivers surveyed, the vast majority believed they were more skilled and safer at driving than the average driver surveyed. Further, Benoît, Dubra, and Moore (2015) find that people overplaced themselves in their performance on quizzes and show it cannot be explained by a model of rational expected utility maximization. Other papers provide a basis for overconfidence like Anderson, Brion, Moore, and Kennedy (2012), who find that when an individual is overconfident, others perceive them as more competent which in turn leads to higher social status for the individual. This can reinforce feelings of overconfidence, despite contrary evidence. Although this channel provides insight into how individuals might have developed overconfidence initially, we do not introduce it into our model, instead assuming agents are already overconfident. This way we can focus on the misperception's effect on the learning process. As a more theoretical basis for the development of overconfidence, Heifetz, Shannon, and Spiegel (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 mechanism is very closely tied with our welfare results in Section 4.4, where agents' overconfidence can be welfare improving because overconfidence is mitigating 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.

Healy, 2008; Moore, Tenney, and Haran, 2015). We focus on overconfidence in abilities.

2 Multi-Agent Learning Environment

Environment There are two agents, indexed by $i \in I \equiv \{1,2\}$ 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 $\mathbb{R}$.[4] The output function Q is deterministic and twice continuously differentiable, with its derivatives having polynomial growth in a^i, a^j and ϕ.[5] 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 the game. 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 $\tilde{a}^i > A^i$. The self-perceptions of the agents, $\tilde{a}^i$ and $\tilde{a}^j$, are common knowledge. Agents realize that their counterpart may be subject to biases but they have a dogmatic optimistic belief about themselves.[6]

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.[7] The psychology literature documents self-serving attribution biases in both environments.[8] For

[4]We assume log-concavity, i.e. the second-order derivative of $\log f(\epsilon)$ is strictly negative and bounded from below. This technical assumption is to ensure the subjective beliefs of any agent are well-defined and have finite moments after any history.

[5]Function $Q\left(e^i, e^j, a^i, a^j, \phi\right)$ is of polynomial growth in ϕ if for any e^i, e^j, a^i, a^j, there are $\kappa, k, b > 0$ such that $|Q\left(e^i, e^j, a^i, a^j, \phi\right)| \leq \kappa|\phi|^k + b$. We require similar properties to hold for a^i and a^j. This ensures that the expected output and its derivatives exist after arbitrary history.

[6]For evidence that agents can maintain overconfidence while faced with peers who demonstrate a belief that the agent is overconfident, consider Kennedy, Anderson, and Moore (2013). In lab experiments, the authors show that overconfident individuals maintained nearly consistent levels of overconfidence in their ability to answer trivia questions after learning others felt they overestimated their ability. Moreover individuals maintained overconfidence after recognizing that their peers were overconfident.

[7]We do not consider the case where agents learn about both the common fundamental and each other's ability simultaneously. In this alternative learning environment, the agents' subjective models are under-identified, giving rise to a continuum of Berk-Nash equilibria, each supported by a different joint belief over ϕ and a^j. A previous version of this paper considers this setup, but we had to assume agents received two different output signals to avoid under-identification.

[8]See Miller and Ross (1975) for a review of the early empirical studies. Subjects were found to make

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 about 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 continuously 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 assumption says that a higher ability and a larger fundamental positively influence the common output. The second assumption guarantees both agents are optimizing and making inference in a predictable direction. For example, consider the engineer who, as a consequence of overconfidence in his ability, underestimates the quality of a project idea (Case 1), then $\psi^i = \phi$. Suppose $Q_{e^i \phi} > 0$ and $Q_{e^i a^i} \leq 0$. Then evidently this agent should lower his effort in response. If instead both cross derivatives are positive, then more structure is needed to determine how he best responds.[9] We assume $Q_{e^i \psi^i} \neq 0$ to rule out

self-serving attributions in interpersonal influence settings (e.g. teacher-student interactions) where they could attribute the outcomes to another person and in skill tasks where they could attribute to the luck or the overall ability of the team they were in.

[9] Heidhues, Kőszegi, and Strack (2018) also make this assumption. We discuss the role of this assumption in more detail at the end of Section 4.3.

the uninteresting case where agents always exert the same amount of efforts. The assumption that the signs of cross derivatives do not change with i is without loss of generality.[10] If Q is symmetric for agents i and j, then this assumption is automatically satisfied. The third assumption guarantees that, given any fixed action profile, agent i can always perfectly justify the distribution of the outputs by attaching probability 1 to a wrong fundamental value or a wrong ability level of his teammate.

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, for agent i to maximize his stage payoff, he must form some beliefs over what action player j is going to play. With dominance solvability it is clear how the conjecture about agent j's action is formed; player 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. Further, this is equivalent to assuming that agents play a Nash equilibrium each period, which boils down to the following restriction: agents choose efforts $\{e^1_t, e^2_t\}$, 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) that players assume they are in a stationary environment, we model players to be a little more sophisticated so that they understand that 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 since

[10] 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.

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—it is hard to isolate how inferences are affected by overconfidence over time when the agents are also misspecified about the game structure.

Timing In period t, each agent $i \in I$ chooses effort e_t^i according to his belief π_{t-1}^i. After observing output q_t, each agent updates his posterior to π_t^i and enters the next period.

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 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 $Q\left(e^i, e^j, a^i, a^j, \phi\right) = \phi\left(e^i + e^j + e^i e^j + a^i + a^j\right)$ and $c^i(e^i) = \frac{1}{2}\kappa(e^i)^2$, where $\kappa \geq \overline{\phi}$. No effort from the engineers lead to no output from the projects and therefore no bonuses. In this example the agents will learn about the fundamental, ϕ, which governs the productivity of the project ($\psi^i = \phi$). Notice that agent i's effort and the fundamental are complements—a higher belief in the fundamental motivates a greater input of effort.[11] The efforts of the agents are complements too as they share knowledge and experience gained from reading articles or testing out different methods.[12]

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 $Q(e^i, e^j, a^i, a^j, \phi) = \log(e^i a^i + e^j a^j + \phi)$. 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.[13]

[11] This differs from the motivating example in the introduction where we assume the common fundamental and effort are substitutes. Assume complementarity here allows us to use a simpler functional form.

[12] Note that Assumption 1(ii) is satisfied since $Q_{e^i a^i} = 0$ and $Q_{e^i \psi^i} = Q_{e^i \phi} > 0$.

[13] Note that Assumption 1(ii) is satisfied since $Q_{e^i a^i} > 0$ and $Q_{e^i \psi^i} = Q_{e^i a^j} < 0$.

Example 3. The legislature passes a law which must be implemented by two federal agencies who each work together to create a series of rules that enforce different aspects of the law.[14] The two agencies learn about the underlying quality of the law, ϕ, while dedicating effort e^i towards writing each rule. Each agency is overconfident in their 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 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) if the other agency also works harder (higher e^j).[15]

3 Steady State

Following the definition developed in Esponda and Pouzo (2016), we now define 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.[16]

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,

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

[14]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 specifically securities based swaps while the CFTC writes the rules for all other types. The agencies have similar policy goals and often share information information in order to create better and more consistent rules (see Bils (2020)).

[15]Note that Assumption 1(ii) is satisfied since $Q_{e^i a^i} = 0$ and $Q_{e^i \psi^i} = Q_{e^i \phi} > 0$.

[16]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.

in Case 2,

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

(ii) For all ψ^i in the support of π^i,

$$\psi^i \in \arg\min_{\psi^{i\prime}} K^i\left(\boldsymbol{e}, \psi^{i\prime}\right) \tag{3}$$

where the KL divergence $K^i\left(\boldsymbol{e}, \psi^{i\prime}\right)$ is given by the following in Case 1,

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

and the following in Case 2,

$$\mathbb{E}\left[\log \frac{f(\epsilon)}{f\left(Q\left(\boldsymbol{e}, A^i, A^j, \Phi\right) - Q\left(\boldsymbol{e}, \tilde{a}^i, \psi^{i\prime}, \Phi\right) + \epsilon\right)}\right], \tag{5}$$

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, they must be best responses to the current belief and the opponent's action; meanwhile, given that efforts converge, an agent must converge to beliefs that best fit the data among all possible beliefs in the long term, which are captured by the KL minimizers. To characterize the equilibrium, let $\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right) \equiv \left(e^{*i}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right), e^{*j}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right)\right)$ be the myopically optimal action profile where each agent i assigns probability 1 to ψ^i. Essentially, 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,

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

and in Case 2,

$$g^i\left(\boldsymbol{e}, \psi^i\right) := Q\left(\boldsymbol{e}, A^i, A^j, \Phi\right) - Q\left(\boldsymbol{e}, \tilde{a}^i, \psi^i, \Phi\right). \tag{7}$$

We will call $\boldsymbol{g}(\boldsymbol{e},\boldsymbol{\psi}) \equiv \left(g^i\left(\boldsymbol{e},\psi^i\right), g^j\left(\boldsymbol{e},\psi^j\right)\right) = 0$ the *no-gap condition*. Intuitively, fixing efforts $\boldsymbol{e}$, the solution to the no-gap condition $\boldsymbol{\psi}$ returns a 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, Kőszegi, and Strack (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.*

We provide a sufficient condition for Assumption 3. Lemma 3 establishes that uniqueness is guaranteed if agents are not too misspecified. We discuss in Section 6 how our insights extend to scenarios where this uniqueness assumption fails.

