OSIPP Discussion Paper: DP-2021-E-001

Multi-Player Bayesian Learning with Misspecified Models

January 15, 2021

Takeshi Murooka*

Osaka School of International Public Policy (OSIPP), Osaka University

Yuichi Yamamoto†

Institute of Economic Research, Hitotsubashi University

Abstract: We consider strategic players who may have a misspecified view about an environment, and investigate their long-run behavior. Each period, players simultaneously take actions, observe a public outcome, and then update own belief about an uncertain economic state by using Bayes' rule. We provide a condition under which players' beliefs and actions converge to a steady state, and then characterize how one's misspecification influences the long-run (steady-state) actions. When a player has a biased view about the physical environment (e.g., overconfidence on own capability or prejudice on an opponent's capability), the presence of strategic interaction influences the size of the impact of misspecification, but not the direction. In particular, when the game is symmetric, the presence of strategic interaction amplifies the deviation of the long-run actions from those in the correctly specified model. When a player misspecifies the opponent's view about the environment (e.g., the player is not aware of the opponent's bias), the strategic interaction generates a directional impact for the long-run actions. We extensively discuss implications to a variety of applications, such as Cournot duopoly, team production, tournaments, and discrimination.

JEL Codes: C73, D83, D90, D91

Keywords: model misspecification, learning, convergence, overconfidence, bias transmission

* Email: murooka@osipp.osaka-u.ac.jp

† Email: yyamamoto@ier.hit-u.ac.jp

Multi-Player Bayesian Learning with Misspecified Models*

Takeshi Murooka — Osaka University
Yuichi Yamamoto — Hitotsubashi University

January 15, 2021

Abstract

We consider strategic players who may have a misspecified view about an environment, and investigate their long-run behavior. Each period, players simultaneously take actions, observe a public outcome, and then update own belief about an uncertain economic state by using Bayes' rule. We provide a condition under which players' beliefs and actions converge to a steady state, and then characterize how one's misspecification influences the long-run (steady-state) actions. When a player has a biased view about the physical environment (e.g., overconfidence on own capability or prejudice on an opponent's capability), the presence of strategic interaction influences the size of the impact of misspecification, but not the direction. In particular, when the game is symmetric, the presence of strategic interaction amplifies the deviation of the long-run actions from those in the correctly specified model. When a player misspecifies the opponent's view about the environment (e.g., the player is not aware of the opponent's bias), the strategic interaction generates a directional impact for the long-run actions. We extensively discuss implications to a variety of applications, such as Cournot duopoly, team production, tournaments, and discrimination.

*We are grateful to seminar and conference audiences for helpful comments. Murooka acknowledges financial support from JSPS KAKENHI (JP16K21740, JP18H03640, JP19K01568, JP20K13451). Yamamoto acknowledges financial support from JSPS KAKENHI (JP20H00070, JP20H01475).

1 Introduction

Economic agents often take actions based on a subjective and potentially misspecified view of the world.[1] Past work has shown that in a static game-theoretic environment, one's misspecification can improve her equilibrium payoff. The reason is that one's misspecification can influence *the opponent's action* through strategic interaction. For example, in a one-shot Cournot game, if a firm is overconfident about the demand and willing to produce more, the rival firm best-responds to it and produces less, which improves the overconfident firm's profit (Kyle and Wang, 1997). We build on this literature by considering a Bayesian-learning feature; we develop a dynamic model in which misspecified players learn an unknown economic state from public signals. Our primary interest is to understand how misspecified players process information, and how it influences their long-run behavior.

We believe that this question is of importance for the following reason. In many economic applications, it is assumed that players know all relevant aspects of the environment. This assumption is often justified by the argument that even if some economic variable is initially unknown, as long as it is fixed over time, players will eventually learn it from observed outcomes. However, when players have a misspecified view about some aspect of the world, they may process information incorrectly and may fail to learn the true economic variable. Thus, to understand the long-run behavior of misspecified players, we need to consider a dynamic model (rather than a one-shot model) and carefully think about how they update the beliefs over time.

More generally, when players have a misspecified view about the world, it is natural to expect that they will eventually change the belief about some economic

[1]As experimental and empirical evidence, people exhibit overconfidence in strategic entries (Camerer and Lovallo, 1999), corporate investments (Malmendier and Tate, 2005), and merger decisions (Malmendier and Tate, 2008). There is also recent evidence that overconfidence is persistent; Hoffman and Burks (2020) find that workers are persistently overconfident about their own productivity, and Huffman, Raymond, and Shvets (2019) find that managers are persistently overconfident about future performance.

variable, as they observe outcomes which are systematically different from the anticipation. For example, if a firm is persistently overconfident about some aspect of the demand function (e.g., the intercept of the inverse demand curve), on average, actual prices are lower than the firm's anticipation. So it is likely that after a long time, this firm becomes (unrealistically) pessimistic about other aspect of the demand (e.g., the slope of the inverse demand curve). Similarly, in tournaments, if an agent is persistently overconfident about her own capability, after a series of unexpected losses, she may start to think that the tournament is unfair. Our framework is useful to understand players' long-run behavior in these cases, and we show that this learning feature has a substantial impact on the equilibrium outcomes in various applications.

Formally, we consider a two-player Bayesian learning problem with model misspecification. In each period t, each player i simultaneously chooses an action x_i and then learns an unknown economic state θ from a public signal y. The distribution of the signal y is influenced by the state θ, the action profile $x = (x_1, x_2)$, and a parameter a which describes the environment/technology. For example, in a Cournot duopoly, the distribution of the market price y depends on the quantity $x_1 + x_2$, as well as the intercept a and the slope θ of the inverse demand curve.

We assume that a player may have a *misspecification*, in that she may have a persistently biased view about the parameter a. In Section 3, we consider the case in which player 1 is unbiased and knows the true environment, but player 2 incorrectly believes that the true parameter is $A \neq a$. We call it *first-order misspecification*, as player 2 has an incorrect first-order belief about the parameter a. This is the simplest form of misspecification, and encompasses a variety of commonly observed biases, including overconfidence about one's own capability, prejudice about the opponent's capability, misestimation about market demand, and so on.

We provide a condition in which players' beliefs (about θ) and actions converge to a steady state after a long time. Then, we characterize how player 2's misspecification influences this steady state. It turns out that the impact of misspecification on the biased player's action can be written as a product of two terms, a *base misspecification effect* and a *multiplier effect*. The base misspecification effect measures how one's misspecification would influence her own action *if there was no strategic interaction between the players*, in that the opponent takes a constant action every period. This essentially measures the impact of misspecification in a *single-agent* Bayesian learning problem, which is studied in specific contexts in the literature (e.g., Fudenberg, Romanyuk, and Strack (2017), Heidhues, Kőszegi, and Strack (2018), and He (2019)).

The multiplier effect is new, and it describes how strategic interaction amplifies/reduces the base misspecification effect. When a biased player changes her action due to the base misspecification effect, the opponent best-responds to it, which in turn influences the biased player's steady-state belief and the optimal action, and this process continues repeatedly. The multiplier effect quantifies the overall size of this feedback process. Under a regularity condition, the multiplier effect is positive, which means that strategic interaction can influence the *size* of the impact of misspecification, but not the direction. In particular, if the game is symmetric as in common-interest games, the multiplier effect is larger than one, so both strategic complements and substitutes amplify the base misspecification effect.

To fix ideas, consider a Cournot duopoly in which some aspect of the demand (e.g., the slope of the inverse demand curve) is unknown. Suppose that one of the firms is persistently overconfident about other aspect of the demand (e.g., the intercept of the inverse demand curve). This overconfidence *boosts* the firm's incentive to produce, and as noted earlier, it improves the overconfident firm's equilibrium payoff in the static game. This is the *direct effect* of overconfidence. However, in our dynamic model, the firm's overconfidence leads to incorrect learning, which *weakens* the incentive to produce; indeed, the overconfident firm finds that realized prices are systematically lower than the anticipation on average, so after a long time, it becomes unrealistically pessimistic about the unknown economic state. This is the *indirect learning effect* of overconfidence, and the base misspecification effect is the sum of these two countervailing forces. If the direct effect outweighs the indirect effect, the base misspecification effect is positive. In this case, the result of Kyle and Wang (1997) persists even in the long run, that is, the overconfident firm produces more and earns a better profit than in the correctly-specified case. However, if the indirect effect is larger, this result is overturned: while overconfidence improves the static equilibrium payoff, in the long run, it is detrimental and the overconfident firm earns a lower profit.

When there are multiple players, they may not only have a bias about a physical environment, but have a bias about the opponent's view about the environment. For example, a junior scholar may incorrectly believe that he is undervalued by a senior colleague (in reality, the senior colleague correctly understands the junior scholar's capability). Or, a worker (she) may be unaware that a co-worker (he) is overconfident about his own capability. In Section 4, we consider how such "second-order misspecification" influences players' behavior. We consider the case in which one of the players has second-order misspecification (i.e., she correctly specifies a but has an incorrect view about the opponent's view about a), the case in which one of the players has "double misspecification" in that she has both first-order and second-order misspecification, and the case in which both players have double misspecification. The analysis is significantly more complicated than in the case of first-order misspecification, because a player with second-order or double misspecification has an incorrect view about how the opponent updates the belief, and we need to keep track of how the belief of such a "hypothetical player" evolves over time. Nonetheless, under an additional assumption, we find that players' actions and beliefs still converge to a steady state. We also show that the impact of one's misspecification on the steady-state outcome is still represented by a product of the base misspecification effect and the multiplier effect.

To illustrate how second-order misspecification influences players' behavior, consider the Cournot model, and as in the case of first-order misspecification, assume that one of the firms (say, firm 2) is overconfident about some aspect of the demand. In addition, assume now that the rival firm (say, firm 1) has second-order misspecification, i.e., firm 1 is not aware of firm 2's overconfidence and incorrectly believes that firm 2 has an unbiased view about the demand. In this setup, firm 2's overconfidence cannot directly affect the opponent's production. Indeed, even though firm 2 is overconfident and willing to produce more, firm 1 is not aware of it and hence does not change the production level. Accordingly, in the one-shot game, firm 2's overconfidence *never* improves the equilibrium payoff. Nonetheless, we find that in the long run, overconfidence may be beneficial. To see why, note that firm 1 is not aware of firm 2's overconfidence, and hence it makes a wrong prediction about firm 2's production. In particular, for some parameter values, firm 1 underestimates firm 2's production, in which case it finds that a market price is systematically lower than the anticipation. Then as time goes, firm 1 becomes pessimistic about an unknown economic state and reduces the production, which benefits the overconfident firm 2.

A key here is that even though firm 1 has an unbiased view about the physical environment, incorrect learning occurs due to second-order misspecification. A similar idea can be used to explain *bias transmission*. In Section 4.4.3, we consider a teacher who has a bias against female students. We show that the teacher's bias can endogenously induce female students' negative self-stereotypes, if the students are not aware of the teacher's bias.

In Section 5, we present a general model which subsumes all of the aforementioned misspecifications as special cases, and characterize the asymptotic behavior of players' actions an beliefs. Section 6 summarizes the related literature.

2 Setup

There are two players $i = 1,2$ and infinitely many periods $t = 1,2,\cdots$. At the beginning of the game, an unobservable economic state θ^* is drawn from a closed interval $\Theta = [\underline{\theta},\overline{\theta}]$, according to a common prior distribution $\mu \in \triangle\Theta$. In each period t, each player i has a belief $\mu_i^t \in \triangle\Theta$ about θ, and chooses an action x_i from a closed interval $X_i = [0,\overline{x}_i]$. Player i's action x_i is not observable by the other player $j \neq i$. Given an action profile $x = (x_1,x_2)$, the players observe a noisy public signal $y = Q(x_1,x_2,a,\theta^*)+\varepsilon$, where $a \in \boldsymbol{R}$ is a fixed parameter which describes a physical environment (e.g., a parameter which determines a market demand, a player's capability, etc) and ε is a random noise whose distribution is $N(0,1)$. Each player i receives a payoff $u_i(x_i,y)$. Both Q and u_i are twice continuously differentiable.

Crucially, players may have bias about the environment, and as a result, and the parameter a is not common knowledge. We allow a variety of bias, including bias about the parameter a, bias about the opponent's bias about the parameter a, and so on. In what follows, we will consider each bias separately.

3 First-Order Misspecification

We first consider the simplest form of misspecification, called *first-order misspecification*.

3.1 First-Order Misspecification: Characterization

Suppose that one of the players (player 2) has a biased view about the parameter a, whereas the other player is unbiased. Specifically, consider the following information structure:

- Player 1 believes that for each parameter θ, the signal y is given by $y = Q(x_1, x_2, a, \theta) + \varepsilon$.
- Player 2 (incorrectly) believes that for each parameter θ, the signal y is given by $y = Q(x_1, x_2, A, \theta) + \varepsilon$, where $A \neq a$.
- The above beliefs are common knowledge (e.g., player 2 believes that player 1 believes that $y = Q(x_1, x_2, a, \theta) + \varepsilon$, and the like).

This is the case in which players "agree to disagree" with the level of the parameter a. Kyle and Wang (1997) consider a short-run effect of this first-order misspecification, i.e., they consider how it influences players' actions in a variant of one-shot Cournot game. In this section, we consider a long-run effect of this misspecification.

Player 1's subjective expected stage-game payoff given an action profile x and a state θ is

$$U_1(x, \theta) = E[u_1(x_1, Q(x, a, \theta) + \varepsilon)]$$

and player 2's subjective expected stage-game payoff is

$$U_2(x,A,\theta) = E[u_2(x_2, Q(x,A,\theta)+\varepsilon)],$$

where the expectation is taken with respect to ε. Note that player 2 evaluates payoffs given her subjective signal distribution $Q(x,A,\theta)+\varepsilon$. To economize notation, we will write $U_2(x,\theta)$ instead of $U_2(x,A,\theta)$ when it does not cause a confusion.

We assume that players are myopic, so that they play a Nash equilibrium each period. In period one, both players have the same belief $\mu_1^1 = \mu_2^1 = \mu$, so a Nash equilibrium (x_1^1, x_2^1) solves the first-order condition $\frac{\partial E[U_i(x,\theta)|\mu]}{\partial x_i} = 0$ for each i, where the expectation is taken with respect to θ. At the end of period one, players observe a public signal y^1, and update the posterior beliefs using Bayes' rule. Assuming that no one has deviated in period one, each player i's posterior belief μ_i^2 in period two is given by

$$\mu_1^2(\theta) = \frac{\mu_1^1(\theta) f(y - Q(x^1,a,\theta))}{\int_\Theta \mu_1^1(\tilde{\theta}) f(y - Q(x^1,a,\tilde{\theta}))d\tilde{\theta}},$$
$$\mu_2^2(\theta) = \frac{\mu_1^1(\theta) f(y - Q(x^1,A,\theta))}{\int_\Theta \mu_1^1(\tilde{\theta}) f(y - Q(x^1,A,\tilde{\theta}))d\tilde{\theta}},$$

where x^1 is the Nash equilibrium played in period one and f is the density function of the noise term ε. Note that player 2's posterior μ_2^2 differs from player 1's posterior μ_1^2, as she incorrectly believes that the mean output is $Q(x^1,A,\theta)$ rather than $Q(x^1,a,\theta)$. These posteriors are common knowledge among players.[2] So in period two, players play a Nash equilibrium given the belief profile $\mu^2 = (\mu_1^2, \mu_2^2)$, which solves $\frac{\partial E[U_i(x,\theta)|\mu_i^2]}{\partial x_i} = 0$ for each i. Likewise, in any subsequent period $t \geq 3$, players play a Nash equilibrium given the belief profile $\mu^t = (\mu_1^t, \mu_2^t)$, where μ^t is computed by Bayes' rule.

Our goal is to characterize the long-run outcome in this Bayesian learning model, i.e., how players' actions x^t and beliefs μ^t look like after a long time. While the evolution of the biased player's belief is governed by a complicated

[2]This is because the players' information structure about the parameter a is common knowledge.

stochastic process, we show in Section 5 that there is a sufficient condition under which players' beliefs and actions converge almost surely to a steady state after a long time. As will be seen, this condition is satisfied in many economic examples, such as Cournot competition and team production. So in what follows, we will assume that players' beliefs and actions converge to a steady state, and investigate the property of the steady state.

Formally, a *steady state* in this Bayesian learning model is a pair $(x_1^*, x_2^*, \mu_1^*, \mu_2^*)$ of an action profile and a belief profile which satisfies the following four conditions:

$$x_1^* \in \arg\max_{x_1} U_1(x_1, x_2^*, \theta^*), \tag{1}$$

$$x_2^* \in \arg\max_{x_2} U_2(x_1^*, x_2, \theta_2), \tag{2}$$

$$\mu_1^* = 1_{\theta^*}, \tag{3}$$

$$\mu_2^* = 1_{\theta_2} \text{ s.t. } Q(x^*, A, \theta_2) = Q(x^*, a, \theta^*). \tag{4}$$

The first two conditions are incentive compatibility conditions, which require that each player maximizes her payoff given some beliefs. The next two conditions require that these beliefs satisfy consistency: (3) asserts that the unbiased player 1 correctly learns the true state θ^* in a steady state. (4) requires that player 2's belief be concentrated on a state θ_2 with which her subjective signal distribution coincides with the true distribution. This condition must be satisfied in a steady state, because otherwise, player 2 is "surprised" by observed signals being different from what she thinks, and changes her belief about θ accordingly.

In general, this steady-state belief θ_2 is different from the true state θ^*, i.e., the biased player 2 cannot learn θ^* correctly. For example, consider the case in which the mean output Q is increasing in a and θ. By the implicit function theorem, we have $\frac{\partial \theta_2}{\partial A} = -\frac{Q_a}{Q_\theta} < 0$. So if player 2 is overconfident about the parameter a, she *underestimates* the state θ. Intuitively, the overconfident player observes signals which are lower than her anticipation, and becomes pessimistic about the state θ. We assume that for each (x, A), there is a unique state θ_2 which solves the consistency condition $Q(x, a, \theta^*) = Q(x, A, \theta)$, and we denote it by $\theta_2(x, A)$. Intuitively, this $\theta_2(x,A)$ is player 2's long-run belief; if players choose the same action profile x every period, then almost surely, player 2's belief will be concentrated on the state $\theta_2(x,A)$ after a long time (Berk (1966)). Player 1's long-run belief is defined as $\theta_1(x,A) = \theta^*$ for all x, because she is unbiased and can learn the true state θ^* regardless of players' play.

The steady-state action profile is different from a Nash equilibrium, because player 2's belief is endogenously determined by the consistency condition (4). Indeed, it is characterized as an intersection of the *asymptotic best response correspondences* (rather than the standard best response correspondences), which is defined as

$$BR_i(x_{-i}) = \left\{ x_i \,\middle|\, x_i \in \arg\max_{x_i'} U_i\left(x_i', x_{-i}, \theta_i(x,A)\right) \right\}. \tag{5}$$

Intuitively, $BR_i(x_{-i})$ describes player i's optimal action *in the long run*, when the opponent chooses the same fixed action x_{-i} each period and player i updates the belief over time. Player 1's asymptotic best response BR_1 coincides with the standard best response correspondence given the state θ^*, because she is unbiased and can learn the true state θ^* regardless of the opponent's play. In contrast, player 2's asymptotic best response BR_2 is different from the standard best response, because her long-run belief $\theta_2(x,A)$ is endogenously determined; if the opponent chooses the same action x_1 every period and if player 2's action converges to some x_2, then the limiting belief must be concentrated on $\theta_2(x,A)$. Player 2's limiting action x_2 must be optimal given this belief, as stated in the definition of BR_i. By a fixed-point theorem, $BR_i(x_{-i})$ is non-empty for all x_{-i}. Also a standard argument shows that BR_i is upper hemi-continuous in x_{-i}.

As is well-known, the slope of the standard best-response function is given by $-\frac{\partial^2 U_i/\partial x_i \partial x_j}{\partial^2 U_i/\partial^2 x_i}$; note that both the denominator and the numerator measure the impact of one's action on player i's marginal utility. It turns out that the slope BR_i' of the asymptotic best response function can be computed by a similar formula,

$$BR_i' = -\frac{M_{ij}}{M_{ii}},$$

where for each i and j (possibly $i = j$),

$$M_{ij} = \frac{\partial}{\partial x_j}\left(\frac{\partial U_i(x,\theta)}{\partial x_i}\bigg|_{\theta=\theta_i(x,A)} \right)$$
$$= \frac{\partial^2 U_i(x,\theta_i)}{\partial x_i \partial x_j}\bigg|_{\theta_i=\theta_i(x,A)} + \frac{\partial^2 U_i(x,\theta)}{\partial x_i \partial \theta}\bigg|_{\theta=\theta_i(x,A)} \frac{\partial \theta_i(x,A)}{\partial x_j}.$$

measures the impact of player j's action on player i's marginal utility *in the long run*, assuming that players choose the same action every period.[3] Intuitively, player j's action x_j influences player i's marginal utility directly and indirectly through the long-run belief $\theta_i(x,A)$. The first term of M_{ij} represents the direct effect, and the second term represents the indirect effect through the belief. The indirect effect is zero for $i = 1$, because player 1's long-run belief is constant and does not depend on the actions (i.e., $\theta_1(x,A) = \theta^*$ for all x). Also, for $i = 2$, the indirect effect disappears in the limit as $A \to a$, because $\theta_2(x,a) = \theta^*$ for all x. So when model misspecification is small (i.e., A is close to a), each M_{ij} is approximated by $\frac{\partial^2 U_i}{\partial x_i \partial x_j}$, so BR_i' is approximately the same as the slope of the standard best-response function.

Our first proposition quantifies the impact of player 2's first-order misspecification on the steady-state action, using this slope of the asymptotic best response curve. We assume that the steady state satisfies the following regularity condition:

Definition 1. A steady state x^* is *regular* if the following conditions are satisfied in x^*: (i) the steady-state action x_i^* is uniquely optimal, i.e., $U_i(x^*, \theta_i(x^*,A)) > U_i(x_i, x_{-i}^*, \theta_i(x^*,A))$ for all i and $x_i \neq x_i^*$, (ii) x^* and $\theta_2(x^*,A)$ are interior points, (iii) $BR_1' BR_2' \neq 1$, and (iv) $M_{ii} \neq 0$ for each i.

The conditions (i) and (ii) are standard. The condition (iii) is satisfied for generic parameters, and it ensures that the multiplier effect appearing in Proposition 1 below is well-defined. The condition (iv) requires two inequalities, $M_{11} = \frac{\partial^2 U_1}{\partial^2 x_1} < 0$ and $M_{22} = \frac{\partial^2 U_2}{\partial^2 x_2} + \frac{\partial \theta_2}{\partial x_2} \frac{\partial^2 U_2}{\partial x_2 \partial \theta} < 0$. The first inequality is simply the second-order condition for player 1's incentive compatibility. The second inequality is different from the second-order condition, as it involves the learning effect $\frac{\partial \theta_2}{\partial x_2} \frac{\partial^2 U_2}{\partial x_2 \partial \theta}$.

[3] Here we assume that $M_{ii} \neq 0$, so that BR_i' is well-defined.

This inequality is satisfied in many economic applications, e.g., it is satisfied if the second-order condition for player 2's incentive compatibility holds and misspecification is small (so that the learning effect is close to zero). Note that Heidhues, Kőszegi, and Strack (2018) impose a similar condition: They consider Bayesian learning by a single agent and assume a unique steady state, which requires $M_{ii} \leq 0$ in the steady state.[4] As will be seen, this condition (iv) is useful when we interpret the base misspecification effect appearing in Proposition 1. This condition also ensures that the slope of the asymptotic best response curve $BR_i' = -\frac{M_{ij}}{M_{ii}}$ is well-defined.[5]

Finally, let

$$M_{2A} := \frac{\partial}{\partial A}\left(\left.\frac{\partial U_2(x,A,\theta)}{\partial x_2}\right|_{\theta=\theta_2(x,A)} \right)$$
$$= \left.\frac{\partial^2 U_2(x,A,\theta)}{\partial x_2 \partial A}\right|_{\theta=\theta_2(x,A)} + \left.\frac{\partial^2 U_2(x,A,\theta)}{\partial x_2 \partial \theta}\right|_{\theta=\theta_2(x,A)} \frac{\partial \theta_2(x,A)}{\partial A} \qquad (6)$$

denote the impact of player 2's bias A on her marginal utility in the long run. Again, the first term $\frac{\partial^2 U_2}{\partial x_2 \partial A}$ measures the direct effect, and the second term $\frac{\partial^2 U_2}{\partial x_2 \partial \theta_2}\frac{\partial \theta_2}{\partial A}$ measures the indirect effect through the belief. Now we are ready to state our first proposition.

Proposition 1 (Steady State under First-Order Misspecification)**.** *Let x^* be a regular steady state for some parameter A^*.*[6] *Then there is an open neighborhood of*

[4]Indeed, in a single-agent problem, if $M_{ii} > 0$ in some steady state, then there are multiple steady states. A proof is akin to that of the second part of Proposition 1, and hence is omitted.

[5]We conjecture that if there is a steady state with $M_{22} > 0$, then it is unstable in the sense that players' actions do not converge there with positive probability. At least, we show that if player 1 chooses the steady-state action every period and player 2 learns an unknown state θ, then player 2's action never converges to a steady state with $M_{22} > 0$. The proof is available upon request.

[6]The regularity conditions (i) and (ii) ensure that the steady state is continuous with respect to the parameter A and the first-order condition for the incentive compatibility is satisfied there. The regularity conditions (iii) is to use the implicit function theorem. $M_{11} \neq 0$ and $M_{22} \neq 0$ in the regularity condition (iv) ensure that the base misspecification effect and the slope of the asymptotic best response curve are well-defined.

A^* such that for any value A in this neighborhood, there is a regular steady state x^* which is continuous with respect to A, and we have

$$\frac{\partial x_2^*}{\partial A} = -\frac{M_{2A}}{M_{22}} \cdot \frac{1}{1 - BR_1' BR_2'}$$
$$\frac{\partial x_1^*}{\partial A} = \frac{\partial x_2^*}{\partial A} \cdot BR_1'.$$

Suppose in addition that given the parameter A^*, the steady state is unique and each asymptotic best response BR_i is a continuous function. Then, $BR_1' BR_2' < 1$.[7]

This proposition shows that the impact of first-order misspecification on the steady-state action is represented as the *base misspecification effect* $-\frac{M_{2A}}{M_{22}}$ times the *multiplier effect* $\frac{1}{1-BR_1' BR_2'}$. The base misspecification effect measures how player 2's bias influences her steady-state action x_2^* *in the absence of strategic interaction*. To see what it means, consider the case in which player 1 chooses the same fixed action each period, so player 2 faces a single-agent problem. Suppose that player 2's bias A increases a bit. This influences player 2's marginal utility by M_{2A} (note that this includes the indirect effect through the steady-state belief), and hence her optimal long-run action changes. The base misspecification effect $-\frac{M_{2A}}{M_{22}}$ measures this change. Since we assume $M_{22} < 0$, the sign of the base misspecification effect coincides with the sign of M_{2A}. That is, player 2's steady-state action increases if and only if a change in A has a positive impact on her marginal utility.

This base misspecification effect is further divided into two parts: Since M_{2A} is a sum of the direct effect and the indirect effect (see (6)), the base misspecification

[7] If these additional assumptions do not hold, there may be a steady state with $BR_1' BR_2' > 1$. But it seems that such a steady state is unstable in an evolutionary sense, especially when misspecification is small. Indeed, in a one-shot game with correctly specified model, a Nash equilibrium with $BR_1' BR_2' > 1$ is not stable under the replicator dynamics (hence it is not an ESS) or the best response dynamics. So in practice, if players' play converge after a long time, it is natural to expect that $BR_1' BR_2' < 1$ in the steady state.

effect is rewritten as

$$-\frac{M_{2A}}{M_{22}} = -\frac{1}{M_{22}} \left(\underbrace{\left.\frac{\partial^2 U_2(x^*, A, \theta)}{\partial x_2 \partial A}\right|_{\theta=\theta_2(x^*,A)}}_{\text{direct effect from } A\text{'s change}} + \underbrace{\frac{\partial \theta_2}{\partial A} \left.\frac{\partial^2 U_2(x^*, A, \theta)}{\partial x_2 \partial \theta}\right|_{\theta=\theta_2(x^*,A)}}_{\text{indirect leargning effect through } \theta_2} \right).$$

The first term in the brackets is the *direct effect*, which measures how player 2's bias influences her action, when her belief θ_2 is fixed. This coincides with the impact of one's bias on her action in a *one-shot game with no learning*, which is studied by past work such as Kyle and Wang (1997). The second term in the brackets is the *(indirect) learning effect*, which measures how player 2's bias influences her action through the steady-state belief θ_2; the term $\frac{\partial \theta_2}{\partial A}$ measures how player 2's bias influences her steady-state belief, and $\frac{\partial^2 U_2}{\partial x_2 \partial A}$ measures how it influences her marginal utility. Intuitively, this learning effect represents how the biased player modifies her view about the world and how it influences her action in the long run. As will be explained later, this effect is the source of the self-defeating feature in Heidhues, Kőszegi, and Strack (2018).

The multiplier effect $\frac{1}{1-BR_1'BR_2'}$ in Proposition 1 measures how strategic interaction between two players amplifies/weakens the base misspecification effect. To better understand the nature of this multiplier effect, suppose that player 2 changes her action by Δ. Then player 1 best-responds to it and changes her action by $BR_1'\Delta$, which in turn has a feedback effect of $BR_1'BR_2'\Delta$ on player 2's steady-state action; note that player 1's action influences player 2's optimal action directly and indirectly through her belief $\theta_2(x,A)$, and both these effects are taken into account in the asymptotic best response BR_2'.

This process continues multiple times; the feedback effect on player 2's action influences player 1's action, which again causes a feedback effect of $(BR_1'BR_2')^2\Delta$ on player 2's action, and so on. Summing all these feedback effects, player 2's action changes by

$$\sum_{k=0}^{\infty} (BR_1'BR_2')^k \Delta = \frac{1}{1-BR_1'BR_2'}\Delta.$$

So the multiplier $\frac{1}{1-BR_1'BR_2'}$ can be seen as a result of the infinite adjustment process between the two strategic players.

The following corollary is an immediate consequence of Proposition 1:

Corollary 1. *Suppose that all the assumptions in Proposition 1 (including the ones in the second part) are satisfied. Then we have the following results:*

(i) The multiplier $\frac{1}{1-BR_1'BR_2'}$ *is positive. So a strategic interaction influences the size of the impact of misspecification, but not the direction.*

(ii) If $sgn(BR_1') = sgn(BR_2')$*, then the multiplier* $\frac{1}{1-BR_1'BR_2'}$ *is greater than one and is increasing in* $|BR_i'|$*. So both strategic substitutes and strategic complements amplify the impact of the first-order misspecification.*

(iii) If $sgn(BR_1') \neq sgn(BR_2')$*, then the multiplier* $\frac{1}{1-BR_1'BR_2'}$ *is less than one and decreasing in* $|BR_i'|$*. So a strategic interaction reduces the impact of the first-order misspecification.*

For the special case in which $A = a$, BR_i reduces to the standard best response curve, so it is fairly easy to compute the slope BR_i'. In particular, when the game is symmetric, we have $\text{sgn}(BR_1') = \text{sgn}(BR_2')$ in any symmetric equilibrium, so a strategic interaction amplifies the impact of first-order misspecification.

So far we have assumed that only player 2 is misspecified, but it is straightforward to see that a similar result holds even when both players have first-order misspecification. Suppose that player 1 believes that the true parameter is $A_1 \neq a$, and player 2 believes that the true parameter is $A_2 \neq a$. Suppose also that these first-order beliefs are common knowledge. For each i, let $\theta_i(x,A_i)$ denote θ which satisfies the consistency condition $Q(x,a,\theta^*) = Q(x,A_i,\theta_i)$. Intuitively, when players play the same action profile x every period, player i's limiting belief (if exists) must be concentrated on $\theta_i(x,A_i)$. Define each player i's asymptotic best response BR_i by (5), and let BR_i' denote its slope. Then the impact of a marginal increase in player 2's misspecification on steady-state actions can be represented just as in Proposition 1, that is,

$$\frac{\partial x_2^*}{\partial A_2} = -\frac{M_{2A}}{M_{22}} \cdot \frac{1}{1 - BR_1' BR_2'},$$
$$\frac{\partial x_1^*}{\partial A_2} = \frac{\partial x_2^*}{\partial A_2} \cdot BR_1'.$$

The only difference from Proposition 1 is that player 1's asymptotic best response is different from the standard best response; now it takes into account the fact that her long-run belief $\theta_1(x, A_1)$ is endogenously determined.

3.2 First-Order Misspecification: Applications

This subsection investigates applications in Cournot duopoly, team production, and tournament, with highlighting how first-order misspecification influences the long-run behavior and welfare.

3.2.1 Cournot duopoly

We first consider a symmetric Cournot duopoly and study how a firm's bias influences the equilibrium outcome. In each period, each firm $i = 1,2$ simultaneously chooses its quantity x_i. Then, each firm observes a market price $y = Q(x_1 + x_2, a, \theta) + \varepsilon$, where a is a parameter which influences the demand and θ is an unknown economic state. Firm i's payoff is $u_i(x_i, y) = x_i y - c(x_i)$, where $x_i y$ is firm i's revenue and $c(x_i)$ is firm i's production cost. Throughout, we assume that the inverse demand function Q is strictly decreasing and weakly concave in the first element, and the cost function c is strictly increasing and weakly convex.[8]

Kyle and Wang (1997), Heifetz, Shannon, and Spiegel (2007), and Englmaier (2010) study (a variant of) one-shot Cournot competition with linear demand, and show that a firm who has overconfidence about the intercept and/or the slope of

[8]These assumptions ensure the concavity of each firm's payoff function, a downward-sloping best response curve, and the uniqueness of Nash equilibrium under the correctly specified model. See, for example, Tirole (1988).

the inverse demand function earns higher profits than the unbiased rival firm. Intuitively, the overconfident firm is willing to produce more than in the correctly specified model. Knowing that, the unbiased firm reduces its production level in equilibrium, which yields higher profits to the overconfident firm. This mechanism, which we call the *strategic effect*, is akin to the commitment to produce more in the context of quantity competition (e.g., being a first mover in the standard Stackelberg duopoly).[9]

A natural question is whether this strategic effect is beneficial for the overconfident firm in the long run. So long as the game is repeated, the overconfident firm is persistently "surprised" by a realized price being lower than its anticipation. This suggests that the overconfident firm's subjective view about the world would change after a long time, in order to better explain such lower prices. Accordingly, the overconfident firm changes the behavior, and the rival firm best-responds to it. In what follows, we will study how this influences the equilibrium outcome.

Suppose that firm 2 incorrectly believes that the true parameter is $A > a$, where $Q_A > 0$ and $Q_{xA} \geq 0$ for all x with $x_1 + x_2 > 0$. Intuitively, firm 2 is overconfident about the price level Q and (weakly) overconfident about the slope of the inverse demand curve Q_x. In the special case of linear demand, this assumption reduces to the one imposed in the literature (e.g., Kyle and Wang (1997)). Firm 1 knows that the true parameter is a, and the firms' first-order beliefs are common knowledge. We also assume that $Q_\theta > 0$ and $Q_{x\theta} \geq 0$ for all x with $x_1 + x_2 > 0$, i.e., the state θ has positive impacts on the price level and the slope of the inverse demand function.

Here are two examples which satisfy the assumptions above:[10]

$$Q(x_1 + x_2, a, \theta) = a - (1 - \theta)(x_1 + x_2), \tag{7}$$

and

$$Q(x_1 + x_2, a, \theta) = \theta - (1 - a)(x_1 + x_2). \tag{8}$$

[9] See also Fershtman and Judd (1987) who investigate the strategic effect induced by managerial compensation contracts under a correctly specified model.

[10] In these examples, as Q is linear in θ, the identifiability condition in Section 5 is satisfied. Hence, from Proposition 13, the actions and the beliefs converge to a steady state almost surely.

