The Ronald O. Perelman Center for Political Science and Economics (PCPSE)
133 South 36th Street
Philadelphia, PA 19104-6297

pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper
21-019

A Multi-Agent Model of Misspecified Learning with Overconfidence

CUIMIN BA
University of Pennsylvania

ALICE GINDIN
University of Pennsylvania

January 28, 2021

https://ssrn.com/abstract=3914110

A Multi-Agent Model of Misspecified Learning with Overconfidence*

Cuimin Ba †　　Alice Gindin ‡

First version: Aug 27, 2019
This version: Jan 28, 2021

Abstract

Overconfidence has been extensively documented in psychology and economics. This paper studies the long-term interaction between two overconfident agents who learn about common payoff-relevant fundamentals, such as the quality of a joint project or their working environment, and choose how much effort to exert. Overconfidence causes agents to underestimate the fundamental 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: when one agent best responds to his own overconfidence, the other agent underestimates the fundamental more severely and takes an more extreme action, generating a positive feedback loop. The opposite pattern, mutually-limiting learning, arises when informational externalities are negative. Additionally, overconfidence can lead to Pareto improvement in welfare as it corrects the inefficiencies that arise in public good provision problems. This contrasts with the analogous single-agent environment, in which there is no scope for informational externalities and overconfidence can only decrease 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, Annie Liang, and George Mailath for their guidance and support at every stage of this project. We would like to thank Alvaro Sandroni, Andrew Postlewaite, Takuo Sugaya, Steven Matthews, Kevin He, and participants of the Young Economists Symposium and seminars at 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 underestimates how talented his team is at performing joint tasks, leading to a welfare loss. However, when working on a task within a team, all members of the team learn and adjust their effort simultaneously. In this paper, we consider what happens when *multiple* persistently overconfident agents learn about their environment from observing payoffs. We show that this can change the learning dynamics as well as the resulting welfare effects.

To ground this idea, consider two engineers who work together on different components of an overarching project assigned by their shared supervisor. Both engineers are overconfident in their research skills, yet neither knows the underlying quality of the projects their supervisor designs. Both engineers learn about the quality of the overall project over time by working on, and being evaluated for, their individual components. However, both share knowledge and experience gained from reading articles or testing out different methods, so the time one puts into his own project component will affect the others' progress. 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 overarching project's underlying 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 will converge 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 is motivated to work harder, but finds himself considerably more disappointed by the new output that corresponds to the higher levels of both efforts. The extra disappointment exacerbates the drop in his inference about the project quality and encourages him to exert even more effort. 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 process by which the presence of two engineers simultaneously adjusting their effort makes beliefs more extreme "mutually-reinforcing" learning 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 due to the common-good nature of joint research efforts.

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 determined by their own effort, the effort of the other agent, the agent's own ability, as well as some unknown fundamental. Each agent chooses an effort to myopically maximize his payoff given the effort of the other agent and his beliefs. We assume that an agent has a degenerate belief about the value of their own ability and study the case where that point belief is incorrect. The unknown fundamental corresponds to the quality of the research idea in the example above. The agents have non-degenerate priors about the value of the fundamental and each other's ability and update these beliefs over time.

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 payoff externality for agent i, it also provides *an informational externality* by affecting agent i's inference problem over the fundamental through two channels. The first channel is a direct one, in which a change in agent j's effort changes the signal structure for agent i whenever the marginal product of agent i's ability or the fundamental is changed by agent j's effort. The second channel takes effect whenever 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. With proper assumptions on the payoff function which ensure 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.

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 only results in distorted inferences and 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 almost always fails to be socially efficient due to externalities, Pareto improvement can be obtained by perturbing agents' effort within a small neighborhood to reduce such externalities—overconfidence can serve as a tool for perturbation. To achieve a Pareto improvement, intuitively, the misspecification should induce the agents to move in the same direction as the externalities. For example, if agents' efforts exert positive externalities on each other's payoff, both agents end up with higher payoffs if their efforts are distorted slightly upwards. We also show that when agents can be over or underconfident, there always exists a range of self-perception levels under which agents enjoy higher payoffs in the long run than they would if correctly specified.

The above insights are demonstrated by analyzing the long-run beliefs and effort choices. We show that under certain conditions, agents converge to a Berk-Nash equilibrium, meaning 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. The informational externalities must be either both positive or both 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 exacerbated as the agent re-optimizes his effort—a self-defeating learning pattern arises. 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. In recent work, Murooka and Yamamoto (2021) consider how multiple misspecified agents learn from a signal of common output when they each have fixed beliefs about a total team capability.[1]

There is a growing literature that explores the implications of model misspecification on learning.[2] Esponda and Pouzo (2016) propose the solution concept, Berk-Nash equilibrium, for such games with misspecification. In recent years, there have been substantial progress in showing the convergence of beliefs to the Berk-Nash equilibrium in general environments. 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 (2019) instead assume a finite-state but otherwise general setting and propose stronger conditions than Kullback-Leibler divergence dominance; Bohren and Hauser (2019) characterize conditions under which correct learning, incorrect learning, or cyclical learning arise in a binary-state learning environment. However, none of their techniques are directly applicable to our multi-agent model that assumes continuous actions and states (i.e. the value of the fundamental). 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 overcon-

[1] We became aware of Murooka and Yamamoto (2021) as we were completing this revision.

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

fident about their abilities, which is well supported by the psychology literature on overconfidence.[3] One illustration of this is 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. Much of the literature discusses the "better-than-average effect", as most 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.

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 explores the patterns of mutually-reinforcing and mutually-limiting learning in the equilibrium, and analyze the welfare implications. Section 5 shows that in the presence of unambiguously positive or negative informational externalities, the two-agent learning process will converge to the Berk-Nash equilibrium. Section 6 provides extensions including allowing for underconfidence. Section 7 concludes.

2 Multi-Agent Learning Environment

Environment There are two agents, indexed by $i \in I \equiv \{1, 2\}$. In each period $t \in \{1, 2, ...\}$, each agent i simultaneously chooses an effort level e^i_t from a compact set $[\underline{e}, \overline{e}] \subset \mathbb{R}$. Agent i then obtains a payoff q^i_t that is determined by his own effort e^i, the effort of the other agent e^j, his own ability a^i, a common unknown fundamental ϕ, and random noise. We write agent i's payoffs as $q^i_t = Q^i(e^i, e^j, a^i, \phi) + \epsilon^i_t$, where ϵ^i_t is a zero-mean i.i.d. random variable drawn from some continuous distribution with a positive and log-concave density

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

f.[4] The payoff function Q^i is twice continuously differentiable, with its derivatives having polynomial growth in ϕ.[5] All past payoffs and efforts are publicly observable.

Each agent's ability a^i and the fundamental ϕ are independently drawn before the game starts from c.d.f. M_0 with potentially unbounded support $(\underline{a}, \overline{a}) \subseteq \mathbb{R}$ and c.d.f. Π_0 with potentially unbounded support $\left(\underline{\phi}, \overline{\phi}\right) \subseteq \mathbb{R}$, respectively, and remain fixed throughout. Denote their realizations as A^1, A^2, and Φ. We assume M_0 and Π_0 each have finite moments and a bounded strictly positive continuous densities μ_0 and π_0.

Misspecification Agents are overconfident in their own ability. In particular, agent i believes that his true ability is actually $\tilde{a}^i \in (\underline{a}, \overline{a})$ and $A^i \leq \tilde{a}^i$. The self-perceptions of the agents , i.e. $\tilde{a}^1$ and $\tilde{a}^2$, are common knowledge. However, agents realize that their counterpart may be subject to bias and are uncertain about the other agent's true ability. Hence, agent i uses a misspecified model to learn about the common fundamental ϕ as well as the true ability of the other agent, a^j. We use π_t^i and μ_t^i to denote agent i's belief about the fundamental and the other agent's ability at time t. In addition, let the agents start with the correct priors about ϕ^i and a^j, i.e. $\pi_0^i = \pi_0$ and $\mu_t^i = \mu_0$, which ensures that any mislearning is a result of overconfidence rather than misspecified priors.

We make the following assumptions about the payoff functions.

Assumption 1. *For all i and $j \neq i$: (i) $Q_a^i := \partial Q^i / \partial a^i$ and $Q_\phi^i := \partial Q^i / \partial \phi^i$ are strictly bounded and positive; (ii) the signs of $Q_{e^i a}^i := \partial^2 Q^i / \partial a^i \partial e^i$, and $Q_{e^i \phi}^i := \partial^2 Q^i / \partial \phi^i \partial e^i$ are different, $Q_{e^i\phi}^i \neq 0$, and the signs do not vary with i;* [6] *(iii) $\forall e^i, e^j$, there always exists a solution $\phi^i \in \left(\underline{\phi}_i, \overline{\phi}^i\right)$ to $Q^i\left(e^i, e^j, \tilde{a}^i, \phi^i\right) = Q^i\left(e^i, e^j, A^i, \Phi\right)$.*

The first assumption says that the ability and the fundamental positively influence one's payoff. The second assumption guarantees both agents are optimizing 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. Suppose $Q_{e^i\phi}^i > 0$ and $Q_{e^i a}^i \leq 0$. Then evidently this agent should decrease his effort in response. If instead both cross derivatives

[4] We assume log-concavity, i.e. the second-order derivative of $f(\epsilon)$ 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 $Q^i\left(e^i, e^j, a^i, \phi\right)$ is of polynomial growth in ϕ if for any e^i, e^j, a^i, there are $\kappa, k, b > 0$ such that $|Q^i\left(e^i, e^j, a^i, \phi\right)| \leq \kappa |\phi|^k + b$. This ensures the expected payoff and its derivatives exist after arbitrary history.

[6] That is, $sgn(Q_{e^i a}^i) = sgn(Q_{e^j a}^j) \neq sgn(Q_{e^i\phi}^i) = sgn(Q_{e^j\phi}^j)$. Also we assume $Q_{e^i\phi}^i \neq 0$ to rule out the uninteresting case where agents always exert the same amount of efforts.

are positive, then more structure is needed to determine how the agent best responds.[7] Finally, the third assumption guarantees that the agent can always identify a point belief that perfectly explains the distribution of the payoffs he will observe given any fixed action profile.

Actions The agents are myopic and maximize their payoff in the current period. Since the history of payoffs and efforts is public, agents' posteriors $\left\{\pi_{t-1}^1, \pi_{t-1}^2\right\}$ are common knowledge. We assume that they use the 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 of effort is diminishing, $Q^i_{e^ie^i} < 0$; (ii) $Q^i_{e^i}\left(\underline{e}, e^j, \tilde{a}^i, \phi^i\right) > 0 > Q^i_{e^i}\left(\overline{e}, e^j, \tilde{a}^i, \phi^i\right), \forall e^j, \phi^i$; (iii) the diminishing return dominates any complementarity or substitutability between efforts, $|Q^i_{e^ie^i}| > |Q^i_{e^ie^j}|$, with $Q^i_{e^ie^j} \geq 0$ for all values or ≤ 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 common knowledge of rationality. 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}.[8] 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.

Timing At time t, agent i chooses effort e^i_t according to his beliefs π^i_{t-1}. Then after observing his payoff q^i_t, and the other agent's payoff, q^j_t, the agent updates his posteriors π^i_t and μ^i_t and then enters the next period.