Lemma 3. *Suppose Assumptions 1 and 2 hold. There exist $\Delta^1, \Delta^2 > 0$, such that whenever $|\tilde{a}^i - A^i| < \Delta^i, \forall i = 1, 2$, 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 (common fundamental or teammate's ability) 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

The well-known notion of payoff externality describes the direct influence of agent j' actions on agent i's utility, such as in 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 choice; 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 $Q_{e^i e^j} \geq 0$, $\operatorname{sgn}(Q_{e^i \psi^i}) = \operatorname{sgn}(Q_{e^j \psi^i})$, and $\operatorname{sgn}(Q_{e^i a^i}) = \operatorname{sgn}(Q_{e^j a^i})$; it is *negative* if $Q_{e^i e^j} \leq 0$, $\operatorname{sgn}(Q_{e^i \psi^i}) \neq \operatorname{sgn}(Q_{e^j \psi^i})$, and $\operatorname{sgn}(Q_{e^i a^i}) \neq \operatorname{sgn}(Q_{e^j a^i})$; otherwise, it is *neither positive or negative.*

The direction of 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 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, 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), we can readily verify for both Case 1 and Case 2 that the informational externalities between the two agents must point to 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 $Q(e^i, e^j, a^i, a^j, \phi) = \phi(e^i + e^j + e^i e^j + a^i + a^j)$. Simple calculations establish that $Q_{e^i e^j} > 0$, $Q_{e^i \psi^i} = Q_{e^i \phi} > 0$, $Q_{e^j \psi^i} = Q_{e^j \phi} > 0$, and $Q_{e^i a^i} = Q_{e^j a^i} = 0$. 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 $Q(e^i, e^j, a^i, a^j, \phi) = \log(e^i a^i + e^j a^j + \phi)$. Simple calculations tell us that $Q_{e^i e^j} < 0$, $Q_{e^i \psi^i} = Q_{e^i a^j} < 0$, $Q_{e^j \psi^i} = Q_{e^j a^j} > 0$, $Q_{e^i a^i} > 0$, $Q_{e^j a^i} < 0$. 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 *iff* efforts are complements; in addition, the agents *cannot* exhibit negative information 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 *iff* efforts are substitutes; in addition, agents *cannot* exhibit positive information externalities. Since we require overconfidence and the resulted underestimation of the unknown to affect agent j's optimal effort in the same direction for both agents, $Q_{e^j a^j}$ and $Q_{e^i a^i}$ have the same sign. By Assumption 1, the sign of $Q_{e^j a^j}$ must disagree with the sign of $Q_{e^j \psi^j}$, or equivalently, the sign of $Q_{e^j a^i}$.[17] Therefore, $sgn(Q_{e^i a^i}) \neq sgn(Q_{e^j a^i})$, ruling out positive informational externalities.

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

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, Kőszegi, and Strack (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 more extreme belief in the unknown variable, thereby encouraging more extreme actions and leading 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^j_S but allow agent i to optimize his action in each period. Maintaining the assumption that there exists a unique steady-state, let's denote the steady-state inferences as $\boldsymbol{\psi_S} = \left(\psi^i_S, \psi^j_S\right)$ and actions as $\boldsymbol{e_S} = \left(e^i_S, e^j_S\right)$.[18] We now compare this steady state with $(\boldsymbol{e_\infty}, \boldsymbol{\psi_\infty})$, 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 is **reinforced** when agent j is free to optimize than when agent j's action is fixed at the level e^j_S, where e^j_S is picked from $[\underline{e}, \overline{e}]$ such that*

(i) if $Q_{e^i\psi^i} > 0$, then $e^j_S > e^j_\infty(\tilde{\boldsymbol{a}})$;

(ii) if $Q_{e^i\psi^i} < 0$, then $e^j_S < e^j_\infty(\tilde{\boldsymbol{a}})$.

In other words, fixing e^j at a level e^j_S satisfying conditions (i) and (ii) implies $\boldsymbol{\psi_\infty}(\tilde{\boldsymbol{a}}) < \boldsymbol{\psi_S} < \boldsymbol{\Phi}$.

The requirement that $e^j_S > e^j_\infty(\tilde{\boldsymbol{a}})$ or $e^j_S < e^j_\infty(\tilde{\boldsymbol{a}})$—they ensures agent j's action is less distorted at the fixed level. In fact, we could replace the condition by $e^j_S = e^j_\infty(\boldsymbol{A})$ or $e^j_S = e^j_\infty(\tilde{a}^i, A^j)$,[19] and then interpret the constraint as a suggestion from an outside analyst who tries to mitigate the distortion due to overconfidence: if agents were to fix their

[18] We can ensure the uniqueness of the steady state with a similar sufficient condition as in Lemma 3.

[19] By Lemma 6, when $Q_{e^i\psi^i}$ is positive (negative) for both i, we have $e^j_\infty\left(\tilde{a}^i, A^j\right)$ and $e^j_\infty(\boldsymbol{A})$ are both larger (smaller) than $(\tilde{\boldsymbol{a}})$.

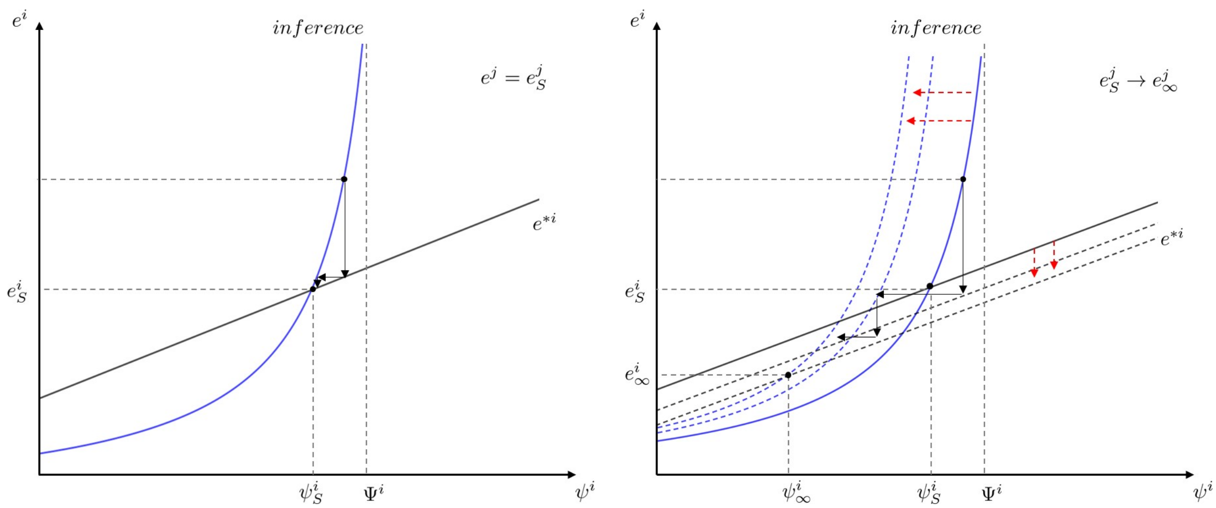

Figure 1: Mutually-reinforcing learning. The left panel shows how allowing agent i to change her action induces a lower inference for agent i, with e^j fixed at e^j_S, and the right panel shows how agent i gets an even lower inference when agent j is also allowed to revise actions. The black straight 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.

actions at a level which best responds to a correct self-perception, then both agents would understand the environment better.