In the first example (7), firm 2 is overconfident about the intercept of the inverse demand function, and learns the slope of the inverse demand function. This happens, for example, when the firm has overconfidence about the preference of the representative customers and learns their number.[11] In the second example (8), firm 2 is overconfident about the slope of the inverse demand function, and learns the intercept. This happens, for example, when the firm has overconfidence about the number of the customers and learns their preference.[12]

In both these examples, if the firms interact only once, the overconfident firm earns a higher equilibrium profit than the rival firm (Kyle and Wang, 1997). However, it turns out that these two examples give qualitatively different long-run outcomes; as will be explained, a firm's overconfidence improves its long-run payoff in the second example, but it is detrimental in the first example.

Recall from Proposition 1 that the impact of firm 2's overconfidence on its own action is represented as the base misspecification effect $-\frac{M_{2A}}{M_{22}}$ times the multiplier. For ease of exposition, we assume that misspecification is small (i.e., A is close to a) so that $M_{22} < 0$ and the multiplier is positive. Simple algebra shows that the base misspecification effect in our Cournot model is written as

$$-\frac{1}{M_{22}}\left[\underbrace{\overbrace{Q_A(x_1^*+x_2^*,A,\theta_2)}^{\text{direct effect}}+\overbrace{\frac{\partial\theta_2}{\partial A}Q_\theta(x_1^*+x_2^*,A,\theta_2)}^{\text{indirect effect}}}_{\text{on the price level}}+x_2^*\left(\underbrace{\overbrace{Q_{xA}(x_1^*+x_2^*,A,\theta_2)}^{\text{direct effect}}+\overbrace{\frac{\partial\theta_2}{\partial A}Q_{x\theta}(x_1^*+x_2^*,A,\theta_2)}^{\text{indirect effect}}}_{\text{on the slope}}\right)\right]. \tag{9}$$

By the implicit function theorem, we have $\frac{\partial\theta_2}{\partial A} = -\frac{Q_A(x_1^*+x_2^*,A,\theta_2)}{Q_\theta(x_1^*+x_2^*,A,\theta_2)} < 0$. Plugging this into (9), the first two terms in the brackets cancel out. Intuitively, this is a consequence of the consistency condition (4); in the long run, the overconfidence about the price level is offset by a pessimistic view about the parameter θ, and the firm correctly predicts the average price level. Accordingly, the base misspecifi-

[11] Suppose that there are $\frac{1}{1-\theta}$ customers, and each of them purchases $a-p$ units of products, where p denotes a price. Then the total demand is $x = \frac{a-p}{1-\theta}$, which results in the inverse demand function $p = a-(1-\theta)x$.

[12] Suppose that there are $\frac{1}{1-a}$ customers, and each of them purchases $\theta - p$ units of products. Then the total demand is $x = \frac{\theta-p}{1-a}$, which results in the inverse demand function $p = \theta-(1-a)x$.

cation effect is simplified to

$$-\frac{x_2^*}{M_{22}}\left[Q_{xA}(x_1^*+x_2^*,A,\theta_2)-\frac{Q_A(x_1^*+x_2^*,A,\theta_2)}{Q_\theta(x_1^*+x_2^*,A,\theta_2)}Q_{x\theta}(x_1^*+x_2^*,A,\theta_2)\right]. \tag{10}$$

(10) implies that the overconfident firm's long-run action crucially relies on its steady-state belief about the slope of the inverse demand curve. Suppose first that $Q_{xA}-\frac{Q_A}{Q_\theta}Q_{x\theta}<0$, i.e., the firm is pessimistic about the demand slope in the steady state. In this case, the base misspecification effect is negative, and the overconfident firm produces less in equilibrium. The rival firm best-responds to it and produces more, which decreases the overconfident firm's equilibrium payoff. Hence, the firm's overconfidence is detrimental in the long run. It corresponds to the first example (7), where $Q_{xA}-\frac{Q_A}{Q_\theta}Q_{x\theta}=-\frac{1}{x_1+x_2}<0$. Intuitively, after observing unexpectedly low prices, the overconfident firm becomes pessimistic about the demand slope and reduces the production level, which hurts own long-run profit.

Next, suppose that $Q_{xA}-\frac{Q_A}{Q_\theta}Q_{x\theta}>0$, i.e., the overconfident firm is optimistic about the slope in the steady state. In this case, the base misspecification effect is positive, and hence the overconfident firm produces more than in the correctly specified model. The rival firm best-responds to it and produces less, which improves the overconfident firm's equilibrium payoff. Hence, the firm's overconfidence is beneficial even in the long run. It corresponds to the second example (8), where $Q_{xA}-\frac{Q_A}{Q_\theta}Q_{x\theta}=1>0$. In this example, the overconfident firm becomes pessimistic about the demand intercept over time, but its overconfidence about the demand slope is persistent. Given this optimistic view about the demand slope, in the steady state, the overconfident firm keeps producing more than the rival firm, and earns a higher long-run profit.

Remark 1. From (9), the indirect learning effect in this Cournot model is

$$-\frac{1}{M_{22}}\left(\frac{\partial\theta_2}{\partial A}Q_\theta(x_1^*+x_2^*,A,\theta_2)+x_2^*\frac{\partial\theta_2}{\partial A}Q_{x\theta}(x_1^*+x_2^*,A,\theta_2)\right).$$

Since we assume $Q_\theta>0$ and $Q_{x\theta}\geq 0$, this effect is negative. This implies that learning mitigates the short-run effect of overconfidence; i.e., in the long run, the overconfident firm produces less and earns a lower profit than in the corresponding one-shot model. When $Q_{xA} - \frac{Q_A}{Q_\theta} Q_{x\theta} > 0$ or equivalently $\frac{Q_{x\theta}}{Q_\theta} < \frac{Q_{xA}}{Q_A}$, this learning effect is relatively small, so overconfidence is still beneficial in the long run. By contrast, when $\frac{Q_{x\theta}}{Q_\theta} > \frac{Q_{xA}}{Q_A}$, the learning effect outweighs the short-run strategic effect, so overconfidence is detrimental in the long run.

3.2.2 Team production

Consider two players working on a joint project. Each period, each player i chooses an effort level x_i, and observes a stochastic output $y = Q(x, a, \theta^*) + \varepsilon$ where a is the total capability of the players and θ^* is an unknown fundamental. We assume that Q is twice-continuously differentiable and symmetric in that $Q(x_1, x_2, a, \theta) = Q(x_2, x_1, a, \theta)$. Assume $Q_{x_i} > 0$, $Q_a > 0$, and $Q_\theta > 0$. Player i's payoff is $y - c(x_i)$, where $c(x_i)$ is the effort cost. Assume $c' > 0$. Assume also that there is a unique Nash equilibrium x^{correct}, and it is symmetric in that $x_1^{\text{correct}} = x_2^{\text{correct}}$.

Suppose that player 2 has first-order misspecification, in that she incorrectly believes that the capability is $A \neq a$. When $A > a$, it represents player 2's overconfidence about her own capability. When $A < a$, it represents player 2's underconfidence about her own capability, or prejudice about the opponent's capability. Proposition 1 shows that the impact of the misspecification is represented by the base misspecification effect times the multiplier. In the team production problem, the base misspecification effect is written as

$$-\frac{1}{M_{22}} \left(\underbrace{Q_{x_i A}}_{\text{direct effect}} + \underbrace{\frac{\partial \theta_2}{\partial A} Q_{x_i \theta}}_{\text{indirect effect}} \right). \tag{11}$$

The first term in the brackets is the direct effect of player 2's bias A on the marginal productivity, and the second term is the indirect effect through the steady-state belief θ_2. As is clear from this formula, the sign of the base misspecification effect crucially depends on the signs of $Q_{x_i a}$ and $Q_{x_i \theta}$. In what follows, we assume $Q_{x_i a} \leq 0$ and $Q_{x_i \theta} > 0$, which requires that the marginal return Q_{x_i} be negatively correlated with the capability a, and positively correlated with the fundamental

θ.[13] The same assumption is imposed in Heidhues, Kőszegi, and Strack (2018), so by focusing on this case, we can illustrate what is new in our multi-agent setup.[14]

Assume $M_{22} < 0$ as usual. Then, by the assumption $Q_{x_i a} \leq 0$, the direct effect appearing in (11) is non-positive; in the one-shot game with no learning, the overconfident player's effort level is lower than in the correctly specified model. Intuitively, when $Q_{x_i a} \leq 0$, the overconfident player thinks that the marginal productivity is lower than the truth, and thus reduces the effort level.

Also, since $Q_{x_i\theta} > 0$ and $\frac{\partial \theta_2}{\partial A} = -\frac{Q_a}{Q_\theta} < 0$, the indirect learning effect in (11) is negative. This implies that in the long run, the overconfident player's effort level is even lower than in the one-shot game. Intuitively, the overconfident player finds that the outputs are lower than what she thinks, and becomes pessimistic about θ over time, which lowers her effort level. This effect is exactly the self-defeating property discussed in Heidhues, Kőszegi, and Strack (2018), which states that an ability to adjust an action depending on an updated belief distorts the action more.

A difference from Heidhues, Kőszegi, and Strack (2018) is that this self-defeating effect is further amplified by the multiplier $\frac{1}{1-BR_1'BR_2'} > 1$ in our multi-player setup. Intuitively, if the overconfident player becomes pessimistic about θ and reduces the effort level, then the opponent best-respond to it; she increases the effort if $Q_{x_1x_2} < 0$, and decreases the effort if $Q_{x_1x_2} > 0$. This influences the overconfident player's optimal action, and in *both* cases, the overconfident player

[13]With this assumption, when $A > a$, the monotonicity condition stated in Proposition 5 is satisfied, so both the actions and the beliefs indeed converge to a steady state. When $A < a$, the monotonicity condition fails; but if the identifiability condition stated in Section 5 is satisfied, then the actions and the beliefs converge to a steady state. This identifiability condition is satisfied when Q is linear in θ.

[14]It is not difficult to apply the technique here for other cases. For example, suppose that $Q < 0$ is the damage from drought and agents invest to irrigation which mitigate the damage. Suppose that Q takes a form of

$$Q = -\frac{1}{\theta}\left(\frac{1}{x_1+x_2} + \frac{1}{a}\right).$$

In this case, $Q_{x_i a} \geq 0$ and $Q_{x_i\theta} < 0$, so both the direct effect and the indirect effect are positive. Hence, Corollary 3 holds if we reverse all the inequalities. For example, player 2's overconfidence about her capability *increases* her effort in the one-shot game, and she makes even more effort in the long run.

reduces the effort level further. Due to this mechanism, the impact of misspecification on the steady-state action is larger in our strategic setup than in the single-agent setup.

Our long-run welfare effects are also different from Heidhues, Kőszegi, and Strack (2018). When the players' actions are strategic substitutes (i.e., $Q_{x_1x_2} < 0$), the opponent increases own effort level in response to the overconfident player's action, so the overconfident player can free ride. Indeed, if $Q_{x_1x_2} < 0$, small overconfidence strictly improves own long-run payoff: the above free-riding effect outweighs the direct cost of misperception.

When $A < a$, a misspecified player underestimates the total capability of the players. For example, the player is underconfident about own capability or has prejudice about the opponent's capability. In this case, a player's small underestimation improves social surplus and the opponent's long-run payoff. Further, if $Q_{x_1x_2} < 0$, it also improves the misspecified player's long-run payoff. Intuitively, given such a pessimistic view, the misspecified player becomes optimistic about the fundamental θ over time. This learning effect outweighs the direct effect of underestimation, and the misspecified player contributes more in the long run.

3.2.3 Tournaments

As a third application, we discuss a standard tournament model based on Lazear and Rosen (1981). Suppose that there are two players. In each period t, each player i chooses an effort level x_i and observes a stochastic output $y \in \{w, l\}$, where $y = w$ means "player 2 wins" and $y = l$ means "player 1 wins." The probability of $y = w$ (i.e., the probability of player 2 being a winner) is $Q(x_1, x_2, a_1, a_2, \theta)$, where a_i denotes player i's capability and θ is an unknown economic state. We assume that $Q_{x_1} < 0$, $Q_{x_2} > 0$, $Q_{a_1} < 0$, $Q_{a_2} > 0$, and $Q_\theta > 0$; i.e., player i has a better chance of winning if she exerts more effort and/or has a better skill. These assumptions are satisfied, for example, if

$$Q(x_1, x_2, a_1, a_2, \theta) = \theta \frac{x_2 + a_2}{x_1 + a_1 + x_2 + a_2}. \tag{12}$$

This functional form is commonly used in the literature since Tullock (1980). The parameter θ represents players' uncertainty about fairness of the evaluation system: it is a fair contest if $\theta = 1$, but player 1 is favored if $\theta < 1$, and player 2 is favored if $\theta > 1$.[15] Players' beliefs about this parameter θ changes over time, depending on the observed output. A winner receives a payoff $W = 1$, and a loser receives a payoff $L = 0$. Each agent's effort cost is $c(x_i)$, and we assume that $c' > 0$. Player 1's payoff is $u_1(x_1, y) = Prob(y = l) - c(x_1) = [1 - Q(x_1, x_2, a_1, a_2, \theta)] - c(x_1)$, while player 2's payoff is $u_2(x_2, y) = Prob(y = w) - c(x_2) = Q(x_1, x_2, a_1, a_2, \theta) - c(x_2)$.

Suppose that player 2 has first-order misspecification in that she incorrectly believes that her capability is $A \neq a_2$. When $A > a_2$, it represents player 2's overconfidence about her own capability or prejudice about the opponent's capability. When $A < a_2$, it represents player 2's underconfidence about her own capability.

This setup is slightly different from the one we have studied so far; we consider the binary signal space $Y = \{w, l\}$ instead of the continuous signal space. However, this does not change the steady-state conditions at all, i.e., the conditions (3)-(2) must be satisfied in a steady state in the tournament model with binary signals. Accordingly, Proposition 1 applies to the tournament model, and the impact of misspecification is represented by the base misspecification effect times the multiplier.

Simple algebra shows that the base misspecification effect in this tournament model is written as

$$-\frac{1}{M_{22}}\left(\underbrace{Q_{x_2A}(x_1^*, x_2^*, a_1, A, \theta_2)}_{\text{direct effect}} - \underbrace{\frac{\partial \theta_2}{\partial A} Q_{x_2\theta}(x_1^*, x_2^*, a_1, A, \theta_2)}_{\text{indirect learning effect}}\right). \tag{13}$$

This is exactly the same as the base misspecification effect (11) in the team production, so the results in Section 3.2.2 continue to hold. For example, in Tullock-type tournament (12), we have $Q_{x_2A} = -2\theta \frac{x_1+a_1}{(x_1+a_1+x_2+A)^3} < 0$ and $Q_{x_2\theta} = \frac{x_1+a_1}{(x_1+a_1+x_2+A)^2} >$

[15] Another example is $Q(x_1, x_2, a_1, a_2, \theta) = \theta + \frac{x_2+a_2}{x_1+a_1+x_2+a_2}$. In this example, it is a fair contest if $\theta = 0$, but player 1 is favored if $\theta < 0$, and player 2 is favored if $\theta > 0$. All of the following discussions hold except that $Q_{x_i\theta} = 0$ in this example.

0, so assuming $M_{22} < 0$, both the direct effect and the indirect effect in the base misspecification effect are negative. This means that the overconfident player does not work hard in the one-shot game, and in the long run, her effort level is even lower than that. Intuitively, the overconfident player incorrectly believes that the marginal return of effort is low ($Q_{x_2 A} < 0$) and does not work hard in the one-shot game. On top of that, since she wins less frequently than what she thinks, after a long time, she becomes pessimistic about θ and incorrectly believes that the contest is unfair. This learning effect further reduces her effort.

However, this base misspecification effect is *reduced* by the multiplier effect. Indeed, in this tournament model, the multiplier is

$$\frac{1}{1 - BR_1' BR_2'} = \frac{1}{1 - \frac{M_{12}}{M_{11}} \frac{M_{21}}{M_{22}}} = \frac{1}{1 + \frac{Q_{x_1 x_2}^2}{M_{11} M_{22}}} \leq 1.$$

Here the inequality follows from the fact that $M_{11} < 0$ and $M_{22} < 0$ with small misspecification. Note that this inequality is strict whenever $Q_{x_1 x_2} \neq 0$.[16] Intuitively, when the overconfident player 2 reduces the effort due to the base misspecification effect, the opponent best-responds to it; she increases the effort if $Q_{x_1 x_2} < 0$, and decreases the effort if $Q_{x_1 x_2} > 0$. This in turn influences player 2's optimal action, and in both cases, she *increases* the effort; this mitigates the base misspecification effect.

This result is quite different from that in the team production, where the multiplier is greater than one and *amplifies* the base misspecification effect. A crucial difference is that players in the tournament have conflicting interests about the output y; player 2 prefers $y = w$ while player 1 prefers $y = l$. Accordingly we have $\text{sgn}(BR_1') \neq \text{sgn}(BR_2')$, which implies that the multiplier is less than one and strategic interaction weakens the impact of misspecification. By contrast, in the team production, players have a common preference on y, and accordingly we have $\text{sgn}(BR_1') = \text{sgn}(BR_2')$. In this case, the multiplier is larger than one, and strategic interaction strengthens the impact of misspecification. The same argument

[16] For example, in the Tullock-type tournament, we have $Q_{x_1 x_2} = 0$ only when $x_1 = x_2 = 0$, so the multiplier is less than one for all parameter A with which $x_1^{\text{first}} \neq x_2^{\text{first}}$.

applies to a more general setup; in a common interest game, we should expect a larger deviation of long-run actions from a correctly specified model than that in the single-agent model.

4 Higher-Order Misspecification

This section analyzes higher-order misspecification. Section 4.1 covers second-order misspecification where a player has an incorrect view about the opponent's view. Section 4.2 covers double misspecification. It is a combination of first- and second-order misspecification: either one or both players has an incorrect view about the physical environment and believes that the opponent shares the same view. Section 4.3 focuses on the case in which a player's misspecification is small and derives a simple characterization. Section 4.4 investigates applications in Cournot duopoly and team production with higher-order misspecification.

4.1 Second-Order Misspecification

In the previous subsection, we consider the situation in which a player has an incorrect view about the physical environment, but correctly understands what the opponent thinks about the physical environment. However, economic agents often have various forms of bias about what the opponent thinks. For example, Madarász (2012) and Gagnon-Bartsch (2016) study the cases in which a player misperceives the other players' knowledge/preference, which can be represented by a in our model. Players may also misperceive other players' depth of reasoning (see Eyster (2019) for an extensive review).

In this subsection, we consider a long-run impact of a player's bias about the opponent's view about the world. We assume that player 2 has *second-order misspecification*, in that she correctly understands the physical environment, but has an incorrect view about the opponent's view about a.[17] Formally, we consider

[17] As evidence from laboratory experiments, subjects often systematically mispredict other subjects' preferences and actions (Van Boven, Dunning, and Loewenstein, 2000, for example). Lud-

the following information structure:

- Both players believe that for each parameter θ, the signal y is given by $y = Q(x_1, x_2, a, \theta) + \varepsilon$.

- Player 2 (incorrectly) believes that it is common knowledge that "for each parameter θ, player 1 believes that the signal y is given by $y = Q(x_1, x_2, A, \theta) + \varepsilon$ and player 2 believes that the signal y is given by $y = Q(x_1, x_2, a, \theta) + \varepsilon$," where $A \neq a$.

- Player 1 knows player 2's information structure above.

Intuitively, this is the case in which player 2 has prejudice, in the sense that she incorrectly believes that she has better information than the opponent does. Player 1 is unbiased, because she knows both the physical environment and the opponent's information.

In this setup, player 2 faces *inferential naivety*. She believes that the opponent takes an action based on a misspecified model, so she makes an incorrect prediction about the opponent's play. It turns out that this inferential naivety influences player 2's action in two ways. First, player 2 best-responds to this incorrectly predicted action of the opponent. Second, player 2 interprets an observed signal conditional on the incorrectly predicted action, which leads to misguided learning.

Assume that players are myopic so that they maximize the expected stage-game payoffs each period. To characterize equilibrium actions when player 2 has second-order misspecification, it is useful to introduce *hypothetical player* 1 who incorrectly believes that the true parameter is $A \neq a$. Player 2 believes that the opponent is this hypothetical player 1, so each period, she chooses a Nash equilibrium action against this hypothetical player. The true player 1 correctly understands player 2's reasoning, and best responds to player 2's action.

Formally, let $(\hat{\mu}_1, \hat{x}_t)$ denote the action and the belief of the hypothetical player, and let $x = (x_1, x_2, \hat{x}_1)$ denote an action profile in the three-player game. The

wig and Nafziger (2011) report that most subjects in their experiments are not aware of or underestimate overconfidence of other subjects.

hypothetical player 1's expected stage-game payoff given θ is

$$\hat{U}_1(x,\theta,A) = E[u_1(\hat{x}_1, Q(\hat{x}_1,x_2,A,\theta)+\varepsilon)],$$

because she thinks that the parameter is $A \neq a$. Player 2's expected stage-game payoff is

$$U_2(x,\theta) = E[u_2(x_2, Q(\hat{x}_1,x_2,a,\theta)+\varepsilon)],$$

because she thinks that the opponent is the hypothetical player who chooses $\hat{x}_1$. Player 1's subjective expected stage-game payoff is

$$U_1(x,\theta) = E[u_1(x_1, Q(x_1,x_2,a,\theta)+\varepsilon)].$$

Using these notations, the equilibrium strategy in the infinite-horizon game is described as follows. In period one, all players have the same belief $\mu_1^1 = \mu_2^1 = \hat{\mu}_1^1 = \mu$. So they play a Nash equilibrium $(x_1^1, x_2^1, \hat{x}_1^1)$, which solves the first-order conditions $\frac{\partial E[U_1(x,\theta)|\mu]}{\partial x_1} = 0$, $\frac{\partial E[U_2(x,\theta)|\mu]}{\partial x_2} = 0$, and $\frac{\partial E[\hat{U}_1(x,\theta)|\mu]}{\partial \hat{x}_1} = 0$. At the end of period one, players observe a public signal $y^1 = Q(x_1^1, x_2^1, a, \theta^*) + \varepsilon$, and updates the posterior beliefs using Bayes' rule. So each player's belief in period two is

$$\mu_1^2(\theta) = \frac{\mu_1^1(\theta) f(y - Q(x_1^1, x_2^1, a, \theta))}{\int_\Theta \mu_1^1(\tilde{\theta}) f(y - Q(x_1^1, x_2^1, a, \tilde{\theta}))d\tilde{\theta}},$$
$$\mu_2^2(\theta) = \frac{\mu_2^1(\theta) f(y - Q(\hat{x}_1^1, x_2^1, a, \theta))}{\int_\Theta \mu_2^1(\tilde{\theta}) f(y - Q(\hat{x}_1^1, x_2^1, a, \tilde{\theta}))d\tilde{\theta}},$$
$$\hat{\mu}_1^2(\theta) = \frac{\hat{\mu}_1^1(\theta) f(y - Q(\hat{x}_1^1, x_2^1, A, \theta))}{\int_\Theta \hat{\mu}_1^1(\tilde{\theta}) f(y - Q(\hat{x}_1^1, x_2^1, A, \tilde{\theta}))d\tilde{\theta}}.$$

As is clear from this formula, while player 2 correctly knows the parameter a, her posterior μ_2^2 differs from player 1's posterior μ_1^2 because she uses Bayes' rule based on the wrong prediction $\hat{x}_1^1 \neq x_1^1$ about player 1's action. Since actions are not observable, on the equilibrium path, the beliefs $(\mu_2^2, \hat{\mu}_2^2)$ are common knowledge between player 2 and the hypothetical player 1. Also player 1 knows the belief profile $\mu^2 = (\mu_1^2, \mu_2^2, \hat{\mu}_1^2)$. So in period two, players play a Nash equilibrium given this belief profile $\mu^2 = (\mu_1^2, \mu_2^2, \hat{\mu}_1^2)$. Likewise, in any subsequent period

$t \geq 3$, players play a Nash equilibrium given the belief profile $\mu^t = (\mu_1^t, \mu_2^t, \hat{\mu}_1^t)$, where μ^t is computed by Bayes' rule.

As will be shown in Section 5, under a mild sufficient condition, players' beliefs and actions almost surely converge to a *steady state* $(x_1^*, x_2^*, \hat{x}_1^*, \mu_1^*, \mu_2^*, \hat{\mu}_1^*)$ which satisfies the following conditions:

$$x_1^* \in \arg\max_{x_1} U_1(x_1, x_2^*, \hat{x}_1^*, \theta^*), \tag{14}$$

$$x_2^* \in \arg\max_{x_2} U_2(x_2, x_1^*, \hat{x}_1^*, \theta_2), \tag{15}$$

$$\hat{x}_1^* \in \arg\max_{\hat{x}_1} \hat{U}_1(\hat{x}_1, x_1^*, x_2^*, \hat{\theta}_1), \tag{16}$$

$$\mu_1^* = 1_{\theta^*}, \tag{17}$$

$$\mu_2^* = 1_{\theta_2} \text{ s.t. } Q(\hat{x}_1^*, x_2^*, a, \theta_2) = Q(x_1^*, x_2^*, a, \theta^*), \tag{18}$$

$$\hat{\mu}_1^* = 1_{\hat{\theta}_1} \text{ s.t. } Q(\hat{x}_1^*, x_2^*, A, \hat{\theta}_1) = Q(x_1^*, x_2^*, a, \theta^*). \tag{19}$$

The first three conditions (14), (15), and (16) are the incentive compatibility conditions, which require that each player maximize her own payoff given some beliefs. The next three conditions (17), (18), and (19) require that these beliefs satisfy consistency, in that each (actual and hypothetical) player's belief is concentrated on a state with which each player's subjective signal distribution coincides with the objective distribution.

As in the case with first-order misspecification, we assume that for each action profile x, there is a unique state which solves the consistency condition (18), and we denote it by $\theta_2(x,A)$. This $\theta_2(x,A)$ can be interpreted as player 2's long-run belief when players choose the same action x each period. Similarly, we assume that for each x, there is a unique state which solves (19), and we denote it by $\hat{\theta}_1(x,A)$. Player 1's long-run belief is defined as $\theta_1(x,A) = \theta^*$ for all x.

We will characterize how player 2's misspecification influences the steady-state actions, and to do so, the following notation is useful. For each $i, j = 1,2,3$ (possibly $i = j$), let

$$M_{ij} = \left.\frac{\partial^2 U_i(x,\theta)}{\partial x_i \partial x_j}\right|_{\theta=\theta_i(x,A)} + \frac{\partial \theta_i(x,A)}{\partial x_j} \cdot \left.\frac{\partial^2 U_i(x,\theta)}{\partial x_i \partial \theta}\right|_{\theta=\theta_i(x,A)}.$$

denote the impact of player j's action on player i's marginal utility in the long run. (Here player 3 refers to the hypothetical player, and x_3, U_3, and θ_3 are $\hat{x}_1$, $\hat{U}_1$, and $\hat{\theta}_1$, respectively.) The first term is the direct effect, and the second term is the indirect effect through the steady-state belief θ_i. Then define the slope of player i's asymptotic best response curve with respect to player j's action as

$$BR'_{ij} = -\frac{M_{ij}}{M_{ii}}.$$

Intuitively, BR'_{ij} measures how player j's action influences player i's optimal long-run action, while the action of $l \neq i, j$ being fixed. The slope of player 1's asymptotic best response curve, BR'_{12} and BR'_{13}, coincides with that of the standard best response curve. This is so because she can learn the true state regardless of the opponents' play, and the indirect effects in M_{11}, M_{12}, and M_{13} are zero. In particular, $BR'_{13} = 0$, because player 3 is not player 1's opponent and the direct effect in M_{13} is zero. On the other hand, the slopes of the other players' asymptotic best response curves are different from those of the standard best response, due to the indirect effect. For example, BR'_{21} and BR'_{31} need not be zero, even though players 2 and 3 do not think that player 1 is the opponent. Importantly, the indirect effects in M_{21}, M_{23}, M_{31}, and M_{33} do not disappear even in the limit as $A \to a$. This is so because there is inferential naivety, and $\theta_2(x,a)$ and $\hat{\theta}_1(x,a)$ can be different from θ^* if $x_1 \neq \hat{x}_1$. This is in a sharp contrast with the case with first-order misspecification, where all the indirect effects disappear in the limit as $A \to a$.

Let

$$M_{3A} := \left.\frac{\partial^2 \hat{U}_1(x,\theta,A)}{\partial \hat{x}_1 \partial A}\right|_{\theta=\hat{\theta}_1(x,A)} + \left.\frac{\partial^2 \hat{U}_1(x,\theta,A)}{\partial \hat{x}_1 \partial \theta}\right|_{\theta=\hat{\theta}_1(x,A)} \frac{\partial \hat{\theta}_1(x,A)}{\partial A}$$

denote the impact of the hypothetical player's bias A on her own marginal utility. Now we are ready to state the result:

Definition 2. A steady state x^* is *regular* if the following conditions are satisfied in x^*: (i) the steady-state action x_i^* is uniquely optimal, (ii) x^*, $\theta_2(x^*,A)$, and $\hat{\theta}_1(x^*,A)$ are interior points, (iii) $BR'_{23}BR'_{32} \neq 1$ and $BR'_{12}BR'_{21} + BR'_{23}BR'_{32} +$

$BR'_{12}BR'_{23}BR'_{31} \neq 1$, and (iv) $M_{ii} < 0$ for each i.[18]

Proposition 2 (Steady State under Second-Order Misspecification)**.** *Let x^* be a regular steady state for some parameter A^*.*[19] *Then there is an open neighborhood of A^* such that for any value A in this neighborhood, there is a regular steady state x^* which is continuous with respect to A, and we have*

$$\frac{\partial x_2^*}{\partial A} = -\frac{M_{3A}}{M_{33}}\left(\frac{BR'_{23}}{1-BR'_{23}BR'_{32}}\right)\left(\frac{1}{1-BR'_{12}NE'_2}\right)$$
$$\frac{\partial x_1^*}{\partial A} = \frac{\partial x_2^*}{\partial A}\cdot BR'_{12}$$

where

$$NE'_2 = \frac{BR'_{21}+BR'_{23}BR'_{31}}{1-BR'_{23}BR'_{32}}. \tag{20}$$

The first equation in this proposition describes how player 2's second-order misspecification influences her own steady-state action x_2, and the second equation states that the rational player 1 simply best-responds to player 2's play. To interpret the first equation, recall that the parameter A represents the first-order belief (about the physical environment) of the hypothetical player. So when this parameter A changes, it influences the hypothetical player's optimal action $\hat{x}_1$ directly and indirectly through the steady-state belief. The first term $-\frac{M_{3A}}{M_{33}}$ in the equation measures this impact, holding the other players' actions fixed. Note that

[18]As in the case with first-order misspecification, the regularity conditions (i) and (ii) ensure that the steady state is continuous with respect to the parameter A and the first-order condition for the incentive compatibility is satisfied there. The condition (iii) is needed for the multiplier effect to be well-defined. The condition (iv) ensures that the base misspecification effect and the slope of the asymptotic best response curve are well-defined. This condition is also useful when we interpret the base misspecification effect.

[19]Under the following additional assumption, we can also show $\frac{BR'_{23}(BR'_{32}+BR'_{12}BR'_{31})}{1-BR'_{12}BR'_{21}} < 1$. Specifically, given $\hat{x}_1$, let $NE(\hat{x}_1)$ denote the set of (x_1,x_2) satisfying (18), (14), and (15) for some θ_2. Also, given (x_1,x_2), let $BR_3(x_1,x_2)$ denote the set of $\hat{x}_1$ satisfying (19) and (16) for some , $\hat{\theta}_1$. If NE and BR_3 are continuous functions (rather than correspondences) and if a steady state is unique, then $\frac{BR'_{23}(BR'_{32}+BR'_{12}BR'_{31})}{1-BR'_{12}BR'_{21}} < 1$. A proof is available upon request.

this term is very similar to the base misspecification effect appearing in Proposition 1.

The second term in the equation, $\frac{BR'_{23}}{1-BR'_{23}BR'_{32}}$, measures how the hypothetical player's action $\hat{x}_1$ influences player 2's action, holding player 1's action being fixed. When the hypothetical player's action $\hat{x}_1$ changes by $-\frac{M_{3A}}{M_{33}}$, player 2 best-responds to it, and her steady-state belief is affected. Accordingly, player 2's optimal long-run action changes by $-\frac{M_{3A}}{M_{33}}BR'_{23}$. Also, holding player 1's action fixed, this effect is amplified by the strategic interaction between player 2 and the hypothetical player; a change in player 2's action influences the hypothetical player's action and belief, which in turn influences player 2's action and belief, and so on. As in the case with first-order misspecification, this effect is represented by the multiplier $\frac{1}{1-BR'_{23}BR'_{23}}$. So in total, when player 1's action is fixed, player 2's second-order misspecification influences her own steady-state action by $-\frac{M_{3A}}{M_{33}}\left(\frac{BR'_{23}}{1-BR'_{23}BR'_{32}}\right)$.

The last term in the equation, $\frac{1}{1-BR'_{12}NE'_2}$, measures how player 1's strategic play amplifies/reduces the impact of misspecification. To see what it means, it is useful to define *player 2's asymptotic Nash equilibrium correspondence* as

$$NE_2(x_1) = \{x_2 | \exists \hat{x}_1 \text{ satisfying (15), (16), (18), (19)}\}$$

for each x_1. Intuitively, $NE_2(x_1)$ denotes player 2's steady-state action, when player 1 chooses the same action x_1 every period while the other players learn the state and adjust actions. Then the term NE'_2 appearing in the proposition can be interpreted as the slope of this Nash equilibrium correspondence NE_2, i.e., it measures how a marginal change in player 1's (constant) action x_1 influences player 2's steady-state action.[20]

[20]To see that $NE'_2 = \frac{BR'_{21}+BR'_{23}BR'_{31}}{1-BR'_{23}BR'_{32}}$ is the slope of NE_2, suppose that the steady-state action $(x_2, \hat{x}_1)$ is an interior solution for every x_1. Then the following first-order conditions must be satisfied in any steady state:

$$\left.\frac{\partial U_2}{\partial x_2}\right|_{\theta_2=\theta_2(x,A)} = 0, \quad \left.\frac{\partial \hat{U}_1}{\partial \hat{x}_1}\right|_{\hat{\theta}_1=\hat{\theta}_1(x,A)} = 0.$$

Applying the implicit function theorem to this system of equations (here we regard $(x_2, \hat{x}_1)$ as a

With this interpretation in mind, suppose that player 2's action changes by Δ. This influences player 1's optimal action by $BR'_{12}\Delta$, which in turn influences player 2's (and the hypothetical player's) steady-state beliefs and actions. This feedback effect on player 2's action is $BR'_{12}NE'_2\Delta$. This process continues infinitely, which results in the multiplier effect $\frac{1}{1-BR'_{12}NE'_2}$.[21]

While NE'_2 is somewhat similar to BR'_2 appearing in the case of first-order misspecification, there are two important differences. First, in NE'_2, we consider the case in which both player 2 and the hypothetical player adjust actions (and play a Nash equilibrium) every period. In BR'_2 (and in BR'_{21}), we consider the case in which only player 2 adjusts actions. Second, since player 2 does not think that player 1 is the opponent in the case of second-order misspecification, $NE'_2 = \frac{BR'_{21}+BR'_{23}BR'_{31}}{1-BR'_{23}BR'_{32}}$ involves only the indirect effect; the first term BR'_{21} in the numerator represents how player 1's action influences player 2's action through the steady-state belief, and the second term $BR'_{23}BR'_{31}$ represents how player 1's action influences player 2's action through the hypothetical player's action. These effects are amplified by the strategic interaction between player 2 and the hypothetical player, and hence we have $1-BR'_{23}BR'_{32}$ in the denominator.

4.2 Double Misspecification

In many economic situations of our interest, a biased player naively thinks that her own view about the world is absolutely correct and the opponent share the same view about the world. This happens, for example, when a player is completely positive about own capability.[22] Alternatively, a player may be unaware of or

function depending on the parameter x_1), we indeed have $\frac{\partial x_2}{\partial x_1} = \frac{BR'_{21}+BR'_{23}BR'_{31}}{1-BR'_{23}BR'_{32}}$.

[21] In this argument, we implicitly use the fact that player 1's optimal action is not affected by the hypothetical player's action.