[7]Indeed, as shown in the Appendix (see the proof of Proposition 4), whether the agent exerts more effort critically depends on the sign of $Q^i_{e^ia} - Q^i_{e^i\phi}\frac{Q^i_a}{Q^i_\phi}$, which can be signed if Assumption 1 holds. Heidhues, Kőszegi, and Strack (2018) also made this assumption.

[8]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 payoffs 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 in their model if players start with conjectures on each other's actions that are correct in equilibrium since 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.

2.1 Examples

We present a few parametric applications 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 different parts of a joint project. Each team member's payoff depends on a common fundamental representing the the quality of their supervisor's overall project idea. They are both overconfident in their research ability and are periodically evaluated on their progress. For a concrete functional form, let $Q^i\left(e^i, e^j, a^i, \phi\right) = \left(e^i + ke^j\right)\phi + a^i + \lambda e^i e^j - c\left(e^i\right)^2$, where $k, \lambda > 0, c > \lambda/2$. Notice that agent i's effort and the fundamental are complements—a higher belief in the fundamental motivates a greater input of effort.[9] The efforts of the agents are complements too as they share knowledge and experience gained from reading articles or testing out different methods.

Example 2. Two VC firms simultaneously choose how much to invest in a growing industry, where the marginal return is decreasing in the total amount of investment. Each is overconfident in their ability to identify the best startups within the industry, but both are unsure of the prospects of the industry as a whole, captured by ϕ. In each period, firm i's payoff is given by $Q^i(e^i, e^j, a^i, \phi) = e^i\phi - \lambda e^i e^j + a^i - c(e^i)^2$, where $\lambda > 0, c > \lambda/2$.

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.[10] 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 agency's utility is given by $Q^i(e^i, e^j, a^i, \phi) = a^i + \phi - L(\phi - e^i) + \lambda e^i e^j - c\left(e^i\right)^2$, where $\lambda > 0$, $c > \lambda/2$, and L is a positive 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).

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

[10] For instance consider two US agencies: the SEC and CFTC. Both agencies are tasked with regulating financial products. In the case of regulating financial swaps, the SEC writes rules pertaining to 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)).

3 Steady State

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

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

(i) e^i is optimal given π^i and e^j, i.e.

$$e^i \in \arg\max_{\tilde{e}^i} \mathbb{E}_{\pi^i}\left[Q^i\left(\tilde{e}^i, e^j, \tilde{a}^i, \phi^i\right)\right]. \tag{1}$$

(ii) For all ϕ^i in the support of π^i and all a^j in the support of μ^i,

$$\left(\phi^i, a^j\right) \in \arg\min_{\hat{\phi}^i, \hat{a}^j} K^i\left(\boldsymbol{e}, \hat{\phi}^i, \hat{a}^j\right) \tag{2}$$

where $K^i\left(\boldsymbol{e}, \hat{\phi}^i, \hat{a}^j\right)$ represents the KL divergence, given by

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

By strict concavity of the payoff function, mixed strategy equilibria are ruled out. 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; on the other hand, 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 $\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \boldsymbol{\phi}\right) \equiv \left(e^{*i}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\phi}\right), e^{*j}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\phi}\right)\right)$ be the solution to Eq. (1) where each agent i assigns probability 1 to ϕ^i. We can alternatively define it as the

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

solution to the first-order condition:

$$Q^i_{e^i}\left(\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}},\boldsymbol{\phi}\right),\tilde{a}^i,\phi^i\right)=0,\forall i. \tag{3}$$

Essentially, this is the Nash equilibrium of a one-shot game when we fix the beliefs in the fundamental to a Dirac measure at $\boldsymbol{\phi}$. It is straightforward to show the existence and uniqueness of $\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}},\boldsymbol{\phi}\right)$ for all $\boldsymbol{\phi}$ under Assumption 2.

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

Next, we define the *gap* function g^i for each player i,

$$g^i\left(\boldsymbol{e},\phi^i\right)\equiv Q^i\left(\boldsymbol{e},A^i,\Phi\right)-Q^i\left(\boldsymbol{e},\tilde{a}^i,\phi^i\right), \tag{4}$$

and let $\boldsymbol{g}\left(\boldsymbol{e},\boldsymbol{\phi}\right)\equiv\left(g^i\left(\boldsymbol{e},\phi^i\right),g^j\left(\boldsymbol{e},\phi^j\right)\right)=0$ denote the *no-gap condition for oneself.* This captures the discrepancy between the actual average payoff and agent i's expected average payoff when agents choose $\boldsymbol{e}$ and the agent i holds the belief that is concentrated at ϕ^i. Intuitively, fix the efforts $\boldsymbol{e}$, the solution to the no-gap condition gives the fundamental value that agent i finds most likely because it perfectly matches the distribution of payoffs. Analogously, the best guess that agent i has about agent j's true ability is captured by a similar condition,

$$\gamma^i\left(\boldsymbol{e},\phi^i,a^j\right)\equiv Q^j\left(\boldsymbol{e},A^j,\Phi\right)-Q^j\left(\boldsymbol{e},a^j,\phi^i\right)=0. \tag{5}$$

Analogously, let $\boldsymbol{\gamma}\left(\boldsymbol{e},\boldsymbol{\phi},\boldsymbol{a}\right)\equiv\left(\gamma^i\left(\boldsymbol{e},\phi^i,a^j\right),\gamma^j\left(\boldsymbol{e},\phi^j,a^i\right)\right)=0$ denote the *no-gap condition for the opponent.* Notice that agent i relies on his own guess about the common fundamental to update about the other agent's ability, thereby preserving and transmitting errors in inferences. The two sets of no-gap conditions exactly characterize the point where the weighted Kullback-Leibler divergence is minimized to 0 for both agents. We thus obtain the following lemma.

Lemma 2. *Fix any $\tilde{\boldsymbol{a}}\in\times_I\left(\underline{a},\overline{a}\right)$. Under Assumptions 1 and 2, there exists at least one Berk-Nash equilibrium. Moreover, each equilibrium $\mathbf{e}_\infty$ is associated with a supporting belief that is a Dirac measure at $\left(\boldsymbol{\phi}_\infty,\hat{\boldsymbol{a}}_\infty\right)$, which satisfies the following:*

(i) Optimality: $\mathbf{e}_\infty=\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}},\boldsymbol{\phi}_\infty\right)$.

(ii) Consistency: $\boldsymbol{g}\left(\mathbf{e}_\infty,\boldsymbol{\phi}_\infty\right)=\boldsymbol{\gamma}\left(\mathbf{e}_\infty,\boldsymbol{\phi}_\infty,\hat{\boldsymbol{a}}_\infty\right)=0.$

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

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

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

We now 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 $\boldsymbol{\phi}_{\infty}$ and the true value of the fundamental Φ 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. 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.*

Definition 2. We say agent j creates an *informational externality* for agent i when agent j's action changes agent i's inference about ϕ^i, or equivalently, at least one of $Q^i_{e^j a}, Q^i_{e^j \phi}$, or $Q^i_{e^i e^j}$ is nonzero.

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 q^i, 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 $Q^i_{e^j a}$ and $Q^i_{e^j \phi}$. Meanwhile, the indirect effect operates through agent i's optimization process. When $Q^i_{e^i e^j} \neq 0$, 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 further changing the underlying distribution of q^i.[12] 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 over i is *positive* if $Q^i_{e^i e^j} \geq 0$, $\text{sgn}(Q^i_{e^i \phi}) = \text{sgn}(Q^i_{e^j \phi})$, and $\text{sgn}(Q^i_{e^i a}) = \text{sgn}(Q^i_{e^j a})$; it is *negative* if $Q^i_{e^i e^j} \leq 0$, $\text{sgn}(Q^i_{e^i \phi}) \neq \text{sgn}(Q^i_{e^j \phi})$, and $\text{sgn}(Q^i_{e^i a}) \neq \text{sgn}(Q^i_{e^j a})$; otherwise, it is *neither positive or negative.*

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 creates positive informational externality, the agents' efforts are complements. Furthermore, the parameter value ϕ^i that solves $g^i(\boldsymbol{e}, \phi^i) = 0$ 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 way and are mutually reinforcing. Analogously, negative informational externality of agent j amounts to substitutable efforts with opposite influences on agent i's belief.

We use two examples to illustrate how the positive or negative informational externalities lead to distortions of one's belief in opposite ways.

Example 1 (cont.)**.** Consider the example with the engineers where agents are learning about the quality of the overall project their PI has assigned them to. Simple calculations establish that $Q^i_{e^i e^j}, Q^i_{e^i \phi}, Q^i_{e^j \phi} > 0$ and $Q^i_{e^i a}, Q^i_{e^j a} = 0$. Therefore, the engineers have positive informational externalities over one another. The gap function, $g^i(\boldsymbol{e}, \phi) = (e^i + ke^j)(\Phi - \phi) + A^i - \tilde{a}^i$, is an increasing function in e^j and e^i whenever an engineer is both overconfident and underestimates the fundamental. First notice that, ceteris paribus, a lower e^j implies a lower belief over ϕ in order to compensate for a decreased g^i, or a negative gap. More intuitively, since the marginal return to the unknown quality of the project, ϕ is increasing in the teammate's effort ($Q^i_{e^j \phi} > 0$), the marginal return on ϕ decreases in response to a lower e^j. Hence, engineer i believes the quality of the overarching project has to be be much worse to justify his own underperformance. Moreover, since $Q^i_{e^i e^j} > 0$, a lower e^j motivates engineer i to play a even lower action, amplifying the downward effect on his belief in the project quality, ϕ.

[12] Notably, the notion of informational externality is distinct from informative actions in social learning environments. The agents do not infer the fundamental from each other's effort choice but face a signal structure that varies with the efforts.

Example 2 (cont.)**.** Now consider the venture capital example where firms are unsure of the prospects of the whole industry. It is straightforward to verify that the informational externalities are negative. Overconfident managers tend to attribute underperformance to the industry and underestimate the marginal return to investment, thereby underinvesting in the industry ($Q^i_{e^i a} = 0$, $Q^i_{e^i\phi} > 0$). Notice that, upon observing a lower e^j, since investments are substitutes ($Q^i_{e^i e^j} < 0$), firm i is motivated to seek a larger investment, which feeds back to the no gap condition and generate a higher belief in ϕ: intuitively, as the gap function is given by $g^i(\boldsymbol{e}, \phi) = e^i(\Phi - \phi) + A^i - \tilde{a}^i$, firm i need not underestimate the growth of the industry as much as before to justify the constant gap introduced by overconfidence.

4.2 Mutually-Reinforcing Learning

Motivated by these observations, we consider the following question: how do the agents mutually influence their beliefs in the unknown fundamental 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 fundamental, thereby encouraging more extreme actions and even lower payoff. We first show that this pattern is reinforced 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. Denote the steady-state inferences as $\boldsymbol{\phi_S} = \left(\phi^i_S, \phi^j_S\right)$ and actions as $\boldsymbol{e_S} = \left(e^i_S, e^j_S\right)$. We now compare the steady state with our two-agent environment in which we allow both agents to adjust actions, which leads to the steady state $(\boldsymbol{e_\infty}, \boldsymbol{\phi_\infty})$ as defined earlier. The following proposition shows that, as more agents actively participate in action optimization, the steady-state underestimation becomes more severe. We thus say that their learning processes are *mutually-reinforcing*.