The message conveyed by Proposition 1 is twofold. First, a self-defeating pattern emerges since $\psi^j_\infty(\tilde{\boldsymbol{a}}) < \psi^j_S < \Psi^j$. That is, agent j underestimates the unknown variable more when he is allowed to optimize. More importantly, with a positive informational externality from agent j, a mutually-reinforcing pattern can be observed by noting that $\psi^i_\infty(\tilde{\boldsymbol{a}}) < \psi^i_S < \Psi^i$, which means agent i's inference also turns out to be more extreme when agent j can freely change actions.

To illustrate the mechanism, we describe the learning dynamics heuristically for Case 1 with complementary efforts. For our purpose, we assume $Q_{e^i\psi^i} > 0$ (that is, $Q_{e^i\phi} > 0$) and $Q_{e^ia^i} \leq 0$ 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.

If agent j is forced to take a relatively high effort, e_S^j, and agent i starts to optimize against the high e_S^j as shown in the figure, 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 is going to hold a belief ψ_S^i, which is lower than the belief he started with. This process is shown in the left panel of Fig. 1. 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 exerts lower effort than he did when his coworker was constrained to play a fixed 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

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

Hence, we now have a larger decline in agent i's evaluation of the common fundamental, i.e. the belief formation curve shifts to the left. More intuitively, since the common fundamental ϕ and the teammate's effort e^j are complements, 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 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, which are potentially much more extreme than e_S^i and ψ_S^i.

We now examine mutually-reinforcing learning through another lens. Now that the presence of an actively-optimizing second agent reinforces one agent's mislearning, it should be intuitive that mislearning becomes more severe when the second agent is more biased. Proposition 2 confirms this intuition. As one or both agents become more confident, i.e. have higher self-perceptions, positive informational externalities imply that they each have 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 $\boldsymbol{\psi}_{\boldsymbol{\infty}}(\tilde{\boldsymbol{a}}) < \boldsymbol{\psi}_{\boldsymbol{\infty}}(\tilde{\boldsymbol{a}}') < \boldsymbol{\Psi}$.*

The pattern of mutually-reinforcing learning generates novel policy implications. First of all, 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 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 the informational externalities are negative the 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. The observation in Section 4.1 tells us that mutually-limiting learning arises when agents learn about each other 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_S^j (chosen by the same rule in Proposition 1), i.e. $\psi_S^i < \psi_\infty^i(\tilde{\boldsymbol{a}}) < \Psi$.

(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 $\psi_\infty^i(\tilde{\boldsymbol{a}}') < \psi_\infty^i(\tilde{\boldsymbol{a}}) < \Psi$.

We illustrate mutually-limiting learning in the context of Case 2 with substitutable efforts. For a better comparison 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). Once we allow agent j to freely optimize, however, 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

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

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 ends at action e_∞^i

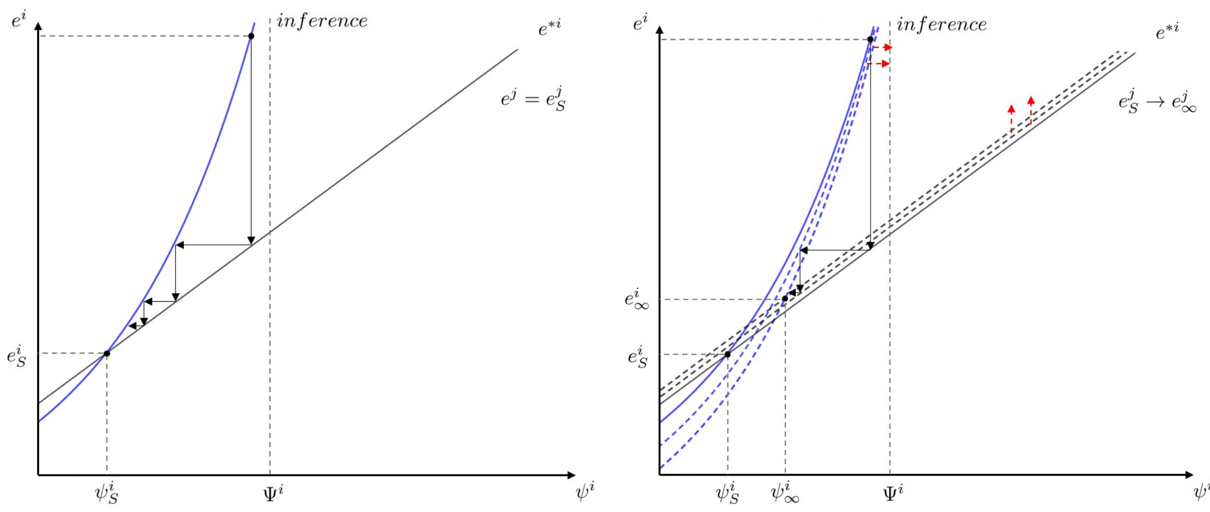

Figure 2: Mutually-limiting learning. The left panel shows the steady state with e^j fixed at e^j_S, and the right panel shows how agent i gets a relatively higher inference when agent j is also allowed to revise actions.

and belief ψ^i_∞ that are less extreme than the single-agent steady state.

In sum, when learning about a common fundamental the agents may exhibit mutually-reinforcing learning but not mutually-limiting learning, while the opposite occurs if agents learn about each other's ability. How much does this observation hinge on the parametric conditions in Assumption 1, i.e. $Q_{e^i a^i}$ shares the same sign as $Q_{e^j a^j}$ but has the opposite sign as $Q_{e^i \psi^i}$? As we mentioned earlier, the former assumption only serves as a normalization. The latter assumption, however, is critically important to mutual learning. Consider an alternate environment in which $sgn(Q_{e^i a^i}) = sgn(Q_{e^i \psi^i}) > 0$, then both the optimal effort curve and the inference formation curve could be non-monotone, and thus the effect of e^j on i's inference could also be non-monotone. Without imposing more structure over the functional form of Q, positive or negative informational externalities in this environment do not directly imply mutually-reinforcing learning or mutually-limiting learning.

4.4 Welfare Analysis

In this subsection, we analyze the welfare implications of overconfidence. We first discuss different sources of welfare impact and briefly analyze the single-agent case where overconfidence almost always leads to utility loss. Then we analyze the two-agent case and obtain 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 when the bias is small by checking a few conditions.

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

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

where $e^i_\infty(\tilde{\boldsymbol{a}}) = e^{*i}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}_{\boldsymbol{\infty}}\left(\tilde{\boldsymbol{a}}\right)\right)$ depend on $\tilde{a}^j$ and $\psi^j_\infty\left(\tilde{\boldsymbol{a}}\right)$ 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*, i.e. the deviation of $\tilde{a}^i$ from A^i; (ii) the distortion of e^i due to *false inference*, i.e. 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 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.[20] We summarize these observations below.

Claim 1. If the output function has the following form

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

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

When there are multiple agents, effects (iii) and (iv) start to kick in. 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) efforts. 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 small amounts of overconfidence. Since e^i has been optimized, the change in agent i's payoff will be dictated by the

[20] One exception in which outputs are unchanged is when the two fundamentals can be summarized by a new variable $\theta \equiv h\left(a^i, \psi^i\right)$. In this case, the agent is correctly specified as long as $\Theta = h\left(A^i, \Psi^i\right)$ is inside the support of his prior. In this setting effect (i) and effect (ii) exactly offset each other. This exception is also noted by Heidhues, Kőszegi, and Strack (2018).

change of e^j and the derivative of agent i's payoff with respect to e^j, rendering the effect of a distorted e^i secondary compared to agent j's payoff externality. That is, effects (i) to (iii) are secondary compared to effect (iv). Hence, the task of determining how welfare changes with overconfidence reduces to determining how steady state efforts change.

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 .*[21]

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.*

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

In a simple setting where agents do not impose informational externalities over each other, their effort choices depend only on incorrect self-perceptions and the resulting underestimation of the unknown variable. In particular, it is straightforward that mild overconfidence increases both agent's efforts and corrects the inefficiency if $Q_{e^i a^i} \geq 0$ and $Q_{e^i \psi^i} < 0$, while an opposite pattern arises if $Q_{e^i a^i} \leq 0$ and $Q_{e^i \psi^i} > 0$. With positive informational externalities the welfare change becomes even larger towards the same direction since the agents mutually reinforce their mislearning and their efforts deviate further from the correctly-specified benchmark. By contrast, under negative informational externalities and the assumption that $Q_{e^i \psi^i} > 0$, the decrease in agent j's effort creates upward pressure on agent i's effort, which will counteract the effect of agent i's overconfidence. When the agents are equally overconfident and Q is symmetric with respect to i and j, the agents still exert lower efforts—the existence of negative informational externalities only mitigates the welfare loss. In the extreme case where only agent j is overconfident, agent j works less hard but agent i works harder in response to a lower e^j. This makes agent j better off and agent i worse off. Continuity ensures that this continues to hold when agent i is slightly overconfident. We compute the steady states and plot the direction of the welfare changes for Examples 1 and 2 in Fig. 3. The results are consistent with the predictions of Proposition 4.