[22] For example, Camerer and Lovallo (1999) provide lab-experimental evidence of overconfidence in strategic entry settings, and Benoît, Dubra, and Moore (2015) provide lab-experimental evidence of overconfidence which excludes the possibility of rational Bayesian reasoning. See Kőszegi (2014) and Grubb (2015) for theoretical studies applying overconfidence in strategic settings.

underestimate a particular aspect of the environment.[23]

In our framework, such a situation can be described by considering a *doubly misspecified* agent, who has both the first-order misspecification (she incorrectly believes that the true parameter is $A \neq a$) and the second-order misspecification (she incorrectly believes that the opponent thinks that the true parameter is A). We first consider the case in which only player 2 is misspecified. Then we consider the case in which both players are misspecified.

4.2.1 One-Sided Double Misspecification

We consider the case in which only player 2 is misspecified. Specifically, we assume that:

- Player 2 (incorrectly) believes that for each parameter θ, the signal y is given by $y = Q(x_1, x_2, A, \theta) + \varepsilon$, where $A \neq a$.
- Player 2 (incorrectly) believes that it is common knowledge that "the signal y is given by $y = Q(x_1, x_2, A, \theta) + \varepsilon$."
- Player 1 knows player 2's information structure above.

With this information structure, player 2 has an incorrect view about the parameter a, and in addition, she has inferential naivety in that she incorrectly believes that player 1 takes an action based on a misspecified model. Player 1 is unbiased, in the sense that she correctly understands the true parameter a and she knows player 2's information structure (which allows her to make a correct prediction about player 2's action).

Assume again that the agents are myopic, so that they maximize the expected stage-game payoff each period. As in the case of second-order misspecification, we consider a hypothetical player 1 who thinks that it is common knowledge that

[23] For example, consumers often systematically ignore or underestimate a specific type of fee/risk; see Heidhues and Kőszegi (2018) for theoretical applications and evidence.

the true parameter is $A \neq a$. Let $x = (x_1, x_2, \hat{x}_1)$ denote an action profile in the three-player game, and let $\hat{U}_1(x, \theta, A)$ denote the hypothetical player's stage-game payoff, $U_2(x, \theta, A)$ denote player 2's stage-game payoff, and $U_1(x, \theta)$ denote player 1's stage-game payoff. Note that player 2 and the hypothetical player evaluates the expected payoff assuming that the signal is given by $y = Q(\hat{x}_1, x_2, A, \theta) + \varepsilon$. The equilibrium strategy in the infinite-horizon game is very similar to that in the case of the second-order misspecification; we only need to replace the parameter a which appears in player 2's expected payoff and Bayes' formula with the biased parameter A.

In this environment, the following conditions must be satisfied in a steady state:

$$x_1^* \in \arg\max_{x_1} U_1(x_1, x_2^*, a, \theta^*), \tag{21}$$

$$x_2^* \in \arg\max_{x_2} U_2(\hat{x}_1^*, x_2, A, \theta_2), \tag{22}$$

$$\hat{x}_1^* \in \arg\max_{\hat{x}_1} \hat{U}_1(\hat{x}_1, x_2^*, A, \theta_2), \tag{23}$$

$$\mu_1^* = 1_{\theta^*}, \tag{24}$$

$$\mu_2^* = \hat{\mu}_1^* = 1_{\theta_2} \text{ s.t. } Q(\hat{x}_1^*, x_2^*, A, \theta_2) = Q(x_1^*, x_2^*, a, \theta^*). \tag{25}$$

The first three conditions (21), (22), and (23) are incentive compatibility conditions, which require that each player maximizes her payoff given some beliefs. The next two conditions (24) and (25) assert that these beliefs satisfy consistency, in that each player's belief is concentrated on a state under which each player's subjective signal distribution coincides with the objective distribution. Note that the hypothetical player's belief is exactly the same as player 2's belief, as they both believe that it is common knowledge that the true parameter is A. We assume that for each action profile x and parameter A, there is a unique state $\theta_2(x, A)$ which solves (25). Intuitively, this $\theta_2(x, A)$ is player 2's long-run belief when players choose the same action profile x every period. Player 1's long-run belief is $\theta_1(x, A) = \theta^*$.

Define the slope of player i's asymptotic best response curve with respect to player j's action as

$$BR'_{ij} := -\frac{M_{ij}}{M_{ii}}$$

where for each $i, j = 1,2,3$ (possibly $i = j$),

$$M_{ij} = \left.\frac{\partial^2 U_i(x,\theta)}{\partial x_i \partial x_j}\right|_{\theta=\theta_i(x,A)} + \left.\frac{\partial^2 U_i(x,\theta)}{\partial x_i \partial \theta}\right|_{\theta=\theta_i(x,A)} \frac{\partial \theta_i(x,A)}{\partial x_j}.$$

measures the impact of player j's action on player i's marginal utility in the long run. Here again, player 3 refers to the hypothetical player, and her action, belief, and utility are denoted by x_3, θ_3, and U_3 rather than $\hat{x}_1$, $\hat{\theta}_1$, and $\hat{U}_1$.

For each $i = 2,3$, let

$$M_{iA} := \left.\frac{\partial^2 U_i(x,\theta,A)}{\partial x_i \partial A}\right|_{\theta=\theta_i(x,A)} + \left.\frac{\partial^2 U_i(x,\theta,A)}{\partial x_i \partial \theta}\right|_{\theta=\theta_i(x,A)} \frac{\partial \theta_i(x,A)}{\partial A}$$

denote the impact of player i's first-order misspecification on her marginal utility. The following proposition characterizes how player 2's double misspecification influences the steady-state actions.

Definition 3. A steady state x^* is *regular* if the following conditions are satisfied in x^*: (i) the steady-state action x_i^* is uniquely optimal, (ii) x^* and $\theta_2(x^*,A) = \hat{\theta}_1(x^*,A)$ are interior points, (iii) $BR'_{23}BR'_{32} \neq 1$ and $BR'_{12}BR'_{21} + BR'_{23}BR'_{32} + BR'_{12}BR'_{23}BR'_{31} \neq 1$, and (iv) $M_{ii} < 0$ for each i.

Proposition 3 (Steady State under One-Sided Double Misspecification)**.** *Let x^* be a regular steady state for some parameter A^*. Then there is an open neighborhood of A^* such that for any value A in this neighborhood, there is a regular steady state x^* which is continuous with respect to A, and we have*

$$\frac{\partial x_2^*}{\partial A} = -\left(\frac{M_{2A}}{M_{22}} + \frac{M_{3A}}{M_{33}} BR'_{23}\right)\left(\frac{1}{1 - BR'_{23}BR'_{32}}\right)\left(\frac{1}{1 - BR'_{12}NE'_2}\right),$$

$$\frac{\partial x_1^*}{\partial A} = \frac{\partial x_2^*}{\partial A} \cdot BR'_{12}$$

where NE'_2 is defined by (20).

The first equation in this proposition characterizes how player 2's double misspecification influences her own steady-state action. This is very similar to the first equation in Proposition 2. The only difference is that here player 2 has first-order misspecification about the parameter a, which influences her optimal action by the base misspecification effect $-\frac{M_{2A}}{M_{22}}$. All other terms are the same as those in Proposition 2. The second equation in the proposition simply states that the rational player 1 best-responds to player 2's action.

4.2.2 Two-Sided Double Misspecification

Now consider the case in which both players are misspecified. We assume:

- Each player i (incorrectly) believes that for each parameter θ, the signal y is given by $y = Q(x_1, x_2, A_i, \theta) + \varepsilon$, where $A_i \neq a$.

- Each player i (incorrectly) believes that it is common knowledge that "the signal y is given by $y = Q(x_1, x_2, A_i, \theta) + \varepsilon$."

We allow $A_1 \neq A_2$, so the dufferent players may have different levels of misspecification. Note that even when $A_i = a$, player i may be biased, in the sense that she may not know the opponent's bias about the technology a. That is, when $A_i = a$, player i believes that the opponent believes that the technology is a, but in reality, the opponent may believe $A_j \neq a$.

For find an equilibrium in this environment, it is useful to consider *two* hypothetical players. Hypothetical player 1 is player 1 who thinks that it is common knowledge that the true technology is A_2. Hypothetical player 2 is player 2 who thinks that it is common knowledge that the true technology is A_1. Note that player 1 and hypothetical player 2 think that they play the game each other, and the same is for player 2 and hypothetical player 1. Let $\hat{x}_i$, $\hat{\mu}_i$, and $\hat{U}_i$ denote hypothetical player i's action and belief.

Then a steady state in this environment satisfies the consistency and incentive-compatibility conditions:

$$x_i^* \in \arg\max_{x_i} U_i(x_i, \hat{x}_j^*, A_i, \theta_i) \quad \forall i, \;\; j \neq i \tag{26}$$

$$\hat{x}_i^* \in \arg\max_{\hat{x}_i} \hat{U}_i(\hat{x}_i, x_j^*, A_i, \theta_j) \quad \forall i, \;\; j \neq i, \tag{27}$$

$$\mu_1^* = \hat{\mu}_2^* = 1_{\theta_1} \;\text{ s.t. }\; Q(x_1^*, \hat{x}_2^*, A_1, \theta_1) = Q(x_1^*, x_2^*, a, \theta^*) \tag{28}$$

$$\mu_2^* = \hat{\mu}_1^* = 1_{\theta_2} \;\text{ s.t. }\; Q(\hat{x}_1^*, x_2^*, A_2, \theta_2) = Q(x_1^*, x_2^*, a, \theta^*). \tag{29}$$

We assume that for each action profile x and parameter A_i, there is a unique state $\theta_i(x, A_i) = \hat{\theta}_j(x, A_i)$ which solves the consistency condition $Q(x_i, \hat{x}_{-i}, A_i, \theta_i) = Q(x_1, x_2, a, \theta^*)$. This is player i's (and the hypothetical player j's) long-run belief when the same action profile x is chosen every period.

Define the slope of player i's asymptotic best response curve with respect to player j's action as

$$BR'_{ij} := -\frac{M_{ij}}{M_{ii}}$$

where for each $i, j = 1, 2, 3, 4$ (possibly $i = j$),

$$M_{ij} = \left.\frac{\partial^2 U_i(x, \theta)}{\partial x_i \partial x_j}\right|_{\theta = \theta_i(x, A)} + \left.\frac{\partial^2 U_i(x, \theta)}{\partial x_i \partial \theta}\right|_{\theta = \theta_i(x, A)} \frac{\partial \theta_i(x, A)}{\partial x_j}.$$

measures the impact of player j's action on player i's marginal utility in the long run. Here, players 3 and 4 refer to the hypothetical players 1 and 2, respectively. Their actions, beliefs, and utilities are denoted by x_3, x_4, θ_3, θ_4, U_3, and U_4, rather than $\hat{x}_1$, $\hat{x}_2$, $\hat{\theta}_1$, $\hat{\theta}_2$, $\hat{U}_1$, and $\hat{U}_2$. Note that $M_{13} = M_{24} = M_{34} = M_{43} = 0$, and thus $BR'_{13} = BR'_{24} = BR'_{34} = BR'_{43} = 0$. All other M_{ij} involve indirect effects.

For each $i = 2, 3$, let

$$M_{iA} := \left.\frac{\partial^2 U_i(x, \theta, A)}{\partial x_i \partial A}\right|_{\theta = \theta_i(x, A)} + \left.\frac{\partial^2 U_i(x, \theta, A)}{\partial x_i \partial \theta}\right|_{\theta = \theta_i(x, A)} \frac{\partial \theta_i(x, A)}{\partial A}$$

denote the impact of player i's first-order misspecification on her marginal utility. The following proposition characterizes how player 2's double misspecification influences the steady-state actions.

Definition 4. A steady state x^* is *regular* if the following conditions are satisfied in x^*: (i) the steady-state action x_i^* is uniquely optimal, (ii) x^* and $\theta_2(x^*,A) = \hat{\theta}_1(x^*,A)$ are interior points, (iii) $BR'_{14}BR'_{41} \neq 1$, $BR'_{23}BR'_{32} \neq 1$, and $BR'_{14}BR'_{41} + BR'_{23}BR'_{32} + (BR'_{21} + BR'_{23}BR'_{31})(BR'_{12} + BR'_{14}BR'_{42}) \neq 1$, and (iv) $M_{ii} < 0$ for each i.

Proposition 4 (Steady State under Two-Sided Double Misspecification)**.** *Let x^* be a regular steady state for some parameter $A^* = (A_1^*, A_2^*)$. Then there is an open neighborhood of A_2^* such that for any value A_2 in this neighborhood, there is a regular steady state x^* which is continuous with respect to A_2, and we have*

$$\frac{\partial x_2^*}{\partial A_2} = -\left(\frac{M_{2A}}{M_{22}} + BR'_{23}\frac{M_{3A}}{M_{33}}\right)\left(\frac{1}{1 - BR'_{23}BR'_{32}}\right)\left(\frac{1}{1 - NE'_1 NE'_2}\right),$$

$$\frac{\partial x_1^*}{\partial A_2} = \frac{\partial x_2^*}{\partial A_2} \cdot NE'_1$$

where NE'_2 is defined by (20) and NE'_1 is similarly defined as

$$NE'_1 = \frac{BR'_{12} + BR'_{14}BR'_{42}}{1 - BR'_{14}BR'_{41}}.$$

The equations in this proposition are very similar to those in Proposition 3, and the only difference is that the term BR'_{12} in Proposition 3 is replaced with NE'_1. Recall that under one-sided double misspecification, player 1 is rational and best responds to player 2's play. The term BR'_{12} in Proposition 3 quantifies this effect. On the other hand, under two-sided double misspecification, player 1 is also misspecified, and when player 2's action changes, it influences player 1's steady-state action in a more complicated way. The term NE'_1 quantifies this effect. More precisely, NE'_1 can be regarded as the slope of *player* 1*'s asymptotic Nash equilibrium correspondence* defined as

$$NE_1(x_2) = \{x_1 | \exists \hat{x}_2 \text{ satisfying (26) for } i = 1, \text{ (27) for } i = 2, (28)\}.$$

That is, NE'_1 considers the case in which player 2 chooses the same action x_2 every period, and measures how a marginal change in x_2 influences player 1's steady-state action. Just like NE'_2, this NE'_1 involves only the indirect effect, because player 1 does not think that player 2 is the opponent.

4.3 Small Misspecification: Limit as $A \to a$

So far we have characterized how a player's bias influences actions for a general misspecified parameter A. In this subsection, we will consider the case in which player 2 has small misspecification, in the sense that the misspecified parameter A is close to the true value a. We will show that in this special case, the multipliers derived in Propositions 1-4 can be replaced with much simpler forms, which allows us to make a clean comparison of the impacts of first-order, second-order, and double misspecification.

Let x^{first}, x^{second}, x^{double}, and x^{correct} denote the steady-state action profile with first-order misspecification, second-order misspecification, one-sided double misspecification, and the correctly specified model, respectively. Also, let x^{double2} denote the steady-state action profile with two-sided double misspecification when $A_1 = a$ and $A_2 = A$; i.e., player 1 knows that the true technology is a, but does not recognize player 2's bias A. A game is *symmetric* if $X_1 = X_2$, $u_1(x_1, y) = u_2(x_2, y)$ for all x_1 and x_2 with $x_1 = x_2$, and $Q(x_1, x_2, a, \theta) = Q(x_2, x_1, a, \theta)$.

Proposition 5. *Suppose that in each misspecification, there is a unique steady state at $A = a$ and it is regular. Then we have the following result:*

(i) At $A = a$, $x_1^{first} = x_1^{second} = \hat{x}_1^{second} = x_1^{double} = \hat{x}_1^{double} = x_1^{double2} = \hat{x}_1^{double2} = x_1^{correct}$ and $x_2^{first} = x_2^{second} = x_2^{double} = x_2^{double2} = \hat{x}_2^{double2} = x_2^{correct}$.

(ii) At $A = a$,

$$\frac{\partial x_2^{double}}{\partial A} = \frac{\partial x_2^{first}}{\partial A} + \frac{\partial x_2^{second}}{\partial A}.$$

(iii) Assume in addition that the game is symmetric and that $x_1^{correct} = x_2^{correct}$.

Then at $A = a$,

$$\frac{\partial x_2^{second}}{\partial A} = -\frac{U_{ij} - L}{U_{ii} - L} \cdot \frac{\partial x_2^{first}}{\partial A},$$
$$\frac{\partial x_2^{double}}{\partial A} = \frac{\partial x_2^{first}}{\partial A} + \frac{\partial x_2^{second}}{\partial A} = \left(1 - \frac{U_{ij} - L}{U_{ii} - L}\right) \frac{\partial x_2^{first}}{\partial A},$$
$$\frac{\partial x_2^{double2}}{\partial A} = \frac{(U_{ii} - L)(U_{ii} + U_{ij} - L)}{U_{ii}(U_{ii} + U_{ij} - 2L)} \frac{\partial x_2^{double}}{\partial A},$$
$$\frac{\partial x_1^{double2}}{\partial A} = -\frac{L}{U_{ii}} \left(1 - \frac{U_{ij}}{U_{ii}}\right) \frac{\partial x_2^{double2}}{\partial A},$$

where $U_{ii} = \frac{\partial^2 U_1(x^{correct}, \theta^*)}{\partial x_1^2}$, $U_{ij} = \frac{\partial^2 U_1(x^{correct}, \theta^*)}{\partial x_1 \partial x_2}$, *and* $L = \frac{\partial \hat{\theta}_1}{\partial x_1} \frac{\hat{U}_1}{\partial \hat{x}_1 \partial \theta} = \frac{Q_{x_1}}{Q_\theta} \cdot \frac{\partial^2 U_1(x^{correct}, \theta)}{\partial x_1 \partial \theta}$.

Part (i) shows that if there is no misspecification in that $A = a$, the steady state actions are the same for all cases. Part (ii) shows that when misspecification is small, the impact of one-sided double misspecification can be decomposed into that of first-order misspecification and that of second-order misspecification. This result essentially follows from the fact that double misspecification is a combination of first-order misspecification and second-order misspecification.

Part (iii) considers a symmetric game, and compares the impact of various misspecification, and all these equations follow from simple algebra. The term L measures the indirect learning effect on the hypothetical player's marginal utility, when the real player (secretly) increases the action.

4.4 Higher-Order Misspecification: Applications

To highlight the effects of higher-order misspecification, this subsection investigates applications to Cournot duopoly and team production, with higher-order misspecification. We also discuss an application to gender bias, with a focus on how a teacher's bias can be endogenously transmitted to a student under two-sided double misspecification.

4.4.1 Cournot duopoly

Consider the Cournot duopoly model introduced in Section 3.2.1. Let x_i^{correct} and π^{correct} denote firm i's steady-state action and payoff in the correctly specified model. The following corollary is an immediate consequence of Proposition 5. It shows that the long-run impact of misspecification crucially depends on the parameter $\frac{Q_{x\theta}}{Q_\theta}$, which measures the normalized impact of the state θ on the demand slope. We assume that steady state is unique in each misspecification at $a = A_1 = A_2$.[24]

Corollary 2. *Suppose that in each misspecification, there is a unique steady state at $A = a$ and it is regular.[25] Suppose also that the base misspecification effect (10) is positive at $A = a$, i.e., $\frac{Q_{x\theta}}{Q_\theta} < \frac{Q_{xA}}{Q_A}$ in the unique Nash equilibrium in the correctly specified model with known θ^*. Then, the following results hold.*

(i) Under first-order misspecification and one-sided double misspecification, we have $\frac{\partial x_1}{\partial A} < 0$, $\frac{\partial x_2}{\partial A} > 0$, $\frac{\partial(x_1+x_2)}{\partial A} > 0$, $\frac{\partial \pi_1}{\partial A} < 0$, $\frac{\partial \pi_2}{\partial A} > 0$, and $\frac{\partial(\pi_1+\pi_2)}{\partial A} < 0$ at $A = a$ both in the one-shot game with known θ^* and in the long-run steady state.

(ii) Under second-order misspecification, all inequalities in (i) are reversed in the one-shot game with known θ^*. In the long-run steady state, all inequalities in (i) remain true if $\frac{Q_{x\theta}}{Q_\theta} > \frac{Q_{xx}}{Q_x}$, but are reversed if $\frac{Q_{x\theta}}{Q_\theta} < \frac{Q_{xx}}{Q_x}$.

(iii) Under two-sided double misspecification with $A_1 = a$, there is $\overline{A}_2 > A_2$ such that for all $A_2 \in (a, \overline{A}_2)$, we have $x_1 = x_1^{correct}$, $x_2 > x_2^{correct}$, and $\pi_i < \pi_i^{correct}$ for each i in the one-shot game with known θ^*. In the long-run steady state, all inequalities in (i) remain true.

---

[24] In Appendix B, we provide a sufficient condition for the unique steady state in each misspecification.

[25] Regularity in the case of second-order misspecification requires $M_{33} = U_{ii} - L < 0$, which is satisfied if and only if $\frac{Q_{x\theta}}{Q_\theta} < \frac{Q_{xx}}{Q_x} + \frac{1}{x} - \frac{c''}{xQ_x}$. Other conditions stated in the definition of regularity are satisfied for generic parameters.

If the base misspecification effect is negative (i.e., $\frac{Q_{x\theta}}{Q_\theta} > \frac{Q_{xA}}{Q_A}$), all the inequalities remain true in the one-shot game with known θ^*, but are reversed in the long-run steady state.

Part (ii) of the corollary considers the case with second-order misspecification, in the sense that firm 2 incorrectly believes that firm 1 is overconfident. If the base misspecification effect is positive, in the long run, firm 2 incorrectly believes that firm 1 produces more than in the correctly specified model, i.e., $\frac{\partial \hat{x}_1}{\partial A} > 0$.[26] This incorrect prediction has two countervailing forces on firm 2's action. First, there is a direct effect, in that firm 2 best-responds to this incorrect prediction and produces less. Second, firm 2 observes prices higher than the anticipation and becomes optimistic about the state θ; this indirect learning effect increases the incentive to produce. More formally, simple algebra shows that

$$\frac{\partial x_2}{\partial A} = BR'_{23}\frac{\partial \hat{x}_1}{\partial A} = -\frac{U_{ij} - L}{U_{ii}}\frac{\partial \hat{x}_1}{\partial A}.$$

In this Cournot model, $U_{ij} = Q_x + x_2 Q_{xx}$ and $L = \frac{Q_x}{Q_\theta}(Q_\theta + x_i Q_{x\theta})$. Plugging this into the above equation,

$$\frac{\partial x_2}{\partial A} = -\frac{1}{U_{ii}}\left\{ \underbrace{\overbrace{Q_x}^{\text{direct effect}} - \overbrace{\frac{Q_x}{Q_\theta}Q_\theta}^{\text{indirect effect}}}_{\text{on the price level}} + x_2 \left(\underbrace{\overbrace{Q_{xx}}^{\text{direct effect}} - \overbrace{\frac{Q_x}{Q_\theta}Q_{x\theta}}^{\text{indirect effect}}}_{\text{on the slope}} \right) \right\} \frac{\partial \hat{x}_1}{\partial A}.$$

Note that the first two terms in the curly brackets cancel out; this happens because firm 2 correctly predicts the price in the steady state. Hence, we have

$$\frac{\partial x_2}{\partial A} = -\frac{x_2}{U_{ii}}\left(Q_{xx} - \frac{Q_x}{Q_\theta}Q_{x\theta}\right)\frac{\partial \hat{x}_1}{\partial A}.$$

That is, firm 2's steady-state action is determined by its subjective view about the slope in the steady state. As stated in the corollary, if $Q_{xx} - \frac{Q_x}{Q_\theta}Q_{x_i\theta} > 0$ so that firm 2 is optimistic about the slope, then it produces more than in the correctly

[26] Formally, this follows from Proposition 5 (iii), $U_{ii} < 0$, $U_{ii} - L < 0$, and the base misspecification effect being positive.

specified model, i.e., $\frac{\partial x_2}{\partial A} > 0$. On the other hand, firm 2 produces less if it is pessimistic about the slope.

Part (i) of the corollary considers the case with one-sided double misspecification, and shows that its impact on the steady state is similar to that of first-order mispecification. That is, when firm 2 is overconfident and the opponent knows it, firm 2's higher-order misspecification does not have a substantial impact on the steady state. To see the intuition, suppose first that $Q_{xx} - \frac{Q_x}{Q_\theta} Q_{x_i\theta} > 0$. In this case, both first-order misspecification and second-order misspecification increase firm 2's production. Since double-misspecification is a combination of these two misspecifications (formally, see Proposition 5(ii)), it is immediate that firm 2 produces even more under one-sided double misspecification. Suppose next that $Q_{xx} - \frac{Q_x}{Q_\theta} Q_{x_i\theta} < 0$. In this case, second-order misspecification has an opposite impact on firm 2's production, that is, firm 2 produces less under second-order misspecification. However, as we show in the proof, this negative impact (of second-order misspecification) is outweighed by the positive impact of first-order misspecification. So double misspecification increases firm 2's production, although its impact is smaller than that of first-order misspecification.

Part (iii) of the corollary considers the case with two-sided double misspecification, where firm 1 is not aware of firm 2's overconfidence. In this case, firm 2's overconfidence hurts *both* firms in the short-run. This is because the strategic effect does not exist in this environment; although the overconfident firm 2 produces more, firm 1 is not aware of it and chooses the Nash equilibrium action in the correctly-specified model. So firm 2's overconfidence simply makes her action suboptimal, just as any bias does in the single-agent problem.

However, in the long run, firm 2's overconfidence can still improve the equilibrium payoff. A key is the learning effect on firm 1: Since firm 1 is not aware of overproduction by the opponent, it observes a market price lower than the anticipation. Accordingly, firm 1 becomes pessimistic about θ, and in particular about the demand slope. This means that firm 1 produces less in the long run, which yields a higher profit to the overconfident firm 2 *even in the absence of the strategic effect*. This is an important difference from the case with first-order misspecification or one-sided double misspecification, in which the overconfident firm benefits from the strategic effect. Also, under two-sided double misspecification, firm 2's overconfidence may improves the total welfare $\pi_1 + \pi_2$ if this learning effect is large enough. As stated in parts (i) and (ii) of Corollary 2, this cannot happen in first-order misspecification and one-sided double misspecification, in which there is no learning effect on firm 1. These results highlight how the awareness about an opponent's bias makes qualitative differences in multi-player learning models.

4.4.2 Team production

In Section 3.2.2, we have focused on team production first-order misspecification. However, various forms of higher-order misspecifications are relevant in practice. For example, a player may have second-order misspecification and incorrectly believes that the opponent has overconfidence, underconfidence, or prejudice. Or, a player may have double misspecification; she has overconfidence, underconfidence, or prejudice as in the case of first-order misspecification, and on top of that, she naively thinks that the opponent shares the same view. Using Proposition 5, we can quantify the impact of such misspecification. Again, we assume that steady state is unique in each misspecification at $a = A_1 = A_2$.

Corollary 3. *Suppose that in each misspecification, there is a unique steady state at $A = a$ and it is regular. Let $L = \frac{Q_{x_i} Q_{x_j \theta}}{Q_\theta} > 0$ be as in Proposition 5.*

(i) Under first-order misspecification and one-sided double misspecification,

(a) If $Q_{x_1 x_2} > 0$, then in the one-shot game with known θ^*, $\frac{\partial x_2}{\partial A} \leq \frac{\partial x_1}{\partial A} \leq 0$ and $\frac{\partial \pi_1}{\partial A} \leq \frac{\partial \pi_2}{\partial A} \leq 0$ at $A = a$ with strict inequalities if $Q_{x_i a} \neq 0$. In the long-run steady state, all the inequalities are strict for any $Q_{x_i a} \leq 0$.

(b) If $Q_{x_1 x_2} < 0$, then in the one-shot game with known θ^*, $\frac{\partial x_2}{\partial A} \leq 0 \leq \frac{\partial x_1}{\partial A}$, $\frac{\partial (x_1 + x_2)}{\partial A} \leq 0$, $\frac{\partial \pi_1}{\partial A} \leq 0 \leq \frac{\partial \pi_2}{\partial A}$, and $\frac{\partial (\pi_1 + \pi_2)}{\partial A} \leq 0$ at $A = a$ with strict inequalities if $Q_{x_i a} \neq 0$. In the long-run steady state, all the inequalities are strict for any $Q_{x_i a} \leq 0$.

(ii) Under second-order misspecification,

(a) If $Q_{x_1x_2} > L$, then in the one-shot game with known θ^*, $\frac{\partial x_2}{\partial A} \leq \frac{\partial x_1}{\partial A} \leq 0$ and $\frac{\partial \pi_1}{\partial A} \leq \frac{\partial \pi_2}{\partial A} \leq 0$ at $A = a$ with strict inequalities if $Q_{x_ia} \neq 0$. In the long-run steady state, all the inequalities are strict for any $Q_{x_ia} \leq 0$.

(b) If $Q_{x_1x_2} \in (0, L)$, then in the one-shot game with known θ^*, $\frac{\partial x_2}{\partial A} \leq \frac{\partial x_1}{\partial A} \leq 0$ and $\frac{\partial \pi_1}{\partial A} \leq \frac{\partial \pi_2}{\partial A} \leq 0$ at $A = a$ with strict inequalities if $Q_{x_ia} \neq 0$. In the long-run steady state, $\frac{\partial x_2}{\partial A} > \frac{\partial x_1}{\partial A} > 0$ and $\frac{\partial \pi_1}{\partial A} > \frac{\partial \pi_2}{\partial A} > 0$ at $A = a$.

(c) If $Q_{x_1x_2} < 0$, then in the one-shot game with known θ^*, $\frac{\partial x_1}{\partial A} \leq 0 \leq \frac{\partial x_2}{\partial A}$, $\frac{\partial (x_1+x_2)}{\partial A} \geq 0$, $\frac{\partial \pi_2}{\partial A} \leq 0 \leq \frac{\partial \pi_1}{\partial A}$, and $\frac{\partial (\pi_1+\pi_2)}{\partial A} \geq 0$ at $A = a$ with strict inequalities if $Q_{x_ia} \neq 0$. In the long-run steady state, all the inequalities are strict for any $Q_{x_ia} \leq 0$.

(iii) Under two-sided double misspecification with $A_1 = a$, in the one-shot game with known θ^*, there is $\overline{A}_2 > a$ such that for any $A_2 \in (a, \overline{A}_2)$, $x_1 = x_1^{correct}$, $x_2 \leq x_2^{correct}$, $\pi_1 \leq \pi_2 \leq \pi_i^{correct}$ with strict inequalities if $Q_{x_ia} \neq 0$. In the long-run steady state, $\frac{\partial x_i}{\partial A} < 0$ and $\frac{\partial \pi_i}{\partial A} < 0$ at $A = a$.

Part (i) considers the case in which player 2 is overconfident and player 1 knows it. It shows that the overconfident player reduces the effort, regardless of whether she knows the opponent's information or not. This result follows from the fact that the base misspecification effect is negative. It also shows that player 1 simply best-responds to this player 2's action; she increases the effort under strategic substitutes, and reduces the effort under strategic complements. Overconfidence always decreases the total surplus, and the overconfident player 2 always obtains a better payoff than the unbiased player 1.

Part (ii) considers the case in which both players know the true parameter a, but player 2 incorrectly believes that the opponent is overconfident. The result depends on the parameter $Q_{x_1x_2}$, and notably, when there is weak complementarity, i.e., $Q_{x_1x_2} \in (0, L)$, both players work less in the short run, but work harder in the long run. The intuition is as follows. Due to the bias, player 2 underestimates player 1's effort. In the one-shot game, player 2 simply best-responds to it by reducing the effort. But in the long run, there is a learning effect; player 2 is surprised by the actual output being better than the anticipation, and over time, she becomes optimistic about θ and provides an extra incentive to work hard. When complementarity is weak, this effect dominates the negative impact in the one-shot game, so player 2 increases the effort.

Part (iii) considers the case in which player 1 is not aware of player 2's overconfidence. An important difference from part (i) is that player 1 reduces the effort in the long run, regardless of the parameter $Q_{x_1 x_2}$. This is due to the learning effect. Since player 1 overestimates player 2's effort, she is disappointed by low output. Hence, over time, she becomes pessimistic about θ and reduces the effort. In this case, both players reduce the effort, and both players earn a lower long-run payoff.

4.4.3 Transmission of Bias

There is a long-standing debate on whether the gender gap in math achievement arises from biological reasons (such as brain functioning) as opposed to culture and social conditioning. There are recent works which support the latter: For example, Carlana (2019) finds that the gender gap in performance in math exam substantially increases when students are assigned to math teachers with stronger gender stereotypes. She argues that this effect is at least partially driven by lower self-confidence on math ability of girls exposed to gender-biased teachers. We will show that our framework is useful to explain such a *bias transmission* from teachers to students.

Consider a variant of our team-production model in two-sided double misspecification. Suppose that player 1 is a student and player 2 is a teacher. The student's achievement (e.g., math test performance) is given by $y = a(x_1 + x_2 + b) + \varepsilon$, where a represents a gender-specific capability and b is the student's capability. The student (player 1) knows her own capability b, but does not know the gender-specific capability. So she thinks that the outcome is given by $y = \theta_1(x_1 + x_2 + b) + \varepsilon$ and learns θ_1 over time. On the other hand, the teacher has a biased view about the gender specific capability, and he thinks that the outcome is given by $y =$

$A_2(x_1+x_2+\theta_2)+\varepsilon$, where $A_2 < a$ represents his bias. He does not know the student's individual capability θ_2, and learns it over time. We assume that each player (incorrectly) thinks that the opponent has the same view about the world, This means that the student is not aware of the teacher's gender-stereotype $A_2 < a$.

This setup is different from the one presented in Section 4.2.2, in that different players learn different parameters. However, this does not has a substantial impact on the property of the steady state. Indeed, it is straightforward to show that the result similar to Corollary 3 still holds, and in the steady state, *both* the teacher and the student work less than in the correctly specified model.

The intuition behind this result is as follows. Since the teacher has a gender-stereotype $A_2 < a$, each period, he exerts less effort than in the correctly specified model. Because the student is not aware of it, from the student's point of view, the realized outcomes are systematically lower than the anticipation. Accordingly, after a long time, she becomes unrealistically pessimistic about the gender-specific capability θ_1, which weakens the incentive to work. So in this framework, even though the student initially has an unbiased view about the environment, the teacher's gender bias is eventually transmitted to the student, due to the second-order misspecification (i.e., the student's unawareness of the teacher's bias).

Recent work by Heidhues, Kőszegi, and Strack (2020b) also argue that a gender bias (more generally, a group discrimination) can endogenously arise as a consequence of misspecified learning. Formally, they develop a single-agent learning model, and show that an underconfident (resp. overconfident) agent tends to underestimate (resp. overestimate) the capability of her in-group members. So in their setup, the source of a group discrimination is one's misconfidence about her own capability. Our result complements their work by considering the case in which an agent (a student in our context) does not have underconfidence, or more generally, any bias about the physical environment. We find that a group discrimination can still arise due to a bias transmission; one's existing prejudice may induce *other players'* negative self-stereotypes through learning.

5 Theorems for Convergence

This section generalizes Esponda, Pouzo, and Yamamoto (2019) by allowing multiple players and continuous actions. We show that the asymptotic motion of players' action frequency is approximated by a differential inclusion, and we use it to derive a sufficient condition for convergence.

5.1 General Setup

For each compact set $A \subset \boldsymbol{R}^n$ (or more generally, separable metric space A), let $\triangle A$ denote the set of probability measures over the set A. We consider the *dual bounded-Lipschitz norm* on $\triangle A$, that is, for each $\mu \in \triangle A$, let

$$\|\mu\| = \sup_{f \in BL(A)} \int_A f d\mu$$

where $BL(A)$ is the set of bounded Lipschitz continuous functions f on A with $\sup_{x \in A} |f(x)| + \sup_{x \neq y} \frac{|f(x)-f(y)|}{|x-y|} \leq 1$. This norm has two nice properties. First, it metrizes the weak topology, that is, the topology induced by the dual bounded-Lipschitz norm coincides with the weak topology on $\triangle A$. Second, with this norm, $\triangle A$ is a compact subset of a Banach space, i.e., the set of finite signed measures on A is a Banach space when paired with the dual bounded-Lipschitz norm, and $\triangle A$ is a compact subset in it. See Dudley (1966) and Billingsley (1999) for references. The first property is needed to obtain our Propisition 8. The second property is crucial in order to use a stochastic approximation technique in the proof of Proposition 9. The dual bounded-Lipschitz norm is used in Hofbauer, Oechssler, and Riedel (2009) and Perkins and Leslie (2014), who study learning dynamics in games with continuous actions.