Proposition 1. *Suppose Assumptions 1 to 3 hold and agent j has a positive informational externality over agent i, then both agents' underestimation of their fundamentals 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 $Q^i_{e^i\phi} > 0$, then $e^j_S > e^j_\infty(\tilde{\boldsymbol{a}})$;

(ii) if $Q^i_{e^i\phi} < 0$, 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 $\boldsymbol{\phi_\infty}(\tilde{\boldsymbol{a}}) < \boldsymbol{\phi_S} < \boldsymbol{\Phi}$.

One may wonder what role is played by the requirement that $e^j_S > e^j_\infty(\tilde{\boldsymbol{a}})$ or $e^j_S < e^j_\infty(\tilde{\boldsymbol{a}})$—they ensure 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 $e^j_S = e^j_\infty(\tilde{a}^i, A^j)$,[13] and then interpret the constraint as a suggestion from an outside analyst who tries to mitigate the distortion due to overconfidence: fix your action at a level which best responds to a correct self-perception, then both of you will understand the environment better and have better performance.

The message conveyed by Proposition 1 is twofold. First, a self-defeating pattern emerges since $\phi^j_\infty(\tilde{\boldsymbol{a}}) < \phi^j_S < \Phi^j$. That is, agent j underestimates the fundamental more when he is allowed to optimize. More importantly, a mutually-reinforcing pattern can be observed by noting that $\phi^i_\infty(\tilde{\boldsymbol{a}}) < \phi^i_S < \Phi$, which means agent i's inference also turns out to be more extreme when agent j can freely change actions. The key driving force is the positive informational externality. To illustrate the mechanism, we describe the learning dynamics heuristically using Example 1.

Example 1 (cont.)**.** When engineer i holds a degenerate belief at ϕ^i and his coworker's effort is fixed at e^j_S, engineer i optimally chooses $e^{*i} = \left(\phi^i + \lambda e^j_S\right)/2c$, which increases in both ϕ^i and e^j. We can plot this function in the $e^i - \phi^i$ domain as in Figure 1. In addition, the no-gap equation yields $\phi^i = \Phi - (\tilde{a}^i - A^i)/\left(e^i + \lambda e^j_S\right)$, which is the blue curve we plot in Figure 1.

If engineer j is forced to take a relatively high effort e^j_S and engineer i starts to optimize against the high e^j_S as shown in the figure, engineer i scales down the effort because he underestimates the overall project quality, ϕ. This decrease in effort results in lower payoff in the following period, resulting in an even lower belief from engineer i. Eventually, engineer i is going to hold a belief ϕ^i_S, which is lower than the belief he started with. This process is shown in the left panel of Figure 1. Suppose now engineer j is also given the chance to optimize; the dynamics change dramatically. Since both engineers have the tendency to scale down their effort and the payoff function admits complementarity between efforts, engineer i exerts lower effort than he did when his coworker was constrained to play a fixed action. In the right panel of Figure 1, this is captured by an downward shift of the optimal action curve. Next, because $Q^i_{e^k\phi} > 0, \forall k$, the decrease in effort leads to a more negative gap between the true and expected payoffs and hence a larger decline in engineer i's evaluation of ϕ, i.e. the

[13] It can be shown that when $Q^i_{e^i\phi} > (<)0$, we have $e^j_\infty\left(\tilde{a}^i, A^j\right), e^j_\infty(\boldsymbol{A}) > (<) e^j_\infty(\tilde{\boldsymbol{a}})$.

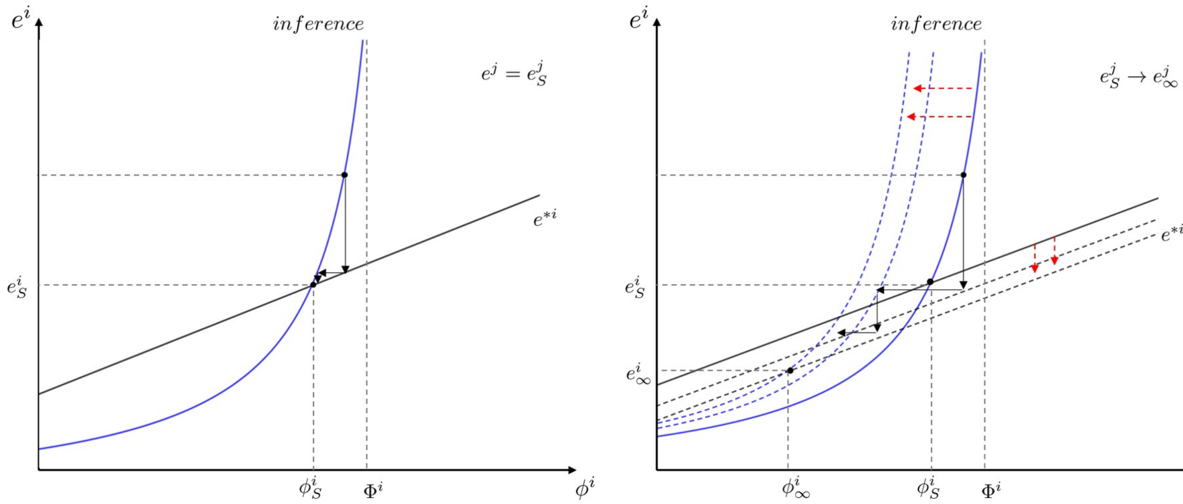

Figure 1: Mutually-reinforcing learning. The left panel shows how allowing agent i to change her action induces a lower inference for agent i, with e^j fixed at e^j_S, and the right panel shows how agent i gets an even lower inference when agent j is also allowed to revise actions. The black straight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

belief formation curve shifts to the left. The same process repeats until both engineers reach the steady state with actions e^i_∞ and beliefs ϕ^i_∞, which are potentially much more extreme than e^i_S and ϕ^i_S.

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, it follows from positive informational externalities that they each have a worse evaluation of the unknown. Similar to Proposition 1, it also embeds the one-agent self-defeating result, as increasing one's confidence level worsens one's own underestimation.

Proposition 2. *Suppose Assumptions 1 to 3 hold and both agents create positive informational externalities. When any of the agents is more overconfident, both agents' underestimation of the fundamental is more severe. That is, let* $\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{a}}' \in \times_I (\underline{a}, \overline{a})$ *and* $\tilde{\boldsymbol{a}} > \tilde{\boldsymbol{a}}' > \boldsymbol{A}$*, then* $\boldsymbol{\phi}_{\boldsymbol{\infty}}(\tilde{\boldsymbol{a}}) < \boldsymbol{\phi}_{\boldsymbol{\infty}}(\tilde{\boldsymbol{a}}') < \boldsymbol{\Phi}$*.*

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 take the form of increasing information sharing to eliminate overconfidence and underestimation, or simply restricting the action choices of certain agents. Last but not least, 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, as one may extrapolate from the previous result, the learning processes become *mutually-limiting*. In particular, allowing another agent to freely optimize will make the original agent's belief distortion less severe. Similarly, increase a second agent's overconfidence will cause the first agent's inferences to be closer to the true value of the unknown.

Proposition 3. *Suppose agent j has a negative informational externality over agent i.*

(i) Agent i's underestimation of the fundamental 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 as before), i.e. $\phi_S^i < \phi_\infty^i(\tilde{\boldsymbol{a}}) < \Phi$.

(ii) As agent j becomes more overconfident, agent i's underestimation of the fundamental 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 $\phi_\infty^i(\tilde{\boldsymbol{a}}') < \phi_\infty^i(\tilde{\boldsymbol{a}}) < \Phi$.

Example 2 (cont.)**.** Firm j immediately scales down its investment when it is free to optimize or if the manager becomes more overconfident. Then as illustrated earlier, firm i is incentivized to incrementally invest more, which pushes up the former firm's belief in the fundamental, correcting the underestimation stemming from overconfidence. If we were to plot the dynamics as we did in Figure 1, as firm j chooses its investment, the belief formation curve stays still but the optimal action curve shifts upwards, moving the intersection point to the right.

4.4 Welfare Analysis

In this subsection, we analyze the welfare implications of overconfidence by studying how the payoffs in the steady state. $(\boldsymbol{e}_\infty(\tilde{\boldsymbol{a}}), \boldsymbol{\phi}_\infty(\tilde{\boldsymbol{a}}))$ vary with $\tilde{\boldsymbol{a}}$, and in particular, whether the payoffs increase or decrease compared to when agents have correct perceptions. 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, and even sometimes good for everyone, but only in a multi-agent environment; second, we could easily tell the direction of the change in welfare when the bias is small by checking a few simple conditions.

The actual average payoff can be rewritten as follows,

$$Q^i\left(\boldsymbol{e}_{\boldsymbol{\infty}}\left(\tilde{\boldsymbol{a}}\right), A^i, \Phi\right) = Q^i\left(e^{*i}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\phi}_{\boldsymbol{\infty}}\left(\tilde{\boldsymbol{a}}\right)\right), e^{*j}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\phi}_{\boldsymbol{\infty}}\left(\tilde{\boldsymbol{a}}\right)\right), A^i, \Phi\right), \forall i.$$

Since optimal actions are determined simultaneously, $e^{*i}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\phi}_{\boldsymbol{\infty}}\left(\tilde{\boldsymbol{a}}\right)\right)$ depend on $\tilde{a}^j$ and $\phi^j_\infty\left(\tilde{\boldsymbol{a}}\right)$ indirectly through $e^{*j}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\phi}_{\boldsymbol{\infty}}\left(\tilde{\boldsymbol{a}}\right)\right)$. 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 deviation of ϕ^i from Φ; (iii) the distortion of e^i due to the distortion of e^j *(complementarity/substitutability)*; (iv) the direct effect of e^j on Q^i (*payoff externality*).

An easy observation is that the sum of effect (i) and effect (ii) is almost always negative: misconceptions always impair the agent's ability to choose the correct actions. Besides, when there is only a single actively-optimizing agent or when Q^i 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.[14] We summarize these observations below.

Claim 1. If the payoff function for any agent i has the following form

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

misspecification in a^i leads to lower or equal average payoff for agent i in the steady state.

When there are multiple agents, the payoff function becomes more complicated and effects (iii) and (iv) start to kick in. The extent to which effects (iii) and (iv) harm or benefit the agents depends on specific parametric assumptions made about Q^i and how much $\tilde{\boldsymbol{a}}$ deviates from $\boldsymbol{A}$.

In Proposition 4, we characterize the welfare impact of small amount of overconfidence.

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

The focus on a small misspecification facilitates the characterization in two ways. For one thing, since e^i has been optimized, the payoff change will be dictated by the change of e^j and the derivative of Q^i 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. For another thing, the change in e^j will be determined by $\tilde{a}^j$ as the influence of e^i is of smaller size.