[21] $B_\delta^+(x) = \{y : y > x, \|y - x\| < \delta\}$ is defined to be the upper right area inside an x-centered circle with radius δ.

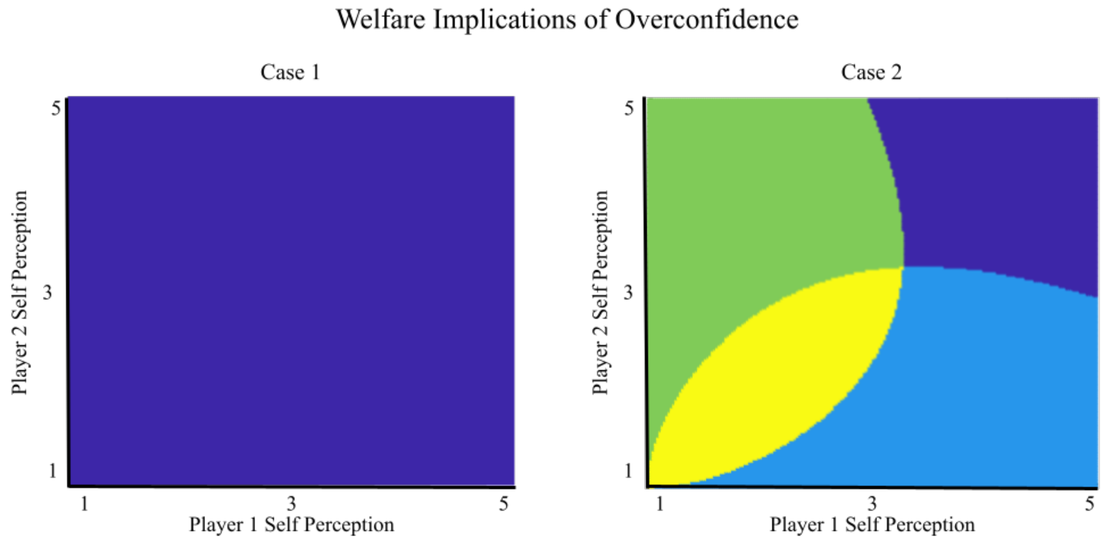

Figure 3: Welfare changes for Examples 1 and 2. Both panels represent the change in welfare to both agents when both agents have true ability $A^1 = A^2 = 1$ and self-perception $\tilde{\boldsymbol{a}}$ as indicated by the graphs. The purple areas represent specifications of confidence where the welfare of both agents decreases compared to the case with correctly specified beliefs ($\tilde{a}^i = A^i$ for both agents). In green areas player 1 is better off and player 2 is worse off. In blue areas player 1 is worse off and player 2 is better off, and in yellow areas both players are better off. The payoff functions used are symmetric for the players and the specific functional forms are $Q(e^1, e^2, a^1, a^2, \phi) = \phi(e^1 + e^2 + e^1e^2 + a^1 + a^2)$ for the first panel, and $Q(e^1, e^2, a^1, a^2, \phi) = log(e^1a^1 + e^2a^2 + \phi)$ for the second panel, with a cost function $c_i(e^i) = 5(e^i)^2$ for both panels. $\Phi = 3$ for both panels.

In sum, our findings highlight the fact that the welfare impact of overconfidence on a pair of agents highly depends on informational externalities. When they are positive, the direction of the welfare change could be easily determined from the primitives. On the contrary, with negative informational externalities, the welfare impact could be sensitive to how much the agents differ in their self-perceptions.

5 Convergence

In this section, we prove that under positive or negative informational externalities, the multi-agent learning processes converge to the steady state (the unique Berk-Nash equilibrium). We make use of a simple and intuitive lemma from Heidhues, Kőszegi, and Strack (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 in the sense that the subjective expectation of outputs should not be consistently lower or higher than the objective expectation. We then use a contraction argument similar to theirs: the structural properties of the output function enable us 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 optimization, we have to impose additional structure to control for the agents' mutual influence.

Assumption 4. *Agents are both overconfident with $\tilde{a}^i \geq A$, and the informational externalities are either both positive or both negative.*

Now we are ready to state the theorem for 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. As shown in the left panel, by our assumption, the belief formation curve intersects with the optimal action curve only once, and the former must be steeper than the latter at the point of intersection. Hence, in a single-agent environment, an iterated elimination of dominated actions and infeasible beliefs eliminates all but the crossing point—the Berk-Nash equilibrium profile. However, in a two-agent environment, their mutual influence must be taken into account—the iterated elimination has to be run simultaneously for both agents. When information externalities are neither positive or negative, agent i's inference, computed from the no-gap condition, and his optimal action may be non-monotone functions of e^j, creating the possibility of cycles in which agents have jointly oscillating actions and beliefs, and never converge.

6 Extensions

6.1 Underconfidence

We begin this section by discussing the implications of underconfidence. The assumption that $\tilde{\boldsymbol{a}} > \boldsymbol{A}$ is important to the direction of mutual learning. In fact, as pointed out by Heidhues, Kőszegi, and Strack (2018), when the agents are underconfident, the single-agent learning process is self-limiting; similarly in this model, with positive informational externalities (Case 1 with complementary efforts) 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 now downward sloping: as agents exert more effort, the marginal return of the unknown variable increases, inducing the agents to overestimate the unknown to a lesser extent. In this case, fixing agent j's effort at some $e^j_S < e^j_\infty(\tilde{\boldsymbol{a}})$ induces a steady state belief higher than when he can freely optimize since a higher effort of agent j lowers agent i's evaluation of ψ^i. Consequently, positive informational externalities help correct the overestimation of ψ^i. With negative informational externalities (Case 2 with substitute efforts), allowing the second agent to freely optimize now increases the overestimation of the first agent. The reinforcement is not as severe as in the overconfidence case, because agent j's effort change is itself constrained through self-limiting learning. Nevertheless, depending on the output function, the resulting belief of agent i on ψ^i could still be arbitrarily far away from truth, contrasting the single-agent underconfidence case where misinference is self-correcting and limited (Heidhues, Kőszegi, and Strack, 2018).

These phenomena are illustrated by Fig. 4. Notably, however, 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 is now ambiguous.

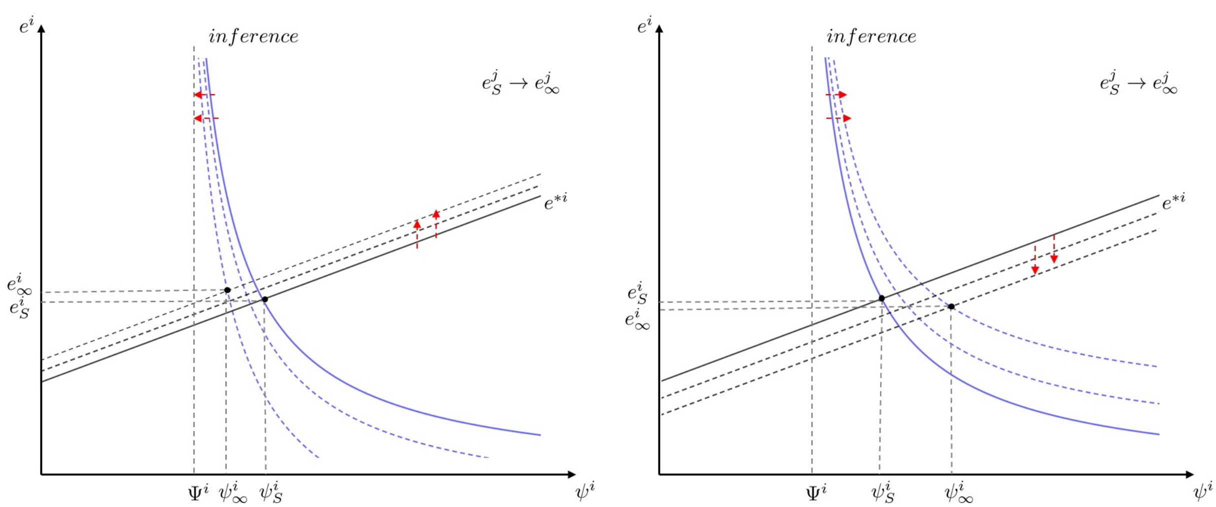

Figure 4: Mutual learning with underconfidence. The left panel shows the heuristic learning dynamics of an underconfident agent when informational externalities are positive. Steady state inferences are lower than when agent j's effort choice is fixed. The right panel shows a reverse pattern when informational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