5.1.1 Objective World

There are two players $i = 1, 2$ and infinitely many periods $t = 1, 2, \cdots$. In each period t, each player i chooses an action x_i from a compact set $X_i \subset \boldsymbol{R}$. These actions are not observable. Then they observe a noisy public output $y \in Y$ which is distributed according to a probability measure $Q(\cdot|x) \in \triangle Y$, where $x = (x_1, x_2)$ denotes the chosen action profile. Each player i's payoff is $u_i(x_i, y)$.

In the infinite-horizon game, each player i's t-period history is $h_i^t = (x_i^\tau, y^\tau)_{\tau=1}^t$, where (x_i^t, y^t) is player i's action and the public outcome in period t. Let H_i^t denote the set of all t-period history, and let $H_i^0 = \{\emptyset\}$. Player i's *pure strategy* in the infinite-horizon game is a mapping $s_i : \bigcup_{t=0}^\infty H_i^t \to X_i$. Let S_i denote the set of player i's pure strategies. Let $h_Y^t = (y^\tau)_{\tau=1}^t$ denote the t-period public history. A strategy is *public* if it depends only on public histories.

5.1.2 Subjective World and Model Hierarchy

We assume that the output distribution Q is not common knowledge among players. Instead, each player i has a set $\Theta_{i,1}$ of subjective models, and in each model $\theta_{i,1} \in \Theta_{i,1}$, the output distribution given an action profile x is $Q_{\theta_{i,1}}(\cdot|x)$. Player i thinks that the true world is described by one of these models, and her initial prior about the model is $\mu_{i,1} \in \triangle\Theta_{i,1}$. Player i's models are *correctly specified* if there is $\theta_{i,1}$ such that $Q(\cdot|x) = Q_{\theta_{i,1}}(\cdot|x)$ for all x. Otherwise her models are *misspecified*. Player i also has models about the opponent j's model, that is, player i believes that the opponent j has an initial prior $\mu_{i,2}$ over a model set $\Theta_{i,2}$, where each model $\theta_{i,2}$ induces the output distribution $Q_{\theta_{i,2}}(\cdot|x)$ for each action profile x. This triplet $M_{i,2} = (\mu_{i,2}, \Theta_{i,2}, (Q_{\theta_{i,2}}(\cdot|x))_{(x,\theta_{i,2})})$ is player i's *second-order model* in that it is her subjective view about player j's subjective model. More generally, we assume that each player i has a *model hierarchy* $M_i = (M_{i,1}, M_{i,2}, \cdots)$ where each $M_{i,k} = (\mu_{i,k}, \Theta_{i,k}, (Q_{\theta_{i,k}}(\cdot|x))_{(x,\theta_{i,k})})$ is player i's *kth-order model*. That is, player i believes that player j believes that player i's model is $M_{i,3}$, player i believes that player j believes that player i believes that player j's model is $M_{i,4}$, and so on.

This framework is flexible and allows us to study a variety of information structures, including first-order misspecification, second-order misspecification, and double misspecification which we studied in Sections 3 and 4. In what follows, we will maintain the following technical assumptions.

Assumption 1. The following conditions hold:

(i) Y and Θ are Borel subsets of the Euclidean space, and Θ is compact.

(ii) There is a Borel probability measure $\nu \in \triangle Y$ such that $Q(\cdot|x)$ and $Q_{\theta_{i,k}}(\cdot|x)$ are absolutely continuous with respect to ν for all x and i, k, and $\theta_{i,k}$. (An implication is that there are densities $q(\cdot|x)$ and $q_{\theta_{i,k}}(\cdot|x)$ such that $\int_A q(y|x)\nu(dy) = Q(A|x)$ and $\int_A q_{\theta_{i,k}}(y|x)\nu(dy) = Q_{\theta_{i,k}}(A|x)$ for any $A \subseteq Y$ Borel.)

(iii) $q(\cdot|x)$ and $q_{\theta_{i,k}}(\cdot|x)$ are continuous in θ and x.

(iv) There is a function $g : X \times Y \to \boldsymbol{R}$ such that (a) for each y, $g(x,y)$ is continuous in x, (b) $g(x,\cdot) \in L^2(Y, Q(\cdot|x))$ for each x, and (c) for all x, $\hat{x}$ i, k, and $\theta_{i,k}$, $\log \frac{q(\cdot|x)}{q_{\theta_{i,k}}(\cdot|\hat{x})} \leq g(x,\cdot)$ $Q(\cdot|x)$-a.s..

The parts (i)-(iii) are fairly standard. The part (iv) implies that every outcome y is generated by each player i's model, which is useful to establish a uniform version of the law of large numbers. The assumption above is similar to Assumptions 1 and 2 of Esponda, Pouzo, and Yamamoto (2019), but there are two differences. First, we allow the action set X_i to be continuous, in which case we require continuity of q, as described in parts (iii) and (iv-a). Second, we allow inferential naivety, so when we consider the log-likelihood $\log \frac{q(\cdot|x)}{q_{\theta_{i,k}}(\cdot|\hat{x})}$ of the true output probability and the subjective probability, we distinguish the actual action profile x from the inferred action profile $\hat{x}$.

We also assume that each player i believes that the models become common knowledge at higher levels, in the following sense:

Assumption 2. Player i believes that the models (M_{i,k_i}, M_{i,k_i+1}) are common knowledge after level $k_i < \infty$, that is, for each i, there is $k_i < \infty$ such that $(M_{i,k_i}, M_{i,k_i+1}) = (M_{i,k_i+2n}, M_{i,k_i+1+2n})$ for each $n = 1, 2, \cdots$.

To interpret this assumption, suppose that $k_i = 1$, so that

$$M_{i,1} = M_{i,3} = M_{i,5} = \cdots, \text{ and } M_{i,2} = M_{i,4} = M_{i,6} = \cdots. \tag{30}$$

Then player i believes that...[27]

[27] Here we use the fact that if player i believes E, then she believes that she believes E.

- Her own model is $M_{i,1}$.
- She believes that her own model is $M_{i,1}$.
- The opponent j believes that i's model is $M_{i,3} = M_{i,1}$. And so on.

In this sense, player i believes that her model $M_{i,1}$ is common knowledge. Likewise, $M_{i,2} = M_{i,4} = M_{i,6} = \cdots$ implies that player i believes that the opponent j's model $M_{i,2}$ is common knowledge. So (30) indeed implies that player i believes that the models $(M_{i,1}, M_{i,2})$ are common knowledge. Note that this assumption is about whether *player i thinks that* the models are common knowledge, and *not* about whether the models are common knowledge in the objective sense When $k_i > 1$, the condition stated in the assumption above implies that player i believes that the opponent j believes that $\cdots$ that the models (M_{i,k_i}, M_{i,k_i+1}) are common knowledge. We believe that Assumption 2 is satisfied in most applications.[28]

Pick k_i as stated in Assumption 2. Then player i's problem is strategically equivalent to solving the following hypothetical game with $k_i + 1$ agents:

- Each period, each agent $k = 1, 2, \cdots, k_i + 1$ chooses an action $\hat{x}_{i,k}$ from a set $\hat{X}_{i,k}$, where $\hat{X}_{i,k} = X_i$ for odd k, and $\hat{X}_{i,k} = X_j$ for even k.
- Agent 1 is player i herself. She has the model $M_{i,1}$, and thinks that her opponent is agent 2. That is, she thinks that the distribution of the public outcome is $Q_{\theta_{i,1}}(\hat{x}_{i,1}, \hat{x}_{i,2})$ for some $\theta_{i,1}$, where $(\hat{x}_{i,1}, \hat{x}_{i,2})$ is the action chosen by agents 1 and 2.
- Other agents are hypothetical players which describe player i's reasoning. Each agent $k = 2, 3, \cdots, k_i - 1$ has the model $M_{i,k}$, and thinks that her opponent is agent $k + 1$. That is, she thinks that the distribution of the public outcome is $Q_{\theta_{i,k}}(\hat{x}_{i,k}, \hat{x}_{i,k+1})$ for some $\theta_{i,k}$

[28] This assumption is needed to establish Propositions 8 and 9. Indeed, if this assumption is not satisfied, then we need infinite agents to describe player i's reasoning, so the set $\hat{X}$ becomes the product of infinitely many X_1 and X_2. This set $\hat{X}$ is not separable (it is well-known that the l^∞-space is not separable), so the dual bounded-Lipschitz norm on $\triangle \hat{X}$ may not coincide with the topology of weak convergence.

- Agents k_i and k_i+1 has the models M_{i,k_i} and M_{i,k_i+1}, respectively, and they play the game with each other. That is, they think that the output distribution is given by $Q_{\theta_{i,k_i}}(\hat{x}_{i,k_i},\hat{x}_{i,k_i+1})$ for some θ_{i,k_i} and $Q_{\theta_{i,k_i+1}}(\hat{x}_{i,k_i},\hat{x}_{i,k_i+1})$ for some θ_{i,k_i+1}, respectively.

- All the information structure above is common knowledge among the agents.

Intuitively, agent 1's action $\hat{x}_{i,1}$ in this hypothetical game is player i's actual action, agent 2's action $\hat{x}_{i,2}$ is player i's prediction about the opponent j's action, agent 3's action $\hat{x}_{i,3}$ is player i's prediction about j's prediction about i's action, and so on. So the action profile $\hat{x}_i = (\hat{x}_{i,k})_{k=1}^{k_i+1}$ in this hypothetical game is essentially player i's *prediction hierarchy*. Let $\hat{X}_i = \times_{k=1}^{k_i+1} X_{i,k}$ denote the set of all these action profiles.

In what follows, each agent k in this hypothetical game is labelled as (i,k), because they are agents which describe player i's reasoning. The opponent j has a different model hierarchy $M_j \neq M_i$, and hence her reasoning is represented by a different set of agents (j,k).

Let $\hat{s}_{i,k}$ denote agent (i,k)'s strategy in the infinite-horizon hypothetical game, and let $\hat{s}_i = (\hat{s}_{i,k})_{k=1}^{k_i+1}$ denote a strategy profile. This profile $\hat{s}_i$ is also interpreted as player i's *prediction hierarchy* about strategies in the infinite-horizon game. That is, $\hat{s}_{i,1}$ is player i's actual strategy, $\hat{s}_{i,2}$ is player i's prediction about player j's strategy, and so on. So $\hat{s}_{i,k} \in S_i$ for odd k, and $\hat{s}_{i,k} \in S_j$ for even k. We assume that each $\hat{s}_{i,k}$ is pure and public.

Given a pure strategy profile $\hat{s}_i = (\hat{s}_{i,k})$ in the hypothetical game, each agent k's posterior belief $\hat{\mu}_{i,k}^{t+1} \in \triangle\Theta_{i,k}$ can be computed using Bayes' rule, after every public history h_Y^t. Formally, for each t and k, we have

$$\hat{\mu}_{i,k}^{t+1}(\theta_{i,k}) = \frac{\hat{\mu}_{i,k}^{t}(\theta_{i,k}) Q_{\theta_{i,k}}(y^t | \hat{s}_{i,k}(h_Y^{t-1}), \hat{s}_{i,k+1}(h_Y^{t-1}))}{\int_{\Theta_{i,k}} \hat{\mu}_{i,k}^{t}(\theta_{i,k}) Q_{\theta_{i,k}}(y^t | \hat{s}_{i,k}(h_Y^{t-1}), \hat{s}_{i,k+1}(h_Y^{t-1})) d\theta_{i,k}}$$

where $\hat{s}_{i,k_i+2} = \hat{s}_{i,k_i}$. Here we use the fact that agent k thinks that the signal y^t in period t is drawn given the action profile $(\hat{s}_{i,k}(h_Y^{t-1}), \hat{s}_{i,k+1}(h_Y^{t-1}))$, where $\hat{s}_{i,k}(h_Y^{t-1})$ is her own action, and $\hat{s}_{i,k+1}(h_Y^{t-1})$ is the opponent $k+1$'s action. The above formula is valid only if no one deviates from the profile $\hat{s}_i$; if some agent k deviates, then her posterior belief must be computed using a different formula. A strategy profile $\hat{s}_i$ is *Markov* if each agent's strategy depends only on the belief hierarchy $\hat{\mu}_i^t$, i.e., for each k and t, $\hat{s}_{i,k}(h_Y^t)$ depends on h_Y^t only through $\hat{\mu}_i^{t+1}$.

Example 1. (Myopically optimal agents) Suppose that the agents are myopic and maximize their expected stage-game payoffs each period. In such a case, they play a one-shot equilibrium given a belief-hierarchy $\hat{\mu}^t$ in each period t. Recall that each agent (i,k) thinks that her opponent is agent $(i,k+1)$, so her subjective expected stage-game payoff given a model $\theta_{i,k}$ is

$$U_{\theta_{i,k}}(\hat{x}_{i,k},\hat{x}_{i,k+1}) = \int_Y u_{i,k}(\hat{x}_{i,k},y)Q_{\theta_{i,k}}(dy|\hat{x}_{i,k},\hat{x}_{i,k+1})$$

where $u_{i,k} = u_1$ when $i+k$ is even, and $u_{i,k} = u_2$ when $i+k$ is odd. So the strategy profile $\hat{s}_i$ must satisfy the following equilibrium condition:

$$\hat{s}_{i,k}(\hat{\mu}_i) \in \arg\max_{\hat{x}_{i,k}\in\hat{X}_{i,k}} \int_{\Theta_{i,k}} U_{\theta_{i,k}}(\hat{x}_{i,k},\hat{s}_{i,k+1}(\hat{\mu}_i))\hat{\mu}_{i,k}(d\theta_{i,k}) \quad \forall k \forall \hat{\mu}_i. \qquad (31)$$

In words, the agents k_i and k_i+1 choose actions which are best response to each other, the agent k_i-1 best responds to the agent k_i's action, the agent k_i-2 best responds to that action, and so on. It is obvious that such a strategy profile $\hat{s}_i$ is Markov.

Example 2. (Dynamically optimal agents) Now consider dynamically optimal agents, who maximize the expectation of the discounted sum of the stage-game payoffs, $\sum_{t=1}^{\infty}\delta^{t-1}u_{i,k}(\hat{x}_{i,k},y)$. Many applied papers use Markov perfect equilibria as a solution concept. In our context, $\hat{s}_i$ is a Markov perfect equilibrium if given any belief hierarchy $\hat{\mu}_i$, the continuation strategy profile $\hat{s}_i|_{\hat{\mu}_i}$ satisfies

$$\hat{s}_{i,k}|_{\hat{\mu}_i} \in \arg\max_{\hat{s}_{i,k}} \int_{\Theta_{i,k}} \sum_{t=1}^{\infty} \delta^{t-1} E[U_{\theta_{i,k}}(\hat{x}_{i,k}^t, x_{i,k+1}^t)|\hat{s}_{i,k},\hat{s}_{i,k+1}|_{\hat{\mu}_i}]\hat{\mu}_{i,k}(d\theta_{i,k})$$

for each k, where the expectation is taken over $(\hat{x}_{i,k}^t, x_{i,k+1}^t)$. Note that this condition reduces to (31) when $\delta = 0$ so that the agents are myopic.

Let $h=(x^t,y^t)_{t=1}^{\infty}$ denote a sample path (a history in the infinite-horizon game). Also, let $\hat{X}=\hat{X}_1\times\hat{X}_2$ be the product of the sets of all action profiles of the two hypothetical games. Given a sample path h and given strategy profiles $\hat{s}=(\hat{s}_1,\hat{s}_2)$ of the two hypothetical games (for players 1 and 2), let $\sigma^t(h)\in\triangle\hat{X}$ denote the action frequency up to period t, that is,

$$\sigma^t(h)[(\hat{x}_1,\hat{x}_2)]=\frac{1}{t}\sum_{\tau=1}^{t}1_{\{\hat{s}_{i,k}(h_Y^{\tau-1})=\hat{x}_{i,k}\ \ \forall i\forall k\}}.$$

Intuitively, $\sigma^t(h)[(\hat{x}_1,\hat{x}_2)]$ describes how often the action profile $\hat{x}_i$ was chosen in each hypothetical game. (In other words, it describes how often each player i made a prediction hierarchy $\hat{x}_i$.) Note that we cannot directly observe the actions $\hat{x}_{i,k}$ of the higher-level agents (i,k) with $k\geq 2$, as they are hypothetical agents. However, since each agent uses a public strategy $\hat{s}_{i,k}$, we can back it up from the past public history; given a history $h_Y^{\tau-1}$, the hypothetical agent k's action in period τ must be $\hat{s}_{i,k}(h_Y^{\tau-1})$. This allows us to define the action frequency in the hypothetical game as a function of the observed history h.

5.2 Posterior Beliefs and Kullback-Leibler Divergence

We first show that after a long time t, the posterior belief is concentrated on the models which best explain the data. Specifically, we show that the belief is concentrated on the models which minimize the Kullback-Leibler divergence, which is defined as follows. Let $\sigma\in\triangle\hat{X}$ be a probability measure over $\hat{X}$. For each σ, the *Kullback-Leibler divergence* of model $\theta_{i,k}$ for agent k is defined as

$$K_{i,k}(\theta_{i,k},\sigma)=\int_{\hat{X}}\int_Y\log\frac{q(y|\hat{x}_{1,1},\hat{x}_{2,1})}{q_{\theta_{i,k}}(y|\hat{x}_{i,k},\hat{x}_{i,k+1})}Q(dy|\hat{x}_{1,1},\hat{x}_{2,1})\sigma(d\hat{x}).$$

Intuitively, $K_{i,k}(\theta_{i,k},\sigma)$ measures the distance between the true output distribution and the subjective distribution induced by agent k's model $\theta_{i,k}$. To see this, think about the special case in which σ is a degenerate distribution $1_{\hat{x}_1,\hat{x}_2}$. Then the Kullback-Leibler divergence of model $\theta_{i,k}$ can be rewritten as

$$\int_Y\log\frac{q(y|\hat{x}_{1,1},\hat{x}_{2,1})}{q_{\theta_{i,k}}(y|\hat{x}_{i,k},\hat{x}_{i,k+1})}Q(dy|\hat{x}_{1,1},\hat{x}_{2,1}).$$

This measures the distance between the true distribution $q(\cdot|\hat{x}_{1,1},\hat{x}_{2,1})$ and the subjective distribution $q_{\theta_{i,k}}(\cdot|\hat{x}_{i,k},\hat{x}_{i,k+1})$ induced by the model $\theta_{i,k}$. Indeed, this value is always non-negative, and equals zero if and only if the true and subjective distributions are the same. When σ is not a degenerate distribution, we take a weighted sum of the Kullback-Leibler divergence over $\hat{x}=(\hat{x}_1,\hat{x}_2)$, which leads to the definition of $K_{i,k}(\theta_{i,k},\sigma)$ above.

As is clear from this formula, agent k's subjective signal distribution $q_{\theta_{i,k}}(y|\hat{x}_{i,k},\hat{x}_{i,k+1})$ is potentially different from the true distribution $q(y|\hat{x}_{1,1},\hat{x}_{2,1})$ in two ways. First, agent k's model $\theta_{i,k}$ can be *misspecified* in that the distribution $q_{\theta_{i,k}}$ as a function of the chosen action can be different from the true distribution q. Second, agent k can have an *inferential naivety*. That is, while the true distribution is determined by the actual actions chosen by players 1 and 2 (which is denoted by $(\hat{x}_{1,1},\hat{x}_{2,1})$ in our setup), agent k thinks that the output distribution is determined by the actions chosen by agents k and $k+1$.

For each measure $\sigma\in\triangle\hat{X}$, let $\Theta_{i,k}(\sigma)$ denote the minimizers of the Kullback-Leibler divergence, that is,

$$\Theta_{i,k}(\sigma)=\arg\min_{\theta_{i,k}\in\Theta_{i,k}} K_{i,k}(\theta_{i,k},\sigma).$$

Intuitively, this is the set of models which best explains the data when the past action frequency was σ. The minimized Kullback-Leibler divergence is $K^*_{i,k}(\sigma)=\min_{\theta_{i,k}\in\Theta_{i,k}} K_{i,k}(\theta_{i,k},\sigma)$. We first show that these minimizers have useful properties:

Lemma 1. *For each i and k, (i) $K_{i,k}(\theta_{i,k},\sigma)-K^*_{i,k}(\sigma)$ is continuous in $(\theta_{i,k},\sigma)$, and (ii) $\Theta_{i,k}(\sigma)$ is upper hemi-continuous, non-empty, and compact-valued.*

The following proposition shows that after a long time t, the posterior is concentrated on the best models $\Theta_{i,k}(\sigma^t)$. This extends Theorem 1 of Esponda, Pouzo, and Yamamoto (2019) to the case with continuous action set X_i and with multiple players. Let H denote the set of all sample paths $h=(x^t,y^t)_{t=1}^{\infty}$. Given strategy profiles $\hat{s}$, let $P^{\hat{s}}\in\triangle X$ denote the probability distribution of the sample path h. Given a sample path h, let $\hat{\mu}_i^t(h)$ denote the belief hierarchy in period t.

Proposition 6. *Given any i, k, and $\hat{s}$, $P^{\hat{s}}$-almost surely, we have*

$$\lim_{t\to\infty}\int_{\Theta_{i,k}}(K_{i,k}(\theta_{i,k},\sigma^t(h))-K^*_{i,k}(\sigma^t(h)))\hat{\mu}^{t+1}_{i,k}(h)[d\theta_{i,k}]=0. \tag{32}$$

Let $\mathscr{H}$ denote the set of sample paths h which satisfy (32). By Proposition 6, $P^{\hat{s}}(\mathscr{H})=1$.

5.3 Asymptotic Motion of Action Frequency

5.3.1 Stochastic Approximation and Differential Inclusion

In this subsection, we show that given any Markov strategy $\hat{s}$, the asymptotic motion of the action frequency σ^t is represented by a recursive formula. So pick a Markov strategy $\hat{s}$, and pick a sample path $h\in\mathscr{H}$. Then by the definition, the action frequency in each period is written as

$$\sigma^{t+1}(h)=\frac{t}{t+1}\sigma^t(h)+\frac{1}{t+1}1_{\hat{s}(\hat{\mu}^{t+1}(h))}.$$

That is, the action frequency in period $t+1$ is a weighted average of the past action frequency σ^t and today's action $1_{\hat{s}(\hat{\mu}^{t+1}(h))}$. In what follows, we will show that this second term $1_{\hat{s}(\hat{\mu}^{t+1}(h))}$ can be written as a function of σ^t, so that σ^{t+1} is determined recursively.

Pick an arbitrary small $\varepsilon>0$. Then let $B_\varepsilon:\triangle\hat{X}\to\prod_{i=1}^2\prod_{k=1}^{k_i+1}\triangle\Theta_{i,k}$ be the *ε-perturbed belief correspondence* defined as

$$B_\varepsilon(\sigma)=\left\{\hat{\mu}\,\middle|\,\forall i\forall k\int_{\Theta_{i,k}}(K_{i,k}(\theta_{i,k},\sigma)-K^*_{i,k}(\sigma))\hat{\mu}_{i,k}(d\theta_{i,k})\leq\varepsilon\right\}.$$

Roughly, $B_\varepsilon(\sigma)$ is the set of all belief hierarchies $\hat{\mu}$ such that each $\hat{\mu}_{i,k}$ is concentrated on the best models $\Theta_{i,k}(\sigma)$ given the mixture σ in the sense of (32).

Since $h\in\mathscr{H}$, there is T such that for all $t>T$, $\hat{\mu}^{t+1}(h)\in B_\varepsilon(\sigma^t)$. This in turn implies that the action $\hat{s}(\hat{\mu}^{t+1})$ in period $t+1$ must be chosen from the *ε-enlarged policy correspondence* $S_\varepsilon(\sigma^t)$, which is defined as

$$S_\varepsilon(\sigma)=\{\hat{s}(\hat{\mu})|\forall\hat{\mu}\in B_\varepsilon(\sigma)\}$$

for each σ. This immediately implies the following result:

Proposition 7. *Pick a Markov strategy $\hat{s}$. Then given any $h \in \mathscr{H}$, there is a decreasing sequence $\{\varepsilon^t\}_{t=1}^{\infty}$ with $\lim_{t\to\infty} \varepsilon^t = 0$ such that*

$$\sigma^{t+1}(h) \in \frac{t}{t+1}\sigma^t(h) + \frac{1}{t+1} S_{\varepsilon^t}(\sigma^t(h)).$$

This proposition implies that in a later period t, the action chosen in that period is selected from the set $S_\varepsilon(\sigma^t)$ for small ε. Now we ask how this set looks like in the limit as $\varepsilon \to 0$. Given a Markov strategy $\hat{s}$, let

$$\hat{S}(\mu) = \left\{ \hat{x} \middle| \hat{x} = \lim_{n\to\infty} \hat{s}(\hat{\mu}^n) \text{ for some } (\hat{\mu}^n)_{n=1}^{\infty} \text{ with } \lim_{n\to\infty} (\hat{\mu}^n) = \hat{\mu} \right\}$$

for each μ. This $\hat{S}$ is an *upper hemi-continuous policy correspondence induced by $\hat{s}$*. It is obvious that $\hat{s}(\hat{\mu}) \in \hat{S}(\hat{\mu})$ for each $\hat{\mu}$. Also a standard argument shows that $\hat{S}$ is indeed upper hemi-continuous with respect to $\hat{\mu}$. Note that $\hat{S} = \hat{s}$ if $\hat{s}$ is continuous. Then define

$$S_0(\sigma) = \{\hat{x} \in \hat{S}(\hat{\mu}) | \forall \hat{\mu} \in B_0(\sigma)\}$$

where

$$B_0(\sigma) = \{\hat{\mu} | \hat{\mu}_{i,k} \in \Theta_{i,k}(\sigma) \ \ \forall i \forall k\}.$$

The following proposition shows that when $\varepsilon \to 0$, the set $S_\varepsilon(\sigma)$ which appears in the previous proposition is approximated by $S_0(\sigma)$.

Proposition 8. *$S_\varepsilon(\sigma)$ is upper hemi-continuous in (ε, σ) at $\varepsilon = 0$. So with the dual bounded-Lipschitz norm, $\triangle S_\varepsilon(\sigma)$ is upper hemi-continuous at $\varepsilon = 0$.*

Propositions 7 and 8 suggest that after a long time, the motion of the action frequency is approximated by

$$\sigma^{t+1}(h) \in \frac{t}{t+1}\sigma^t(h) + \frac{1}{t+1} S_0(\sigma^t(h)),$$

which is equivalent to

$$\sigma^{t+1}(h) - \sigma^t(h) \in \frac{t}{t+1}(S_0(\sigma^t(h)) - \sigma^t(h))$$

That is, the drift of the action frequency, $\sigma^{t+1}(h) - \sigma^t(h)$, should be proportional to the difference between today's action chosen from $S_0(\sigma^t(h))$ and the current action frequency $\sigma^t(h)$. The next proposition formalizes this idea using the stochastic approximation technique developed by Benaïm, Hofbauer, and Sorin (2005): It shows that the asymptotic motion of the action frequency is described by the differential inclusion

$$\dot{\boldsymbol{\sigma}}(t) \in \triangle S_0(\boldsymbol{\sigma}(t)) - \boldsymbol{\sigma}(t). \tag{33}$$

In this differential inclusion, the drift of the action frequency is $\triangle S_0(\boldsymbol{\sigma}(t)) - \boldsymbol{\sigma}(t)$, rather than $S_0(\boldsymbol{\sigma}(t)) - \boldsymbol{\sigma}(t)$. The reason is as follows. As will be shown in Proposition 9 below, the differential inclusion (33) approximates the motion of the action frequency in the limit as the period length in the discrete-time model shrinks to zero. This means that a small time interval $[t, t+\varepsilon]$ in the continuous-time model should be interpreted as a collection of arbitrarily many periods in the discrete-time model. Suppose now that players' beliefs are in a neighborhood of μ during this time interval $[t, t+\varepsilon]$. In all periods included in this interval, players choose an action profile from the set $S_0(\mu)$, and in particular, if $S_0(\mu)$ contains two or more action profiles, then different action profiles can be chosen in different periods. Accordingly, the action frequency during this interval can take any value in $\triangle S_0(\mu)$, as described by the differential inclusion (33).[29]

To state the result formally, we use the following terminologies, which are standard in the literature on stochastic approximation. Let $\tau_0 = 0$ and $\tau_t = \sum_{n=1}^{t} \frac{1}{n}$ for each $t = 1, 2, \cdots$. Then given a sample path h, the *continuous-time interpolation* of the action frequency σ^t is a mapping $\boldsymbol{w}(h) : [0, \infty) \to \triangle \hat{X}$ such that

$$\boldsymbol{w}(h)[\tau_t + s] = \sigma^t(h) + \frac{\tau}{\tau_{t+1} - \tau_t}(\sigma^{t+1}(h) - \sigma^t(h))$$

for all $t = 0, 1, \cdots$ and $\tau \in [0, \frac{1}{t+1})$. Intuitively, $\boldsymbol{w}$ represents the motion of the action frequency as a piecewise linear path with re-indexed time. A mapping

[29]There is also a technical reason: In the proof of Proposition 9, we apply the stochastic approximation method of Benaïm, Hofbauer, and Sorin (2005), which requires that the drift term be a convex-valued (and upper hemi-continuous) correspondence. So we need to convexify the drift term by taking $\triangle S_0(\boldsymbol{\sigma}(t))$, rather than $S_0(\boldsymbol{\sigma}(t))$.

$\boldsymbol{\sigma} : [0,\infty) \to \triangle\hat{X}$ is a *solution to the differential inclusion (33) with an initial value* $\sigma \in \triangle\hat{X}$ if it is absolutely continuous in all compact intervals, $\boldsymbol{\sigma}(0) = \sigma$, and (33) is satisfied for almost all t. Since $\triangle S_0(\sigma)$ is upper hemi-continuous with closed convex values, given any initial value $\sigma \in \triangle\hat{X}$, the differential inclusion (33) has a solution. (See Theorem 9 of Deimling (1992) on page 117.) Let $Z(\sigma)$ denote the set of all these solutions.

Proposition 9. *Pick a Markov strategy $\hat{s}$. Then for any $T > 0$ and any sample path $h \in \mathscr{H}$,*

$$\lim_{t\to\infty} \inf_{\boldsymbol{\sigma}\in Z(\boldsymbol{w}(h)[t])} \sup_{\tau\in[0,T]} \|\boldsymbol{w}(h)[t+\tau] - \boldsymbol{\sigma}(\tau)\| = 0.$$

5.3.2 Steady State and Generalized Berk-Nash Equilibrium

$\sigma \in \triangle\hat{X}$ is a *steady state* of the differential inclusion (33) if $\sigma \in \triangle S_0(\sigma)$. The following proposition shows that if the action frequency σ^t converges, then its limit point must be a steady state. The proof is exactly the same as Proposition 1 of EPY, and hence we omit it.

Proposition 10. *Pick a Markov strategy s. Then for each sample path $h \in \mathscr{H}$, if the action frequency $\sigma^t(h)$ converges, then its limit point* $\lim_{t\to\infty}\sigma^t(h)$ *is a steady state of (33).*

In all the examples in this paper, we assume that the agents are myopically optimal so that the strategy profile $\hat{s}$ satisfies (31). In this special case, steady states of our differential inclusion are *generalized Berk-Nash equilibria* in the following sense:

Definition 5. A probability measure $\sigma \in \triangle\hat{X}$ is a *generalized Berk-Nash equilibrium* (*GBNE*) if for each pure action profile $\hat{x} = (\hat{x}_1, \hat{x}_2)$ in the support of σ, for each i and for each k, there is a belief $\hat{\mu}_{i,k} \in \triangle\Theta_{i,k}(\sigma)$ such that

$$\hat{x}_{i,k} \in \arg\max_{\hat{x}'_{i,k}} \int_{\Theta_{i,k}} U_{\theta_{i,k}}(\hat{x}'_{i,k}, \hat{x}_{i,k+1})\hat{\mu}_{i,k}(d\theta_{i,k}).$$

A generalized Berk-Nash equilibrium is *degenerate* if it is a point mass on some pure action profile $\hat{x}$.

In words, in a generalized Berk-Nash equilibrium σ, each action profile $\hat{x}$ which has a positive weight in σ is a one-shot equilibrium for some belief $\hat{\mu}$, and this belief $\hat{\mu}$ is concentrated on the models $\Theta_{i,k}(\sigma)$ which minimize the Kullback-Leibular divergence. In a non-degenerate GBNE which assign positive weights on multiple action profiles $\hat{x}$, different action profiles $\hat{x}$ may be supported by different beliefs $\hat{\mu}$. We will discuss more on this later.

Proposition 11. *Suppose that the strategy profile $\hat{s}$ satisfies (31). Then any steady state of our differential inclusion (33) is a generalized Berk-Nash equilibrium. So for each sample path $h \in \mathscr{H}$, if the action frequency $\sigma^t(h)$ converges, then its limit point $\lim_{t\to\infty} \sigma^t(h)$ is a generalized Berk-Nash equilibrium.*

Note that the action frequency may converge to non-degenerate equilibrium σ, which assigns positive probability to multiple action profiles $\hat{x}$. An intuition is as follows. If the action frequency σ^t converges to some σ, then from Proposition 6, the posterior belief $\hat{\mu}^t$ will be concentrated on $\triangle\Theta(\sigma)$ after a long time, that is, $\hat{\mu}^t$ is in a neighborhood of $\triangle\Theta(\sigma)$ for large t. If all the beliefs in this neighborhood induce the same equilibrium action $\hat{x}$ (i.e., $\hat{s}(\hat{\mu}) = \hat{x}$ for all beliefs $\hat{\mu}$ in a neighborhood of $\triangle\Theta(\sigma)$), then the action frequency will eventually converge to a point mass on $\hat{x}$. But in general, this need not be the case; different beliefs $\hat{\mu}$ and $\hat{\mu}'$ in this neighborhood may induce different equilibrium actions $\hat{x}$ and $\hat{x}'$. In such a case, both $\hat{x}$ and $\hat{x}'$ can be chosen infinitely often on the path, and hence have positive weights in the limiting action frequency σ.

Note, however, that in many applications, all GBNE are degenerate. Indeed, if (i) there is a unique equilibrium $\hat{x}$ for each belief $\hat{\mu}$ and (ii) there is a unique minimizer $\theta_{i,k}$ of the Kullback-Leibular divergence for each action frequency σ, then obviously any GBNE is degenerate. All our examples in the paper satisfy these assumptions.

We view GBNE as a natural extension of BNE of Esponda and Pouzo (2016) to our setup. For comparison, let us think about the information structure considered in Esponda and Pouzo (2016), that is, suppose that the subjective signal distribution $Q_{\theta_{i,k}}(\cdot|x)$ is independent of the opponent's action (i.e., $Q_{\theta_{i,k}}(\cdot|x) = Q_{\theta_{i,k}}(\cdot|x_i)$ for each i, k, and x). In this special case, all higher-level agents (i,k)

with $k \geq 2$ are irrelevant, in the sense that they do not influence the actions and the beliefs of the actual player $(i,1)$. This is so because each player i's subjective expected payoff $U_{\theta_{i,1}}(x_i) = \sum_{y \in Y} Q_{\theta_{i,1}}(y|x_i) u_i(x_i, y)$ and her Bayes' formula $\mu_i^{t+1}(\theta_{i,1}) = \frac{\mu_i^t(\theta_{i,1}) q_{\theta_{i,1}}(y^t|x_i^t)}{\int_{\Theta_{i,1}} \mu_i^t(\theta_{i,1}) q_{\theta_{i,1}}(y^t|x_i^t) d\theta_{i,1}}$ are independent of the opponent j's action. Accordingly, GBNE reduces to a probability measure $\sigma \in \triangle(X_1 \times X_2)$ on the set $X_1 \times X_2$ such that for each pure action profile $x = (x_1, x_2)$ in the support of σ and for each i, there is a belief $\mu_i \in \triangle\Theta_{i,1}(\sigma)$ such that

$$x_i \in \arg\max_{x_i'} \int_{\Theta_{i,1}} U_{\theta_{i,1}}(x_i') \mu_i(d\theta_{i,1})$$

where $\Theta_{i,1}(\sigma)$ is the minimizers of the Kullback-Leibular divergence

$$\int_X \int_Y \log \frac{q(y|x)}{q_{\theta_{i,1}}(y|x)} Q(dy|x) \sigma(dx).$$

It is easy to check that GBNE is a weakening of BNE of Esponda and Pouzo (2016), that is, any GBNE σ is a BNE. In particular, degenerate BNE is equivalent to pure-strategy BNE, in that σ is a pure-strategy BNE if and only if it is a degenerate GBNE. However, a non-degenerate GBNE need not be a mixed-strategy BNE, because in a GBNE, (i) different actions x_i may be supported by different beliefs μ_i, and (ii) GBNE distribution σ allows correlation between x_1 and x_2. This difference comes from the fact that GBNE is a limit point of the action frequency, while BNE is a limit point of the action itself. Specifically, in Esponda and Pouzo (2016), there is an i.i.d. payoff perturbation each period, so that each player (independently) mix actions each period. A mixed-strategy BNE σ is regarded as a limit point of this mixed action. In this case, in a steady state, the mixed strategy σ_i must be optimal (with a payoff perturbation) given a single belief, and there is no correlation between actions of different players. In contrast, in our model, each player chooses a pure action, so there is perfect correlation between x_1 and x_2. Also as noted earlier, different action profiles x and x' which appear in non-degenerate GBNE σ are played in different periods t and t'; so they are supported by different beliefs μ^t and $\mu^{t'}$.