Proposition 4. *Suppose Assumptions 1 to 3 hold. There exists $\delta > 0$ such that, if (i) the agents' overconfidence levels are given by $\tilde{\boldsymbol{a}} \in B_\delta^+(\boldsymbol{A})$ and (ii) $Q^j_{e^j\phi}$ has the opposite (same) sign as $Q^i_{e^j}(\boldsymbol{e_\infty}(\boldsymbol{A}), A^i, \Phi)$, then agent i's payoff in the steady state increases (decreases) compared to when agents are correctly specified about their ability.*[15]

Proposition 4 provides a simple condition under which an overconfident agent can be better off. For example, when $Q^j_{e^j\phi} < 0$ and $Q^i_{e^j}(\boldsymbol{e_\infty}(\boldsymbol{A}), A^i, \Phi) > 0$, both agents step up their efforts due to overconfidence and underestimation of the fundamental, which benefit all agents because it corrects the inefficiency of insufficient efforts in common-good problems. In other words, agent j should have paid more effort than he is willing to because of his positive payoff externality—his overconfidence, albeit a perception bias, accidentally leads to a better outcome. It is worth noting that, if the agents enjoy higher payoffs under the above circumstances, positive informational externalities further increase the welfare. In other words, positive informational externalities and mutually-reinforcing learning should not be simply interpreted as negative results when it comes to welfare; instead, they magnify the payoff externalities, and lead to outcomes which are either "even better" or "even worse". By contrast, negative informational externalities reduce the payoff externalities.

We now compare our welfare predictions to the computed steady states in a neighborhood around the agents' true abilities for the two teamwork examples.

Example 1 (cont.)**.** It is simple to verify that $Q^i_{e^i\phi} > 0, Q^i_{e^j} > 0$, for all i and $j \neq i$. By Proposition 4, when the engineers exhibit slight overconfidence and thus underestimation of the unknown, the engineers' payoffs decrease compared to when they are correctly specified, which corresponds to the upper right-hand quadrant of Figure 2's first panel. Here, mutually-reinforcing learning exacerbates the engineers' underestimation of their project quality and distorts their actions, further dropping their welfare.

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

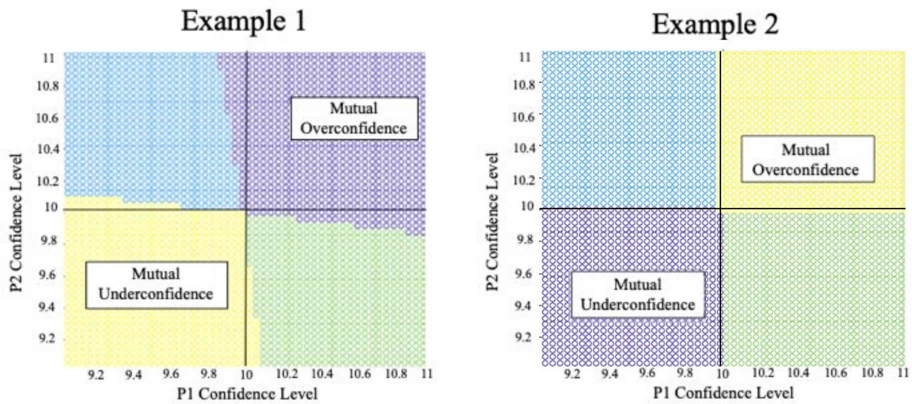

Figure 2: Welfare changes for Examples 1 and 2. Both panels represent the change in welfare to both agents when both agents have true ability $A^1 = A^2 = 10$ and $\tilde{\boldsymbol{a}}$ as indicated by the graphs. The purple areas represent specifications of confidence where the welfare of both agents 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 $Q^i(e^i, e^j, a^i, \phi^i) = (e^i + 2e^j)\phi^i + a^i + e^i e^j - 3(e^i)^2$ for the first panel, and $Q^i(e^i, e^j, a^i, a^j) = e^i\phi - 3e^i e^j + a^i - 10(e^i)^2$ for the second panel.

Example 2 (cont.)**.** Again, it is simple to verify that $Q^i_{e^i\phi} > 0, Q^i_{e^j} < 0$. By Proposition 4, both firms' average payoffs in the steady state increase as they are marginally more overconfident. These changes are reflected from the second panel of Figure 2. The mutual-limiting learning pattern, while alleviating the agents' perception bias, also weakening the incentives for making larger investments thus decreasing their payoffs locally.

In sum, these examples highlight the fact that the welfare impact of overconfidence on a pair of agents highly depends on the directions of payoff externalities and informational externalities, and should be carefully analyzed case by case.

5 Convergence

In this section, we prove that under positive or negative informational externalities, the multi-agent learning processes converge to the steady state, or 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 ϕ^i if its implied distribution of payoffs exhibits systematic mismatch with true distribution in the sense that $Q^i(\boldsymbol{e_t}, \tilde{a}^i, \phi^i)$ should not be consistently lower or higher than $Q^i(\boldsymbol{e_t}, A^i, \Phi)$. We then use a contraction argument similar to theirs: the structural properties of the payoff functions enable us to eliminate a subset of actions given all possible beliefs, which in turn further rules out a subset of beliefs.

However, 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, 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 $\boldsymbol{\phi_\infty}$ and $\boldsymbol{\hat{a}_\infty}$.*

Figure 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 unfeasible 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.

Example 1 (cont.)**.** Figure 3 demonstrates the convergence of beliefs and actions in a simulation. We see that engineer i's effort and inference converge both in the case where his coworker is constrained to a fixed action (which corresponds to the orange paths in the figure) and in the case where engineer j is allowed to adjust his inferences and efforts alongside engineer i (which corresponds to the blue paths in the figure). Note that although the payoff is consistently perturbed by the random noise, it stays centered around the steady state. Additionally, the perturbations do not cause much change in inference and effort levels as t grows larger. Also notice that it takes longer for effort inference and payoff to approach their steady state levels in the case where both engineers are allowed to re-optimize their actions than in the case where engineer j is forced to play a fixed action.

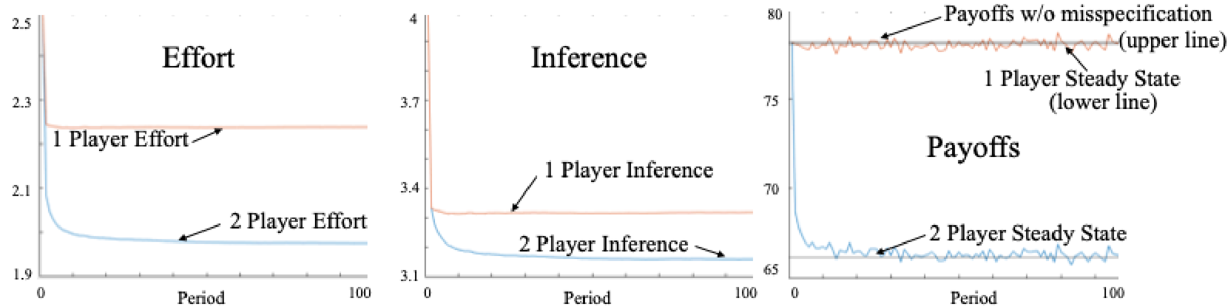

Figure 3: This figure shows the process by which effort, inferences, and average payoffs converge to steady state levels in Example 1. In particular, $Q^i(e^i, e^j, a^i, \phi^i) = (e^i + 5e^j)\phi + a^i + e^i e^j - 1.3(e^i)^2$, $\phi = 4$, $A^1 = A^2 = 20$, and $\tilde{a}^1 = \tilde{a}^2 = 30$. We assume both the distributions of random noises and priors are normal, which enables us to keep track of the mean of each period's posterior distribution. The orange paths indicate what happens to agent 1's actions, inferences (mean of posterior), and payoffs when we force agent 2 to play the action which is optimal when both players hold correct beliefs about ϕ. The blue paths represent agent 1's actions, inferences, and payoffs when both players are allowed to change their actions in response to their changing efforts.

In addition to convergence, Figure 3 highlights our mutually reinforcing learning result. We see how when we allow engineer j to also adjust his efforts, engineer i's efforts and inferences become even lower than they were when his coworker's actions were fixed. Since both engineers exert less effort, engineer i's welfare also suffers as a result of allowing engineer j to adjust his effort.

6 Extensions

6.1 Underconfidence

We begin this section by discussing the implications of underconfidence. The assumption that $\tilde{\boldsymbol{a}} > \boldsymbol{A}$ is critically important to the direction of the mutual learning effect. In fact, when the agents are underconfident, the learning processes are *mutually-limiting* when there are positive informational externalities. The belief formation curve is now downward sloping since as agents exert more effort, the marginal return of the fundamental increases, leading the agents to overestimate the unknown to a lesser extent. Assume for now that $Q^j_{e^j\phi} > 0$, then an underconfident agent chooses lower effort in the steady state than when he is correctly specified, i.e. $e^j_\infty(\tilde{\boldsymbol{a}}) < e^j_\infty(\tilde{a}^i, A^j)$. 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 in the latter case, agent j exerts lower effort, thus worsening agent i's evaluation of ϕ^i. Consequently, positive informational externalities in the underconfidence case help correct the overestimation of ϕ^i. Analogously, negative informational externalities in the underconfidence case generate more extreme overestimation in the steady state, generating a mutually-reinforcing pattern. These phenomena are illustrated by Figure 4 below.

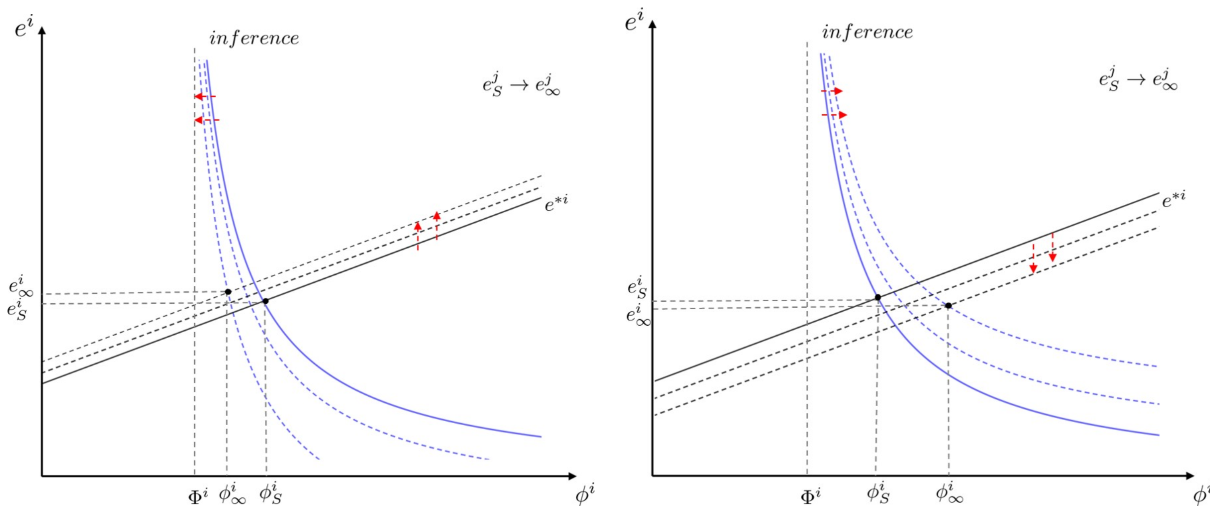

Figure 4: Mutual learning with underconfidence. The left panel shows the heuristic learning dynamics of an underconfident agent when informational externalities are positive. Steady state inferences are lower than when agent j's effort choice is fixed. The right panel shows a reverse pattern when informational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

There seems to be no guarantee that the learning processes will converge: as agents obtain a lower belief in $\boldsymbol{\phi}$, they react by exerting lower effort (in the case of $Q^j_{e^j\phi} > 0$ and $Q^j_{e^ja} \leq 0$), 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.[16] Note that

[16]The assumption of not being too underconfident is essential for the use of the contraction argument,

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 $A^i - \tilde{a}^i$ is small enough for all i, 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 $\boldsymbol{\phi}_\infty$ and $\hat{\boldsymbol{a}}_\infty$.*

Finally, going back to Figure 2, we find that overconfidence and underconfidence impact the agents' steady state payoffs in opposite directions. As a side result, we show that there always exists a certain type of bias in $\tilde{\boldsymbol{a}}$ that is socially beneficial when there are multiple agents in our environment. The key to this result is the Pareto inefficiency of Nash equilibria due to externalities and the flexibility of the self-perception bias which makes it possible to counteract or strengthen the inefficiencies. Therefore, the insight of the possibility of Pareto improvement as a result of biases goes beyond this particular learning framework and the specific form of bias.