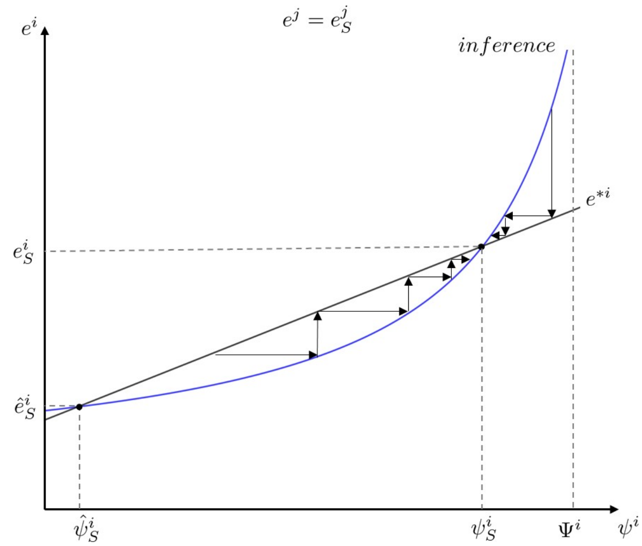

Figure 5: Multiple Berk-Nash equilibria

There seems to be no guarantee that the learning processes will converge: as agents reach a lower belief in $\boldsymbol{\psi}$, they react by exerting lower effort (in the case of $Q_{e^i\psi^i} > 0$ and $Q_{e^ia^i} \leq 0$), pushing up beliefs again. Nevertheless, the following theorem shows that underconfident agents also converge to the Berk-Nash equilibrium when Assumption 4 is replaced by Assumption 5 and—more importantly—the agents are only mildly underconfident. The latter condition ensures that the belief formation curve is steeper than the optimal action curve over the relevant region, enabling the use of the contraction argument.[22] Note that Heidhues, Kőszegi, and Strack (2018), working on a similar but single-agent setting, does not offer a convergence result for the underconfidence case.

Assumption 5. *Agents are both underconfident, and the informational externalities are either both positive or both negative.*

Theorem 2. *Suppose Assumptions 1 to 3 and 5 hold. When for both i, $A^i - \tilde{a}^i$ is sufficiently small, 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 $\boldsymbol{\psi}_\infty$.*

6.2 Multiple Equilibria

The existence of multiple equilibria does not affect our key message, i.e. mutually-reinforcing and mutually-limiting learning, but poses difficulties for the proof of convergence of the learning processes. Fig. 5 shows the single-agent heuristic learning dynamics, where there are two equilibria (e_S^i, ψ_S^i) and $\left(\hat{e}_S^i, \hat{\psi}_S^i\right)$. The contraction argument fails because when, for example, the lower and upper bounds of actions are given by $\hat{e}_S^i$ and e_S^i there can be no further elimination of actions. Hence, this paper does not provide a proof of convergence for such settings. Clearly, the former equilibrium (e_S^i, ψ_S^i) is more plausible since the heuristics make it clear that the agent's belief will drift towards ψ_S^i as he optimizes and updates. In fact, one can use the stochastic approximation tools from Esponda, Pouzo, and Yamamoto (2019) to show convergence to (e_S^i, ψ_S^i) in a finite-action environment—but unfortunately, their techniques do not directly apply to continuous-action settings like ours.[23] Nevertheless, we still feel free to focus on equilibria like (e_S^i, ψ_S^i) over which our analysis of mutual learning patterns remains valid.

6.3 Multiple Agents

Although our paper focuses on the two-player case, the results could be generalized to an arbitrary number of agents. Mutually-reinforcing learning is exacerbated with more overconfident agents exerting positive informational externalities over each other. The parametric assumptions, however, tend to be increasingly more complicated since more cross derivatives are involved. The Pareto improvement when agents are slightly overconfident remains possible and potentially gets stronger.

7 Conclusion

We develop a two-agent learning model with overconfidence and define a new notion of informational externalities to describe how one agent's action can influence the other agent's inference. When positive informational externalities are present, we find a mutually-reinforcing

[22] The assumption of not being too underconfident is essential for the use of the contraction argument, but it may not be a necessary condition for convergence. However, there are no existing results from the literature that can be directly applied here.

[23] The more recent paper Murooka and Yamamoto (2021) generalize 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.

learning pattern that implies strategic interaction exacerbates the underestimation of the common fundamental and makes agents choose more extreme actions; in contrast, learning is mutually-limiting under negative informational externalities. Both patterns are absent in Heidhues, Kőszegi, and Strack (2018) where only one agent actively learns and optimizes. Moreover, our welfare implications starkly contrast with results from single-agent models because in this model there can be Pareto improvement in welfare as a result of overconfidence, and mutually-reinforcing learning can potentially compound this welfare improvement.

One potential future direction is to consider strategic manipulation of informational externalities. For example, in a setting where agents are non-myopic and take into account the influence of their actions over the other agent's beliefs, agents may be incentivized to play actions that are non-optimal in the stage game in order to induce profitable distortions into the other player's long-term beliefs.

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. (6) and (7) could be written as the same condition: for all $i \in I$,

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

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

Proof of Lemma 1. Suppose $e^i \in \arg\max_e Q^i(e, e^j, \tilde{a}^i, \psi^i)$, 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 (e^i, e^j) and $(\hat{e}^i, \hat{e}^j)$, and without loss of generality $e^i > \hat{e}^i$, then

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

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,

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

where equality is obtained if and only if $Q^i(\boldsymbol{e}, A^i, \Psi) = Q^i(\boldsymbol{e}, \tilde{a}^i, \psi^i)$. 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 $(\boldsymbol{e}^*(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}_\infty), \boldsymbol{\psi}_\infty)$ is a pure-strategy Berk-Nash equilibrium if and only if the no-gap conditions hold.

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

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

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 $G^i(\Psi, \psi^j) < 0, \forall \psi^j$, 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,

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

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 $\psi^i = \Psi^i$,

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

Continuity then implies that there exist $\Delta^i > 0, i = 1, 2$, such that for any $a^i \in (A^i, A^i + \Delta^i)$

and for any $\psi^i \in \left[\Psi^i - \frac{\overline{\kappa}_a}{\underline{\kappa}_\psi} \left|\tilde{a}^i - A^i\right|, \Psi^i\right] \subset \left(\underline{\psi}, \overline{\psi}\right)$, the following are true:

$$\frac{\partial G^i\left(\boldsymbol{\psi}\right)}{\partial \psi^i} < 0, \left|\frac{\partial G^i\left(\boldsymbol{\psi}\right)}{\partial \psi^j}\right| < \left|\frac{\partial G^i\left(\boldsymbol{\psi}\right)}{\partial \psi^i}\right|, \forall i. \tag{8}$$

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. (8), if $\boldsymbol{G}\left(\tilde{\boldsymbol{\psi}}\right) = \boldsymbol{G}\left(\hat{\boldsymbol{\psi}}\right) = 0$, 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^i_{t-1} and π^j_{t-1}, there exists a unique rationalizable action profile $\left(e^i_t, e^j_t\right)$.*

Proof. First consider the case of complementary efforts, $Q_{e^ie^j} \geq 0$. Let $\left[\underline{e}^j_0, \overline{e}^j_0\right] = \left[\underline{e}^i_0, \overline{e}^i_0\right] = \left[\underline{e}, \overline{e}\right]$, and recursively define for all i and $\tau \geq 1$,

$$\underline{b}^i_\tau = \arg\max_{e^i} \mathbb{E}_{\pi^i_{\tau-1}}\left[Q^i\left(e^i, \underline{e}^j_{\tau-1}, \tilde{a}^i, \psi^i\right)\right] \equiv \arg\max_{e^i} h^i_\tau\left(e^i, \underline{e}^j_{\tau-1}, \tilde{a}^i\right),$$
$$\overline{b}^i_\tau = \arg\max_{e^i} \mathbb{E}_{\pi^i_{\tau-1}}\left[Q^i\left(e^i, \overline{e}^j_{\tau-1}, \tilde{a}^i, \psi^i\right)\right] \equiv \arg\max_{e^i} h^i_\tau\left(e^i, \underline{e}^j_{\tau-1}, \tilde{a}^i\right).$$

