Online Appendix

Unprecedented

Satoshi Fukuda[†] Yuichiro Kamada[‡]

June 22, 2026

Appendix B provides supplementary discussions for Section 4 of the paper. The proofs of the results in Appendix B.1 to B.5 appear in Appendix B.6. Appendix C offers supplementary discussions for Section 5 of the paper. Appendix D provides detailed discussions on applications that we introduced in Section 6 of the paper.

B Supplementary Discussions for Section 4

B.1 Supplementary Results for Section 4.2

Consider the synchronous model and suppose that a public randomization device is not available. We call such a model the *synchronous model without a public randomization device*. Note that the strategy $\sigma_i^{(T,1)}$ is well defined in such a model as well.

First, we show that a public randomization device is not necessary to obtain multiple equilibria when $x_i \leq 1$ for each $i \in \{1, 2\}$.

Proposition 4. *Suppose $x_i \leq 1$ for each $i \in \{1, 2\}$. For any $p_1, p_2 \in (0, 1]$, there exist $\delta' \in (0, 1)$ and $T' < \infty$ such that if $\delta_i \in (\delta', 1)$ for each $i \in \{1, 2\}$ and $T > T'$, then $\sigma^{(T,1)}$ is a PBE in the synchronous model without a public randomization device.*

Second, we consider the case when $x_i > 1$ for some $i \in \{1, 2\}$ and demonstrate that there is a region of parameter values such that there are multiple equilibria in the synchronous model without a public randomization device while there is a unique PBE in our main model.

Proposition 5. *Suppose $x_i > 1$. For any $p_{-i} \in (0, 1]$, there exists $\delta' \in (0, 1)$ such that, for any $\delta_i \in (\delta', 1)$, the strategy $\sigma_i^{(T,1)}$ is a best response against $\sigma_{-i}^{(T,1)}$ if and only if the following*

[†]Leavey School of Business, Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053, USA, e-mail: sfukuda@scu.edu

[‡]Haas School of Business, University of California Berkeley, 2220 Piedmont Avenue, Berkeley, CA 94720-1900, USA, and University of Tokyo, Faculty of Economics, 7-3-1 Hongo, Bunkyoku, Tokyo, 113-0033, Japan, e-mail: y.cam.24@gmail.com

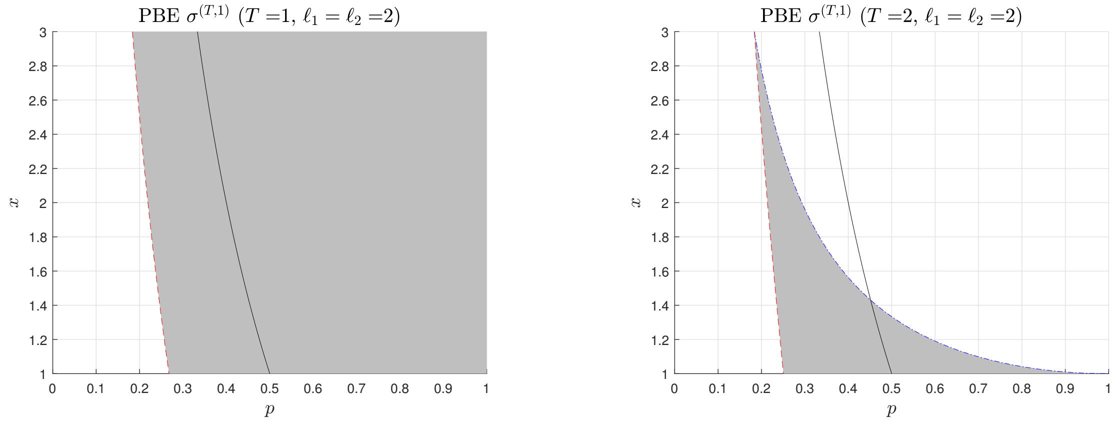

Figure 4: Illustration of Proposition 5 for different values of (p, x). The shaded region depicts the pairs of (p, x) such that $\sigma^{(T,1)}$ is a PBE for sufficiently high discount factors in the synchronous model without a public randomization device. *Left:* $T = 1$. *Right:* $T = 2$. In both panels, we set $\ell_1 = \ell_2 = 2$. The dashed and dashed-dotted curves illustrate the constraints given by (B.1) and (B.2), respectively. The solid curve illustrates the threshold of p below which σ^{G} is a unique PBE for sufficiently high discount factors in our main model with asynchronous moves.

two conditions hold:

$$(1-p_{-i})^T x_i \geq (1-(1-p_{-i})^T)(-\ell_i) + (1-p_{-i})^T \left(T + (1-p_{-i})^T x_i\right); \tag{B.1}$$

$$\frac{x_i - 1}{x_i} < (1-p_{-i})^T. \tag{B.2}$$

Letting $p = p_1 = p_2$ and $x = x_1 = x_2 \geq 1$, Figure 4 uses Proposition 5 to depict the set of (p, x) such that $\sigma^{(T,1)}$ is a PBE for sufficiently high discount factors for a fixed T. Figure 5 then depicts the set of (p, x) such that $\sigma^{(T,1)}$ is a PBE for sufficiently high discount factors for some T. Both figures also depict the threshold probability $p = \frac{2}{1+x+\ell}$ below which σ^{G} is a unique PBE for sufficiently high discount factors in our main model with asynchronous moves.