Proposition 5. *Suppose $Q^i_{e^j}\left(\boldsymbol{e}_\infty\left(\boldsymbol{A}\right), A^i, \Phi\right) \neq 0$. Then there exists $\tilde{\boldsymbol{a}}$ with which both agents have payoffs strictly higher in the steady state than they would have when both are correctly specified.*[17]

6.2 Multiple Equilibria

The existence of multiple equilibria does not affect our key message, i.e. mutually-reinforcing and mutually-limiting learning, but posea difficulties for the proof of convergence of the learning process. Figure 5 shows the single-agent heuristic learning dynamics, where there are two equilibria (e^i_S, ϕ^i_S) and $\left(\hat{e}^i_S, \hat{\phi}^i_S\right)$. The contraction argument fails due to no further elimination when, for example, the lower and upper bounds of actions are given by $\hat{e}^i_S$ and e^i_S. Hence, this paper does not provide a proof of convergence for such settings. Clearly, the former equilibrium (e^i_S, ϕ^i_S) is more attractive since the heuristics make it clear that the agent's

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. For example, the technique developed by Esponda, Pouzo, and Yamamoto (2019), which tackles the problem of convergence by focusing on the asymptotic frequency of actions and using tools from the theory of stochastic approximation, can be adapted to establish convergence only in a *discrete-action* version of our setting.

[17]The proof also establishes that the existence of $\tilde{\boldsymbol{a}}$ with which one agent obtains higher payoff while the other obtains lower payoff and the existence of $\tilde{\boldsymbol{a}}$ such that both agents obtains lower payoff.

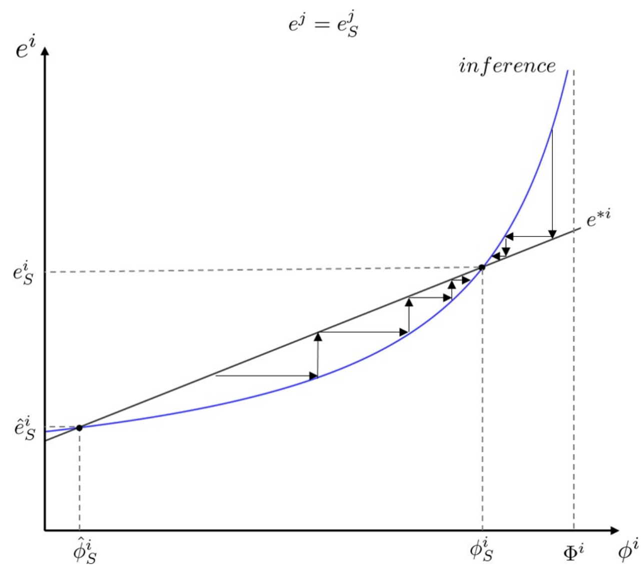

Figure 5: Multiple Berk-Nash equilibria

belief will drift towards ϕ^i_S as he optimizes and updates. In fact, one can use the stochastic approximation tools from Esponda, Pouzo, and Yamamoto (2019) to show convergence in a finite-action environment—but unfortunately, their techniques do not directly apply to continuous-action settings like ours. Nevertheless, we still feel free to focus on equilibria like (e^i_S, ϕ^i_S) 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 easily generalized to an arbitrary number of agents. The pattern of mutually-reinforcing learning exists and becomes exacerbated with more overconfident agents exerting positive informational externalities among each other. The parametric assumptions, however, tend to be increasingly more complicated since more cross derivatives are involved. Moreover, the possibility of Pareto improvement when players are slightly biased always exists. As long as the optimal choice of actions non-trivially depends on self-perceptions and inferences on the unknowns and each agent is exposed to some level of externalities, we could identify small biases with which everyone enjoys strictly high utility.

7 Conclusion

We develop a two-agent learning model with overconfidence. We define a new notion of informational externalities to describe how one agent's action could influence the other agent's inference. When positive informational externalities are present, we find a mutually reinforcing learning pattern that implies strategic interaction exacerbates the underestimation of the common fundamental and make 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 is actively learning and optimizing. Our welfare implications also starkly contrast with welfare results from single-agent models in that there can be Pareto improvement in welfare as a result of overconfidence, and mutually-reinforcing learning potentially improve welfare even further.

One possible 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, agent may be incentivized to play actions that are non-optimal in the stage game, leading to different distortions in long-term beliefs.

A Preliminary Lemmas

For convenience, define $G^i(\boldsymbol{\phi}) \equiv g^i(\boldsymbol{e}^*(\tilde{\boldsymbol{a}}, \boldsymbol{\phi}), \phi^i)$ and denote $\boldsymbol{G}(\boldsymbol{\phi}) = (G^i(\boldsymbol{\Phi}), G^j(\boldsymbol{\Phi}))$. Essentially, this is the gap function when every agent actively optimizes according to a degenerate belief at ϕ^i. The Berk-Nash equilibrium belief satisfies $G^i(\boldsymbol{\phi_\infty}) = 0, \forall i$. Besides, let $\overline{\kappa}_a \geq \max\{Q_a^1, Q_a^2\}$ denote the upper bound on Q_a^i and $\underline{\kappa}_\phi \leq \min\{Q_\phi^1, Q_\phi^2\}$ denote the lower bound on Q_ϕ^i throughout.

Proof of Lemma 1. Suppose $e^i \in \arg\max_e Q^i(e, e^j, \tilde{a}^i, \phi^i)$, 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 $e^*(\tilde{\boldsymbol{a}}, \boldsymbol{\phi})$ 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

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

Combined with Assumption 2, it is implied that $|e^j - \hat{e}^j| > |e^i - \hat{e}^i|, \forall i, j \neq i$. 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,

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

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

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

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

It follows that $G^i\left(\Phi - \frac{\overline{\kappa}_a}{\underline{\kappa}_\phi}\left(\tilde{a}^i - A^i\right), \phi^j\right) \leq 0, \forall \phi^j$. Since $G^i\left(\Phi, \phi^j\right) > 0, \forall \phi^j$, 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. $\square$

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

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

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

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

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

$$\frac{\partial G^i\left(\boldsymbol{\phi}\right)}{\partial \phi^i} < 0, \left|\frac{\partial G^i\left(\boldsymbol{\phi}\right)}{\partial \phi^j}\right| < \left|\frac{\partial G^i\left(\boldsymbol{\phi}\right)}{\partial \phi^i}\right|, \forall i. \tag{6}$$

Suppose there are at least two different roots, $\tilde{\boldsymbol{\phi}}$ and $\hat{\boldsymbol{\phi}}$, and assume without loss of generality $\tilde{\phi}^i < \hat{\phi}^i$. By the second inequality in Eq. (6), if $\boldsymbol{G}\left(\tilde{\boldsymbol{\phi}}\right) = \boldsymbol{G}\left(\hat{\boldsymbol{\phi}}\right) = 0$, then it must be that $|\tilde{\phi}^j - \hat{\phi}^j| > |\tilde{\phi}^i - \hat{\phi}^i|, \forall i, j \neq i$. The statement contradicts itself. $\square$

Lemma 4. *Given $\tilde{a}^i$ and $\pi^i_{t-1}\left(\phi^i\right)$, $\forall i$, the stage game at time t is dominance solvable. That is, there exists a unique rationalizable action profile $\left(e^i_t, e^j_t\right)$.*

Proof. Without loss of generality, assume $Q^i_{e^ie^j} > 0, \forall i, j \neq i$. 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 any $k \in \{1, 2\}$ and $\tau \geq 1$,

$$\begin{aligned}\underline{b}^k_\tau &= \arg\max_{e^k} \mathbb{E}_{\pi^k_{t-1}}\left[Q^k\left(e^k, \underline{e}^k_{\tau-1}, \tilde{a}^k, \phi^k\right)\right] \equiv \arg\max_{e^k} h^k\left(e^k, \underline{e}^k_{\tau-1}, \tilde{a}^k\right), \\ \overline{b}^k_\tau &= \arg\max_{e^k} \mathbb{E}_{\pi^k_{t-1}}\left[Q^k\left(e^k, \overline{e}^k_{\tau-1}, \tilde{a}^k, \phi^k\right)\right] \equiv \arg\max_{e^k} h^k\left(e^k, \underline{e}^k_{\tau-1}, \tilde{a}^k\right),\end{aligned}$$

and $\left[\underline{e}_\tau^k, \overline{e}_\tau^k\right] \equiv \left[\underline{e}_{\tau-1}^k, \overline{e}_{\tau-1}^k\right] \cap \left[\underline{b}_\tau^k, \overline{b}_\tau^k\right]$. By Assumption 2, $\left[\underline{e}_1^k, \overline{e}_1^k\right] \subsetneq \left[\underline{e}_0^k, \overline{e}_0^k\right], \forall k$, which further implies that $\left[\underline{e}_2^k, \overline{e}_2^k\right] \subsetneq \left[\underline{e}_1^k, \overline{e}_1^k\right], \forall k$. The Nested Intervals Theorem implies that each agent's set of rationalizable actions 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, agents converge to $\left(e_t^i, e_t^j\right)$ such that

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

The proof is analogous if $Q_{e^i e^j}^i < 0$ for some i and trivial if $Q_{e^i e^j}^i = 0$. $\square$

B Proofs for Section 4

Lemma 5. *Given $\tilde{\boldsymbol{a}}$, the Berk-Nash equilibrium $(\boldsymbol{e}_\infty, \boldsymbol{\delta}_{\phi_\infty})$ satisfies the following:*

$$Q_{e^i}^i\left(\boldsymbol{e}_\infty, A^i, \Phi\right) + Q_\phi^i\left(\boldsymbol{e}_\infty, \tilde{a}^i, \phi_\infty^i\right) \frac{Q_{e^i e^i}^i\left(\boldsymbol{e}_\infty, \tilde{a}^i, \phi_\infty^i\right)}{Q_{e^i \phi}^i\left(\boldsymbol{e}_\infty, \tilde{a}^i, \phi_\infty^i\right)} < 0, \forall i.$$