The existence of such $\underline{b}^i_\tau$ and $\overline{b}^i_\tau$ follow from Assumption 2 and the continuity of Q^i. Since Q^i is twice continuously differentiable, h^i_τ is also twice continuously differentiable. Note that complementarity between efforts implies $\frac{\partial h^i_\tau}{\partial e^i \partial e^j} \geq 0$, so $\underline{b}^i_\tau \leq \overline{b}^i_\tau$. Let $\left[\underline{e}^i_\tau, \overline{e}^i_\tau\right] \equiv \left[\underline{e}^i_{\tau-1}, \overline{e}^i_{\tau-1}\right] \cap \left[\underline{b}^i_\tau, \overline{b}^i_\tau\right]$. By Assumption 2, $\left[\underline{e}^i_1, \overline{e}^i_1\right] \subsetneq \left[\underline{e}^i_0, \overline{e}^i_0\right] = \left[\underline{e}, \overline{e}\right], \forall i$. Using $\frac{\partial h^i_\tau}{\partial e^i \partial e^j} \geq 0$ again, we know $\left[\underline{e}^j_\tau, \overline{e}^j_\tau\right] \subset \left[\underline{e}^j_{\tau-1}, \overline{e}^j_{\tau-1}\right]$ implies that $\left[\underline{e}^i_{\tau+1}, \overline{e}^i_{\tau+1}\right] \subset \left[\underline{e}^i_\tau, \overline{e}^i_\tau\right], \forall i, j \neq i$ and $\forall \tau > 1$. By the Nested Intervals Theorem, we know that each agent's set of rationalizable actions $\left[\underline{e}^i_\tau, \overline{e}^i_\tau\right]$ 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, $\left(e^i_t, e^j_t\right)$, and it satisfies

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

Next, consider the case of substitute efforts, $Q_{e^ie^j} \leq 0$. Analogously define $\left[\underline{e}^i_0, \overline{e}^i_0\right] = \left[\underline{e}, \overline{e}\right]$ and recursively define for all i and $\tau \geq 1$,

$$\underline{b}^i_\tau = \arg\max_{e^i} \mathbb{E}_{\pi^i_{t-1}}\left[Q^i\left(e^i, \overline{e}^j_{\tau-1}, \tilde{a}^i, \psi^i\right)\right] \equiv \arg\max_{e^i} h^i_\tau\left(e^i, \overline{e}^j_{\tau-1}, \tilde{a}^i\right),$$
$$\overline{b}^i_\tau = \arg\max_{e^i} \mathbb{E}_{\pi^i_{t-1}}\left[Q^i\left(e^i, \underline{e}^j_{\tau-1}, \tilde{a}^i, \psi^i\right)\right] \equiv \arg\max_{e^i} h^i_\tau\left(e^i, \underline{e}^j_{\tau-1}, \tilde{a}^i\right).$$

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

B Proofs for Section 4

Proof of Proposition 1. For this proof, assume $Q_{e^i\psi^i} > 0$ for both $i = 1, 2$. We can accommodate $Q_{e^i\psi^i} < 0, \forall i = 1, 2$ by replacing e^i, e^j with $-e^i, -e^j$ and substituting the constraint with $e^j_S < e^j_\infty(\tilde{\boldsymbol{a}})$.

We start by showing $e^i_S > e^i_\infty$ and $\psi^i_S > \psi^i_\infty$. Fixing agent j's effort at some level e^j, consider the following two equations,

$$g^i(\boldsymbol{e}, \psi^i) = Q^i\left(\boldsymbol{e}, A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right) = 0, \tag{9}$$

$$Q^i_{e^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right) = 0, \tag{10}$$

where e^i and ψ^i are the unknown variables. Notice that both (e^i_S, ψ^i_S) and $(e^i_\infty, \psi^i_\infty)$ can be determined by Eqs. (9) and (10), which correspond to different agent j's efforts, e^j_S and $e^j_\infty(\tilde{\boldsymbol{a}})$.

We denote the set of possible actions and beliefs under a fixed e^j by $D^i_e(e^j)$ and $D^i_\psi(e^j)$ respectively, i.e. $(e^i, \psi^i) \in D^i_e(e^j) \times D^i_\psi(e^j)$ iff it satisfies Eqs. (9) and (10). By Assumptions 1 and 2, both $D^i_e(e^j)$ and $D^i_\psi(e^j)$ are nonempty and compact for any e^j. Further, let $\hat{\psi}^i(e^j)$ represent the largest element in $D^i_\psi(e^j)$, and $\hat{e}^i(e^j)$ represent the corresponding effort in $D^i_e(e^j)$. 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 $(D^i_e(e^j), D^j_e(e^i))$ 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 $e^i_\infty = \hat{e}^i(e^j_\infty)$ and $\psi^i_\infty = \hat{\psi}^i(e^j_\infty)$. By our assumption of a unique steady state at e^j_S, we also have $e^i_S = \hat{e}^i(e^j_S)$ and $\psi^i_S = \hat{\psi}^i(e^j_S)$.

Differentiate Eqs. (9) and (10) with respect to e^j,

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

$$Q^i_{e^ie^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)\frac{\partial e^i}{\partial e^j} + Q^i_{e^i\psi^i}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)\frac{\partial \psi^i}{\partial e^j} + Q^i_{e^ie^j}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right) = 0.$$

Simplify and then we obtain

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

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

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 $e^j_S > e^j_\infty$, we infer that $e^i_S > e^i_\infty$ and $\psi^i_S > \psi^i_\infty$.

Next we prove $\psi^j_S > \psi^j_\infty$. Observe that ψ^j_S is given by

$$Q^j\left(\boldsymbol{e_S},A^j,\Psi^j\right) - Q^j\left(\boldsymbol{e_S},\tilde{a}^j,\psi^j_S\right) = 0.$$

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 than 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 $\psi^i = \hat{\psi}^i\left(e^j\right)$,

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

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

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

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

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

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

□

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

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

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

Now suppose agents create negative informational externalities for each other.

(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 part (i) and (iii) since part (ii) and (iv) follow from analogous arguments. Differentiate the following equations that determine the steady state with respect to a^i and a^j,

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

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

After tedious calculations we obtain the following,

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

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 statements (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 $Q^i_{e^j}\left(\boldsymbol{e},A^i,\Psi^i\right)-Q^i_{e^j}\left(\boldsymbol{e},\tilde{a}^i,\psi^i\right)$ 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, $e^j_S>e^j_\infty\left(\tilde{\boldsymbol{a}}\right)$ implies $\psi^i_S<\psi^i_\infty\left(\tilde{\boldsymbol{a}}\right)<\Psi^i$.

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

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

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 $Q(\boldsymbol{e_\infty}(\boldsymbol{a}),A^i,A^j,\Phi)-c(e^i_\infty(\boldsymbol{a}))$ with respect to a^i and a^j and evaluate the derivative at $\boldsymbol{A}$, then we obtain

$$\begin{aligned}\frac{\partial[Q(\boldsymbol{e_\infty}(\boldsymbol{a}),A^i,A^j,\Phi)-c(e^i_\infty(\boldsymbol{a}))]}{\partial a^i}&=Q_{e^j}(\boldsymbol{e_\infty}(\boldsymbol{a}),A^i,A^j,\Phi)\frac{\partial e^j_\infty(\boldsymbol{a})}{\partial a^i},\\ \frac{\partial[Q(\boldsymbol{e_\infty}(\boldsymbol{a}),A^i,A^j,\Phi)-c(e^i_\infty(\boldsymbol{a}))]}{\partial a^j}&=Q_{e^j}(\boldsymbol{e_\infty}(\boldsymbol{a}),A^i,A^j,\Phi)\frac{\partial e^j_\infty(\boldsymbol{a})}{\partial a^i}.\end{aligned}$$

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

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 yield $\frac{\partial e^j_\infty(\boldsymbol{a})}{\partial a^j} / \frac{\partial e^j_\infty(\boldsymbol{a})}{\partial a^i} = -Q^i_{e^ie^i}/Q^i_{e^ie^j} < -1$, where the inequality follows from Assumption 2. Hence,

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

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

C Proofs for Section 5

Following Heidhues, Kőszegi, and Strack (2018) (HKS), for each agent i, we define $m^i_t(\psi^i)$ 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,

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

Let $\tilde{\mathbb{P}}^i_t$ denote agent i's subjective probability measure conditional on the history up to time t. Define the lowest upper bound and the highest lower bound for any agent i' long-run beliefs as follows,

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

Write the vectors of bounds as $\underline{\boldsymbol{\psi}}_{\boldsymbol{\infty}} = \left(\underline{\psi}^i_\infty, \underline{\psi}^j_\infty\right), \overline{\boldsymbol{\psi}}_{\boldsymbol{\infty}} = \left(\overline{\psi}^i_\infty, \overline{\psi}^j_\infty\right)$. We show that $\underline{\boldsymbol{\psi}}_{\boldsymbol{\infty}}$ and $\overline{\boldsymbol{\psi}}_{\boldsymbol{\infty}}$ 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 implied by ψ^i should not be consistently higher or lower than the one implied by Ψ^i. We replicate the proof to show that this lemma applies to our two-agent environment as well.