5.4 Motion of the KL Minimizer

5.4.1 Identifiability and Differential Inclusion

Our Proposition 9 shows that the asymptotic motion of the action frequency σ^t is described by the differential inclusion (33). However, solving the differential inclusion (33) is not easy in general. For example, in many applications (including the ones in this paper), there are continuous actions, in which case the action frequency σ^t is a probability distribution over an infinite-dimensional (continuous) space, and thus the differential inclusion becomes an infinite-dimensional problem. In this section, we show that this dimensionality problem can be avoided if we look at the asymptotic motion of the belief, rather than that of the action frequency.

We will impose the following *identifiability* assumption, which requires that there be a unique KL minimizer $\theta_{i,k}(\sigma)$ for each measure $\sigma \in \triangle \hat{X}$. This assumption is satisfied in many applications, see Esponda and Pouzo (2016) for more detailed discussions on this assumption.

Assumption 3. For each i, k, and σ, there is a unique minimizer $\theta_{i,k}(\sigma) \in \Theta_{i,k}$ of the Kullback-Leibular divergence $K_{i,k}(\theta_{i,k}, \sigma)$.

Since $\Theta_{i,k}(\sigma)$ is upper hemi-continuous in σ, under the identifiability assumption, each KL minimizer $\theta_{i,k}(\sigma)$ is continuous in σ. The next lemma shows that $\theta(\sigma) = (\theta_{i,k}(\sigma))_{i,k}$ is Lipschitz continuous if some additional assumptions hold. With an abuse of notation, let $K_{i,k}(\theta_{i,k}, \hat{x}) = K_{i,k}(\theta_{i,k}, \sigma)$ for $\sigma = 1_{\hat{x}}$.

Assumption 4. The following conditions hold:

(i) For each i, k, and m, $\frac{\partial K_{i,k}(\theta_{i,k},\hat{x})}{\partial \theta_{i,k,m}} < \infty$, where $\theta_{i,k,m}$ denotes the m-th component of $\theta_{i,k}$. Also for each $\hat{x}$, $K_{i,k}(\theta_{i,k}, \hat{x})$ is twice-continuously differentiable with respect to $\theta_{i,k}$, that is, $\frac{\partial^2 K_{i,k}(\theta_{i,k},\hat{x})}{\partial \theta_{i,k,m} \partial \theta_{i,k,n}}$ is continuous in $\theta_{i,k}$.

(ii) $\frac{\partial K_{i,k}(\theta_{i,k},\hat{x})}{\partial \theta_{i,k,m}}$ is equi-Lipschitz continuous, that is, there is $L > 0$ such that $\left|\frac{\partial K_{i,k}(\theta_{i,k},\hat{x})}{\partial \theta_{i,k,m}} - \frac{\partial K_{i,k}(\theta_{i,k},\hat{x}')}{\partial \theta_{i,k,m}}\right| < L|\hat{x} - \hat{x}'|$ for all i, k, m, $\theta_{i,k}$, $\hat{x}$, and $\hat{x}'$.

(iii) The KL minimizer $\theta(\sigma)$ satisfies both the first-order and second-order conditions for each σ. (An implication is that the inverse of the Hessian matrix exists.)

Lemma 2. *$\theta(\sigma)$ is Lipschitz continuous in σ. That is, there is $L > 0$ such that $|\theta(\sigma) - \theta(\tilde{\sigma})| \leq L\|\sigma - \tilde{\sigma}\|$.*

Now we consider the motion of the KL minimizer $\theta^t = (\theta^t_{i,k})_{i,k}$. Let $\boldsymbol{w}_\theta$ denote the continuous-time interpolation of θ^t. Let $\nabla K_{i,k}(\theta_{i,k}, x) = (\frac{\partial K_{i,k}(\theta_{i,k},x)}{\partial \theta_{i,k,m}})_m$, and $\nabla K(\theta, x) = (\frac{\partial K_{i,k}(\theta_{i,k},x)}{\partial \theta_{i,k,m}})_{i,k,m}$. Also let $\nabla^2 K_{i,k}(\theta_{i,k}, \sigma)$ denote the Hessian matrix of $K_{i,k}(\theta_{i,k}, \sigma)$ with respect to $\theta_{i,k}$, that is, each component of $\nabla^2 K_{i,k}(\theta_{i,k}, \sigma)$ is $\frac{\partial^2 K_{i,k}(\theta_{i,k},\sigma)}{\partial \theta_{i,k,m} \partial \theta_{i,k,n}}$. Let $\nabla^2 K(\theta, \sigma)$ denote a block diagonal matrix whose main diagonal blocks are $\nabla^2 K_{i,k}(\theta_{i,k}, \sigma)$, that is,

$$\nabla^2 K(\theta, \sigma) = \begin{pmatrix} \nabla^2 K_{1,1}(\theta_{1,1}, \sigma) & & 0 \\ & \nabla^2 K_{1,2}(\theta_{1,2}, \sigma) & \\ 0 & & \ddots \end{pmatrix}.$$

With an abuse of notation, let $S_0(\theta)$ denote $S_0(\sigma)$ for σ with $\theta(\sigma) = \theta$. The following proposition shows that the asymptotic motion of the KL minimizer is described by the differential inclusion

$$\dot{\boldsymbol{\theta}}(t) \in \bigcup_{\sigma:\theta(\sigma)=\boldsymbol{\theta}(t)} \bigcup_{\sigma' \in \triangle S_0(\boldsymbol{\theta}(t))} -(\nabla^2 K(\boldsymbol{\theta}(t), \sigma))^{-1} \left(\nabla K(\boldsymbol{\theta}(t)), \sigma'\right). \tag{34}$$

Let $Z_\theta(\boldsymbol{\theta}(0))$ be the set of solutions to the differential inclusion (34) with the initial value $\boldsymbol{\theta}(0)$.

Proposition 12. *Suppose that Assumptions 3 and 4 hold. Then for any $T > 0$ and any sample path $h \in \mathscr{H}$,*

$$\lim_{t\to\infty} \inf_{\theta \in Z_\theta(\boldsymbol{w}_\theta(h)[t])} \sup_{\tau \in [0,T]} |\boldsymbol{w}_\theta(h)[t+\tau] - \boldsymbol{\theta}(\tau)| = 0.$$

To interpret the differential inclusion (34), consider the special case in which $\Theta_{i,k} \subset \boldsymbol{R}$, i.e., assume that agent k's model $\theta_{i,k}$ is one-dimensional. Then from

(33), we have

$$\dot{\boldsymbol{\theta}}_{i,k}(t) \in \bigcup_{\sigma:\theta(\sigma)=\boldsymbol{\theta}(t)} \bigcup_{\sigma' \in \triangle S_0(\boldsymbol{\theta}(t))} -\frac{K'_{i,k}(\boldsymbol{\theta}_{i,k}(t), \sigma')}{K''_{i,k}(\boldsymbol{\theta}_{i,k}(t), \sigma)} \tag{35}$$

for each i and k, where $K'_{i,k}(\theta,\sigma) = \frac{\partial K_{i,k}(\theta,\sigma)}{\partial \theta}$ and $K''_{i,k}(\theta,\sigma) = \frac{\partial^2 K_{i,k}(\theta,\sigma)}{\partial \theta^2}$.

The denominator $K''_{i,k}(\boldsymbol{\theta}_{i,k}(t),\sigma)$ measures the curvature of the Kullback-Leibular divergence. Note that this term is always positive, because the second-order condition must be satisfied (Assumption 4(iii)). So this term influences the absolute value of $\boldsymbol{\theta}(t)$, but not the sign of $\dot{\boldsymbol{\theta}}_{i,k}(t)$; this in turn implies that this denominator influences the speed of $\boldsymbol{\theta}_{i,k}(t)$, but not the direction. Intuitively, when the curve is flatter (i.e., $K''_{i,k}$ is close to zero), all models in a neighborhood of $\boldsymbol{\theta}(t)$ almost equally fit the past data. Hence the KL minimizer $\boldsymbol{\theta}(t)$ is more sensitive to the new data generated by today's action, and it changes quickly.

The numerator $-K'_{i,k}(\boldsymbol{\theta}_{i,k}(t),\sigma')$ measures how much an increase in $\theta_{i,k}$ improves fitness to the new data generated by today's action σ'. This term influences the sign of $\dot{\boldsymbol{\theta}}_{i,k}(t)$, so it determines whether $\boldsymbol{\theta}_{i,k}(t)$ moves up or down. Intuitively, when this numerator is positive, (at least in a neighborhood of $\boldsymbol{\theta}(t)$) higher θ better explains the new data generated by today's action, so $\boldsymbol{\theta}(t)$ moves up. On the other hand, when this numerator is negative, lower θ better explains the new data, so $\boldsymbol{\theta}(t)$ moves down.

When we consider the dynamic of $\theta^t = \theta(\sigma^t)$, the drift of θ^t cannot be uniquely determined, for two reaosns. First, the KL minimizer θ^t may not uniquely determine the agents' actions today, in the sense that $S_0(\theta^t)$ may not be a singleton. (As pointed out by Esponda, Pouzo, and Yamamoto (2019), in the single-agent setup, this happens when the agent is indifferent over multiple actions at a model $\theta = \theta^t$.) In our differential inclusion (35), this multiplicity is captured by taking the union over $\sigma' \in \triangle S_0(\boldsymbol{\theta}(t))$. Note that the same multiplicity problem appears in the differential inclusion (33).

Second, the KL minimizer θ^t may not uniquely determine the past action frequency, in the sense that there may be more than one σ such that $\theta(\sigma) = \theta^t$. Note that even if two action frequencies σ and $\tilde{\sigma}$ yield the same KL minimizer (i.e., $\theta(\sigma) = \theta(\tilde{\sigma})$), they may yield different curvatures of the KL divergence, so they influence the speed of $\theta_{i,k}(t)$ differently. In our differential inclusion, this multiplicity is captured by taking the union over σ with $\theta(\sigma) = \theta(t)$.

5.4.2 Convergence with Identifiability

Using the differential inclusion (34), now we will derive a sufficient condition for the agents' beliefs θ^t to converge for each type of misspecification considered in Sections 3 and 4. First, consider first-order misspecification in Section 3, i.e., set $k_1 = k_2 = 2$, $M_{1,1} = M_{2,2}$, and $M_{2,1} = M_{1,2}$, and assume that agent $(1,1)$ is correctly specified. Assume as in Section 3 that players play a static Nash equilibrium every period and they predict the opponent's strategy correctly, so that $\hat{s}_{1,1} = \hat{s}_{2,2}$ and $\hat{s}_{2,1} = \hat{s}_{1,2}$. Since $k_1 = k_2 = 2$, θ^t is a four-dimensional variable, i.e., $\theta^t = (\theta^t_{1,1}, \theta^t_{1,2}, \theta^t_{2,1}, \theta^t_{2,2})$. However, the differential inclusion (34) is essentially a one-dimensional problem. Indeed,

- Agent $(2,2)$ is redundant in that her belief is identical with the belief of agent $(1,1)$ every period. (This follows from the fact that player 2 knows player 1's posterior belief every period.) Similarly, agent $(1,2)$ is redundant.
- Agent $(1,1)$ is correctly specified, so we have $\theta_{1,1}(\sigma) = \theta^*$ for all σ, which implies that $\theta^t_{1,1} = \theta^*$ for any period t. (This reflects the fact that player 1 eventually learn the true state regardless of players' play.)

So in order to know the asymptotic motion of θ^t, we only need to know the asymptotic motion of player 2's belief, $\theta^t_{2,1}$. Similarly, under one-dimensional double misspecification considered in Section 4.2.1, the differential inclusion (34) reduces to a one-dimensional problem. This is so because player 1 eventually learns the true state and the hypothetical player is redundant, in that her belief is exactly the same as player 2's belief after every history. Our next proposition shows that in these cases, the belief converges almost surely, regardless of the initial prior.

To state the result formally, we use the following terminologies. Let I be the set of all agents. An agent (i,k) is *correctly specified* if there is a *correct model*

$\theta^*_{i,k} \in \Theta_{i,k}$ such that $Q_{\theta^*_{i,k}}(x) = Q(x)$ for all x. An agent (i,k) has *no inferential naivety* if (i) $i+k$ is even and $(\hat{s}_{i,k}(\hat{\mu}), \hat{s}_{i,k+1}(\hat{\mu})) = (\hat{s}_{1,1}(\hat{\mu}), \hat{s}_{2,1}(\hat{\mu}))$ for all $\hat{\mu}$, or (ii) $i+k$ is odd and $(\hat{s}_{i,k}(\hat{\mu}), \hat{s}_{i,k+1}(\hat{\mu})) = (\hat{s}_{2,1}(\hat{\mu}), \hat{s}_{1,1}(\hat{\mu}))$ for all $\hat{\mu}$. An agent (i,k) is *unbiased* if she is correctly specified and has no inferential naivety. Let I^* denote the set of all unbiased agents (i,k).

Given a Markov strategy $\hat{s}$, two agents (i,k) and (j,l) are equivalent if $i+k \equiv j+l \pmod 2$, $M_{i,k} = M_{j,l}$, and $(s_{i,k}, s_{i,k+1}) = (s_{j,l}, s_{j,l+1})$. Consider a partition of the set of agents induced by the equivalence classes, and for each equivalence class, pick a representative agent (i,k). Let I^{**} the set of these representative agents. (Mathematically, I^{**} is a quotient set $I/\sim$, where $(i,k) \sim (j,l)$ if (i,k) and (j,l) are eqivalent.) For each $(i,k) \in I^{**}$, let $I(i,k)$ denote the set of all agents equivalent to (i,k). Note that all agents in the set $I(i,k)$ choose the same action and have the same belief every period.

θ is a *steady state* if it is a steady state of the differential inclusion (34), i.e., $\nabla K(\theta, \sigma') = 0$ for some $\sigma' \in \triangle S_0(\theta)$. Let E denote the set of all steady states.

Proposition 13. *Suppose that Assumptions 3 and 4 hold. Pick a Markov strategy $\hat{s}$, and suppose that $I^{**} \setminus I^* = \{(i,k)\}$ and $\Theta_{i,k} \subset \boldsymbol{R}$. Then for any sample path $h \in \mathscr{H}$, the belief converges to a steady state, i.e., $\lim_{t\to\infty} \theta^t(h) \in E$. In particular, if there is a unique steady state, θ^t converges there almost surely.*

The assumption $I^{**} \setminus I^* = \{(i,k)\}$ implies that there is essentially only one agent whose belief moves in a non-trivial way; all other agents are redundant or learn the true state eventually. This assumption is satisfied, for example, in a game with first-order misspecification and one-sided double misspecification.[30] The proposition shows that in such a case, the belief converges almost surely as long as the identifiability condition (Assumption 3) and Assumption 4 hold. This result is a natural extension of Heidhues, Kőszegi, and Strack (2020a) and Esponda, Pouzo, and Yamamoto (2019), who prove convergence of the belief for a single-agent problem.

[30] Note that this proposition applies even when players are patient; if players play a Markov perfect equilibrium, we have $I^{**} \setminus I^* = \{(i,k)\}$.

Next, consider second-order misspecification studied in Section 4.1. In this case, the *two* beliefs θ_2 and $\hat{\theta}_1$ evolve in a non-trivial way. Similarly, under two-sided double misspecification studied in Section 4.2.2, the two beliefs θ_1 and θ_2 evolve in a non-trivial way. Our next proposition shows that the belief converges even in these cases, provided that some additional conditions hold. Consider the model studied in Section 4.1. Given a misspecified parameter A, let $f_2(\hat{\theta}_1,A)$ denote the set of "steady-state belief" of player 2, when the hypothetical player's belief is fixed at $\hat{\theta}_1$. That is, $f_2(\hat{\theta}_1,A)$ is the set of all θ_2 such that there is $(x_1,x_2,\hat{x}_1)$ satisfying the consistency condition (19) and the incentive-compatibility condition (14) through (16). Similarly, given a belief θ_2 of player 2, let $\hat{f}_1(\theta_2,A)$ be the set of all θ_1 such that there is $(x_1,x_2,\hat{x}_1)$ satisfying the consistency condition (18) and the incentive-compatibility condition (14) through (16). A steady state studied in Section 4.1 can be seen as an intersection of the two graphs $\{(\theta_2,\hat{f}_1(\theta_2,A))|\forall\theta_2\}$ and $\{(f_2(\hat{\theta}_1,A),\hat{\theta}_1)|\forall\hat{\theta}_1\}$.

Proposition 14. *Consider the game with the second-order misspecification in Section 4.1, and suppose that Assumptions 3 and 4 hold. Pick a parameter A such that*

(i) $f_2(\hat{\theta}_1,A)$ *and* $\hat{f}_1(\theta_2,A)$ *are singletons for all* $\hat{\theta}_1$ *and* θ_2 *(so that they are functions of* $\hat{\theta}_1$ *and* θ_2*, rather than correspondences),*

(ii) $f_2(\hat{\theta}_1,A)$ *and* $\hat{f}_1(\theta_2,A)$ *are continuously differentiable in* $\hat{\theta}_1$ *and* θ_2*, respectively, and*

(iii) $\max_{\hat{\theta}_1}|\frac{\partial f_2(\hat{\theta}_1,A)}{\partial\hat{\theta}_1}|\max_{\theta_2}|\frac{\partial\hat{f}_1(\theta_2,A)}{\partial\theta_2}|<1.$

(iv) The KL divergence is single-peaked, in that for each (i,k)*, there is* $\theta_{i,k}$ *such that* $K'_{i,k}(\tilde{\theta}_{i,k},\sigma)<0$ *for all* $\tilde{\theta}_{i,k}<\theta_{i,k}$ *and* $K'_{i,k}(\tilde{\theta}_{i,k},\sigma)>0$ *for all* $\tilde{\theta}_{i,k}>\theta_{i,k}$

Then there is a unique steady state $\theta^*=(\theta_2^*,\hat{\theta}_1^*)$*, and the belief converges there almost surely regardless of the initial prior. The same result holds for two-sided double misspecification by replacing* $\hat{f}_1(\theta_2,A)$ *with* $f_1(\theta_2,A)$*, where* $f_1(\theta_2,A)$ *is a "steady-state belief" of player* 1 *given player* 2*'s belief* θ_2

This proposition gives a sufficient condition for belief convergence. It requires that f_2 and $\hat{f}_1$ are continuously differentiable functions and their derivatives are not too large (i.e., changing $\hat{\theta}_1$ does not influence the steady-state belief of the opponent much, and vice versa).

In the case of second-order misspecification, if the misspecification is small in that the parameter A is close to a, there is a simpler sufficient condition for convergence; we can drop assumption (iii) in Proposition 14. To see why, consider the case in which $A = a$ so that there is no misspecification. We claim that assumption (i) in Proposition 14 automatically implies assumption (iii). Indeed, when $A = a$, we always have $\theta^* \in \hat{f}_1(\theta_2, a)$, so that assumption (i) implies $\hat{f}_1(\theta_2, a) = \theta^*$ for all θ_2. Then assumption (iii) is satisfied as $\frac{\partial \hat{f}_1(\theta_2, A)}{\partial \theta_2} = 0$. By the continuity, the same is true as long as the parameter A is close to a. The following corollary summarizes this discussion:

Corollary 4. *Consider the game with the second-order misspecification in Section 4.1. Suppose that assumptions (i) and (ii) in Proposition 14 are satisfied at $A = a$. Then there is $\overline{A} > a$ such that for any $A \in (a, \overline{A})$, there is a unique steady state, and θ^t converges there almost surely regardless of the initial prior.*

5.4.3 Convergence Without Identifiability

In the previous subsection, we have seen that the motion of the KL minimizer can be described by a differential inclusion if the identifiability condition holds. However, there are some economic examples which do not satisfy the identifiability. For example, the model of overconfidence studied in Heidhues, Kőszegi, and Strack (2018) does not satisfy the identifiability in general.

The following proposition shows that even in such a situation, the belief still converges to a steady state if some additional assumptions on payoffs and information structures are satisfied. For each action frequency σ, let $\underline{\theta}_{i,k}(\sigma)$ denote the minimal KL minimizer, that is, let $\underline{\theta}_{i,k}(\sigma) = \min_{\theta_{i,k} \in \Theta_{i,k}(\sigma)} \theta_{i,k}$. Likewise, let $\overline{\theta}_{i,k}(\sigma)$ denote the maximal KL minimizer. Also, when the problem is one-dimensional (i.e., $I^{**} \setminus I^* = \{(i,k)\}$), for each model $\theta_{i,k}$, let $S_0(\theta_{i,k}) = S_0(\mu)$

where μ is a degenerate belief on θ such that $\theta_{j,l} = \theta_{i,k}$ for all $(j,l) \in I(i,k)$ and $\theta_{j,l} = \theta^*$ for all $(j,l) \in I^*$.

Proposition 15. *Pick any Markov strategy $\hat{s}$. Assume that*

(i) The problem is one-dimensional, i.e., $I^{**} \setminus I^* = \{(i,k)\}$ and $\Theta_{i,k} \subset \boldsymbol{R}$.

(ii) For each pure action profile x, the KL divergence $K_{i,k}(\theta, 1_x)$ is single-peaked, i.e., there is a unique KL minimizer $\theta_{i,k}(x)$, $\frac{\partial K_{i,k}(\theta,1_x)}{\partial \theta_{i,k}} < 0$ for $\theta_{i,k} < \theta_{i,k}(x)$, and $\frac{\partial K_{i,k}(\theta,1_x)}{\partial \theta_{i,k}} > 0$ for $\theta_{i,k} > \theta_{i,k}(x)$.

(iii) There is a unique steady state σ^*, and $\theta_{i,k}(\sigma^*) = \{\theta^*_{i,k}\}$.

(iv) $S_0(\tilde{\theta}_{i,k})$ is a function (rather than a correspondence) of $\tilde{\theta}_{i,k}$, and $\theta_{i,k}(S_0(\tilde{\theta}_{i,k}))$ is increasing in $\tilde{\theta}_{i,k}$.

(v) For each belief $\hat{\mu}$ whose support is compact, $S_0(\mu) \subseteq \bigcup_{\theta_{i,k} \in co(supp \hat{\mu}_{i,k})} S_0(\theta_{i,k})$.

Then for each sample path $h \in \mathscr{H}$, $\lim_{t\to\infty} \underline{\theta}_{i,k}(\sigma^t(h)) = \lim_{t\to\infty} \overline{\theta}_{i,k}(\sigma^t(h)) = \theta^*_{i,k}$.

6 Related Literature

There is a rapidly growing literature on Bayesian learning with model misspecification. Nyarko (1991) presents an example in which the agent's action does not converge. Fudenberg, Romanyuk, and Strack (2017) consider a general two-state model and characterize the agent's asymptotic actions and behavior. Heidhues, Kőszegi, and Strack (2018), Heidhues, Kőszegi, and Strack (2020a), and He (2019) study a continuous-state setup, and they show that the agent's action and belief converge to a Berk-Nash equilibrium of Esponda and Pouzo (2016), under some assumptions on payoffs and information structure. Esponda, Pouzo, and Yamamoto (2019) characterize the agent's asymptotic behavior in a general single-agent model. Fudenberg, Lanzani, and Strack (2020) discuss robustness of steady states. All these papers look at a single-agent problem and focus on first-order misspecification. More recently, Ba and Gindin (2020) consider two-player team production in which both players are overconfident about their own capability. They show that if actions (efforts) are complements and information has a positive externality, then learning is mutually reinforcing, i.e., one's strategic play reduces *both* players' efforts and results in a more suboptimal outcome. Our work strengthens their result, in three ways. First, our Proposition 1 gives a necessary and sufficient condition for mutually-reinforcing learning: Assuming that the base misspecification effect is negative, our proposition shows that player 1's strategic play reduces both players' efforts if and only if the two asymptotic best response curves are upward-sloping (i.e., $BR_1' > 0$ and $BR_2' > 0$).[31] Second, our Proposition 1 can be applied to a more general setup, such as Cournot duopoly and tournaments, which allows us to study a wide range of applications. Third, and perhaps most importantly, we develop a model of higher-order misspecification and study how each type of misspecification influences players' beliefs and actions.

Misspecified learning has also been studied in other settings and applications. A literature on social learning studies how inferential naivety or model misspecification influences the asymptotic outcomes (e.g.,DeMarzo, Vayanos, and Zwiebel (2003), Eyster and Rabin (2010), Gagnon-Bartsch (2016), Gagnon-Bartsch and Rabin (2016), Bohren and Hauser (2020), and Frick, Iijima, and Ishii (2020)).[32] Molavi (2020) considers a general equilibrium model in which a representative agent has a misspecified view about the world. Cho and Kasa (2017) study an asset-pricing model in which an agent incorrectly believes that the environment is not stationary.

A prominent example of the first-order misspecification is overconfidence on own capability. Plenty of experimental and empirical papers report that people exhibit overconfidence on own ability in various economic activities, such as strategic entries (Camerer and Lovallo, 1999), corporate investments (Malmendier and

[31] The assumptions of Ba and Gindin (2020) (strategic complementarity and positive information externality) indeed ensure upward-sloping asymptotic best response curves.

[32] For experimental evidence on how social-learning outcomes depart from a correctly specified learning model, see, e.g., Çelen and Kariv (2005), Kübler and Weizsäcker (2004, 2005).

Tate, 2005), and merger decisions (Malmendier and Tate, 2008).[33][34] Furthermore, recent empirical evidence suggests that overconfidence on a particular aspect of own capability persists even after a long time and a plenty of feedback (Hoffman and Burks, 2020; Huffman, Raymond, and Shvets, 2019), which calls for the analysis of long-run behavior under model misspecifications.

Similarly, there is evidence that people sometimes have systematically incorrect views of their opponents' beliefs and actions, as in our second-order misspecification model. In both laboratory and field studies of disclosure games, departing from the theoretical prediction, people often do not disclose unfavorable (but not the worst) information (Dranove and Jin, 2010; Brown, Camerer, and Lovallo, 2012). Jin, Luca, and Martin (forthcoming) conduct more detailed experiments to test potential mechanisms. They find that receivers systematically make incorrect predictions about their opponents' actions, and the receivers also choose actions based on such incorrect views.. In common-value auctions, classical evidence suggests that bidders often mispredict other bidders' behavior and suffer from a loss, which is known as the "winner's curse" (Kagel and Levin, 2002). Avery and Kagel (1997) report that such suboptimal bidding behavior systematically persists even when bidders have received dozens of feedback.

Finally, there is a literature which studies implications of one's misspecification in a static game theoretic model. As discussed in Section 3.2.1, Kyle and Wang (1997) analyze a variant of one-shot Cournot duopoly with first-order misspecification.[35] Madarász (2012) considers a player who misspecifies other play-

[33] See Malmendier and Tate (2015) for reviewing managerial overconfidence, Daniel and Hirshleifer (2015) for reviewing overconfidence in financial markets, and Grubb (2015) for reviewing consumers' overconfidence.

[34] Benoît and Dubra (2011) point out that most of these works may not be sufficient to conclude that people have overconfidence, because it can also be explained by considering a model in which people do not have overconfidence but face uncertainty about some aspect of the environment. However, this explanation is not supported by Benoît, Dubra, and Moore (2015); they conduct a lab experiment which separates out the effect of uncertainty, and find that subjects still exhibit overconfidence.

[35] See Kőszegi (2014), Grubb (2015), and Heidhues and Kőszegi (2018) for reviews of theoretical works which study overconfident agents in contract theory and industrial organization.

ers' knowledge, which can be described by our double misspecification model. Our paper contributes to this literature in that we characterize how misspecified players learn about own environment in a general public-monitoring setup, and derive a number of economic implications. Indeed, as discussed in Aghion, Bolton, Harris, and Jullien (1991), the theory of imperfect competition with complete information (e.g., its demand function is known) is often defended with the argument that if the economic environment remains fixed over time, then players eventually learn all relevant parameters from past experience. Since a misspecified player may not be able to correctly learn a true parameter, we believe it is important to formally analyze the learning process and the effect of strategic interaction on the long-run outcomes.

A Proofs

A.1 Proof of Proposition 1

Pick x^* and A^* as stated. Since the steady-state actions (x_1^*, x_2^*) are interior points, they must satisfy the first-order conditions

$$\frac{\partial U_1(x_1, x_2^*, \theta^*)}{\partial x_1} = 0, \tag{36}$$

$$\left.\frac{\partial U_2(x_1^*, x_2, \theta)}{\partial x_2}\right|_{\theta=\theta_2(x^*,A)} = 0. \tag{37}$$

Let M be the Jacobian of this system of the equations. Then each ij-component of the matrix coincides with M_{ij} defined in the main text. That is,

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$$

Since $BR_1' BR_2' \neq 1$, we have $\det M \neq 0$, so the implicit function theorem guarantees that for any parameter A close to A^*, there is an action profile x^* which satisfies the first-order conditions (36) and (37). These action profiles are globally optimal (i.e., maximize the expected payoff given the belief $\theta_1 = \theta^*$ and $\theta_2(x^*, A)$), because of the regularity conditions (i) and (ii). So this x^* is a steady state given the parameter A. The implicit function theorem also asserts that

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1^*}{\partial A} \\ \frac{\partial x_2^*}{\partial A} \end{bmatrix} = - \begin{bmatrix} 0 \\ M_{2A} \end{bmatrix},$$

Solving this system of equations,

$$\frac{\partial x_2^*}{\partial A} = -\frac{M_{11} M_{2A}}{\det M},$$

$$\frac{\partial x_1^*}{\partial A} = \frac{M_{12} M_{2A}}{\det M}.$$

Dividing both the numerator and denominator of the first equation by $M_{11}M_{22}$ and using $\det M = M_{11}M_{22} - M_{12}M_{21}$, we have $\frac{\partial x_2^*}{\partial A} = -\frac{1}{1 - BR_1' BR_2'} \frac{M_{2A}}{M_{22}}$. Also by combining the two equations above, we have $\frac{\partial x_1^*}{\partial A} = BR_1 \frac{\partial x_2}{\partial A}$.

Next, we prove $BR_1'BR_2' < 1$ by contradiction. Suppose that $BR_1'BR_2' > 1$. Then we have either (i) $BR_1' > 0$ and $BR_2' > 0$, or (ii) $BR_1' < 0$ and $BR_2' < 0$. Consider case (i). Then we have $BR_2' > \frac{1}{BR_1'} > 0$. This means that if we take x_1 on the horizontal axis and x_2 on the vertical axis, then the two asymptotic best response curves are upward-sloping at the steady state action x^*, and BR_2 is steeper than BR_1. This and the continuity of BR_i immediately imply that BR_1 and BR_2 must intersect at some $x_1 > x_1^*$, but this contradicts with the fact that x^* is a unique steady state. The same argument works for case (ii).

Hence we have $BR_1'BR_2' \leq 1$. Also, dividing both sides of $\det M \neq 0$ by $M_{11}M_{22}$, we have $BR_1'BR_2' \neq 1$.

A.2 Proof of Proposition 2

Pick A^* and x^* as stated. Since x^* is an interior point, it must satisfy the first-order conditions

$$\frac{\partial U_1(x_1, x_2^*, \theta^*)}{\partial x_1} = 0, \tag{38}$$

$$\frac{\partial U_2(\hat{x}_1^*, x_2, \theta_2)}{\partial x_2} = 0, \tag{39}$$

$$\frac{\partial \hat{U}_1(x_1, x_2^*, \hat{\theta}_1)}{\partial x_1} = 0. \tag{40}$$

Let M be the Jacobian of this system of the equations. Then each ij-component of the matrix coincides with M_{ij} defined in the main text.

By the regularity condition (iii), $\det M \neq 0$, so the implicit function theorem guarantees that for any parameter A close to A^*, there is an action profile x^* which satisfies the first-order conditions (38)-(40). These action profiles are globally optimal, because of the regularity conditions (i) and (ii). So this x^* is a steady state given the parameter A. The implicit function theorem also asserts that

$$\begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1^*}{\partial A} \\ \frac{\partial x_2^*}{\partial A} \\ \frac{\partial \hat{x}_1}{\partial A} \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ M_{3A} \end{bmatrix},$$

Solving this system of equations,

$$\frac{\partial \hat{x}_1^*}{\partial A} = -\frac{(M_{11}M_{22} - M_{12}M_{21})M_{3A}}{\det M},$$
$$\frac{\partial x_2^*}{\partial A} = \frac{M_{11}M_{23}M_{3A}}{\det M},$$
$$\frac{\partial x_1^*}{\partial A} = -\frac{M_{12}M_{23}M_{3A}}{\det M}.$$

Dividing both the numerator and the denominator of the second equation by $M_{11}M_{22}M_{33}$ and using $\det M = M_{11}M_{22}M_{33} + M_{12}M_{23}M_{31} - M_{12}M_{21}M_{33} - M_{11}M_{32}M_{23}$, we have

$$\begin{aligned}\frac{\partial x_2^*}{\partial A} &= -\frac{BR'_{23}}{1 - BR'_{12}BR'_{23}BR'_{31} - BR'_{12}BR'_{21} - BR'_{23}BR'_{32}} \cdot \frac{M_{3A}}{M_{33}} \\ &= -\frac{M_{3A}}{M_{33}}\left(\frac{BR'_{23}}{1 - BR'_{23}BR'_{32}}\right)\left(\frac{1 - BR'_{23}BR'_{32}}{1 - BR'_{12}BR'_{23}BR'_{31} - BR'_{12}BR'_{21} - BR'_{23}BR'_{32}}\right) \\ &= -\frac{M_{3A}}{M_{33}}\left(\frac{BR'_{23}}{1 - BR'_{23}BR'_{32}}\right)\left(\frac{1}{1 - BR'_{12}NE'_2}\right).\end{aligned}$$

The second equation in the proposition follows from the second and the third equations.

A.3 Proof of Proposition 3

Pick A^* and x^* as stated. Since x^* is an interior point, it must satisfy the first-order conditions

$$\frac{\partial U_1(x_1, x_2^*, \theta^*)}{\partial x_1} = 0, \tag{41}$$
$$\frac{\partial U_2(\hat{x}_1^*, x_2, \theta_2)}{\partial x_2} = 0, \tag{42}$$
$$\frac{\partial \hat{U}_1(x_1, x_2^*, \theta_2)}{\partial x_1} = 0. \tag{43}$$

Let M be the Jacobian of this system of the equations. Then each ij-component of the matrix coincides with M_{ij} defined in the main text.