Proof. Consider agent i's steady state maximization and inference problem for a fixed e^j. For any ϕ^i, define $\hat{e}^i(\phi^i)$ to be such that $Q_{e^i}^i\left(\hat{e}^i(\phi^i), e^j, \tilde{a}^i, \hat{\phi}^i\right)$. We denote the set of possible actions and beliefs as $l_e^i(e^j)$ and $l_\phi^i(e^j)$ respectively, i.e. $(e^i, \phi^i) \in l_e^i(e^j) \times l_\phi^i(e^j)$ if it satisfies the following equations:

$$Q^i\left(e^i, e^j, A^i, \Phi\right) - Q^i\left(e^i, e^j, \tilde{a}^i, \phi^i\right) = 0,$$
$$Q_{e^i}^i\left(e^i, e^j, \tilde{a}^i, \phi^i\right) = 0.$$

Obviously both $l_e^i(e^j)$ and $l_\phi^i(e^j)$ are nonempty and compact for any e^j. Let $l_\phi^{*i}(e^j)$ represent the largest element in $l_\phi^i(e^j)$, and $l_e^{*i}(e^j)$ represent the effort corresponding to $l_\phi^{*i}(e^j)$. Notice that $Q^i(e^i, e^j, A^i, \Phi) - Q^i(e^i, e^j, \tilde{a}^i, \Phi) < 0, \forall e^i, e^j$. Therefore, for any $\hat{\phi}^i > l_\phi^{*i}(e^j)$, it must be that

$$Q^i\left(\hat{e}^i\left(\hat{\phi}^i\right), e^j, A^i, \Phi\right) - Q^i\left(\hat{e}^i\left(\hat{\phi}^i\right), e^j, \tilde{a}^i, \hat{\phi}^i\right) < 0,$$

otherwise one can always find a larger element in $l_\phi^i(e^j)$ than $l_\phi^{*i}(e^j)$, contradicting the definition of $l_\phi^{*i}(e^j)$. This implies that when $\phi^i = l_\phi^{*i}(e^j)$,

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

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

From $Q^i_{e^i}\left(\hat{e}^i\left(\phi^i\right), e^j, \tilde{a}^i, \phi^i\right) = 0$, we know that $Q^i_{e^ie^i}\left(\hat{e}^i\left(\phi^i\right), e^j, \tilde{a}^i, \phi^i\right)\frac{\partial \hat{e}^i}{\partial \phi^i} = -Q^i_{e^i\phi}\left(\hat{e}^i\left(\phi^i\right), e^j, \tilde{a}^i, \phi^i\right)$. Plug back to the previous inequality, we obtain that

$$Q^i_{e^i}\left(\hat{e}^i\left(\phi^i\right), e^j, A^i, \Phi\right) + Q^i_{\phi}\left(\hat{e}^i\left(\phi^i\right), e^j, \tilde{a}^i, \phi^i\right)\frac{Q^i_{e^ie^i}\left(\hat{e}^i\left(\phi^i\right), e^j, \tilde{a}^i, \phi^i\right)}{Q^i_{e^i\phi}\left(\hat{e}^i\left(\phi^i\right), e^j, \tilde{a}^i, \phi^i\right)} < 0.$$

Finally, notice that since $\left(l^{*i}_e\left(e^j\right), l^{*j}_e\left(e^i\right)\right)$ 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. Since $\boldsymbol{e_\infty}$ is a fixed point of the correspondence $\left(l^i_e\left(e^j\right), l^j_e\left(e^i\right)\right)$ and that it is unique by assumption, $\boldsymbol{e_\infty}$ must also be the fixed point of $\left(l^{*i}_e\left(e^j\right), l^{*j}_e\left(e^i\right)\right)$. Thus the above inequality is satisfied in the equilibrium. □

Proof of Proposition 1. Without loss of generality, assume $Q^i_{e^i\phi} > 0$. We can accommodate $Q^i_{e^i\phi} < 0$ by replacing e^i, e^j with $-e^i, -e^j$ and substituting the constraint with $e^j_S < e^j_\infty\left(\tilde{\boldsymbol{a}}\right)$.

We start by showing $e^i_S > e^i_\infty$ and $\phi^i_S > \phi^i_\infty$. Consider the following two equations,

$$Q^i\left(\boldsymbol{e}, A^i, \Phi\right) - Q^i\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right) = 0, \tag{7}$$

$$Q^i_{e^i}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right) = 0, \tag{8}$$

where e^j is fixed but e^i and ϕ^i are unknown. Differentiate,

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

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

Simplify and then we obtain

$$\frac{\partial e^i}{\partial e^j} = \frac{-Q^i_{\phi}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right)Q^i_{e^ie^j}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right) - \left[Q^i_{e^j}\left(\boldsymbol{e}, A^i, \Phi\right) - Q^i_{e^j}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right)\right]Q^i_{e^i\phi}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right)}{Q^i_{e^i\phi}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right)\left(Q^i_{e^i}\left(\boldsymbol{e}, A^i, \Phi\right) + Q^i_{\phi}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right)\frac{Q^i_{e^ie^i}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right)}{Q^i_{e^i\phi}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right)}\right)} > 0,$$

$$\frac{\partial \phi^i}{\partial e^j} = \frac{Q^i_{e^ie^i}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right)\left[Q^i_{e^j}\left(\boldsymbol{e}, A^i, \Phi\right) - Q^i_{e^j}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right)\right] - Q^i_{e^i}\left(\boldsymbol{e}, A^i, \Phi\right)Q^i_{e^ie^j}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right)}{Q^i_{e^i\phi}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right)\left(Q^i_{e^i}\left(\boldsymbol{e}, A^i, \Phi\right) + Q^i_{\phi}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right)\frac{Q^i_{e^ie^i}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right)}{Q^i_{e^i\phi}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right)}\right)} > 0.$$

The positive signs of the derivatives follow from Lemma 5. Notice that both (e^i_S, ϕ^i_S) and $(e^i_\infty, \phi^i_\infty)$ are determined by two equations, i.e. Eqs. (7) and (8), but they correspond to different agent j's efforts. Since $e^j_S > e^j_\infty$, it is implied that $e^i_S > e^i_\infty$ and $\phi^i_S > \phi^i_\infty$.

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

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

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

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

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

Lemma 6. *Suppose agents both create positive informational externalities. Suppose* $\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{a}}' \in \times_I \left(\underline{a}, \overline{a}\right)$ *and* $\tilde{\boldsymbol{a}}' < \tilde{\boldsymbol{a}}$.

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

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

Proof. Consider part (i) first. Differentiate the following equations that determine the steady state with respect to a^i and a^j.

$$Q^i\left(\boldsymbol{e_\infty}\left(\boldsymbol{a}\right), A^i, \Phi\right) - Q^i\left(\boldsymbol{e_\infty}\left(\boldsymbol{a}\right), a^i, \phi^i_\infty\left(\boldsymbol{a}\right)\right) = 0,$$

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

After tedious calculations we obtain the following,

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

where $I(i,k) = -1$ if $i = k$ and $I(i,k) = 1$ if $i \neq k$, and $\omega^{ik} = Q^{-k}_{e^{-k}e^{-i}} + Q^{-k}_{e^{-k}\phi}\frac{\left(Q^{-k,A}_{e^{-i}} - Q^{-k}_{e^{-i}}\right)}{Q^{-i}_\phi} = \frac{Q^{-k}_{e^{-k}\phi}}{Q^{-i}_\phi}\left(Q^{-k,A}_{e^{-i}} - Q^{-k}_{e^{-i}} + Q^{-i}_\phi\frac{Q^{-k}_{e^{-k}e^{-i}}}{Q^{-k}_{e^{-k}\phi}}\right)$. All derivatives are evaluated at $\boldsymbol{e_\infty}\left(\boldsymbol{a}\right), \boldsymbol{a}, \boldsymbol{\phi_\infty}\left(\boldsymbol{a}\right)$, except $Q_{e^{-k}}^{-k,A}$ which is evaluated at $\boldsymbol{e}_{\boldsymbol{\infty}}(\boldsymbol{a}), \boldsymbol{A}, \boldsymbol{\phi}$. From Lemma 5, $\omega^{ii}, \omega^{jj} < 0$ and $\omega^{ij}, \omega^{ji} > 0$. It follows that that $\frac{\partial e_{\infty}^i(\boldsymbol{a})}{\partial a^i} < 0$ and $\frac{\partial e_{\infty}^i(\boldsymbol{a})}{\partial a^j} < 0$ for all i. Therefore, $\tilde{\boldsymbol{a}}' < \tilde{\boldsymbol{a}} \Rightarrow \boldsymbol{e}_{\boldsymbol{\infty}}(\tilde{\boldsymbol{a}}') > \boldsymbol{e}_{\boldsymbol{\infty}}(\tilde{\boldsymbol{a}})$. Part (ii) can be proven analogously. □

Proof of Proposition 2. Since the function $Q^i(\boldsymbol{e}, A^i, \Phi) - Q^i(\boldsymbol{e}, \tilde{a}^i, \phi^i)$ is increasing in e^i and e^j, Lemma 6 implies that

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

Since $Q^i(\boldsymbol{e}_{\boldsymbol{\infty}}(\tilde{\boldsymbol{a}}'), A^i, \Phi) - Q^i(\boldsymbol{e}_{\boldsymbol{\infty}}(\tilde{\boldsymbol{a}}'), \tilde{a}'^i, \phi_{\infty}^i(\tilde{\boldsymbol{a}}')) = 0$, it is implied that $\phi_{\infty}^i(\tilde{\boldsymbol{a}}') > \phi_{\infty}^i(\tilde{\boldsymbol{a}})$, $\forall i$. □

Proof of Proposition 3. Assume for now that $Q_{e^i\phi}^i > 0$. (i) As in the proof of Proposition 1, we could obtain

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

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

Since the informational externality of agent j is negative, both $Q_{e^ie^j}^i$ and $Q_{e^j}^i(\boldsymbol{e}, A^i, \Phi) - Q_{e^j}^i(\boldsymbol{e}, \tilde{a}^i, \phi^i)$ are non-positive and at least one of them has to be strictly negative. It follows that $\frac{\partial e^i}{\partial e^j}, \frac{\partial \phi^i}{\partial e^j} < 0$. Therefore, $e_S^j > e_{\infty}^j(\tilde{\boldsymbol{a}})$ implies $\phi_S^i < \phi_{\infty}^i(\tilde{\boldsymbol{a}}) < \Phi$.

(ii) The function $Q^i(\boldsymbol{e}, A^i, \Phi) - Q^i(\boldsymbol{e}, \tilde{a}^i, \phi^i)$ is increasing in e^i but decreasing in e^j. Moreover, since $e_{\infty}^i(\boldsymbol{a})$ is increasing in a^j and $e_{\infty}^j(\boldsymbol{a})$ is decreasing in a^j, we infer that $e_{\infty}^i(\tilde{\boldsymbol{a}}) > e_{\infty}^i(\tilde{\boldsymbol{a}}')$ and $e_{\infty}^j(\tilde{\boldsymbol{a}}) < e_{\infty}^j(\tilde{\boldsymbol{a}}')$. It follows that

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

So $\phi_{\infty}^i(\tilde{\boldsymbol{a}}') < \phi_{\infty}^i(\tilde{\boldsymbol{a}})$.