Lemma 7 (Heidhues, Kőszegi, and Strack (2018), Lemma 13)**.** *The following are true:*

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

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

(b) *For all i, if $\limsup_{t\to\infty} m_t^i(\psi^i) \leq \overline{m} < 0$ for all $\psi^i \in (l,h) \subset (\underline{\psi}, \overline{\psi})$, 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 $l_t^i(\psi^i) \coloneqq \sum_{s=1}^t \log f\left(m_t^i(\psi^i) + \epsilon_s\right)$ denote the log-likelihood assigned to ψ^i by agent i and $\mathcal{L}(x) \coloneqq \mathbb{E} \log f(x+\epsilon)$. By Lemma 12 in Heidhues, Kőszegi, and Strack (2018), there exists $r > 0$ such that

$$\begin{aligned}
\liminf_{t\to\infty} \inf_{\psi^i\in(l,h)} \frac{l_t^{i\prime}(\psi^i)}{t} \geq & \liminf_{t\to\infty} \inf_{\psi^i\in(l,h)} \frac{1}{t} \frac{\partial \sum_{s=1}^t \mathbb{E} \log f\left(m_t^i(\psi^i)+\epsilon_s\right)}{\partial \psi^i} \\
= & \liminf_{t\to\infty} \inf_{\psi^i\in(l,h)} \frac{1}{t} \sum_{s=1}^t \mathcal{L}'(m_t^i(\psi^i)) \left[-Q_{\psi^i}^i(\boldsymbol{e_t}, \tilde{a}^i, \psi^i)\right] \geq r,
\end{aligned}$$

where the first inequality is an implication of a generalized law of large numbers and the last 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,

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

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

Lemma 8. *For all i, we have that $\Psi^i - \frac{\kappa_a}{\kappa_\psi}(\tilde{a}^i - A^i) \leq \underline{\psi}_\infty^i$ and $\overline{\psi}_\infty^i \leq \Psi^i$.*

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

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

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

We can obtain similar bounds for efforts. Define

$$\begin{aligned}
\underline{e}^i_\infty &\equiv \sup\left\{e^i : e^i \leq \liminf_{t\to\infty} e^i_t \text{ almost surely}\right\},\\
\overline{e}^i_\infty &\equiv \inf\left\{e^i : e^i \geq \limsup_{t\to\infty} e^i_t \text{ almost surely}\right\}.
\end{aligned}$$

Define $E_\infty \equiv [\underline{e}^1_\infty, \overline{e}^1_\infty] \times [\underline{e}^2_\infty, \overline{e}^2_\infty]$, which is the set of efforts that may be taken by agents in the long run. In addition, define $E^D_\infty \equiv \left\{\boldsymbol{e} : \exists \boldsymbol{\psi} \in \left[\underline{\psi}^1_\infty, \overline{\psi}^1_\infty\right] \times \left[\underline{\psi}^2_\infty, \overline{\psi}^2_\infty\right], \forall i, \text{s.t. } \boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right) = \boldsymbol{e}\right\}$, 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. $E_\infty \subseteq E^D_\infty$.

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

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

Continuity of $Q^i_{e^i}$ implies that there exists $\hat{\boldsymbol{\psi}}_{\boldsymbol{t}} \in \times_I\left(\underline{\psi}, \overline{\psi}\right)$ such that $\forall i$,

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

We know that the support of Π^i_t is contained in $\left[\underline{\psi}^i_\infty, \overline{\psi}^i_\infty\right]$ 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 $Q^i_{e^i\psi^i}$, while $\frac{\partial e^{*j}\left(\tilde{\boldsymbol{a}},\boldsymbol{\psi^i}\right)}{\partial \psi^i}$ has the same sign as $Q^i_{e^i\psi^i}Q^j_{e^ie^j}$.*

Proof. Given $\tilde{\boldsymbol{a}}, \boldsymbol{\psi}$, $e_i^*\left(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}\right)$ satisfy

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

Take partial derivatives, we obtain that

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

Therefore, $\frac{\partial e^{*i}(\tilde{\boldsymbol{a}},\boldsymbol{\psi})}{\partial \psi^i}$ has the same sign as $Q^i_{e^i\psi^i}$ and $\frac{\partial e^{*j}(\tilde{\boldsymbol{a}},\boldsymbol{\psi})}{\partial \psi^i}$ has the same sign as $Q^i_{e^i\psi^i}Q^j_{e^ie^j}$. □

Proof of Theorem 1. First consider the asymptotic behavior of $m_t\left(\underline{\psi}_\infty^i\right)$. 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 $\underline{\psi}_\infty^i$ being the highest lower bound for agent i's long-run beliefs almost surely. 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 $Q^i_{e^i\psi^i} > 0$ and $Q^i_{e^ia^i} \leq 0$. Recall that $g^i\left(\boldsymbol{e}, \psi^i\right) = Q^i\left(\boldsymbol{e}, A^i, \Psi^i\right) - Q^i\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right)$. Differentiate g with respect to e^j, we obtain

$$\frac{\partial g^i\left(\boldsymbol{e}, \psi^i\right)}{\partial e^j} = Q^i_{e^j}\left(\boldsymbol{e}, A^i, \Psi^i\right) - Q^i_{e^j}\left(\boldsymbol{e}, \tilde{a}^i, \psi^i\right). \quad (11)$$

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

Hence,

$$0 \geq \liminf_{t\to\infty} m_t^i\left(\underline{\psi}_\infty^i\right) = \liminf_{t\to\infty} g^i\left(\boldsymbol{e}_t, \underline{\psi}_\infty^i\right) \geq G^i\left(\underline{\boldsymbol{\psi}}_{\boldsymbol{\infty}}\right), \forall i,$$
$$0 \leq \limsup_{t\to\infty} m_t^i\left(\overline{\psi}_\infty^i\right) = \limsup_{t\to\infty} g^i\left(\boldsymbol{e}_t, \overline{\psi}_\infty^i\right) \leq G^i\left(\overline{\boldsymbol{\psi}}_{\boldsymbol{\infty}}\right), \forall i. \tag{12}$$

Since $\boldsymbol{e}^*(\tilde{\boldsymbol{a}}, \boldsymbol{\psi})$ 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 $G^i(\boldsymbol{\psi}) = g^i(\boldsymbol{e}^*(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}), \psi^i)$ 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 $G^i(\Psi^i, \psi^j) < 0$ and $G^i\left(\Psi^i - \frac{\kappa_a}{\kappa_\psi}(\tilde{a}^i - A^i), \psi^j\right) > 0$. For convenience, let $\underline{\psi}^i = \Psi^i - \frac{\kappa_a}{\kappa_\psi}(\tilde{a}^i - A^i)$ and $\overline{\psi}^i = \Psi^i$ for all i. The above results can be summarized by:

$$\begin{aligned}
&\boldsymbol{G}\left(\underline{\boldsymbol{\psi}}_{\boldsymbol{\infty}}\right) \leq 0, \boldsymbol{G}\left(\underline{\boldsymbol{\psi}}\right) > 0, \boldsymbol{G}\left(\overline{\boldsymbol{\psi}}_{\boldsymbol{\infty}}\right) \leq 0, \boldsymbol{G}\left(\overline{\boldsymbol{\psi}}\right) < 0,\\
&G^i\left(\underline{\psi}^i, \underline{\psi}_\infty^j\right) > 0, G^j\left(\underline{\psi}^i, \underline{\psi}_\infty^j\right) \leq 0,\\
&G^i\left(\underline{\psi}_\infty^i, \underline{\psi}^j\right) \leq 0, G^j\left(\underline{\psi}_\infty^i, \underline{\psi}^j\right) > 0,\\
&G^i\left(\overline{\psi}^i, \overline{\psi}_\infty^j\right) < 0, G^j\left(\overline{\psi}^i, \overline{\psi}_\infty^j\right) \geq 0,\\
&G^i\left(\overline{\psi}_\infty^i, \overline{\psi}^j\right) \geq 0, G^j\left(\overline{\psi}_\infty^i, \overline{\psi}^j\right) < 0.
\end{aligned} \tag{13}$$