By the regularity condition (iii), $\det M \neq 0$, so the implicit function theorem guarantees that for any parameter A close to A^*, there is an action profile x^* which satisfies the first-order conditions (41)-(43). These action profiles are globally optimal, because of the regularity conditions (i) and (ii). So this x^* is a steady state given the parameter A. The implicit function theorem also asserts that

$$\begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1^*}{\partial A} \\ \frac{\partial x_2^*}{\partial A} \\ \frac{\partial \hat{x}_1}{\partial A} \end{bmatrix} = - \begin{bmatrix} 0 \\ M_{2A} \\ M_{3A} \end{bmatrix},$$

Solving this system of equations,

$$\frac{\partial \hat{x}_1^*}{\partial A} = -\frac{(M_{11}M_{22} - M_{12}M_{21})M_{3A} - (M_{11}M_{32} - M_{12}M_{31})M_{2A}}{\det M},$$
$$\frac{\partial x_2^*}{\partial A} = -\frac{M_{11}(M_{33}M_{2A} - M_{23}M_{3A})}{\det M},$$
$$\frac{\partial x_1^*}{\partial A} = \frac{M_{12}(M_{33}M_{2A} - M_{23}M_{3A})}{\det M}.$$

The rest of the proof is very similar to that of Proposition 2, and hence omitted.

A.4 Proof of Proposition 4

Pick A^* and x^* as stated. Since x^* is an interior point, it must satisfy the first-order conditions

$$\frac{\partial U_1(x_1, \hat{x}_2^*, \theta_1)}{\partial x_1} = 0, \tag{44}$$
$$\frac{\partial U_2(\hat{x}_1^*, x_2, \theta_2)}{\partial x_2} = 0, \tag{45}$$
$$\frac{\partial \hat{U}_1(\hat{x}_1, x_2^*, \theta_2)}{\partial \hat{x}_1} = 0, \tag{46}$$
$$\frac{\partial \hat{U}_2(x_1^*, \hat{x}_2, \theta_1)}{\partial \hat{x}_2} = 0. \tag{47}$$

Let M be the Jacobian of this system of the equations. Then each ij-component of the matrix coincides with M_{ij} defined in the main text.

By the regularity condition (iii), $\det M \neq 0$, so the implicit function theorem guarantees that for any parameter A_2 close to A_2^*, there is an action profile x^* which satisfies the first-order conditions (44)-(47). These action profiles are globally optimal, because of the regularity conditions (i) and (ii). So this x^* is a steady state given the parameter A. The implicit function theorem also asserts that

$$\begin{bmatrix} M_{11} & M_{12} & M_{13} & M_{14} \\ M_{21} & M_{22} & M_{23} & M_{24} \\ M_{31} & M_{32} & M_{33} & M_{34} \\ M_{41} & M_{42} & M_{43} & M_{44} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1^*}{\partial A_2} \\ \frac{\partial x_2^*}{\partial A_2} \\ \frac{\partial \hat{x}_1^*}{\partial A_2} \\ \frac{\partial \hat{x}_2^*}{\partial A_2} \end{bmatrix} = - \begin{bmatrix} 0 \\ M_{2A} \\ M_{3A} \\ 0 \end{bmatrix},$$

Solving this system and using $M_{13} = M_{24} = M_{34} = M_{43} = 0$,

$$\frac{\partial x_2^*}{\partial A_2} = \frac{(M_{14}M_{41}M_{33} - M_{11}M_{33}M_{44})M_{2A} + (M_{11}M_{23}M_{44} - M_{23}M_{14}M_{41})M_{3A}}{\det M}$$
$$\frac{\partial x_1^*}{\partial A_2} = \frac{(M_{12}M_{33}M_{44} - M_{33}M_{42}M_{14})M_{2A} - (M_{23}M_{44}M_{12} - M_{23}M_{42}M_{14})M_{3A}}{\det M}.$$

Dividing both the numerator and the denominator of the first equation by $M_{11}M_{22}M_{33}M_{44}$,

$$\begin{aligned} \frac{\partial x_2^*}{\partial A_2} &= - \left\{ (1 - BR_{14}BR_{41}) \frac{M_{2A}}{M_{22}} + (BR_{23} - BR_{23}BR_{14}BR_{41}) \frac{M_{3A}}{M_{33}} \right\} \frac{M_{11}M_{22}M_{33}M_{44}}{\det M} \\ &= - (1 - BR_{14}BR_{41}) \left(\frac{M_{2A}}{M_{22}} + BR_{23} \frac{M_{3A}}{M_{33}} \right) \frac{M_{11}M_{22}M_{33}M_{44}}{\det M}. \end{aligned}$$

Note that

$$\begin{aligned} \det M =& M_{11}M_{22}M_{33}M_{44} + M_{14}M_{21}M_{33}M_{42} - M_{14}M_{22}M_{33}M_{41} - M_{12}M_{21}M_{33}M_{44} \\ & - M_{11}M_{23}M_{32}M_{44} - M_{14}M_{23}M_{31}M_{42} + M_{14}M_{23}M_{32}M_{41} + M_{12}M_{23}M_{31}M_{44}, \end{aligned}$$

so

$$\begin{aligned} \frac{M_{11}M_{22}M_{33}M_{44}}{\det M} &= \frac{1}{(1 - BR'_{14}BR'_{41})(1 - BR'_{23}BR'_{32}) - (BR'_{12} + BR'_{14}BR'_{42})(BR'_{21} + BR'_{23}BR'_{31})} \\ &= \left(\frac{1}{(1 - BR'_{14}BR'_{41})(1 - BR'_{23}BR'_{32})} \right) \left(\frac{1}{1 - NE'_1 NE'_2} \right) \end{aligned}$$

Plugging this int the equation above, we obtain the first equation in the proposition. The second equation can be derived in a similar way.

A.5 Proof of Proposition 5

Proof of Part (i). When $A = a$, the one-shot Nash equilibrium x^{correct} is a steady state under first-order misspecification, second-order misspecification, and double misspecification. By the assumption, a steady state is unique, and hence the result follows.

Proof of Part (ii). Let M^{first}, M^{second}, and M^{double} denote the matrix M defined in Sections 3, 4.1, and 4.2, respectively. Define M^{first}_{2A}, M^{double}_{2A}, M^{second}_{3A}, and M^{double}_{3A} in a similar way.

As shown in the proof of Proposition 3,

$$\frac{\partial x_2^{\text{double}}}{\partial A} = -\frac{M_{11}^{\text{double}}(M_{33}^{\text{double}}M_{2A}^{\text{double}} - M_{23}^{\text{double}}M_{3A}^{\text{double}})}{\det M^{\text{double}}}.$$

Since $M^{\text{double}} = M^{\text{second}}$ and $M^{\text{second}}_{3A} = M^{\text{double}}_{3A}$ at $A = a$,

$$\frac{\partial x_2^{\text{double}}}{\partial A} = -\frac{M_{11}^{\text{double}}M_{33}^{\text{double}}M_{2A}^{\text{double}}}{\det M^{\text{double}}} + \frac{M_{11}^{\text{second}}M_{23}^{\text{second}}M_{3A}^{\text{second}}}{\det M^{\text{second}}}.$$

As shown in the proof of Proposition 2, $\frac{\partial x_2^{\text{second}}}{\partial A} = \frac{M_{11}^{\text{second}}M_{23}^{\text{second}}M_{3A}^{\text{second}}}{\det M^{\text{second}}}$, so

$$\frac{\partial x_2^{\text{double}}}{\partial A} = -\frac{M_{11}^{\text{double}}M_{33}^{\text{double}}M_{2A}^{\text{double}}}{\det M^{\text{double}}} + \frac{\partial x_2^{\text{second}}}{\partial A}. \tag{48}$$

As $M^{\text{double}}_{31} = 0$, note that

$$\begin{aligned}\det M^{\text{double}} =& M_{11}^{\text{double}}M_{22}^{\text{double}}M_{33}^{\text{double}} + M_{12}^{\text{double}}M_{23}^{\text{double}}M_{31}^{\text{double}} \\ & - M_{12}^{\text{double}}M_{21}^{\text{double}}M_{33}^{\text{double}} - M_{11}^{\text{double}}M_{23}^{\text{double}}M_{32}^{\text{double}}.\end{aligned}$$

Since $\frac{\partial \theta_2^{double}}{\partial x_2} = \frac{Q_{x_2}(x_1^*,x_2^*,a,\theta^*) - Q_{x_2}(\hat{x}_1^*,x_2^*,A,\theta_2)}{Q_{\theta_2}(\hat{x}_1^*,x_2^*,A,\theta_2)} = 0$ and hence $M^{\text{double}}_{12} = M^{\text{double}}_{32}$ at $A = a$,

$$\begin{aligned}\det M^{\text{double}} =& M_{11}^{\text{double}}M_{22}^{\text{double}}M_{33}^{\text{double}} \\ & + M_{12}^{\text{double}}\{M_{23}^{\text{double}}(M_{31}^{\text{double}} - M_{11}^{\text{double}}) - M_{21}^{\text{double}}M_{33}^{\text{double}}\}.\end{aligned}$$

Since $\frac{\partial \theta_2^{double}}{\partial x_1} = \frac{Q_{x_1}(x_1^*,x_2^*,a,\theta^*)}{Q_{\theta_2}(\hat{x}_1^*,x_2^*,A,\theta_2)}$ and $\frac{\partial \theta_2^{double}}{\partial \hat{x}_1} = -\frac{Q_{x_1}(\hat{x}_1^*,x_2^*,A,\theta_2)}{Q_{\theta_2}(\hat{x}_1^*,x_2^*,A,\theta_2)}$, $M^{\text{double}}_{11} = M^{\text{double}}_{31} + M^{\text{double}}_{33}$ at $A = a$, and hence

$$\det M^{\text{double}} = M_{33}^{\text{double}}\{M_{11}^{\text{double}}M_{22}^{\text{double}} - M_{12}^{\text{double}}(M_{23}^{\text{double}} + M_{21}^{\text{double}})\}.$$

Then since $M_{11}^{\text{double}} = M_{11}^{\text{first}}$, $M_{22}^{\text{double}} = M_{22}^{\text{first}}$, $M_{12}^{\text{double}} = M_{12}^{\text{first}}$, and $M_{23}^{\text{double}} + M_{21}^{\text{double}} = M_{21}^{\text{first}}$ at $A = a$,

$$\det M^{\text{double}} = M_{33}^{\text{double}} \det M^{\text{first}}. \tag{49}$$

Plugging this into (48) and using $M_{11}^{\text{double}} = M_{11}^{\text{first}}$ and $M_{2A}^{\text{double}} = M_{2A}^{\text{first}}$,

$$\frac{\partial x_2^{\text{double}}}{\partial A} = -\frac{M_{11}^{\text{first}} M_{2A}^{\text{first}}}{\det M^{\text{first}}} + \frac{\partial x_2^{\text{second}}}{\partial A} = \frac{\partial x_2^{\text{first}}}{\partial A} + \frac{\partial x_2^{\text{second}}}{\partial A}.$$

Proof of Part (iii). We first prove the second equation. As shown in the proof of Proposition 2,

$$\frac{\partial x_2^{\text{second}}}{\partial A} = \frac{M_{11}^{\text{second}} M_{23}^{\text{second}} M_{3A}^{\text{second}}}{\det M^{\text{second}}}.$$

From (49) and $M^{\text{second}} = M^{\text{double}}$,

$$\frac{\partial x_2^{\text{second}}}{\partial A} = \frac{M_{11}^{\text{second}} M_{23}^{\text{second}} M_{3A}^{\text{second}}}{M_{33}^{\text{second}} \det M^{\text{first}}}.$$

When the game is symmetric, we have $M_{3A}^{\text{second}} = M_{2A}^{\text{first}}$. Also $M_{11}^{\text{second}} = M_{11}^{\text{first}}$, so

$$\frac{\partial x_2^{\text{second}}}{\partial A} = -\frac{M_{23}^{\text{second}}}{M_{33}^{\text{second}}} \frac{\partial x_2^{\text{first}}}{\partial A}.$$

Then the result follows from $M_{23}^{\text{second}} = U_{ij} - L$ and $M_{33}^{\text{second}} = U_{ii} - L$.

The first equation in the proposition follows from this result and Proposition 2. The third equation in the proposition follows from part (ii) of this proposition.

We conclude the proof by showing the remaining two equations in the proposition. Let $L_i = \frac{Q_{x_1}}{Q_\theta} \frac{\partial^2 U_i}{\partial x_i \partial \theta}$. When $A_1 = A_2 = a$, the multiplier effect on $\frac{\partial x_2}{\partial A_2}$ appearing in Proposition 4 can be rewritten as

$$\frac{1 - \frac{U_{12}-L_1}{U_{11}} \frac{U_{21}}{U_{22}-L_2}}{\left(1 - \frac{U_{12}-L_1}{U_{11}} \frac{U_{21}}{U_{22}-L_2}\right)\left(1 - \frac{U_{21}-L_2}{U_{22}} \frac{U_{12}}{U_{11}-L_1}\right) - \left(\frac{L_2}{U_{22}-L_2} \frac{U_{12}-L_1}{U_{11}} - \frac{L_1}{U_{11}}\right)\left(\frac{L_1}{U_{11}-L_1} \frac{U_{21}-L_2}{U_{22}} - \frac{L_2}{U_{22}}\right)}$$
$$= \frac{U_{22}(U_{11}-L_1)(U_{11}U_{22} - U_{11}L_2 - U_{21}U_{12} + U_{21}L_1)}{(U_{11}U_{22} - U_{11}L_2 - U_{21}U_{12} + U_{21}L_1)(U_{11}U_{22} - U_{22}L_1 - U_{12}U_{21} + U_{12}L_1) - (U_{12}L_1 - U_{22}L_1)(U_{21}L_1 - U_{11}L_2)}.$$

When the game is symmetric, this reduces to

$$
\begin{aligned}
&\frac{U_{ii}(U_{ii}-L)(U_{ii}-U_{ij})(U_{ii}+U_{ij}-L)}{(U_{ii}-U_{ij})^2(U_{ii}+U_{ij}-L)^2-L^2(U_{ii}-U_{ij})^2}\\
=&\frac{U_{ii}(U_{ii}-L)(U_{ii}+U_{ij}-L)}{(U_{ii}-U_{ij})(U_{ii}^2+U_{ij}^2+2U_{ii}U_{ij}-2U_{ii}L-2U_{ij}L)}\\
=&\frac{U_{ii}(U_{ii}-L)(U_{ii}+U_{ij}-L)}{(U_{ii}-U_{ij})(U_{ii}+U_{ij})(U_{ii}+U_{ij}-2L)}\\
=&\frac{U_{ii}^2}{U_{ii}^2-U_{ij}^2}\cdot\frac{(U_{ii}-L)(U_{ii}+U_{ij}-L)}{U_{ii}(U_{ii}+U_{ij}-2L)}\\
=&\frac{1}{1-BR_1'BR_2'}\cdot\frac{(U_{ii}-L)(U_{ii}+U_{ij}-L)}{U_{ii}(U_{ii}+U_{ij}-2L)}
\end{aligned}
$$

This and Proposition 4 imply the fourth equation in the proposition. Also, the last equation follows from $BR_{12}'=BR_{42}'=-\frac{L}{U_{ii}}$ and $BR_{14}'=-\frac{U_{ij}}{U_{ii}}$ at $A=a$.

A.6 Proof of Corollary 2

Let U_{ii}, U_{ij}, and L be as stated in Proposition 5. That is,

$$
\begin{aligned}
U_{ii}&=\frac{\partial^2 U_i}{\partial^2 x_i}=2Q_x+x_iQ_{xx}-c''<0,\\
U_{ij}&=\frac{\partial^2 U_i}{\partial x_i\partial x_j}=Q_x+x_iQ_{xx}=Q_x+x_iQ_{xx}<0,\\
L&=\frac{Q_{x_i}}{Q_\theta}\frac{\partial^2 U_i}{\partial x_i\partial\theta}=\frac{Q_x}{Q_\theta}(x_iQ_{x\theta}+Q_\theta)<0.
\end{aligned}
$$

Note that

$$U_{ii}<U_{ij}<0. \tag{50}$$

Proof of part (i). First, consider the one-shot game with first-order misspecification. A Nash equilibrium action solves the first-order condition

$$\frac{\partial U_1(x,a,\theta^*)}{\partial x_1}=0 \text{ and } \frac{\partial U_2(x,A,\theta^*)}{\partial x_2}=0.$$

By the implicit function theorem, we have $\frac{\partial x_2}{\partial A} = \frac{1}{1-BR_1'BR_2'}(-\frac{M_{2A}^d}{U_{22}})$ where $M_{2A}^d = Q_A + x_2 Q_{xA}$ is the direct effect of misspecification on player 2's marginal utility, and BR_i' is the slope of the standard best-response function.[36] As explained in the main text, the direct effect is positive. Also, (50) implies $\frac{1}{1-BR_1'BR_2'} > 1$. Hence, we have $\frac{\partial x_2}{\partial A} > 0$. Firm 1 best-responds to this firm 2's action, so $\frac{\partial x_1}{\partial A} = BR_1' \frac{\partial x_2}{\partial A}$. Since (50) implies $BR_1' = -\frac{U_{12}}{U_{11}} \in (-1, 0)$, we have $\frac{\partial x_1}{\partial A} < 0$ and $\frac{\partial x_1}{\partial A} + \frac{\partial x_2}{\partial A} > 0$. For payoffs, note that at $A = a$, we have

$$\frac{\partial \pi_i}{\partial A} = \frac{\partial \pi_i}{\partial x_i}\frac{\partial x_i}{\partial A} + \frac{\partial \pi_i}{\partial x_{-i}}\frac{\partial x_{-i}}{\partial A} = \frac{\partial \pi_i}{\partial x_{-i}}\frac{\partial x_{-i}}{\partial A} = x_i Q_x \frac{\partial x_{-i}}{\partial A},$$

where the second inequality follows from the fact that π_i is an equilibrium payoff so that $\frac{\partial \pi_i}{\partial x_i} = 0$. This immediately implies $\frac{\partial \pi_1}{\partial A} < 0$, $\frac{\partial \pi_2}{\partial A} > 0$, and $\frac{\partial \pi_1}{\partial A} + \frac{\partial \pi_2}{\partial A} > 0$.

Second, consider the infinite-horizon model with first-order misspecification. Since we assume that the base misspecification effect is positive, from Proposition 1 and $\frac{1}{1-BR_1'BR_2'} > 1$ at $A = a$, we have $\frac{\partial x_2}{\partial A} > 0$. The remaining inequalities can be shown as in the one-shot game.

Third, consider the one-shot game with one-sided double misspecification. Player 2 and the hypothetical player 1 plays a Nash equilibrium, which solves the first-order condition

$$\frac{\partial \hat{U}_1(\hat{x}_1, x_2, A, \theta^*)}{\partial \hat{x}_1} = 0 \text{ and } \frac{\partial U_2(\hat{x}_1, x_2, A, \theta^*)}{\partial x_2} = 0. \tag{51}$$

By the implicit function theorem, we have $\frac{\partial x_2}{\partial A} = \frac{1}{1-BR_1'BR_2'}(-\frac{M_{2A}^d}{U_{22}} - BR_2' \frac{M_{1A}^d}{U_{11}}) = \frac{1}{1-BR_1'BR_2'}(-\frac{M_{2A}^d}{U_{22}})(1 + BR_2') > 0$ where $M_{1A}^d = Q_A + \hat{x}_1 Q_{xA} = M_{2A}^d$. The remaining inequalities follow as in the case with first-order misspecification.

Finally, consider the infinite-horizon game with one-sided double misspecification. From (50) and $U_{ii} - L < 0$, we have $1 - \frac{U_{ij} - L}{U_{ii} - L} = \frac{U_{ii} - U_{ij}}{U_{ii} - L} > 0$. Then Proposition 5 and $\frac{\partial x_2^{\text{first}}}{\partial A} > 0$ imply $\frac{\partial x_2^{\text{double}}}{\partial A} > 0$. The remaining inequalities can be shown as in the case with first-order misspecification.

[36] The asymptotic best response curve coincides with the standard best response curve at $A = a$, so we use the same notation.

Proof of part (ii). Consider the one-shot game with second-order misspecification. Player 2 and the hypothetical player 1 plays a Nash equilibrium, which solves the first-order condition

$$\frac{\partial \hat{U}_1(\hat{x}_1, x_2, A, \theta^*)}{\partial \hat{x}_1} = 0 \text{ and } \frac{\partial U_2(\hat{x}_1, x_2, a, \theta^*)}{\partial x_2} = 0.$$

By the implicit function theorem, we have $\frac{\partial x_2}{\partial A} = \frac{1}{1 - BR_1' BR_2'}(-BR_2' \frac{M_{1A}^d}{U_{11}}) < 0$. Then the remaining inequalities follow as in the case with first-order misspecification.

Next, consider the infinite-horizon game with second-order misspecification. Simple algebra shows that

$$U_{ij} - L = \begin{cases} < 0 & \text{if } \frac{Q_{x\theta}}{Q_\theta} < \frac{Q_{xx}}{Q_x}, \\ = 0 & \text{if } \frac{Q_{x\theta}}{Q_\theta} = \frac{Q_{xx}}{Q_x}, \\ > 0 & \text{if } \frac{Q_{x\theta}}{Q_\theta} > \frac{Q_{xx}}{Q_x}. \end{cases}$$

Then, Proposition 5, $U_{ii} - L < 0$, and $\frac{\partial x_2^{\text{first}}}{\partial A} > 0$ imply the result.

Proof of part (iii). Consider the one-shot game with two-sided double misspecification. Player 1 is not aware of player 2's bias, so she chooses the Nash equilibrium x_1^{correct} regardless of the parameter A_2. Player 2 and the hypothetical player 1 play a Nash equilibrium which solves the first-order condition (51), and as shown in part (i), we have $\frac{\partial x_2}{\partial A} > 0$.

Next, consider the infinite-horizon game. Define NE_i' as in Proposition 4. Note that $x_1 = x_2 = \hat{x}_1 = \hat{x}_2 = x^*$ constitutes the steady state at $A_1 = A_2 = a$. A necessary condition for the unique steady state is $|NE_1' NE_2'| = |NE_i'|^2 \leq 1$ at x^*; otherwise, as discussed in the proof of Proposition 1, there exist multiple steady states. Also, $\det M \neq 0$ implies $|NE_i'|^2 \neq 1$. So for a steady state to be unique, we must have $|NE_i'|^2 < 1$ at x^*, which implies $NE_i' \in (-1, 1)$ at x^*. At x^*,

$$NE_i' = -\frac{\frac{\partial^2 U_i}{\partial x_i \partial \theta_i} \frac{\partial \theta_i}{\partial x_j}}{U_{ii} + U_{ij} + \frac{\partial^2 U_i}{\partial x_i \partial \theta_i}\left(\frac{\partial \theta_i}{\partial x_i} + \frac{\partial \theta_i}{\partial x_j}\right)} = -\frac{L}{U_{ii} + U_{ij} - L}.$$

Because $U_{ii} - L < 0$ and $U_{ij} < 0$, the condition $NE_i' \in (-1, 1)$ is equivalent to $U_{ii} + U_{ij} - L < L < -(U_{ii} + U_{ij} - L)$. The first inequality implies $U_{ii} + U_{ij} - 2L < 0$.

Then, Proposition 5 and $\frac{\partial x_2^{\text{first}}}{\partial A} > 0$ imply the result. *Q.E.D.*

A.7 Proof of Corollary 3

Let U_{ii}, U_{ij}, and L be as stated in Proposition 5. That is,

$$U_{ii} = \frac{\partial^2 U_i}{\partial^2 x_i} = Q_{x_i x_i} - c''(x_i),$$
$$U_{ij} = \frac{\partial^2 U_i}{\partial x_i \partial x_j} = Q_{x_1 x_2},$$
$$L = \frac{Q_{x_i}}{Q_\theta} \frac{\partial^2 U_i}{\partial x_i \partial \theta} = \frac{Q_{x_i}}{Q_\theta} Q_{x_i \theta}.$$

Suppose that $A = a$. From Proposition 1, we have $BR'_1 BR'_2 < 1$ at the unique steady state x^{correct}. Recall that at $A = a$, BR'_i is simply the slope of the best response curve in the standard sense. So by symmetry, we have $BR'_1 = BR'_2$, which in turn implies $|BR'_i| = |\frac{U_{ij}}{U_{ii}}| < 1$. Note also that $U_{ii} < 0$, because the second-order condition for the incentive-compatibility condition must be satisfied according to the regularity condition imposed in Proposition 1. Taken together, we have $U_{ii} < U_{ij}$.

Proof of part (i). The result for the one-shot game can be shown as in the proof of Corollary 2. So we will prove only the result for the infinite-horizon game. Since the base misspecification effect is negative, we have $\frac{\partial x_2}{\partial A} < 0$ under first-order misspecification. Also, since $U_{ii} < 0$, $L > 0$, and $U_{ii} < U_{ij}$, we have

$$1 - \frac{U_{ij} - L}{U_{ii} - L} = \frac{U_{ii} - U_{ij}}{U_{ii} - L} > 0.$$

Then from Proposition 5, we have $\frac{\partial x_2}{\partial A} < 0$ even under one-sided double misspecification. The remaining inequalities can be shown as in the proof of Corollary 2.

Proof of part (ii). Again, we will prove only the result for the infinite-horizon game. If $Q_{x_1 x_2} > L$ as stated in (a), then

$$0 < -\frac{U_{ij} - L}{U_{ii} - L} < 1.$$

Here the first inequality follows from $U_{ii} < 0$, $L > 0$, and $U_{ij} - L = Q_{x_1 x_2} - L > 0$, and the second inequality follows from $U_{ii} < 0$, $L > 0$, and $|\frac{U_{ij}}{U_{ii}}| < 1$. Then from Proposition 5, $\frac{\partial x_2}{\partial A} < 0$ under second-order misspecification.

On the other hand, if $Q_{x_1x_2} < L$ as stated in (b) and (c), then

$$-\frac{U_{ij}-L}{U_{ii}-L} < 0.$$

Hence $\frac{\partial x_2}{\partial A} > 0$ under second-order misspecification. The remaining inequalities can be shown as in the proof of Corollary 2.

The proof of part (iii) is also similar to that of corollary 2. *Q.E.D.*

A.8 Proof of Lemma 1

For the case in which X is finite, this is exactly the same as Lemma 1 of Esponda, Pouzo, and Yamamoto (2019). For the case in which X is continuous, we need a minor modification of the proof. We first prove a preliminary lemma:

Lemma 3. *Assume that X is continuous. Under Assumption 1(iii) and (iv), $\int_Y g(x,y)Q(dy|x)$ is bounded and continuous in x.*

Proof. Take a sequence x^n converging to x. Then

$$\begin{aligned}
&\int_Y g(x^n,y)Q(dy|x^n) - \int_Y g(x,y)Q(dy|x) \\
\leq &\left|\int_Y g(x^n,y)Q(dy|x^n) - \int_Y g(x^n,y)Q(dy|x)\right| \\
&+\left|\int_Y g(x^n,y)Q(dy|x) - \int_Y g(x,y)Q(dy|x)\right|.
\end{aligned}$$

From Assumption 1(iii), $Q(dy|x^n)$ weakly converges to $Q(dy|x)$, so the first term of the right-hand side converges to zero. Also from Assumption 1(iv-a), $g(x^n,y)$ pointwise converges to $g(x,y)$, so the second term converges to zero. *Q.E.D.*

As shown in the display in EPY, we have

$$\begin{aligned}
K_{i,k}(\theta^n_{i,k},\sigma^n) - K_i(\theta^n_{i,k},\sigma) \leq &\int_X\int_Y g(x,y)Q(dy|x)\sigma^n_{\hat{X}_{1,1}\times\hat{X}_{2,1}}(dx) \\
&-\int_X\int_Y g(x,y)Q(dy|x)\sigma_{\hat{X}_{1,1}\times\hat{X}_{2,1}}(dx)
\end{aligned}$$

where $\sigma_{\hat{X}_{1,1}\times\hat{X}_{2,1}}$ and $\sigma^n_{\hat{X}_{1,1}\times\hat{X}_{2,1}}$ are the marginals of σ and σ^n on $\hat{X}_{1,1}\times\hat{X}_{2,1}$, respectively. From Lemma 3, the right-hand side converges to zero as $\sigma^n\to\sigma$. The rest of the proof is exactly the same as in EPY.

A.9 Proof of Proposition 6

For the special case in which X is finite, Theorem 1 of Esponda, Pouzo, and Yamamoto (2019) proves the same result. We need a minor modification to their proof, as they use finiteness of X in Step 2 in the proof of Lemma 2.

Pick i, k, $\theta_{i,k}$. Then let

$$f_l(\hat{x}) = E_{Q(\cdot|\hat{x}_{1,1},\hat{x}_{2,1})}\left[\sup_{\theta'_{i,k}\in O(\theta_{i,k},\frac{1}{l})}\left|\frac{q(y|\hat{x}_{1,1},\hat{x}_{2,1})}{q_{\theta_{i,k}}(y|\hat{x}_{i,k},\hat{x}_{i,k+1})} - \frac{q(y|\hat{x}_{1,1},\hat{x}_{2,1})}{q_{\theta'_{i,k}}(y|\hat{x}_{i,k},\hat{x}_{i,k+1})}\right|\right]$$

where $O(\theta_{i,k},\frac{1}{l})$ is a $\frac{1}{l}$-neighborhood of $\theta_{i,k}$. Then as explained at the end of the the first paragraph in EPY's step 2, $\lim_{l\to\infty} f_l(\hat{x}) \to 0$ for each $\hat{x}$. In what follows, we will show that this convergence is uniform in $\hat{x}$; then there is $\delta(\theta_{i,k},\varepsilon)$ with which (16) of EPY holds, and the rest of the proof is exactly the same as EPY's.

Pick an arbitrary $\varepsilon > 0$. For each $\hat{x}$, let $F(\hat{x}) = \{l \in [0,\infty) | f_l(\hat{x}) \geq \varepsilon\}$. Then we have the following lemma:

Lemma 4. *For each $\hat{x}$, there is $l(\hat{x}) > 0$ such that $F(\hat{x}) = [0, l(\hat{x})]$. Also $F(\hat{x})$ is upper hemi-continuous in $\hat{x}$.*

Proof. The first part follows from the fact that $f_l(\hat{x})$ is continuous and decreasing in l, and $\lim_{l\to\infty} f_l(\hat{x}) = 0$.

To prove the second part, pick $\hat{x}$ and an arbitrary small $\eta > 0$. Then $f_{l(\hat{x})+\eta}(\hat{x}) < \varepsilon$. Since $f_l(\hat{x})$ is continuous in $\hat{x}$, there is an open neighborhood U of $\hat{x}$ such that $f_{l(\hat{x})+\eta}(\hat{x}') < \varepsilon$ for all $\hat{x}' \in U$. This implies that $l(\hat{x}') < l(\hat{x}) + \eta$ for all $\hat{x}' \in U$. *Q.E.D.*

The above lemma implies that $l(\hat{x})$ is an upper hemi-continuous function, and from the Maximum theorem, $l(\hat{x})$ is bounded; $l(\hat{x}) < l^*$ for some l^*. Hence $f_l(\hat{x}) \leq \varepsilon$ for all $\hat{x}$ and $l \geq l^*$, implying uniform convergence.

A.10 Proof of Proposition 8

This is very similar to the first step of the proof of Proposition 2 in EPY. However, we need a minor modification, as X may not be finite in our setup. We first prove upper hemi-continuity of $B_\varepsilon(\sigma)$.

Lemma 5. *$B_\varepsilon(\sigma)$ is upper hemi-continuous in (ε,σ).*

Proof. Since $\prod_{i=1}^{2}\prod_{k=1}^{k_i+1}\triangle\Theta_{i,k}$ is compact, it is sufficient to show that $(\varepsilon^n,\sigma^n,\hat{\mu}^n)\to(\varepsilon,\sigma,\hat{\mu})$ and $\hat{\mu}^n\in B_{\varepsilon^n}(\sigma^n)$ for each n imply $\hat{\mu}\in B_\varepsilon(\sigma)$. Note that

$$\begin{aligned}
&\lim_{n\to\infty}\left(\int_{\Theta_{i,k}}(K_{i,k}(\theta_{i,k},\sigma^n)\hat{\mu}^n_{i,k}(d\theta_{i,k})-\int_{\Theta_{i,k}}(K_{i,k}(\theta_{i,k},\sigma)\hat{\mu}_{i,k}(d\theta_{i,k})\right)\\
&\quad=\lim_{n\to\infty}\left(\int_{\Theta_{i,k}}(K_{i,k}(\theta_{i,k},\sigma^n)\hat{\mu}^n_{i,k}(d\theta_{i,k})-\int_{\Theta_{i,k}}(K_{i,k}(\theta_{i,k},\sigma)\hat{\mu}^n_{i,k}(d\theta_{i,k})\right)\\
&\qquad+\lim_{n\to\infty}\left(\int_{\Theta_{i,k}}(K_{i,k}(\theta_{i,k},\sigma)\hat{\mu}^n_{i,k}(d\theta_{i,k})-\int_{\Theta_{i,k}}(K_{i,k}(\theta_{i,k},\sigma)\hat{\mu}_{i,k}(d\theta_{i,k})\right).
\end{aligned}$$

The first term of the right-hand side is zero, because $K_{i,k}(\cdot,\sigma^n)$ pointwise converges to $K_{i,k}(\cdot,\sigma)$ (which follows from the fact that σ^n weakly converges to σ). Also the second term of the right-hand side is zero, as $\mu^n_{i,k}$ weakly converges to $\mu_{i,k}$.

$$\lim_{n\to\infty}\int_{\Theta_{i,k}}(K_{i,k}(\theta_{i,k},\sigma^n)\hat{\mu}^n_{i,k}(d\theta_{i,k})=\int_{\Theta_{i,k}}(K_{i,k}(\theta_{i,k},\sigma)\hat{\mu}_{i,k}(d\theta_{i,k}).$$

Since $\hat{\mu}^n\in B_{\varepsilon^n}(\sigma^n)$,

$$\int_{\Theta_{i,k}}(K_{i,k}(\theta_{i,k},\sigma^n)-K^*_{i,k}(\sigma^n))\hat{\mu}^n_{i,k}(d\theta_{i,k})\le\varepsilon^n.$$

Taking $n\to\infty$ and using continuity of $K^*_{i,k}(\sigma)$ (which follows from the theory of maximum),

$$\int_{\Theta_{i,k}}(K_{i,k}(\theta_{i,k},\sigma)-K^*_{i,k}(\sigma))\hat{\mu}_{i,k}(d\theta_{i,k})\le\varepsilon.$$

Hence $\mu\in B_\varepsilon(\sigma)$, which implies upper hemi-continuity of $B_\varepsilon(\sigma)$. *Q.E.D.*

Now we show that $S_\varepsilon(\sigma)$ is upper hemi-continuous at $\varepsilon=0$. Since X is compact, it suffices to show that $(\varepsilon^n,\sigma^n,x^n)\to(0,\sigma,x)$ and $x^n\in S_{\varepsilon^n}(\sigma^n)$ for each n, imply $x\in S_\varepsilon(\sigma)$. As noted earlier, we already know that $S_0(\sigma)$ is upper hemi-continuous in σ. So without loss of generality, we assume $\varepsilon^n>0$ for all n.

Since $x^n \in S_{\varepsilon^n}(\sigma^n)$, there is $\hat{\mu}^n \in B_{\varepsilon^n}(\sigma^n)$ with $x^n = \hat{s}(\hat{\mu}^n)$. The sequence $(\varepsilon^n, \sigma^n, x^n, \hat{\mu}^n)$ is in a compact set, so there is a convergent subsequence, still denoted by $(\varepsilon^n, \sigma^n, x^n, \hat{\mu}^n)$. Let $\hat{\mu} = \lim_{n\to\infty} \hat{\mu}^n$. Then $\hat{\mu} \in B_0(\sigma)$, as $B_\varepsilon(\sigma)$ is upper hemi-continuous and $\hat{\mu}^n \in B_{\varepsilon^n}(\sigma^n)$ for each n. Also, we have $x \in \hat{S}(\hat{\mu})$, because $\hat{S}$ is upper hemi-continuous and $x^n \in \hat{S}(\hat{\mu}^n)$ for each n. Hence $x \in S_0(\sigma)$.

A.11 Proof of Proposition 9

The proof is very similar to that of Theorem 2 of EPY. In EPY, the proof consists of three steps. In the first two steps, they show that $\boldsymbol{w}$ is a perturbed solution of the differential inclusion. Then in the last step, they show that a perturbed solution is an asymptotic pseudotrajectory (i.e., it satisfies (33)).