The proof is analogous when $Q^i_{e^i\phi} < 0$. $\square$

Lemma 7. *There exists $\delta > 0$ such that, if agents have misconceptions $\tilde{\boldsymbol{a}} \in B_\delta(\boldsymbol{A})$ and $(e^j_\infty(\tilde{\boldsymbol{a}}) - e^j_\infty(\boldsymbol{A}))$ has the same (opposite) sign as $Q^i_{e^j}(\boldsymbol{e_\infty}(\boldsymbol{A}), A^i, \Phi)$, then agent i's average payoff in the steady state increases (decreases).*

Proof of Lemma 7. First assume $Q^i_{e^j}(\boldsymbol{e_\infty}(\boldsymbol{A}), A^i, \Phi) > 0$. It follows from Taylor's expansion that

$$Q^i\left(\boldsymbol{e_\infty}(\boldsymbol{A}) + (\epsilon, \eta), A^i, \Phi\right) = Q^i\left(\boldsymbol{e_\infty}(\boldsymbol{A}), A^i, \Phi\right) + Q^i_{e^j}\left(\boldsymbol{e_\infty}(\boldsymbol{A}), A^i, \Phi\right)\eta + o(\epsilon) + o(\eta)$$

for $\epsilon \in \mathbb{R}$ and $\eta \in \mathbb{R}_+$. Therefore, there exist positive numbers $\overline{\epsilon}$ and $\overline{\eta}$, such that for any $\hat{e}^i \in B_{\overline{\epsilon}}(e^i_\infty(\boldsymbol{A}))$ and $\hat{e}^j \in (e^j_\infty(\boldsymbol{A}), e^j_\infty(\boldsymbol{A}) + \overline{\eta})$,

$$Q^i\left(\left(\hat{e}^i, \hat{e}^j\right), A^i, \Phi\right) > Q^i\left(\boldsymbol{e_\infty}(\boldsymbol{A}), A^i, \Phi\right).$$

It follows from continuity of $\boldsymbol{e_\infty}$ that there exists $\delta > 0$ such that $\tilde{\boldsymbol{a}} \in B_\delta(\boldsymbol{A})$ implies $e^i_\infty(\tilde{\boldsymbol{a}}) \in B_{\overline{\epsilon}}(e^i_\infty(\boldsymbol{A}))$ and $e^j_\infty(\tilde{\boldsymbol{a}}) \in B_{\overline{\eta}}\left(\boldsymbol{e^i_\infty}(\boldsymbol{A})\right)$. We need $e^j_\infty(\tilde{\boldsymbol{a}}) - e^{*j}(\boldsymbol{A}) > 0$ so that agent i's payoff is higher under $\tilde{\boldsymbol{a}}$. For the case where $Q^i_{e^j}(\boldsymbol{e_\infty}(\boldsymbol{A}), A^i, \Phi) < 0$, we can analogously show that agent i's payoff is higher under $\tilde{\boldsymbol{a}}$ when $\tilde{\boldsymbol{a}} \in B_\delta(\boldsymbol{A})$ and $e^j_\infty(\tilde{\boldsymbol{a}}) - e^{*j}(\boldsymbol{A}) < 0$. $\square$

Proof of Proposition 4. Let δ_1 be such that Lemma 7 applies. Notice that

$$\frac{\partial e^i_\infty(\boldsymbol{A})}{\partial a^i} = \frac{-Q^{-i}_{e^{-i}e^{-i}}\left(Q^i_{e^i a} - Q^i_{e^i\phi}\frac{Q^i_a}{Q^i_\phi}\right)}{Q^{-i}_{e^{-i}e^{-i}}Q^i_{e^ie^i} + Q^{-i}_{e^ie^{-i}}Q^i_{e^ie^{-i}}}$$
$$\frac{\partial e^j_\infty(\boldsymbol{A})}{\partial a^i} = \frac{Q^j_{e^je^i}\left(Q^i_{e^i a} - Q^i_{e^i\phi}\frac{Q^i_a}{Q^i_\phi}\right)}{Q^{-i}_{e^{-i}e^{-i}}Q^i_{e^ie^i} + Q^{-i}_{e^ie^{-i}}Q^i_{e^ie^{-i}}}$$

By Assumption 2, $\frac{\partial e^j_\infty(\boldsymbol{A})}{\partial a^i} < \frac{\partial e^i_\infty(\boldsymbol{A})}{\partial a^i}$. Let δ_2 be such that $\frac{\partial e^j_\infty(\tilde{\boldsymbol{a}})}{\partial a^i} < \frac{\partial e^i_\infty(\tilde{\boldsymbol{a}})}{\partial a^i}$ for all $\tilde{\boldsymbol{a}} \in B_{\delta_2}(\boldsymbol{A})$ and let $\delta = \min\{\delta_1, \delta_2\}$. Suppose $Q^j_{e^j\phi} < 0$, then it follows that when $\tilde{\boldsymbol{a}} \in B^+_\delta(\boldsymbol{A})$, agent j increases his effort, i.e. $e^j_\infty(\boldsymbol{A}) < e^j_\infty(\tilde{\boldsymbol{a}})$, and thus by Lemma 7, agent i's payoff increases if and only if $Q^i_{e^j}(\boldsymbol{e_\infty}(\boldsymbol{A}), A^i, \Phi) > 0$. The proof for the other case is analogous. $\square$

Proof of Proposition 5. Without loss of generality, assume we want to make both agents better off (the proof for other cases follow analogously). By Lemma 7, we know that $\tilde{\boldsymbol{a}}$ has to be such that $\tilde{\boldsymbol{a}} \in B_\delta(\boldsymbol{A})$ and $e^i_\infty(\tilde{\boldsymbol{a}}) > e^i_\infty(\boldsymbol{A}), \forall i$ for some δ. This is possible if there exist $\Delta\tilde{a}^i$ such that $|\Delta\tilde{a}^i| < \delta, \forall i$, and

$$\begin{bmatrix} \frac{\partial e^1_\infty(\boldsymbol{A})}{\partial a^1} & \frac{\partial e^1_\infty(\boldsymbol{A})}{\partial a^2} \\ \frac{\partial e^2_\infty(\boldsymbol{A})}{\partial a^1} & \frac{\partial e^2_\infty(\boldsymbol{A})}{\partial a^2} \end{bmatrix} \begin{bmatrix} \Delta\tilde{a}^1 \\ \Delta\tilde{a}^2 \end{bmatrix} > \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

That is, we only need the coefficient matrix to be invertible. Its determinant is given by $\frac{\partial e^1_\infty(\boldsymbol{A})}{\partial a^1} \cdot \frac{\partial e^2_\infty(\boldsymbol{A})}{\partial a^2} - \frac{\partial e^1_\infty(\boldsymbol{A})}{\partial a^2} \cdot \frac{\partial e^2_\infty(\boldsymbol{A})}{\partial a^1}$. By Lemma 6,

$$\frac{\partial e^i_\infty(\boldsymbol{A})}{\partial a^k} = \frac{I(i,k) Q^{-k}_{e^{-k}e^{-i}} \left(Q^k_{e^k a} - Q^k_{e^k \phi} \frac{Q^k_a}{Q^k_\phi} \right)}{Q^{-i}_{e^{-i}e^{-i}} Q^i_{e^i e^i} + Q^{-i}_{e^i e^{-i}} Q^i_{e^i e^{-i}}}, \forall i, k,$$

where $I(i,k) = -1$ if $i = k$ and $I(i,k) = 1$ if $i \neq k$. All derivatives are evaluated at $e^i_\infty(\boldsymbol{A}), \boldsymbol{A}$, and $\boldsymbol{\Phi}$. Non-zero determinant is equivalent to

$$\frac{\prod_{k=1,2} \left(Q^k_{e^k a} - Q^k_{e^k \phi} \frac{Q^k_a}{Q^k_\phi} \right)}{\left(Q^{-i}_{e^{-i}e^{-i}} Q^i_{e^i e^i} + Q^{-i}_{e^i e^{-i}} Q^i_{e^i e^{-i}} \right)^2} \left[Q^1_{e^1 e^1} Q^2_{e^2 e^2} - Q^1_{e^1 e^2} Q^2_{e^1 e^2} \right] \neq 0.$$

By Assumption 1, $Q^k_{e^k a} - Q^k_{e^k \phi} \frac{Q^k_a}{Q^k_\phi} \neq 0$, so the above condition holds. □

C Proofs for Section 5

Following Heidhues, Kőszegi, and Strack (2018), let us define $m^i_t(\phi^i)$ to keep track of the actual gap in average payoffs when the fundamental is ϕ^i,

$$m^i_t(\phi^i) = Q^i(\boldsymbol{e_t}, A^i, \Phi) - Q^i(\boldsymbol{e_t}, \tilde{a}^i, \phi^i), \forall i.$$

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

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

Write the vectors of bounds as $\underline{\boldsymbol{\phi}}_{\boldsymbol{\infty}} = \left(\underline{\phi}^1_\infty, \underline{\phi}^2_\infty \right), \overline{\boldsymbol{\phi}}_{\boldsymbol{\infty}} = \left(\overline{\phi}^1_\infty, \overline{\phi}^2_\infty \right)$. We show that $\underline{\boldsymbol{\phi}}_{\boldsymbol{\infty}}$ and $\overline{\boldsymbol{\phi}}_{\boldsymbol{\infty}}$ are bounded in Lemma 9.

Next, we proceed by stating an important lemma established in Heidhues, Kőszegi, and Strack (2018) that could be easily reformulated in a two-agent environment. Lemma 8 shows that if some fundamental level ϕ^i is in the support of the long-run beliefs, the average payoff implied by ϕ^i should not be consistently higher or lower than the one implied by Φ.

Lemma 8 (Heidhues, Kőszegi, and Strack (2018), Lemma 13)**.** *(a) For all i, if $\liminf_{t\to\infty} m_t^i\left(\phi^i\right) \geq \underline{m} > 0$ for all $\phi^i \in (l,h) \subset \left(\underline{\phi}, \overline{\phi}\right)$, then*

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

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

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

Lemma 9. *For all i, we have that $\Phi - \frac{\overline{\kappa}_a}{\underline{\kappa}_\phi}\left(\tilde{a}^i - A^i\right) \leq \underline{\phi}_\infty^i$ and $\overline{\phi}_\infty^i \leq \Phi$.*

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

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

It then follows from Lemma 8 that a small neighborhood of $\underline{\phi}_\infty^i$ will be assigned zero probability by the agent almost surely in the long run, which is a contradiction to the definition of $\underline{\phi}_\infty^i$. Hence, $\Phi - \frac{\overline{\kappa}_a}{\underline{\kappa}_\phi}\left(\tilde{a}^i - A^i\right) \leq \underline{\phi}_\infty^i$. Analogously we can prove $\overline{\phi}_\infty^i \leq \Phi$. $\square$

We can obtain similar bounds for efforts. Define

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

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

Lemma 10. $E_\infty \subseteq E_\infty^D$.

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

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

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

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

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

Lemma 11. $\frac{\partial e^{*i}(\tilde{\boldsymbol{a}},\boldsymbol{\phi})}{\partial \phi^i}$ *has the same sign as* $Q_{e^i\phi}^i$*, while* $\frac{\partial e^{*j}(\tilde{\boldsymbol{a}},\boldsymbol{\phi})}{\partial \phi^i}$*has the same sign as* $Q_{e^i\phi}^i Q_{e^ie^j}^j$*.*

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

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

Take partial derivatives, we obtain that

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

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

Proof of Theorem 1. It is sufficient to show that actions and beliefs about the common fundamental converge, with which agents' beliefs about each other's ability must also converge to the solution characterized by the no gap condition. First, consider the asymptotic behavior of $m_t\left(\underline{\phi}_\infty^i\right)$. By Lemma 8 and the continuity of Q^i, it must be true that $\liminf_{t\to\infty} m_t^i\left(\underline{\phi}_\infty^i\right) \leq 0$ almost surely, otherwise it contradicts the assumption of $\underline{\phi}_\infty^i$ being the infimum of the agent's long-run beliefs almost surely. Similarly, it must be that $\limsup_{t\to\infty} m_t^i\left(\overline{\phi}_\infty^i\right) \leq 0$.