By Brouwer's fixed point theorem, there exist $\hat{\boldsymbol{\psi}}, \tilde{\boldsymbol{\psi}}$ such that $\tilde{\boldsymbol{\psi}} \in \left[\boldsymbol{\Psi} - \frac{\kappa_a}{\kappa_\psi}(\tilde{\boldsymbol{a}} - \boldsymbol{A}), \underline{\boldsymbol{\psi}}_{\boldsymbol{\infty}}\right]$, $\hat{\boldsymbol{\psi}} \in \left[\overline{\boldsymbol{\psi}}_{\boldsymbol{\infty}}, \boldsymbol{\Psi}\right]$, 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 $\hat{\boldsymbol{\psi}} = \tilde{\boldsymbol{\psi}} = \underline{\boldsymbol{\psi}}_{\boldsymbol{\infty}} = \overline{\boldsymbol{\psi}}_{\boldsymbol{\infty}} = \boldsymbol{\psi}_{\boldsymbol{\infty}}$ and thus $E_\infty = E_\infty^D = \{\boldsymbol{e}_{\boldsymbol{\infty}}\}$.

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 $\underline{\boldsymbol{\psi}}_{\boldsymbol{\infty}} \neq \overline{\boldsymbol{\psi}}_{\boldsymbol{\infty}}$. Since informational externalities are negative, the signs of Eq. (11) 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, $e^{*i}\left(\tilde{\boldsymbol{a}}, \left(\underline{\psi}_\infty^i, \overline{\psi}_\infty^j\right)\right) \leq e_t^i \leq e^{*i}\left(\tilde{\boldsymbol{a}}, \left(\overline{\psi}_\infty^i, \underline{\psi}_\infty^j\right)\right), \forall i$. Hence,

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

In addition, $G^i(\boldsymbol{\psi}) = g^i(\boldsymbol{e}(\tilde{\boldsymbol{a}}, \boldsymbol{\psi}), \psi^i)$ is decreasing in ψ^j. Therefore, we obtain some inequalities different from those in Eq. (13):

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

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

Again, there exist two different roots to $\boldsymbol{G}(\boldsymbol{\psi}) = 0$ if $\underline{\boldsymbol{\psi}}_{\boldsymbol{\infty}} \neq \overline{\boldsymbol{\psi}}_{\boldsymbol{\infty}}$, 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. □

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

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

Therefore, $g^i\left(\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \overline{\boldsymbol{\psi}}_{\boldsymbol{\infty}}\right), \underline{\psi}^i_\infty\right) \leq 0 \leq g^i\left(\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \underline{\boldsymbol{\psi}}_{\boldsymbol{\infty}}\right), \overline{\psi}^i_\infty\right), \forall i$. For each i, define $h^i\left(\boldsymbol{\zeta}, \boldsymbol{\psi}\right) :=$

$g^i\left(\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}},\boldsymbol{\zeta}\right),\psi^i\right)$. Differentiate h^i with respect to ζ^k and ψ^k, we obtain

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

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

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

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

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

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

References

Anderson, C., S. Brion, D. A. Moore, and J. A. Kennedy (2012): "A Status-Enhancement Account of Overconfidence," *Journal of Personality and Social Psychology*, 103(4), 718.

Benoît, J.-P., J. Dubra, and D. A. Moore (2015): "Does the Better-than-average Effect Show that People are Overconfident?: Two Experiments," *Journal of the European Economic Association*, 13(2), 293–329.

Bils, P. (2020): "Policymaking with Multiple Agencies," *American Journal of Political Science*, 64(3), 634–648.

Bohren, J. A. (2016): "Informational Herding with Model Misspecification," *Journal of Economic Theory*, 163, 222–247.

Bohren, J. A., and D. N. Hauser (2019): "Bounded Rationality and Learning: A Framework and a Robustness Result," *PIER Working Paper No. 18-017*.

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

Campbell, W. K., and C. Sedikides (1999): "Self-threat Magnifies the Self-serving Bias: A Meta-Analytic Integration," *Review of General Psychology*, 3(1), 23–43.

Dunne, T., M. J. Roberts, and L. Samuelson (1988): "Patterns of Firm Entry and Exit in US Manufacturing Industries," *RAND Journal of Economics*, pp. 495–515.

Esponda, I., and D. Pouzo (2016): "Berk-Nash Equilibrium: A Framework for Modeling Agents with Misspecified Models," *Econometrica*, 84(3), 1093–1130.

Esponda, I., D. Pouzo, and Y. Yamamoto (2019): "Asymptotic Behavior of Bayesian Learners with Misspecified Models," *arXiv:1904.08551*.

Eyster, E., and M. Rabin (2010): "Naive Herding in Rich-information Settings," *American Economic Journal: Microeconomics*, 2(4), 221–43.

Frick, M., R. Iijima, and Y. Ishii (2019): "Misinterpreting Others and the Fragility of Social Learning," *Cowles Foundation Discussion Paper*.

——— (2021): "Belief convergence under misspecified learning: A martingale approach," *Cowles Foundation Discussion Paper*.

Fudenberg, D., G. Lanzani, and P. Strack (2020): "Limits Points of Endogenous Misspecified Learning," *Available at SSRN: http://dx.doi.org/10.2139/ssrn.3553363*.

Fudenberg, D., G. Romanyuk, and P. Strack (2017): "Active Learning with a Misspecified Prior," *Theoretical Economics*, 12(3), 1155–1189.

Gagnon-Bartsch, T. (2017): "Taste Projection in Models of Social Learning," *Working paper*.

Gagnon-Bartsch, T., and M. Rabin (2017): "Naive Social Learning, Mislearning, and Unlearning," *Working paper*.

Gervais, S., and I. Goldstein (2007): "The Positive Effects of Biased Self-perceptions in Firms," *Review of Finance*, 11(3), 453–496.

He, K. (2018): "Mislearning from Censored Data: The Gambler's Fallacy in Optimal-Stopping Problems," *arXiv:1803.08170*.

Heidhues, P., B. Kőszegi, and P. Strack (2018): "Unrealistic Expectations and Misguided Learning," *Econometrica*, 86(4), 1159–1214.

——— (2021): "Convergence in models of misspecified learning," *Theoretical Economics*, 16(1), 73–99.

Heifetz, A., C. Shannon, and Y. Spiegel (2007): "What to maximize if you must," *Journal of Economic Theory*, 133(1), 31–57.

Kennedy, J. A., C. Anderson, and D. A. Moore (2013): "When overconfidence is revealed to others: Testing the status-enhancement theory of overconfidence," *Organizational Behavior and Human Decision Processes*, 122(2), 266–279.

Langer, E. J., and J. Roth (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," *Journal of Personality and Social Psychology*, 32(6), 951.

Miller, D. T., and M. Ross (1975): "Self-Serving Biases in the Attribution of Causality: Fact Or Fiction?," *Psychological Bulletin*, 82(2), 213.

Moore, D. A., and P. J. Healy (2008): "The Trouble with Overconfidence," *Psychological Review*, 115(2), 502.

Moore, D. A., E. R. Tenney, and U. Haran (2015): "Overprecision in Judgment," *The Wiley Blackwell Handbook of Judgment and Decision Making*, 2, 182–209.

Murooka, T., and Y. Yamamoto (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," *Journal of Economic Theory*, 55(2), 416–427.

Rabin, M., and J. L. Schrag (1999): "First Impressions Matter: A Model of Confirmatory Bias," *The Quarterly Journal of Economics*, 114(1), 37–82.

Ross, M., and F. Sicoly (1979): "Egocentric Biases in Availability and Attribution.," *Journal of Personality and Social Psychology*, 37(3), 322.

Svenson, O. (1981): "Are We All Less Risky and More Skillful Than Our Fellow Drivers?," *Acta psychologica*, 47(2), 143–148.