Our Propositions 7 and 8 imply that $\boldsymbol{w}$ is indeed a perturbed solution in the sense of EPY. We can also show that a perturbed solution is indeed an asymptotic pseudotrajectory. The proof is omitted because, other than replacing the Euclidean norm with the dual bounded-Lipschitz norm, it is exactly the same as the last step of EPY.[37]

A.12 Proof of Lemma 2

We will show that $\theta(\sigma)$ is Lipschitz continuous in σ. Under Assumptions 4(i) and (iii), the inverse $(\nabla^2 K_{i,k}(\theta_{i,k}(\sigma),\sigma))^{-1}$ of the Hessian matrix exists for each σ, and is continuous in σ. This means that $\|(\nabla^2 K_{i,k}(\theta_{i,k}(\sigma),\sigma))^{-1}\|$ is bounded and continuous in σ, where $\|C\| = \max_{ij} |c_{ij}|$ denotes the max norm of a matrix $C = \{c_{ij}\}$, . Since $\triangle\hat{X}$ is compact, there is L_1 such that $\|(\nabla^2 K_{i,k}(\theta_{i,k}(\sigma),\sigma))^{-1}\| < L_1$ for all i, k, and σ. Pick such L_1.

[37]This parallels Perkins and Leslie (2014), who show that the stochastic approximation technique of Benaïm (1999) for the Euclidean space extends to Banach spaces with the same proof. Our result differs from Perkins and Leslie (2014) in that we consider a differential inclusion, rather than a differential equation. But this does not cause any technical difficluty, because (i) $\triangle\hat{X}$ is a compact subset of a banach space with the dual bounded Lipschitz norm and (ii) Mazur's lemma, which is used to establish the result for differential inclusions in Euclidean spaces (Benaïm, Hofbauer, and Sorin (2005) and EPY), is valid even in Banach spaces.

Under Assumption 4(ii), there is $L_2 > 1$ such that

$$\left|\frac{\partial K_{i,k}(\theta_{i,k},\hat{x})}{\partial \theta_{i,k,m}} - \frac{\partial K(\theta_{i,k},\hat{x}')}{\partial \theta_{i,k,m}}\right| < L_2|\hat{x}-\hat{x}'|$$

for all i, k, m, $\theta_{i,k}$, $\hat{x}$, and $\hat{x}'$. Also, under Assumption 4(i), there is $L_3 > 1$ such that

$$\left|\frac{\partial K_{i,k}(\theta_{i,k},\hat{x})}{\partial \theta_{i,k,m}}\right| < L_3$$

for all i, k, m, $\theta_{i,k}$, and $\hat{x}$. Then for each σ and σ', we have

$$\begin{aligned}&\left|\frac{\partial K_{i,k}(\theta_{i,k},\sigma)}{\partial \theta_{i,k,m}} - \frac{\partial K_{i,k}(\theta_{i,k},\sigma')}{\partial \theta_{i,k,m}}\right| \\ &= \left|\int \frac{\partial K_{i,k}(\theta_{i,k},\hat{x})}{\partial \theta_{i,k,m}}\sigma(d\hat{x}) - \int \frac{\partial K_{i,k}(\theta_{i,k},\hat{x})}{\partial \theta_{i,k,m}}\sigma'(d\hat{x})\right| \leq 4L_2L_3\|\sigma-\sigma'\|\end{aligned}$$

where the inequality follows from the definition of the dual bounded-Lipschitz norm and the fact that $\frac{1}{4L_2L_3}\frac{\partial K_{i,k}(\theta_{i,k},\hat{x})}{\partial \theta_{i,k,m}} \in BL(\hat{X})$. This in turn implies that $\nabla K_{i,k}(\theta_{i,k},\sigma)$ is equi-Lipschitz continuous, that is, there is $L_4 > 0$ such that $|\nabla K_{i,k}(\theta_{i,k},\sigma) - \nabla K_{i,k}(\theta_{i,k},\sigma')| < L_4\|\sigma-\sigma'\|$ for all i, k, $\theta_{i,k}$, σ, and σ'.

Let $L = L_1L_4$. We will show that $\theta(\sigma)$ is Lipschitz continuous with the constant L. To do so, pick two action frequencies σ and $\sigma' \neq \sigma$ arbitrarily. For each $\beta \in [0,1]$, let $\sigma_\beta = \beta\sigma + (1-\beta)\sigma'$ denote a convex combination of σ and σ'. From Assumption 4(iii), the KL minimizer $\theta_{i,k}(\sigma_\beta)$ must solve the first-order condition

$$\nabla K_{i,k}(\theta_{i,k},\sigma_\beta) = 0,$$

which is equivalent to

$$\beta\nabla K_{i,k}(\theta_{i,k},\sigma) + (1-\beta)\nabla K_{i,k}(\theta_{i,k},\sigma') = 0.$$

Then by the implicit function theorem,

$$\frac{d\theta(\sigma_\beta)}{d\beta} = -(\nabla^2 K_{i,k}(\theta_{i,k}(\sigma_\beta),\sigma_\beta))^{-1}\left(\nabla K_{i,k}(\theta_{i,k}(\sigma_\beta),\sigma) - \nabla K_{i,k}(\theta_{i,k}(\sigma_\beta),\sigma')\right). \tag{52}$$

Using the fundamental theorem of calculus, we have

$$\begin{aligned}
&\theta(\sigma)-\theta(\sigma')\\
&=\theta(\sigma_1)-\theta(\sigma_0)\\
&=-\int_0^1 (\nabla^2 K_{i,k}(\theta_{i,k}(\sigma_\beta),\sigma_\beta))^{-1}\left(\nabla K_{i,k}(\theta_{i,k}(\sigma_\beta),\sigma)-\nabla K_{i,k}(\theta_{i,k}(\sigma_\beta),\sigma')\right)d\beta.
\end{aligned}$$

Then by the definition of L_1 and L_4,

$$|\theta(\sigma)-\theta(\tilde{\sigma})|\leq\int_0^1 L_1L_4\|\sigma-\sigma'\|d\beta=L\|\sigma-\sigma'\|,$$

as desired.

A.13 Proof of Proposition 12

We will first present a preliminary lemma. Pick an arbitrary action frequency $\boldsymbol{\sigma}(0)\in\triangle\hat{X}$ and a solution $\boldsymbol{\sigma}\in Z(\boldsymbol{\sigma}(0))$ to the differential inclusion (33) starting from this $\boldsymbol{\sigma}(0)$. Let $\boldsymbol{\theta}(t)=\theta(\boldsymbol{\sigma}(t))$ for each t. The following lemma shows that $\{\boldsymbol{\theta}(t)\}_{t\geq 0}$ solves (34).

Lemma 6. *Pick $t\geq 0$ such that (33) holds. Then $\dot{\boldsymbol{\theta}}(t)$ exists and satisfies (34).*

Proof. Pick t as stated, and pick $\boldsymbol{\sigma}^*\in\triangle S_0(\boldsymbol{\sigma}(t))$ such that $\dot{\boldsymbol{\sigma}}(t)=\boldsymbol{\sigma}^*-\boldsymbol{\sigma}(t)$. Let $\sigma_\beta=\beta\boldsymbol{\sigma}^*+(1-\beta)\boldsymbol{\sigma}(t)$ for each $\beta\in[0,1]$. Then we have

$$\frac{\theta(\boldsymbol{\sigma}(t+\varepsilon))-\theta(\boldsymbol{\sigma}(t))}{\varepsilon}=\left(\frac{\theta(\sigma_\varepsilon)-\theta(\sigma_0)}{\varepsilon}+\frac{\theta(\boldsymbol{\sigma}(t+\varepsilon))-\theta(\sigma_\varepsilon)}{\varepsilon}\right).$$

All we need to show is that the right-hand side has a limit as $\varepsilon\to 0$, and the limit is in the right-hand side of (34). Then $\frac{\theta(\sigma(t+\varepsilon))-\theta(\sigma(t))}{\varepsilon}$ also has a limit $\dot{\boldsymbol{\theta}}(t)$ and this limit value satisfies (34).

Note first that $\lim_{\varepsilon\to 0}\frac{\theta(\sigma_\varepsilon)-\theta(\sigma_0)}{\varepsilon}$ exists and is in the right-hand side of (34).

Indeed, from (52),

$$
\begin{aligned}
&\lim_{\varepsilon\to 0}\frac{\theta(\sigma_\varepsilon)-\theta(\sigma_0)}{\varepsilon}\\
&=\left.\frac{d\theta(\sigma_\beta)}{d\beta}\right|_{\beta=0}\\
&=-(\nabla^2 K_{i,k}(\theta_{i,k}(\sigma_0),\sigma_0))^{-1}\left(\nabla K_{i,k}(\theta_{i,k}(\sigma_0),\sigma_1)-\nabla K_{i,k}(\theta_{i,k}(\sigma_0),\sigma_0)\right)\\
&=-(\nabla^2 K_{i,k}(\theta_{i,k}(\boldsymbol{\sigma}(t)),\boldsymbol{\sigma}(t)))^{-1}\left(\nabla K_{i,k}(\theta_{i,k}(\boldsymbol{\sigma}(t)),\sigma^*)\right)
\end{aligned}
$$

where the second equality follows from the fact that $\theta_{i,k}(\sigma_0)$ solves the first-order condition.

We conclude the proof by showing that $\lim_{\varepsilon\to 0}\frac{\theta(\boldsymbol{\sigma}(t+\varepsilon))-\theta(\sigma_\varepsilon)}{\varepsilon}=0$. Since $\theta(\sigma)$ is Lipschitz continuous, there is $L>0$ such that

$$
\begin{aligned}
\left|\frac{\theta(\boldsymbol{\sigma}(t+\varepsilon))-\theta(\sigma_\varepsilon)}{\varepsilon}\right| &\le L\left\|\frac{\boldsymbol{\sigma}(t+\varepsilon))-\sigma_\varepsilon}{\varepsilon}\right\|\\
&=L\left\|\frac{(\boldsymbol{\sigma}(t+\varepsilon))-\boldsymbol{\sigma}(t))-(\sigma_\varepsilon-\sigma_0)}{\varepsilon}\right\|.
\end{aligned}
$$

Taking $\varepsilon\to 0$,

$$
\begin{aligned}
\lim_{\varepsilon\to 0}\left|\frac{\theta(\boldsymbol{\sigma}(t+\varepsilon))-\theta(\sigma_\varepsilon)}{\varepsilon}\right| &= L\left\|\lim_{\varepsilon\to 0}\frac{\boldsymbol{\sigma}(t+\varepsilon)-\boldsymbol{\sigma}(t)}{\varepsilon}-\lim_{\varepsilon\to 0}\frac{\sigma_\varepsilon-\sigma_0}{\varepsilon}\right\|\\
&=L\left\|\frac{d\boldsymbol{\sigma}(t)}{dt}-\left.\frac{d\sigma_\beta}{d\beta}\right|_{\beta=0}\right\|=0
\end{aligned}
$$

Q.E.D.

Now we prove the proposition. Pick $T>0$ and $h\in\mathscr{H}$ arbitrary. Pick any small $\varepsilon>0$. Since $\theta(\sigma)$ is uniformly continuous in σ (this follows from the continuity of θ and the compactness of $\triangle\hat{X}$), there is $\eta>0$ such that $|\theta(\sigma)-\theta(\tilde{\sigma})|<\varepsilon$ for any σ and $\tilde{\sigma}$ with $\|\sigma-\tilde{\sigma}\|<\eta$. From Proposition 9, there is t^* such that for any $t>t^*$, there is $\boldsymbol{\sigma}\in Z(\boldsymbol{w}(h)[t])$ such that

$$
\|\boldsymbol{w}(h)[t+\tau]-\boldsymbol{\sigma}(\tau)\|<\eta
$$

for all $\tau \in [0,T]$. Pick such $\boldsymbol{\sigma}$, and consider the corresponding $\boldsymbol{\theta}$, i.e., let $\boldsymbol{\theta}(t) = \theta(\boldsymbol{\sigma}(t))$ for each t. Then by the definition of η, we have

$$\|\boldsymbol{w}_{\theta}(h)[t+\tau] - \boldsymbol{\theta}(\tau)\| < \varepsilon$$

for all $\tau \in [0,T]$. Also this $\boldsymbol{\theta}$ solves (34).[38] This implies the result we want.

A.14 Proof of Proposition 13

Pick an arbitrary sample path $h \in \mathscr{H}$. Let $(i,k) \in I^{**} \setminus I^{*}$. Throughout the proof, for each $\theta_{i,k}$, let $\theta(\theta_{i,k})$ denote θ such that $\theta_{j,l} = \theta_{i,k}$ for all $(j,l) \in I(i,k)$ and $\theta_{j,l} = \theta^{*}_{j,l}$ for all $(j,l) \in I^{*}$.

We will show that $\lim_{t\to\infty} d(\theta^{t}_{i,k}(h), E_{i,k}) = 0$, where $E_{i,k}$ is the set of agent (i,k)'s steady state model, i.e., $E_{i,k}$ is the set of all $\theta_{i,k}$ such that $K'_{i,k}(\theta_{i,k}, \sigma') = 0$ for some $\sigma' \in \triangle S_0(\theta(\theta_{i,k}))$. There are two cases to be considered.

A.14.1 Case 1: $\liminf_{t\to\infty} \theta^{t}_{i,k}(h) \neq \limsup_{t\to\infty} \theta^{t}_{i,k}(h)$.

In this case, we will show that $[\liminf_{t\to\infty} \theta^{t}_{i,k}(h), \limsup_{t\to\infty} \theta^{t}_{i,k}(h)] \subseteq E_{i,k}$. This immediately implies $\lim_{t\to\infty} d(\theta^{t}_{i,k}(h), E_{i,k}) = 0$.

Suppose not, so that there is a model $\theta_{i,k} \in [\liminf_{t\to\infty} \theta^{t}_{i,k}(h), \limsup_{t\to\infty} \theta^{t}_{i,k}(h)]$ such that $\theta^{*}_{i,k} \notin E_{i,k}$. Then (i) $K'_{i,k}(\theta^{*}_{i,k}, \sigma') > 0$ for all $\sigma' \in \triangle S_0(\theta(\theta^{*}_{i,k}))$, or (ii) $K'_{i,k}(\theta^{*}_{i,k}, \sigma') < 0$ for all $\sigma' \in \triangle S_0(\theta(\theta^{*}_{i,k}))$. Indeed, if not and there are two measures $\sigma', \sigma'' \in \triangle S_0(\theta(\theta^{*}_{i,k}))$ with $K'_{i,k}(\theta^{*}_{i,k}, \sigma') > 0$ and $K'_{i,k}(\theta^{*}_{i,k}, \sigma'') < 0$, then by mixing σ' and σ'' we can construct $\sigma^{*} \in \triangle S_0(\theta(\theta^{*}_{i,k}))$ with $K'_{i,k}(\theta^{*}_{i,k}, \sigma^{*}) = 0$, which is a contradiction. In what follows, we will focus on the case (i), where $K'_{i,k}(\theta^{*}_{i,k}, \sigma') > 0$ for all $\sigma' \in \triangle S_0(\theta(\theta^{*}_{i,k}))$. The proof for the case (ii) is symmetric.

Since $K'_{i,k}(\theta_{i,k}, \sigma)$ is continuous in $(\theta_{i,k}, \sigma)$ and $\triangle S_0(\theta(\theta_{i,k}))$ is upper hemicontinuous in $\theta_{i,k}$, there is $\varepsilon > 0$ such that $K'_{i,k}(\theta_{i,k}, \sigma') > 0$ for any $\theta_{i,k}$ with $|\theta_{i,k} -$

[38] Note that $\boldsymbol{\theta}$ is absolutely continuous because $\boldsymbol{\sigma}$ is absolutely continuous and $\theta(\sigma)$ is Lipschitz continuous. Also from Lemma 6, $\boldsymbol{\theta}$ satisfies (34) for almost all t.

$\theta^*_{i,k}| \le \varepsilon$ and any $\sigma' \in \triangle S_0(\theta(\theta_{i,k}))$. Pick such $\varepsilon > 0$. Then pick T such that

$$\boldsymbol{\theta}_{i,k}(t) \ge \theta^*_{i,k} + \varepsilon \tag{53}$$

for any $t \ge T$ and for any solution $\boldsymbol{\theta} \in Z_\theta(\theta(\theta_{i,k}))$ with any $\theta_{i,k}$ with $\theta_{i,k} \ge \theta^*_{i,k} - \varepsilon$.

From Proposition 12, there is t^* such that for any $t > t^*$, $\boldsymbol{\theta} \in Z_\theta(\boldsymbol{w}_\theta(t))$, and $s \in [0, 2T]$,

$$|\boldsymbol{w}_{\theta,i,k}(t+s) - \boldsymbol{\theta}_{i,k}(s)| < \frac{\varepsilon}{2} \tag{54}$$

where $\boldsymbol{w}_{\theta,i,k}$ denote the (i,k)-component of $\boldsymbol{w}_\theta$. Pick such t^*. Since $\theta^*_{i,k} \le \limsup_{t\to\infty} \theta^t(h)$, there is $t^{**} > t^*$ such that $\boldsymbol{w}_{\theta,i,k}(t^{**}) \ge \theta^*_{i,k} - \varepsilon$. Pick such t^{**}. Then from (53), we have

$$\boldsymbol{\theta}_{i,k}(s) \ge \theta^*_{i,k} + \varepsilon$$

for any $s \ge T$ and for any solution $\boldsymbol{\theta} \in Z_\theta(\boldsymbol{w}_\theta(t^{**}))$. This inequality and (54) implies

$$\boldsymbol{w}_{\theta,i,k}(t^{**} + s) \ge \theta^*_{i,k} + \frac{\varepsilon}{2} \quad \forall s \in [T, 2T].$$

Likewise, since $\boldsymbol{w}_{\theta,i,k}(t^{**} + T) \ge \theta^*_{i,k} + \frac{\varepsilon}{2}$, it follows from (53) that

$$\boldsymbol{\theta}_{i,k}(s) \ge \theta^*_{i,k} + \varepsilon$$

for any $s \ge T$ and for any solution $\boldsymbol{\theta} \in Z_\theta(\boldsymbol{w}_\theta(t^{**} + T))$. This inequality and (54) implies

$$\boldsymbol{w}_{\theta,i,k}(t^{**} + s) \ge \theta^*_{i,k} + \frac{\varepsilon}{2} \quad \forall s \in [2T, 3T].$$

Iterating this argument, we can show that

$$\boldsymbol{w}_{\theta,i,k}(t^{**} + s) \ge \theta^*_{i,k} + \frac{\varepsilon}{2} \quad \forall s \in [T, \infty).$$

But this means that $\liminf_{t\to\infty} \theta^t_{i,k}(h) \ge \theta^*_{i,k} + \frac{\varepsilon}{2}$, which is a contradiction.

A.14.2 Case 2: $\liminf_{t\to\infty} \theta^t_{i,k}(h) = \limsup_{t\to\infty} \theta^t_{i,k}(h)$.

In this case, $\lim_{t\to\infty} \theta^t_{i,k}(h)$ exists. Let $\theta^*_{i,k} = \lim_{t\to\infty} \theta^t_{i,k}(h)$. We will show that $\theta^*_{i,k} \in E$.

Suppose not so that $\theta^* \notin E$. Then as in the previous case, (i) $K'_{i,k}(\theta^*_{i,k}, \sigma') > 0$ for all $\sigma' \in \triangle S_0(\theta(\theta^*_{i,k}))$, or (ii) $K'_{i,k}(\theta^*_{i,k}, \sigma') < 0$ for all $\sigma' \in \triangle S_0(\theta(\theta^*_{i,k}))$. We will focus on the case (i).

As in the previous case, there is $\varepsilon > 0$ such that $K'_{i,k}(\theta_{i,k}, \sigma') > 0$ for any $\theta_{i,k}$ with $|\theta_{i,k} - \theta^*_{i,k}| \leq \varepsilon$ and any $\sigma' \in \triangle S_0(\theta(\theta_{i,k}))$. Pick such $\varepsilon > 0$. Then pick T such that (53) holds for any $t \geq T$ and for any solution $\boldsymbol{\theta} \in Z_\theta(\theta(\theta_{i,k}))$ with any $\theta_{i,k}$ with $\theta_{i,k} \geq \theta^*_{i,k} - \varepsilon$.

From Proposition 12, there is t^* such that (54) holds for any $t > t^*$, $\boldsymbol{\theta} \in Z'_\theta(\boldsymbol{w}_\theta(t))$, and $s \in [0, 2T]$. Pick such t^*. Since $\theta^*_{i,k} = \lim_{t\to\infty} \theta^t(h)$, there is $t^{**} > t^*$ such that $\boldsymbol{w}_{\theta,i,k}(t^{**}) \geq \theta^*_{i,k} - \varepsilon$. Pick such t^{**}. Then as in the previous case, we can show that

$$\boldsymbol{w}_{\theta,i,k}(t^{**} + s) \geq \theta^*_{i,k} + \frac{\varepsilon}{2} \quad \forall s \in [T, \infty).$$

But this means that $\lim_{t\to\infty} \theta^t_{i,k}(h) \geq \theta^*_{i,k} + \frac{\varepsilon}{2}$, which is a contradiction.

A.15 Proof of Proposition 14

By assumptions (i)-(iii), it is obvious that the graphs $\{(\theta_2, \hat{f}_1(\theta_2, A)) | \forall \theta_2\}$ and $\{(\hat{\theta}_1, f_2(\hat{\theta}_1, A)) | \forall \hat{\theta}_1\}$ intersect once. This immediately implies that there is a steady state and it is unique. So what remains is to show that the belief converges to this steady state almost surely.

Recall that under second-order misspecification, there are three players: Player 1, player 2, and the hypothetical player.[39] Let $\sigma = (\sigma_1, \sigma_2, \hat{\sigma}_1)$ denote an action frequency of these players. Given σ, let $K_1(\theta_1, \sigma)$ denote the weighted KL divergence between player 1's subjective signal distribution and the true distribution, and let $\theta_1(\sigma)$ denote the minimizer of this KL divergence. Similarly define $K_2(\theta_2, \sigma)$, $\theta_2(\sigma)$, $\hat{K}_1(\hat{\theta}_1, \sigma)$, and $\hat{\theta}_1(\sigma)$.

Since player 1 is unbiased, we have $\theta_1(\sigma) = \theta^*$ for all σ, which implies that θ^t_1 is constant over time and $\theta^t_1 = \theta^*$ for all t. On the other hand, $\theta^t = (\theta^t_2, \hat{\theta}^t_1)$

[39]Formally, in the game with second-order misspecification, we have $k_1 = 3$, $k_2 = 2$, $M_{2,1} = M_{1,2}$, and $M_{2,2} = M_{1,3}$. Agents $(1,2)$ and $(1,3)$ are redundant, so $I^{**} = \{(1,1), (2,1), (2,2)\}$. Agent $(1,1)$ is player 1, agent $(2,1)$ is player 2, and agent $(2,2)$ is the hypothetical player.

changes as time goes, and our Proposition 12 ensures that the motion of this θ^t is approximated by the differential inclusion (35), which can be rewritten as the two dimensional problem

$$\left(\frac{d\boldsymbol{\theta}_2(t)}{dt}, \frac{d\hat{\boldsymbol{\theta}}_1(t)}{dt}\right) \in \bigcup_{\sigma:\theta(\sigma)=\boldsymbol{\theta}(t)} \bigcup_{\sigma' \in \triangle S_0^*(\boldsymbol{\theta}(t))} \left(-\frac{K_2'(\boldsymbol{\theta}_2(t),\sigma')}{K_2''(\boldsymbol{\theta}_2(t),\sigma)}, -\frac{\hat{K}_1'(\hat{\boldsymbol{\theta}}_1(t),\sigma')}{\hat{K}_1''(\hat{\boldsymbol{\theta}}_1(t),\sigma)}\right) \tag{55}$$

where $S_0^*(\theta)$ denote the set of all one-shot equilibrium $x = (x_1, x_2, \hat{x}_1)$ given the beliefs $\theta_1 = \theta^*$ and $\theta = (\theta_2, \hat{\theta}_1)$.

In what follows, we will show that regardless of the initial value, any solution to the differential inclusion (55) converges to the steady state after a long time. This implies that the steady state is globally attracting in the sense of Esponda, Pouzo, and Yamamoto (2019), and their Proposition 2 ensures that θ^t converges to the steady state almost surely, as desired.

The following lemma partially characterizes the solution to the differential inclusion (55): It shows that $\boldsymbol{\theta}_2(t)$ moves toward $f_2(\hat{\boldsymbol{\theta}}_1(t), A)$ at any time t.

Lemma 7. *Pick any initial value $\boldsymbol{\theta}(0)$ and any solution $\boldsymbol{\theta}$ to the differential inclusion (55). Then for any $t \geq 0$ with $\boldsymbol{\theta}_2(t) > f_2(\hat{\boldsymbol{\theta}}_1(t))$, we have $\dot{\boldsymbol{\theta}}_2(t) < 0$. Similarly, for any $t \geq 0$ with $\boldsymbol{\theta}_2(t) < f_2(\hat{\boldsymbol{\theta}}_1(t))$, we have $\dot{\boldsymbol{\theta}}_2(t) > 0$*

Proof. Suppose that $\boldsymbol{\theta}_2(t) > f_2(\hat{\boldsymbol{\theta}}_1(t), A)$ at some time t. To prove $\dot{\boldsymbol{\theta}}_2(t) < 0$, it suffices to show that $K_2'(\boldsymbol{\theta}_2(t), \sigma') > 0$ for any $\sigma' \in \triangle S_0^*(\boldsymbol{\theta}(t))$.

Suppose not and there is $\sigma' \in \triangle S_0^*(\boldsymbol{\theta}(t))$ such that $K_2'(\boldsymbol{\theta}_2(t), \sigma') < 0$. (We ignore the case with $K_2'(\boldsymbol{\theta}_2(t), \sigma') = 0$, because in such a case, $\boldsymbol{\theta}_2(t) \in f_2(\hat{\boldsymbol{\theta}}_1(t), A)$, which contradicts with the uniqueness of $f_2(\hat{\boldsymbol{\theta}}_1(t), A)$.) We consider the following two cases:

Case 1: $\boldsymbol{\theta}_2(t) = \overline{\theta}$. In this case, the KL minimizer given this σ' is $\theta_2(\sigma') = \overline{\theta} = \boldsymbol{\theta}_2(t)$ (this follows from the fact that the KL divergence is single-peaked w.r.t. θ_2). Hence $\boldsymbol{\theta}_2(t) = \overline{\theta}$ is a steady state, i.e., $\boldsymbol{\theta}_2(t) \in f_2(\hat{\boldsymbol{\theta}}_1(t), A)$. But this contradicts with the uniqueness of $f_2(\hat{\boldsymbol{\theta}}_1(t), A)$.

Case 2: $\boldsymbol{\theta}_2(t) < \overline{\theta}$. An argument similar to that in Case 1 shows that at $\theta_2 = \overline{\theta}$, we have $K_2'(\overline{\theta}, \sigma') > 0$ for all $\sigma' \in \triangle S_0^*(\overline{\theta}, \hat{\boldsymbol{\theta}}_1(t))$. On the other hand, by the assumption, at $\theta_2 = \boldsymbol{\theta}_2(t)$, $K_2'(\boldsymbol{\theta}_2(t), \boldsymbol{\sigma}') < 0$ for some $\boldsymbol{\sigma}' \in \triangle S_0^*(\boldsymbol{\theta}(t))$. Then since the set $\{K_2'(\theta_2, \boldsymbol{\sigma}') | \forall \boldsymbol{\sigma}' \in \triangle S_0^*(\theta_2, \hat{\boldsymbol{\theta}}_1(t))\}$ is convex and upper hemi-continuous in θ_2, there must be $\theta_2 \in [\boldsymbol{\theta}_2(t), \overline{\theta}]$ and $\boldsymbol{\sigma}' \in \triangle S_0^*(\theta_2, \hat{\boldsymbol{\theta}}_1(t))$ such that $K_2'(\boldsymbol{\theta}_2(t), \boldsymbol{\sigma}') = 0$. This implies that $\theta_2 \in f_2(\hat{\boldsymbol{\theta}}_1, A)$, but it contradicts with the uniqueness of $f_2(\hat{\boldsymbol{\theta}}_1, A)$.

Next, suppose that $\boldsymbol{\theta}_2(t) < f_2(\hat{\boldsymbol{\theta}}_1(t), A)$ at some time t. Then an argument similar to the one above shows that $K_2'(\boldsymbol{\theta}_2(t), \boldsymbol{\sigma}') < 0$ for any $\boldsymbol{\sigma}' \in \triangle S_0^*(\boldsymbol{\theta}(t))$, which implies $\dot{\boldsymbol{\theta}}_2(t) > 0$. *Q.E.D.*

Now we will construct a Lyapunov function V to show that any solution to the differential inclusion (55) converges to the steady state. Without loss of generality, assume that the steady state is $\boldsymbol{\theta}^* = (\theta_2^*, \hat{\theta}_1^*) = (0,0)$. From assumption (iii), there is $\kappa > 0$ such that $\max_{\hat{\theta}_1} |\frac{f_2(\hat{\theta}_1, A)}{\partial \hat{\theta}_1}| < \kappa < \frac{1}{\max_{\theta_2} |\frac{\hat{f}_1(\theta_2, A)}{\partial \theta_2}|}$. Pick such κ, and for each $\boldsymbol{\theta} = (\theta_2, \hat{\theta}_1)$, let

$$V(\boldsymbol{\theta}) = \max\left\{|\theta_2|, |\kappa \hat{\theta}_1|\right\}.$$

We will show that given any initial value $\boldsymbol{\theta}(0)$ and given any solution $\boldsymbol{\theta}$ to the differential inclusion (34),

$$\dot{V}(\boldsymbol{\theta}(t)) < 0$$

for all t with $\boldsymbol{\theta}(t) \neq (0,0)$. We will consider the following cases separately:

Case 1: $|\boldsymbol{\theta}_2(t)| > |\kappa \hat{\boldsymbol{\theta}}_1(t)|$. Assume first that $\boldsymbol{\theta}_2(t) > 0$. Then by the definition of κ and $f_2(0) = 0$, we have $f_2(\hat{\boldsymbol{\theta}}_1(t)) < |\kappa \hat{\boldsymbol{\theta}}_1(t)| < \boldsymbol{\theta}_2(t)$. Then from Lemma 7 and $\boldsymbol{\theta}_2(t) > 0$, we have $\dot{V}(\boldsymbol{\theta}(t)) = \dot{\boldsymbol{\theta}}_2(t) < 0$.

Assume next that $\boldsymbol{\theta}_2(t) < 0$. By the definition of κ and $f_2(0) = 0$, we have $f_2(\hat{\boldsymbol{\theta}}_1(t)) > -|\kappa \hat{\boldsymbol{\theta}}_1(t)| > \boldsymbol{\theta}_2(t)$. Then from Lemma 7 and $\boldsymbol{\theta}_2(t) < 0$, we have $\dot{V}(\boldsymbol{\theta}(t)) = -\dot{\boldsymbol{\theta}}_2(t) < 0$.

Case 2: $|\boldsymbol{\theta}_2(t)| < |\kappa \hat{\boldsymbol{\theta}}_1(t)|$. An argument similar to those for Case 1 shows that $\dot{V}(\boldsymbol{\theta}(t)) < 0$.

Case 3: $|\boldsymbol{\theta}_2(t)| = |\kappa \hat{\boldsymbol{\theta}}_1(t)|$. We will focus on the case with $\boldsymbol{\theta}_2(t) > 0$ and $\hat{\boldsymbol{\theta}}_1(t) > 0$, because a similar argument applies to all other cases. Then as in the first half of Case 1, we have $\dot{\boldsymbol{\theta}}_2(t) < 0$. Also, a similar argument shows that $\dot{\hat{\boldsymbol{\theta}}}_1(t) < 0$. Hence we have $\dot{V}(\boldsymbol{\theta}(t)) = \{\dot{\boldsymbol{\theta}}_2(t), \kappa \dot{\hat{\boldsymbol{\theta}}}_1(t)\} < 0$.

A.16 Proof of Proposition 15

The result immediately follows from the following lemma:

Lemma 8. *Suppose that all the assumptions stated in Proposition 15 are satisfied. Then for any sample path $h \in \mathscr{H}$,*

(i) $\liminf_{t\to\infty} \underline{\theta}_{i,k}(\sigma^t(h)) \geq \theta^*_{i,k}$.

(ii) $\limsup_{t\to\infty} \overline{\theta}_{i,k}(\sigma^t(h)) \leq \theta^*_{i,k}$.

In what follows, we will prove this lemma. We will focus on part (i), because the proof of part (ii) is symmetric.

We begin with stating two preliminary lemmas. The first lemma considers the case in which the current action frequency has a unique KL minimizer $\theta^t_{i,k}$, and shows that if the current KL minimizer $\theta^t_{i,k}$ is lower than the steady state belief $\theta^*_{i,k}$, then today's action $S_0(\theta^t_{i,k})$ induces a higher KL minimizer. This implies that the KL minimizer tomorrow will be closer to the steady state belief than the current one. Likewise, if the current KL minimizer is higher than the steady state belief, then today's action induces a lower KL minimizer.

Lemma 9. $\theta_{i,k}(S_0(\tilde{\theta}_{i,k})) > \tilde{\theta}_{i,k}$ *for all* $\tilde{\theta}_{i,k} < \theta^*_{i,k}$*, and* $\theta_{i,k}(S_0(\tilde{\theta}_{i,k})) < \tilde{\theta}_{i,k}$ *for all* $\tilde{\theta}_{i,k} > \theta^*_{i,k}$.

Proof. Note that $\theta_{i,k}(S_0(\cdot))$ is a continuous mapping from $\Theta_{i,k} \subseteq \boldsymbol{R}$ to itself, and its fixed point is a steady state. Since there is a unique steady state, the result follows from a standard argument. *Q.E.D.*

The next lemma considers the case in which the current action frequency need not have a unique minimizer, and shows that the result similar to the previous lemma holds; very roughly, if the smallest KL minimizer $\underline{\theta}(\theta^t_{i,k})$ is lower than the steady state belief, then it will move up and approaches the steady state belief. The proof is omitted, as it is very similar to Lemma 4 of Esponda, Pouzo, and Yamamoto (2019).

Lemma 10. *Pick any* $\theta_{i,k} < \theta^*_{i,k}$ *and any* σ *such that* $\sigma_{j,l} = \sigma_{\tilde{j},\tilde{l}}$ *for each* $(j,l) \in I^{**}$ *and* $(\tilde{j},\tilde{l}) \in I(j,l)$ *and such that* $K_{i,k}(\theta_{i,k},\sigma) < K_{i,k}(\tilde{\theta}_{i,k},\sigma)$ *for all* $\tilde{\theta}_{i,k} < \theta_{i,k}$*. Then for any solution* $\boldsymbol{\sigma} \in Z(\sigma)$ *starting from this* σ*, we have* $\underline{\theta}_{i,k}(\boldsymbol{\sigma}(t)) > \theta$ *for all* $t > 0$*.*

Now we will prove Lemma 8. Suppose not, so that there is a sample path $h \in \mathscr{H}$ such that $\liminf_{t\to\infty} \underline{\theta}_{i,k}(\sigma^t(h)) < \theta^*_{i,k}$. Pick such h, and let $\theta^0_{i,k} = \liminf_{t\to\infty} \underline{\theta}_{i,k}(\sigma^t(h))$. Let $\boldsymbol{w} : [0,\infty) \to \triangle X$ denote the continuous-time interpolation of the action frequency $(\sigma^t(h))_{t=1}^{\infty}$.