Case (i): Both agents are overconfident and create positive externalities. Without loss of generality, it is sufficient to show convergence under the assumption that $Q^i_{e^i\phi} > 0$ and $Q^i_{e^i a} \leq 0$.

Recall that $g^i(\boldsymbol{e}, \phi^i) = Q^i(\boldsymbol{e}, A^i, \Phi) - Q^i(\boldsymbol{e}, \tilde{a}^i, \phi^i)$. Differentiate g with respect to e^j, we obtain

$$\frac{\partial g^i(\boldsymbol{e}, \phi^i)}{\partial e^j} = Q^i_{e^j}\left(\boldsymbol{e}, A^i, \Phi\right) - Q^i_{e^j}\left(\boldsymbol{e}, \tilde{a}^i, \phi^i\right). \tag{9}$$

Since $Q^i_{e^k\phi} > 0, Q^i_{e^k a} \leq 0, \forall i, k$, we know that $\frac{\partial g^i\left(\boldsymbol{e}, \phi^i\right)}{\partial e^k} > 0$, $\forall i, k$. By Lemma 11, since $E_\infty \subseteq E^D_\infty$, it must be that almost surely, when t is large enough, $\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \underline{\boldsymbol{\phi}}_{\boldsymbol{\infty}}\right) \leq \boldsymbol{e_t} \leq \boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \overline{\boldsymbol{\phi}}_{\boldsymbol{\infty}}\right)$. Hence,

$$\begin{aligned}
&0 \geq \liminf_{t\to\infty} g^i\left(\boldsymbol{e}_t, \underline{\phi}^i_\infty\right) = \liminf_{t\to\infty} m^i_t\left(\underline{\phi}^i_\infty\right) \geq G^i\left(\underline{\boldsymbol{\phi}}_{\boldsymbol{\infty}}\right), \forall i,\\
&0 \leq \limsup g^i\left(\boldsymbol{e}_t, \overline{\phi}^i_\infty\right) = \limsup_{t\to\infty} m^i_t\left(\overline{\phi}^i_\infty\right) \leq G^i\left(\overline{\boldsymbol{\phi}}_{\boldsymbol{\infty}}\right), \forall i.
\end{aligned} \tag{10}$$

Notice that $G^i(\boldsymbol{\phi}) = g^i\left(\boldsymbol{e}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\phi}\right), \phi^i\right)$ is increasing in ϕ^j, since $\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \boldsymbol{\phi}\right)$ is increasing in ϕ^j and $\frac{\partial g^i\left(\boldsymbol{e}, \phi^i\right)}{\partial e^k} \geq 0$, $\forall i, k$. Therefore, $G^i\left(\underline{\phi}^i_\infty, \phi^j\right) \leq 0$ for any $\phi^j \leq \underline{\phi}^j_\infty$ and $G^i\left(\overline{\phi}^i_\infty, \phi^j\right) \geq 0$ for any $\phi^j \geq \overline{\phi}^j_\infty$. Moreover, notice that $\forall i, \phi^j, G^i\left(\Phi, \phi^j\right) < 0$ and $G^i\left(\Phi - \frac{\overline{\kappa}_a}{\underline{\kappa}_\phi}\left(\tilde{a}^i - A^i\right), \phi^j\right) > 0$. For notational convenience, let $\underline{\psi}^i = \Phi - \frac{\overline{\kappa}_a}{\underline{\kappa}_\phi}\left(\tilde{a}^i - A^i\right)$ and $\overline{\psi}^i = \Phi$ for all i. The above results can be summarized by:

$$\begin{aligned}
&\boldsymbol{G}\left(\underline{\boldsymbol{\phi}}_{\boldsymbol{\infty}}\right) \leq 0, \boldsymbol{G}\left(\underline{\boldsymbol{\psi}}\right) > 0, \boldsymbol{G}\left(\overline{\boldsymbol{\phi}}_{\boldsymbol{\infty}}\right) \leq 0, \boldsymbol{G}\left(\overline{\boldsymbol{\psi}}\right) < 0,\\
&G^i\left(\underline{\psi}^i, \underline{\phi}^j_\infty\right) > 0, G^j\left(\underline{\psi}^i, \underline{\phi}^j_\infty\right) \leq 0,\\
&G^i\left(\underline{\phi}^i_\infty, \underline{\psi}^j\right) \leq 0, G^j\left(\underline{\phi}^i_\infty, \underline{\psi}^j\right) > 0,\\
&G^i\left(\overline{\psi}^i, \overline{\phi}^j_\infty\right) < 0, G^j\left(\overline{\psi}^i, \overline{\phi}^j_\infty\right) \geq 0,\\
&G^i\left(\overline{\phi}^i_\infty, \overline{\psi}^j\right) \geq 0, G^j\left(\overline{\phi}^i_\infty, \overline{\psi}^j\right) < 0.
\end{aligned} \tag{11}$$

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

Case (ii): Both agents are overconfident create negative informational externalities. Without loss of generality, assume that $Q^i_{e^i\phi} > 0$ and $Q^i_{e^i a} \leq 0$. Analogous to Case (i), we will derive a contradiction if $\underline{\boldsymbol{\phi}}_{\boldsymbol{\infty}} \neq \overline{\boldsymbol{\phi}}_{\boldsymbol{\infty}}$. Since informational externalities are negative, the signs of Eq. (9) are different: $\frac{\partial g^i\left(\boldsymbol{e},\phi^i\right)}{\partial e^i} > 0$ and $\frac{\partial g^i\left(\boldsymbol{e},\phi^i\right)}{\partial e^j} < 0$. Again, by Lemma 11, when t is large enough, $e^{*i}\left(\tilde{\boldsymbol{a}}, \left(\underline{\phi}^i_\infty, \overline{\phi}^j_\infty\right)\right) \le e^i_t \le e^{*i}\left(\tilde{\boldsymbol{a}}, \left(\overline{\phi}^i_\infty, \underline{\phi}^j_\infty\right)\right), \forall i$. Hence,

$$
\begin{aligned}
0 &\ge \liminf_{t\to\infty} g^i\left(\boldsymbol{e}_t, \underline{\phi}^i_\infty\right) = \liminf_{t\to\infty} m^i_t\left(\underline{\phi}^i_\infty\right) \ge G^i\left(\underline{\phi}^i_\infty, \overline{\phi}^j_\infty\right), \forall i,\\
0 &\le \limsup g^i\left(\boldsymbol{e}_t, \overline{\phi}^i_\infty\right) = \limsup_{t\to\infty} m^i_t\left(\overline{\phi}^i_\infty\right) \le G^i\left(\overline{\phi}^i_\infty, \underline{\phi}^j_\infty\right), \forall i.
\end{aligned}
$$

In addition, $G^i(\boldsymbol{\phi}) = g^i\left(\boldsymbol{e}\left(\tilde{\boldsymbol{a}}, \boldsymbol{\phi}\right), \phi^i\right)$ is decreasing in ϕ^j. Therefore, we have some different inequalities than those in Eq. (11):

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

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

Therefore, again, there exist two different roots to $\boldsymbol{G}(\boldsymbol{\phi}) = 0$ if $\underline{\boldsymbol{\phi}}_{\boldsymbol{\infty}} \neq \overline{\boldsymbol{\phi}}_{\boldsymbol{\infty}}$, contradicting Assumption 3. □

D Proofs for Section 6

Proof of Theorem 2. Similar to Theorem 1, we only need to show actions and beliefs about the common fundamental converge. We prove the result when there are positive externalities, $Q^i_{e^i\phi} > 0$ and $Q^i_{e^ia} \le 0$; the proof for other cases is analogous.

Similarly define $\overline{\psi}^i = \Phi - \frac{\kappa_a}{\kappa_\phi}\left(\tilde{a}^i - A^i\right)$ and $\underline{\psi}^i = \Phi$ for all i. Since the agents are underconfident, $\frac{\partial g^i\left(\boldsymbol{e},\phi^i\right)}{\partial e^k} < 0, \forall i, k$. In addition, we again have $\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \underline{\boldsymbol{\phi}}_{\boldsymbol{\infty}}\right) \le \boldsymbol{e_t} \le \boldsymbol{e}^*\left(\tilde{\boldsymbol{a}}, \overline{\boldsymbol{\phi}}_{\boldsymbol{\infty}}\right)$

when t is large enough. Therefore,

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

Therefore, $g^i\left(\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}},\overline{\boldsymbol{\phi}}_{\boldsymbol{\infty}}\right),\underline{\phi}_\infty^i\right) \leq 0 \leq g^i\left(\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}},\underline{\boldsymbol{\phi}}_{\boldsymbol{\infty}}\right),\overline{\phi}_\infty^i\right), \forall i$. Rewrite $g^i\left(\boldsymbol{e}^*\left(\tilde{\boldsymbol{a}},\boldsymbol{\psi}\right),\phi^i\right)$ as $h^i\left(\boldsymbol{\psi},\boldsymbol{\phi}\right)$ for all i. Differentiate, $\forall k \in \{i,j\} = \{1,2\}$,

$$\frac{\partial h^i\left(\boldsymbol{\psi},\boldsymbol{\phi}\right)}{\partial \psi^k} = \left(Q_{e^i}^{i,A} - Q_{e^i}^i\right)\frac{\partial e^{*i}\left(\tilde{\boldsymbol{a}},\boldsymbol{\psi}\right)}{\partial \psi^k} + \left(Q_{e^j}^{i,A} - Q_{e^j}^i\right)\frac{\partial e^{*j}\left(\tilde{\boldsymbol{a}},\boldsymbol{\psi}\right)}{\partial \psi^k}$$
$$\frac{\partial h^i\left(\boldsymbol{\psi},\boldsymbol{\phi}\right)}{\partial \phi^i} = -Q_\phi^i, \frac{\partial h^i\left(\boldsymbol{\psi},\boldsymbol{\phi}\right)}{\partial \phi^j} = 0,$$

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

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

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 $\boldsymbol{\Phi}$; (ii) for all i and $j \neq i$, $\frac{\partial h^i(\boldsymbol{\psi},\boldsymbol{\phi})}{\partial \phi^i} < -\frac{1}{2}\underline{\kappa}_\phi < 0$, and $\left|\frac{\partial h^i(\boldsymbol{\psi},\boldsymbol{\phi})}{\partial \psi^i}\right|, \left|\frac{\partial h^i(\boldsymbol{\psi},\boldsymbol{\phi})}{\partial \psi^j}\right|, \left|\frac{\partial h^i(\boldsymbol{\psi},\boldsymbol{\phi})}{\partial \phi^j}\right| < \frac{1}{4}\underline{\kappa}_\phi$.

Therefore,

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

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

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*.

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.

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.