Pick $\varepsilon > 0$ such that

$$\frac{\partial K_{i,k}(\theta_{i,k},\sigma)}{\partial \theta_{i,k}} < 0 \quad \forall \theta_{i,k} \leq \theta^0_{i,k} + \varepsilon \tag{56}$$

for all σ such that

$$\sigma\left(\bigcup_{\tilde{\sigma}:\underline{\theta}_{i,k}(\tilde{\sigma}) \geq \theta^0_{i,k} - \varepsilon} S_0(\tilde{\sigma})\right) > 1 - 2\varepsilon.$$

To see why such ε exists, note first that from Lemma 9 and Assumption (iv) of Proposition 15, $\theta_{i,k}(S_0(\tilde{\theta}_{i,k})) > \theta^0_{i,k}$ for all $\tilde{\theta}_{i,k} \geq \theta^0_{i,k}$. Since $\theta_{i,k}(S_0(\cdot))$ is continuous, for any small ε, we have $\theta_{i,k}(S_0(\tilde{\theta}_{i,k})) > \theta^0_{i,k} + 2\varepsilon$ for all $\tilde{\theta}_{i,k} \geq \theta^0_{i,k} - \varepsilon$. Then from Assumptions (iv) and (v) of Proposition 15, we have $\theta_{i,k}(x) > \theta^0_{i,k} + 2\varepsilon$ for all σ such that $\underline{\theta}_{i,k}(\sigma) \geq \theta^0_{i,k} - \varepsilon$ and for all $x \in S_0(\sigma)$. Then from the single-peakedness assumption, (56) holds for all σ such that

$$\sigma\left(\bigcup_{\tilde{\sigma}:\underline{\theta}_{i,k}(\tilde{\sigma}) \geq \theta^0_{i,k} - \varepsilon} S_0(\tilde{\sigma})\right) = 1.$$

This inequality does not change even if σ is perturbed, so ε satisfies the desired property. (Take ε small, if necessary.)

Pick $T > 0$ such that $\frac{1}{1+T} < \varepsilon$. Then pick $t^* > 0$ such that for all $t > t^*$,

$$\sup_{s\in[0,2T]} \inf_{\boldsymbol{\sigma}\in Z(\boldsymbol{w}(t))} |\boldsymbol{\sigma}(s) - \boldsymbol{w}(t+s)| < \varepsilon. \tag{57}$$

Pick $t > t^*$ such that $\underline{\theta}_{i,k}(\boldsymbol{w}(t))$ is in the ε-neighborhood of $\theta^0_{i,k}$. Pick any solution $\boldsymbol{\sigma} \in Z(\boldsymbol{w}(t))$ to the differential inclusion starting from this $\boldsymbol{w}(t)$. Then from Lemma 10 (we set $\theta_{i,k} = \underline{\theta}_{i,k}(\boldsymbol{w}(t))$), we have $\underline{\theta}_{i,k}(\boldsymbol{\sigma}(s)) > \underline{\theta}_{i,k}(\boldsymbol{w}(t)) > \theta^0_{i,k} - \varepsilon$ for all $s > 0$. So in this solution $\boldsymbol{\sigma}$, the share of the set of action profiles $\bigcup_{\tilde{\sigma}:\underline{\theta}_{i,k}(\tilde{\sigma})\geq\theta^0_{i,k}-\varepsilon} S_0(\tilde{\boldsymbol{\sigma}})$ increases over time. In particular, by the definition of T, we have

$$\boldsymbol{\sigma}(s)\left[\bigcup_{\tilde{\sigma}:\underline{\theta}_{i,k}(\tilde{\sigma})\geq\theta^0_{i,k}-\varepsilon} S_0(\tilde{\boldsymbol{\sigma}})\right] > 1-\varepsilon \quad \forall s \geq T.$$

Then from (57), we have

$$\boldsymbol{w}(t+s)\left[\bigcup_{\tilde{\sigma}:\underline{\theta}_{i,k}(\tilde{\sigma})\geq\theta^0_{i,k}-\varepsilon} S_0(\tilde{\boldsymbol{\sigma}})\right] > 1-2\varepsilon \quad \forall s \in [T,2T].$$

This and (56) imply

$$\frac{\partial K_{i,k}(\theta_{i,k},\boldsymbol{w}(t+s))}{\partial \theta_{i,k}} < 0 \quad \forall \theta_{i,k} \leq \theta^0_{i,k} + \varepsilon \forall s \in [T,2T].$$

Now consider a solution $\boldsymbol{\sigma}'$ to the differential inclusion starting from $\boldsymbol{w}(t_0+T)$. Then again from Lemma 10 (we set $\theta_{i,k} = \theta^0_{i,k}+\varepsilon$), we have $\underline{\theta}_{i,k}(\boldsymbol{\sigma}(s)) > \theta^0_{i,k}+\varepsilon > \theta^0_{i,k} - \varepsilon$ for all $s > 0$. Hence

$$\boldsymbol{\sigma}'(s)\left[\bigcup_{\tilde{\sigma}:\underline{\theta}_{i,k}(\tilde{\sigma})\geq\theta^0_{i,k}-\varepsilon} S_0(\tilde{\boldsymbol{\sigma}})\right] > 1-\varepsilon \quad \forall s \geq T,$$

which implies

$$\boldsymbol{w}(t+s)\left[\bigcup_{\tilde{\sigma}:\underline{\theta}_{i,k}(\tilde{\sigma})\geq\theta^0_{i,k}-\varepsilon} S_0(\tilde{\boldsymbol{\sigma}})\right] > 1-2\varepsilon \quad \forall s \in [2T,3T]$$

and thus

$$\frac{\partial K_{i,k}(\theta_{i,k},\boldsymbol{w}(t+s))}{\partial \theta_{i,k}} < 0 \quad \forall \theta_{i,k} \leq \theta^0_{i,k} + \varepsilon \forall t \in [T,3T].$$

Iterating the same argument, we can show that

$$\frac{\partial K_{i,k}(\theta_{i,k}, \boldsymbol{w}(t+s))}{\partial \theta_{i,k}} < 0 \quad \forall \theta_{i,k} \leq \theta_{i,k}^{0} + \varepsilon \forall s \geq T.$$

This implies $\underline{\theta}_{i,k}(\boldsymbol{w}(t+s)) > \theta_{i,k}^{0} + \varepsilon$ for all $s \geq T$, which is a contradiction.

B Uniqueness of Steady State

We provide a sufficient condition for the unique steady state in symmetric games at $A_1 = A_2 = a$ in each misspecification.

First-order misspecification. Assume that for any fixed A and θ_i, Nash equilibrium of the one-shot game is unique. Assume also that $\frac{\partial U_i}{\partial x_i} > 0$ at $x_1 = x_2 = 0$ and that $\frac{\partial U_i}{\partial x_i} < 0$ at some $x_1 = x_2 > 0$; this condition ensures that the unique Nash equilibrium of the one-shot game is interior. By the continuity of $\frac{\partial U_i}{\partial x_i}$, there exists $x_1 = x_2 => 0$ such that $\frac{\partial U_i}{\partial x_i} = 0$; that is, the unique Nash equilibrium of the one-shot game is symmetric.

Note that steady state is an intersection of asymptotic best response correspondences BR_1 and BR_2. Hence, the steady state is unique if $-1 < BR_i'(x_j) < 1$ for all i and x_j where $j \neq i$.

Second-order and one-sided double misspecification. Assume that for any fixed A and θ_i, Nash equilibrium of the one-shot game is unique. Assume also that $\frac{\partial U_i}{\partial x_i} > 0$ at $x_1 = \hat{x}_1 = x_2 = 0$ and that $\frac{\partial U_i}{\partial x_i} < 0$ at some $x_1 = \hat{x}_1 = x_2 > 0$; this condition ensures that the unique Nash equilibrium of the one-shot game is interior. By the continuity of $\frac{\partial U_i}{\partial x_i}$, there exists $x_1 = \hat{x}_1 = x_2 => 0$ such that $\frac{\partial U_i}{\partial x_i} = 0$; that is, the unique Nash equilibrium of the one-shot game is symmetric.

Note that steady state is an intersection of BR_{12} and $NE_2(x_1)$. Hence, the steady state is unique if $-1 < BR_{12}' < 1$ for all $(x_1 \hat{;} x_2)$ and $-1 < NE_2'(x_1) < 1$ for all x_1.

The conditions and derivations of the uniqueness of one-sided double misspecification is the same as that of two-sided misspecification.

Two-sided double misspecification. Assume that for any fixed A_i and θ_i, Nash equilibrium of the one-shot game is unique. Assume also that $\frac{\partial U_i}{\partial x_i} > 0$ at $x_1 = \hat{x}_2 = x_2 = \hat{x}_1 = 0$ and that $\frac{\partial U_i}{\partial x_i} < 0$ at some $x_1 = \hat{x}_2 = x_2 = \hat{x}_1 > 0$; this condition ensures that the unique Nash equilibrium of the one-shot game is interior. By the continuity of $\frac{\partial U_i}{\partial x_i}$, there exists $x_1 = \hat{x}_2 = x_2 = \hat{x}_1 > 0$ such that $\frac{\partial U_i}{\partial x_i} = 0$; that is, the unique Nash equilibrium of the one-shot game is symmetric. It implies that, at any steady state, $\frac{\partial U_i}{\partial x_i} = \frac{\partial \hat{U}_j}{\partial \hat{x}_j} = 0$ holds only at $x_1 = \hat{x}_2 > 0$. Hence, $x_1 = \hat{x}_2 > 0$ and $x_2 = \hat{x}_1 > 0$ hold at any steady state.

Let $NE_1(x_2)$ denote the set of steady-state action $(x_1, \hat{x}_2)$ given x_2. Note that we omit $\hat{x}_1$ because it does not influence $(x_1, \hat{x}_2)$. Since $x_1 = \hat{x}_2$, we omit the second component, and view NE_1 as the set of steady-state action x_1. That is, $NE_1(x_2)$ is the set of x_1 such that $x_1 = \hat{x}_2$ is a steady-state action. Define $NE_2(x_1)$ in a similar way.

Note that steady state is an intersection of NE_1 and NE_2, and a sufficient condition for the unique steady state is $-1 < NE_i'(x_j) < 1$ for all i and x_j where $j \neq i$. For $x_i = NE_i(x_j)$, FOC must be satisfied:

$$\frac{\partial U_i(x_i, \hat{x}_j, A_i, \theta_i(x_i, x_j, \hat{x}_j))}{\partial x_i} = 0.$$

Using the implicit function theorem with $x_i = \hat{x}_j = NE_i(x_j)$, the slope of $NE_i(x_j)$ is

$$-\frac{\frac{\partial^2 U_i}{\partial x_i \partial \theta_i}\frac{\partial \theta_i}{\partial x_j}}{U_{ii} + U_{ij} + \frac{\partial^2 U_i}{\partial x_i \partial \theta_i}\left(\frac{\partial \theta_i}{\partial x_i} + \frac{\partial \theta_i}{\partial x_j}\right)}$$

Here the numerator measures how much $\frac{\partial U_i}{\partial x_i}$ changes when x_j changes. Since j is not the opponent of player i, there is only an indirect effect. The denominator measures how much $\frac{\partial U_i}{\partial x_i}$ changes when $x_i = \hat{x}_j$ changes. So a sufficient condition for unique steady state is

$$-\frac{\frac{\partial^2 U_i}{\partial x_i \partial \theta_i}\frac{\partial \theta_i}{\partial x_j}}{U_{ii} + U_{ij} + \frac{\partial^2 U_i}{\partial x_i \partial \theta_i}\left(\frac{\partial \theta_i}{\partial x_i} + \frac{\partial \theta_i}{\partial x_j}\right)} \in (-1, 1)$$

for all i, $j \neq i$, x_j, and $x_i \in NE_i(x_j)$.

C Convergence for Examples in Main Text

In this appendix, we check beliefs in each example covered in the main text converge to a steady state.

Example: Equation 7.

Consider the Cournot model in Section 3.2.1. Suppose that the inverse demand function is given by

$$Q(x_1+x_2,a,\theta) = a-(1-\theta)(x_1+x_2)$$

and the cost function is linear, i.e., $c(x_i) = cx_i$ where $c \in (0,a)$. Suppose also that $\Theta = [-d,d]$ where $d \in (0,\frac{1}{3})$ is a fixed parameter. We will show that for each misspecification, the belief converges to a steady state as long as misspecification is small (i.e., A is sufficiently close to a).

First-order misspecification and one-sided double misspecification. Since the inverse demand function Q is linear in θ, the identifiability condition holds, and hence Proposition 13 ensures that the belief converges almost surely under first-order misspecification and one-sided double misspecification. In particular, when the steady state is unique, the belief converges there almost surely regardless of the initial prior.

Second-order misspecification. To prove convergence for small misspecification, it suffices to check the conditions stated in Corollary 4.

Given a misspecified parameter A, let $f_2(\hat{\theta}_1,A)$ denote the set of "steady-state belief" of player 2, when the hypothetical player's belief is fixed at $\hat{\theta}_1$. In such a steady state, the incentive-compatibility conditions and the consistency condition must be satisfied:[40]

$$\begin{aligned}
x_1 &= \frac{a-c}{2(1-\theta^*)} - \frac{x_2}{2},\\
x_2 &= \frac{a-c}{2(1-\theta_2)} - \frac{\hat{x}_1}{2},\\
\hat{x}_1 &= \frac{A-c}{2(1-\hat{\theta}_1)} - \frac{x_2}{2},\\
a-(1-\theta_2)(\hat{x}_1+x_2) &= a-(1-\theta^*)(x_1+x_2).
\end{aligned}$$

From the second and the third equations, we have $x_2 = \frac{2(a-c)}{3(1-\theta_2)} - \frac{A-c}{3(1-\hat{\theta}_1)}$. Combining it into the first and the third equations,

$$\begin{aligned}
x_1+x_2 &= \frac{a-c}{2(1-\theta^*)} + \frac{x_2}{2} = \frac{a-c}{2(1-\theta^*)} + \frac{a-c}{3(1-\theta_2)} - \frac{A-c}{6(1-\hat{\theta}_1)},\\
\hat{x}_1+x_2 &= \frac{A-c}{2(1-\hat{\theta}_1)} + \frac{x_2}{2} = \frac{a-c}{3(1-\theta_2)} + \frac{A-c}{3(1-\hat{\theta}_1)}.
\end{aligned}$$

Plugging them into the last consistency condition,

$$a-(1-\theta_2)\left(\frac{a-c}{3(1-\theta_2)} + \frac{A-c}{3(1-\hat{\theta}_1)}\right) - a + (1-\theta^*)\left(\frac{a-c}{2(1-\theta^*)} + \frac{a-c}{3(1-\theta_2)} - \frac{A-c}{6(1-\hat{\theta}_1)}\right) = 0,$$

which is equivalent to

$$-\frac{(1-\theta_2)(A-c)}{3(1-\hat{\theta}_1)} + \frac{(1-\theta^*)(a-c)}{3(1-\theta_2)} + \frac{a-c}{6} - \frac{(1-\theta^*)(A-c)}{6(1-\hat{\theta}_1)} = 0. \tag{58}$$

[40]Here we implicitly assume that steady states actions and beliefs are interior points, and this assumption is without loss of generality. Indeed, it is straightforward to see that player i's optimal action is an interior solution given any belief θ_i and given any action x_{-i} of the opponent; this immediately implies that steady state actions are interior points. Regarding beliefs, suppose not and there is $f_2(\hat{\theta}_1, A)$ which contains a boundary point, say $\theta_2 = -d$, for some $\theta^*, \hat{\theta}_1 \in (-d,d)$. Then, from the incentive-compatibility conditions, we have $x_1 + x_2 = \frac{a-c}{2(1-\theta^*)} + \frac{a-c}{3(1+d)} - \frac{A-c}{6(1-\hat{\theta}_1)}$ and $x_1 + \hat{x}_2 = \frac{a-c}{3(1+d)} + \frac{A-c}{3(1-\hat{\theta}_1)}$. Given these $(x_1, x_2, \hat{x}_1)$, $\theta_2 = -d$ does not minimize the KL metric; because there is $\theta_2 \in (-d,d)$ which makes the KL metric equal to zero (note that such θ_2 can be found by solving $a-(1-\theta_2)(\hat{x}_1+x_2) = a-(1-\theta^*)(x_1+x_2)$). This implies that the boundary point $\theta_2 = -d$ is not a steady state belief. The same argument is applied to $\theta_2 = d$.

In the rest of this appendix, we will assume that other steady states (such as $\hat{f}_1$) are also interior points, and an argument similar to the one above shows that this assumption is without loss of generality.

A steady-state belief $f_2(\hat{\theta}_1, A)$ must be a solution to this equation (58). At $A = a$, (58) reduces to

$$2(1-\theta^*)(1-\hat{\theta}_1) - (1-\theta_2)(2-2\theta_2+\hat{\theta}_1-\theta^*) = 0. \tag{59}$$

Note that for any $\theta_2, \hat{\theta}_1, \theta^* \in (-d,d)$, the left hand side of (59) is strictly decreasing in $\hat{\theta}_1$ and θ^*. When $\hat{\theta}_1 = \theta^* = d$, there is a unique solution $\theta_2 = d$. When $\hat{\theta}_1 = \theta^* = -d$, there is a unique solution $\theta_2 = -d$. When $\hat{\theta}_1 \in (-d,d)$ or $\theta^* \in (-d,d)$, the left hand side of (59) is strictly increasing in θ_2, negative at $\theta_2 = -d$ and positive at $\theta_2 = d$. Hence, the intermediate value theorem ensures that given any $\hat{\theta}_1, \theta^* \in (-d,d)$, (59) has an interior solution. Also this solution is unique, as the left hand side of (59) is strictly increasing in θ_2. So in sum, for each $\hat{\theta}_1$, the set $f_2(\hat{\theta}_1, A)$ of steady-state beliefs is non-empty and is a singleton.

Similarly, given a misspecified parameter A, let $\hat{f}_1(\theta_2, A)$ denote the set of "steady-state belief" of hypothetical player 1, when player 2's belief is fixed at θ_2. In such a steady state, the incentive compatibility condition and the consistency condition must be satisfied:

$$\begin{aligned}
x_1 &= \frac{a-c}{2(1-\theta^*)} - \frac{x_2}{2}, \\
x_2 &= \frac{a-c}{2(1-\theta_2)} - \frac{\hat{x}_1}{2}, \\
\hat{x}_1 &= \frac{A-c}{2(1-\hat{\theta}_1)} - \frac{x_2}{2}, \\
A - (1-\hat{\theta}_1)(\hat{x}_1 + x_2) &= a - (1-\theta^*)(x_1 + x_2).
\end{aligned}$$

The first three equations yield

$$\begin{aligned}
x_1 + x_2 &= \frac{a-c}{2(1-\theta^*)} + \frac{a-c}{3(1-\theta_2)} - \frac{A-c}{6(1-\hat{\theta}_1)}, \\
\hat{x}_1 + x_2 &= \frac{a-c}{3(1-\theta_2)} + \frac{A-c}{3(1-\hat{\theta}_1)}.
\end{aligned}$$

Plugging them into the last equation,

$$A - (1-\hat{\theta}_1)\left(\frac{a-c}{3(1-\theta_2)} + \frac{A-c}{3(1-\hat{\theta}_1)}\right) - a + (1-\theta^*)\left(\frac{a-c}{2(1-\theta^*)} + \frac{a-c}{3(1-\theta_2)} - \frac{A-c}{6(1-\hat{\theta}_1)}\right) = 0,$$

which is equivalent to

$$A-a+\frac{a-c}{2}-\frac{A-c}{3}-\frac{(1-\hat{\theta}_1)(a-c)}{3(1-\theta_2)}+\frac{(1-\theta^*)(a-c)}{3(1-\theta_2)}-\frac{(1-\theta^*)(A-c)}{6(1-\hat{\theta}_1)}=0.$$

A steady-state belief $\hat{f}_1(\theta_2, A)$ must be a solution to this equation. When $A = a$, the equation above reduces to

$$\frac{(a-c)(\hat{\theta}_1-\theta^*)(1+\theta_2-2\hat{\theta}_1)}{6(1-\hat{\theta}_1)(1-\theta_2)}=0.$$

Hence, when $\Theta = [-d, d]$ where $x \in (0, \frac{1}{3})$, $\hat{\theta}_1 = \theta^*$ is a unique solution, and the conditions stated in Corollary 4 are satisfied.

Two-sided double misspecification. To prove convergence for small misspecification, it suffices to check the conditions stated in Proposition 14.

Given misspecified parameters A_1, A_2, let $f_2(\theta_1)$ denote the set of "steady-state belief" of player 2, when player 1's belief is fixed at θ_1. Note that the incentive-compatibility conditions and the consistency condition are:

$$\begin{aligned}
x_1 &= \frac{A_1-c}{2(1-\theta_1)}-\frac{\hat{x}_2}{2},\\
x_2 &= \frac{A_2-c}{2(1-\theta_2)}-\frac{\hat{x}_1}{2},\\
\hat{x}_1 &= \frac{A_2-c}{2(1-\theta_2)}-\frac{x_2}{2},\\
\hat{x}_2 &= \frac{A_1-c}{2(1-\theta_1)}-\frac{x_1}{2},\\
A_2-(1-\theta_2)(\hat{x}_1+x_2) &= a-(1-\theta^*)(x_1+x_2).
\end{aligned}$$

The first four equations imply $\hat{x}_1 = x_2 = \frac{A_2-c}{3(1-\theta_2)}$ and $x_1 = \hat{x}_2 = \frac{A_1-c}{3(1-\theta_1)}$. Plugging them into the last equation,

$$A_2-2(1-\theta_2)\frac{A_2-c}{3(1-\theta_2)}-a+(1-\theta^*)(\frac{A_1-c}{3(1-\theta_1)}+\frac{A_2-c}{3(1-\theta_2)})=0,$$

which is equivalent to

$$\theta_2 = 1 - \frac{(1-\theta_1)(1-\theta^*)(A_2 - c)}{(1-\theta_1)(A_2 + 3a - 4c) - (1-\theta^*)(A_1 - c)}.$$

So for any θ_1, $f_2(\theta_1)$ is a singleton. Also, at $A_1 = A_2 = a$,

$$\frac{\partial f_2(\theta_1)}{\partial \theta_1} = -\frac{(1-\theta^*)^2}{(3-4\theta_1+\theta^*)^2}.$$

Note that this derivative is negative, and is larger than -1 as $\theta_1, \theta^* \in [-\frac{1}{3}, \frac{1}{3}]$. Hence, if A_1 and A_2 are sufficiently close to a, then $|\frac{\partial f_2(\theta_1)}{\partial \theta_1}| \in (0,1)$.

Similarly, given misspecified parameters, let $f_1(\theta_2)$ denote the set of "steady-state belief" of player 1, when player 2's belief is fixed at θ_2. Then we can show that $f_1(\theta_2)$ is a singleton for all θ_2 and $|\frac{\partial f_1(\theta_2)}{\partial \theta_2}| \in (0,1)$. A proof is similar to that for f_2, and hence omitted.

Example: Equation 8.

Suppose the Cournot model with $c(x_i) = cx_i$ where $c \geq 0$, $a < 1$, $\Theta = [k-d, k+d]$ where $k, d > 0$ are fixed parameters which satisfy $k - d > c$, and

$$Q(x_1 + x_2, a, \theta) = \theta - (1-a)(x_1 + x_2).$$

First-order misspecification and one-sided double misspecification. Note that the identifiability condition is satisfied, and hence Proposition 13 implies that the belief converges almost surely to a steady state.

Second-order misspecification. To prove convergence for small misspecification, it suffices to check the conditions stated in Corollary 4.

Given a misspecified parameter A, let $f_2(\hat{\theta}_1, A)$ denote the set of "steady-state belief" of player 2, when the hypothetical player's belief is fixed at $\hat{\theta}_1$. Note that the incentive-compatibility conditions and the consistency condition are:

$$\begin{aligned} x_1 &= \frac{\theta^* - c}{2(1-a)} - \frac{x_2}{2}, \\ x_2 &= \frac{\theta_2 - c}{2(1-a)} - \frac{\hat{x}_1}{2}, \\ \hat{x}_1 &= \frac{\hat{\theta}_1 - c}{2(1-A)} - \frac{x_2}{2}, \\ \theta_2 - (1-a)(\hat{x}_1 + x_2) &= \theta^* - (1-a)(x_1 + x_2). \end{aligned}$$

From the first three equations, we have $\hat{x}_1 + x_2 = \frac{\theta_2 - c}{3(1-a)} + \frac{\hat{\theta}_1 - c}{3(1-A)}$ and $x_1 + x_2 = \frac{\theta^* - c}{2(1-a)} + \frac{x_2}{2} = \frac{\theta^* - c}{2(1-a)} + \frac{\theta_2 - c}{3(1-a)} - \frac{\hat{\theta}_1 - c}{6(1-A)}$. Plugging them into the last equation and arranging,

$$\theta_2 = \frac{\theta^* + c}{2} + \frac{(1-a)(\hat{\theta}_1 - c)}{2(1-A)}.$$

Hence, for any $\hat{\theta}_1$, $f_2(\hat{\theta}_1, A)$ is a singleton.

Similarly, given a misspecified parameter A, let $\hat{f}_1(\theta_2, A)$ denote the set of "steady-state belief" of hypothetical player 1, when player 2's belief is fixed at θ_2. Note that the incentive-compatibility conditions and the consistency condition are:

$$\begin{aligned} x_1 &= \frac{\theta^* - c}{2(1-a)} - \frac{x_2}{2}, \\ x_2 &= \frac{\theta_2 - c}{2(1-a)} - \frac{\hat{x}_1}{2}, \\ \hat{x}_1 &= \frac{\hat{\theta}_1 - c}{2(1-A)} - \frac{x_2}{2}, \\ \hat{\theta}_1 - (1-A)(\hat{x}_1 + x_2) &= \theta^* - (1-a)(x_1 + x_2). \end{aligned}$$

The first three equations imply $\hat{x}_1 + x_2 = \frac{\theta_2 - c}{3(1-a)} + \frac{\hat{\theta}_1 - c}{3(1-A)}$ and $x_1 + x_2 = \frac{\theta^* - c}{2(1-a)} + \frac{x_2}{2} = \frac{\theta^* - c}{2(1-a)} + \frac{\theta_2 - c}{3(1-a)} - \frac{\hat{\theta}_1 - c}{6(1-A)}$. Plugging them into the last equation and arranging, From them, $\hat{f}_1(\theta_2, A)$ is a solution to

$$\hat{\theta}_1 = \frac{6(1-A)}{3 - 4A + a} \left(\frac{\theta^*}{2} - \frac{(A-a)c}{6(1-A)} - \frac{(A-a)(\theta_2 - c)}{3(1-a)} \right).$$

Hence, for any θ_2, $\hat{f}_1(\theta_2,A)$ is a singleton (also $\hat{f}_1(\theta_2,A)=\theta^*$ at $A=a$), and the conditions stated in Corollary 4 are satisfied.

Two-sided double misspecification. To prove convergence for small misspecification, it suffices to check the conditions stated in Proposition 14.

Given misspecified parameters A_1,A_2, let $f_2(\theta_1)$ denote the set of "steady-state belief" of player 2, when player 1's belief is fixed at θ_1. Note that the incentive-compatibility conditions and the consistency condition are:

$$\begin{aligned}
x_1 &= \frac{\theta_1-c}{2(1-A_1)}-\frac{\hat{x}_2}{2},\\
x_2 &= \frac{\theta_2-c}{2(1-A_2)}-\frac{\hat{x}_1}{2},\\
\hat{x}_1 &= \frac{\theta_2-c}{2(1-A_2)}-\frac{x_2}{2},\\
\hat{x}_2 &= \frac{\theta_1-c}{2(1-A_1)}-\frac{x_1}{2},\\
\theta_2-(1-A_2)(\hat{x}_1+x_2) &= \theta^*-(1-a)(x_1+x_2).
\end{aligned}$$

From the first four equations, $\hat{x}_1=x_2=\frac{\theta_2-c}{3(1-A_2)}$ and $x_1=\hat{x}_2=\frac{\theta_1-c}{3(1-A_1)}$. Plugging them into the last equation and arranging,

$$\theta_2=\frac{1-A_2}{2-a-A_2}\left(-2c+\frac{(1-a)c}{1-A_2}+3\theta^*-\frac{(1-a)(\theta_1-c)}{1-A_1}\right).$$

Hence, for any θ_1, $f_2(\theta_1)$ is a singleton. Also,

$$\frac{\partial f_2(\theta_1)}{\partial\theta_1}=-\frac{(1-a)(1-A_2)}{(1-A_1)(2-a-A_2)}.$$

Note that this derivative is negative, and is larger than -1 if A_1 and A_2 are sufficiently close to a. Hence, $|\frac{\partial f_2(\theta_1)}{\partial\theta_1}|\in(0,1)$ for any θ_1, θ_2, and θ^*.

Similarly, given misspecified parameters, let $f_1(\theta_2)$ denote the set of "steady-state belief" of player 1, when player 2's belief is fixed at θ_2. Then an argument similar to the one above shows that $f_1(\theta_2)$ is a continuous function and $|\frac{\partial f_1(\theta_2)}{\partial\theta_2}|\in(0,1)$.

References

Aghion, P., Bolton, P., Harris, C., and Jullien, B. (1991): "Optimal Learning by Experimentation" *Review of Economic Studies*, 58 (4), 621-654.

Avery, C. and J. Kagel (1997): "Second-price auctions with asymmetric payoffs: an experimental investigation," *Journal of Economics and Management Strategy*, 6, 573-603.

Ba, C. and A. Gindin (2020): "A Multi-Agent Model of Misspecified Learning with Overconfidence," Working Paper.

Benaïm, M. (1999): "Dynamics of Stochastic Approximation Algorithms," Séminaire de Probabilités XXXIII, Lecture Notes in Math. 1709, Springer.

Benaïm, M., J. Hofbauer, and S. Sorin (2005): "Stochastic approximations and differential inclusions," *SIAM Journal on Control and Optimization* 44, 328-348.

Benoît, J.P. and Dubra, J. (2011): "Apparent Overconfidence," *Econometrica* 79, 1591-1625.

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

Berk, R.H. (1966): "Limiting Behavior of Posterior Distributions when the Model is Incorrect," *Annals of Mathematical Statistics* 37, 51-58.

Billingsley, P. (1999): *Convergence of Probability Measures*, Wiley.

Bohren, J.A. and D.N. Hauser (2020): "Learning with Model Misspecification: Characterization and Robustness," Working Paper.

Brown, A.L., Camerer, C.F., and Lovallo, D. (2012): "To review or not to review? Limited strategic thinking at the movie box office," *American Economic Journal: Microeconomics* 4 (2), 1-26.

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

Carlana, M. (2019): "Implicit Stereotypes: Evidence from Teachers' Gender Bias," *Quarterly Journal of Economics*, 134 (3), 1163-1224.

Çelen, B. and S. Kariv (2005): "An experimental test of observational learning under imperfect information," *Economic Theory* 26, 677-699.

Cho, I-K. and K. Kasa (2017): "Gresham's Law of Model Averaging," *American Economic Review*, 107 (11), 3589-3616.

Daniel, K. and D. Hirshleifer (2015): "Overconfident Investors, Predictable Returns, and Excessive Trading," *Journal of Economic Perspectives*, 29 (4), 61-88.

Deimling, K. (1992): *Multivalued Differential Equations*, Walter de Gruyter.

DeMarzo, P.M., D. Vayanos, and J. Zwiebel (2015): "Persuasion Bias, Social Influence, and Unidimensional Opinions," *Quarterly Journal of Economics*, 118 (3), 909-968.

Dranove, D. and Jin, G.Z. (2010): "Quality disclosure and certification: theory and practice," *Journal of Economic Literature*, 48 (4), 935-963.

Dudley, R. (1966): "Convergence of Baire measures," *Studia Mathematica* 27, 251-268.

Englmaier, F. (2010): "Managerial Optimism and Investment Choice," *Managerial and Decision Economics*, 31 (4), 303-310.

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

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

Eyster, E. (2019): "Errors in Strategic Reasoning," In D.B. Bernheim, S. DellaVigna & D. Laibson, (Eds.), *Handbook of Behavioral Economics: Foundations and Applications 2*, 187-259, North Holland.

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

Fershtman, C. and K.L. Judd (1987): "Equilibrium Incentives in Oligopoly," *American Economic Review*, 77 (5), 927-940.

Frick, M., R. Iijima, and Y. Ishii (2020): "Misinterpreting Others and the Fragility of Social Learning," forthcoming in *Econometrica*.

Fudenberg, D., G. Lanzani, and P. Strack (2020): "Limit Points of Endogenous Misspecified Learning," working paper.

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

Gagnon-Bartsch, T. (2016): "Taste Projection in Models of Social Learning," Working Paper.

Gagnon-Bartsch, T. and M. Rabin (2016): "Naive Social Learning, Mislearning, and Unlearning," Working Paper.

Grubb, M. (2015): "Overconfident Consumers in the Marketplace," *Journal of Economic Perspectives*, 29 (4), 9-36.

He, K. (2019): "Mislearning from Censored Data: The Gambler's Fallacy in Optimal-Stopping Problems," Working Paper.

Heidhues, P. and B. Kőszegi (2018): "Behavioral Industrial Organization," In D.B. Bernheim, S. DellaVigna & D. Laibson, (Eds.), *Handbook of Behavioral Economics: Foundations and Applications 1*, 517-612, North Holland.

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

Heidhues, P., B. Kőszegi, and P. Strack (2020a): "Convergence in Models of Misspecified Learning," forthcoming in *Theoretical Economics*.

Heidhues, P., B. Kőszegi, and P. Strack (2020b): "Overconfidence and Prejudice," Working Paper.

Heifetz, A., C. Shannon, and Y. Spiegel (2007): "The Dynamic Evolution of Preferences," *Economic Theory*, 32 (2), 251-286.

Hofbauer, J., J. Oechssler, and F. Riedel (2009): "Brown-von Neumann-Nash Dynamics: the Continuous Strategy Case," *Games and Economic Behavior* 65, 406-429.

Hoffman, M. and S.V. Burks (2020): "Worker Overconfidence: Field Evidence and Implications for Employee Turnover and Firm Profits," *Quantitative Economics* 11 (1), 315-348.

Huffman, D., C. Raymond, and J. Shvets (2019): "Persistent overconfidence and biased memory: Evidence from managers," Working Paper.

Jin, G., Luca, M., Martin, D., (forthcoming): "Is No News (Perceived As) Bad News? An Experimental Investigation of Information Disclosure," *American Economic Journal: Microeconomics*, forthcoming.

Kagel, J.H. and Levin, D. (2002): 'Common Value Auctions and the Winner's Curse.' Princeton University Press, Princeton.

Kőszegi, B. (2014): "Behavioral Contract Theory.," *Journal of Economic Literature*, 52 (4), 1075-1118.

Kübler, D. and G. Weizsäcker (2004): "Limited Depth of Reasoning and Failure of Cascade Formation in the Laboratory," *Review of Economic Studies*, 71 (2), 425-441.

Kübler, D. and G. Weizsäcker (2005): "Are Longer Cascades More Stable?," *Journal of the European Economic Association*, 3 (2-3), 330-339.

Kyle, A. S. and F. A. Wang (1997): "Speculation Duopoly with Agreement to Disagree: Can Overconfidence Survive the Market Test?," *Journal of Finance*, 52 (5), 2073–2090.

Lazear, E.P. and S. Rosen (1981): "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89 (5), 841-864.

Ludwig, S. and J. Nafziger (2011): " Beliefs about Overconfidence," *Theory and Decision*, 70, 475-500.

Madarász, K. (2012): "Information Projection: Model and Applications," *Review of Economic Studies*, 79 (3), 961-985.

Malmendier, U. and T. Geoffrey (2005): "CEO Overconfidence and Corporate Investment," *Journal of Finance*, 60 (6), 2661-2700.

Malmendier, U. and T. Geoffrey (2008): "Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction," *Journal of Financial Economics*, 89 (1), 20-43.

Malmendier, U. and T. Geoffrey (2015): "Behavioral CEOs: The Role of Managerial Overconfidence," *Journal of Economic Perspectives*, 29 (4), 37-60.

Molavi, P. (2020): “Macroeconomics with Learning and Misspecification: A General Theory and Applications,” Working paper.

Nyarko, Y. (1991): “Learning in Mis-Specified Models and the Possibility of Cycles,” *Journal of Economic Theory*, 55, 416-427.

Perkins, S. and D.S. Leslie (2014): “Stochastic Fictitious Play with Continuous Action Sets,” *Journal of Economic Theory* 152, 179-213.

Tirole, J. (1988): *The Theory of Industrial Organization*, MIT Press.

Tullock, G. (1980): “Efficient rent seeking.” In J. M. Buchanan, R. D. Tollison, & G. Tullock (Eds.), *Toward a Theory of the Rent-Seeking Society*, 97-112. College Station: Texas A&M University Press.

Van Boven, L., D. Dunning, and G. Loewenstein (2000): “Egocentric empathy gaps between owners and buyers: Misperceptions of the endowment effect,” *Journal of Personality and Social Psychology*, 79 (1), 66-76.