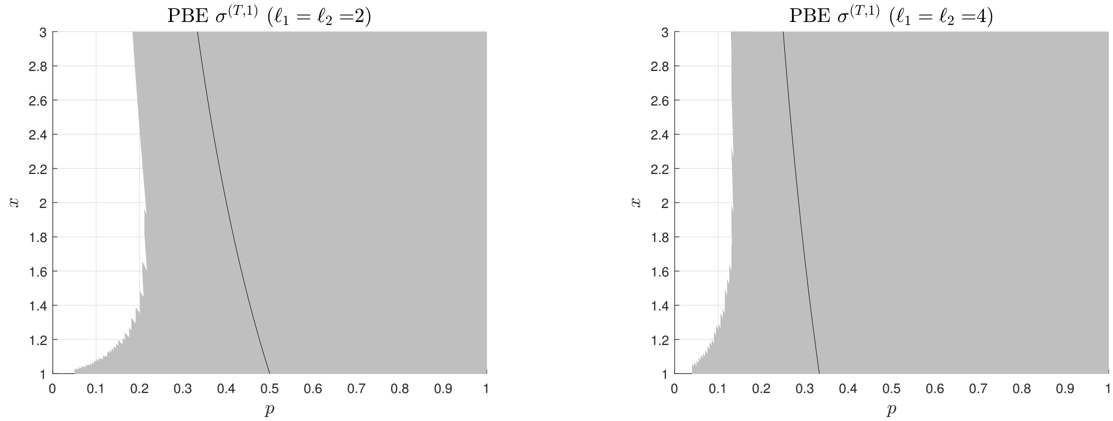

Figure 5: Illustration of Proposition 5 for different values of (p, x). The shaded region depicts the pairs of (p, x) such that $\sigma^{(T,1)}$ is a PBE for sufficiently high discount factors for some T in the synchronous model without a public randomization device. *Left:* $\ell_1 = \ell_2 = 2$. *Right:* $\ell_1 = \ell_2 = 4$. In both panels, the solid curve illustrates the threshold of p below which σ^{G} is a unique PBE for sufficiently high discount factors in our main model with asynchronous moves. By Proposition 4, the shaded area would occupy the entire region (except at $p = 0$) when $x \leq 1$.

Finally, consider again the synchronous model with public randomization. Let us make a formal statement of the result that we mentioned at the end of Section 4.2.

Proposition 6. *For any $p_1, p_2 \in (0, 1]$ and $\overline{q} \in (0, 1)$, there exist $\delta' \in (0, 1)$ and $T' < \infty$ such that, for any sequence $((\delta_i^k)_{i\in\{1,2\}})_{k=1}^{\infty}$ such that $\delta_i^k \in (\delta', 1)$ for each k and $\delta_i^k \to 1$ as $k \to \infty$ for each $i \in \{1, 2\}$, the following hold:*

1. *$\sigma^{(T,\overline{q})}$ is a PBE of the synchronous model with parameters $(p_1, p_2, \overline{q}, \delta_1^k, \delta_2^k)$; and*
2. *Each i's expected payoff from $\sigma^{(T,\overline{q})}$ converges to 0 as $k \to \infty$.*

B.2 The Case with More Than Two Actions

In the main analysis, we assumed that each player has two actions in the stage game. In this subsection, we consider the case that allows for more than two actions, to show the robustness of our main result.

Formally, each player i's action set is $\{a^1, \ldots, a^K\}$. We suppose the following:

(A) For each $k \in \{1, \ldots, K-1\}$, an action profile (a^k, a^k) Pareto dominates (a^{k+1}, a^{k+1}).

(B) $u_i(a^l, a^k) < u_i(a^k, a^k)$ if $l < k$.

Note that, by (B), action profile (a^K, a^K) is a Nash equilibrium of the stage game.

The initial action set is $\{a^1\}$ for each $i \in \{1,2\}$. Once player i chooses a^k, then she cannot take actions in $\{a^1, \ldots, a^{k-1}\}$. For each $k, k' = 1, \ldots, K$ with $k < k'$, there is $p^i_{k,k'} \in [0,1]$. In each period player i moves, when her current action set is $\{a^l, \ldots, a^k\}$ (where $1 \leq l \leq k \leq K$) with $k < k'$, (i) i privately learns $a^{k'}$ with probability $p^i_{k,k'}$ and (ii) she learns $a^{k'}$ if her opponent's latest action is $a^{k'}$. Under any one of these events, her action set becomes $\{a^l, \ldots, a^{k'}\}$. In other words, learning $a^{k'}$ makes all actions a^m with $k < m < k'$ available as well. We call this model the *ordered-action model.*

The set of private histories $H_{i,t}$, strategies, and PBE are defined analogously to Section 2. We extend the notion of grim-trigger strategy profile σ^{G} as follows: At a history $h_{i,t} \in H_{i,t}$ at which each player j's latest action is a^{k_j}, player i plays $\sigma^{\mathrm{G}}_i(h_{i,t}) = a^l$ where $l = \max\{k_1, k_2\}$. Note that, in the two-action case, this reduces to the grim-trigger strategy profile σ^{G} that we defined in Section 3.

Theorem 3. *Consider the ordered-action model. There exist $\overline{p} \in (0,1)$ and $\overline{\delta} \in (0,1)$ such that for any $(p^i_{k,k'})_{i,k,k'}$ with $p^i_{k,k'} < \overline{p}$ for each k, k' such that $k < k'$ and each $i \in \{1,2\}$ and $\delta_i > \overline{\delta}$ for each $i \in \{1,2\}$, σ^{G} is a unique PBE.*

Proof Sketch of Theorem 3. We prove the result by induction. To do this, consider the following claim for each $k \in \{1, \ldots, K\}$:

Claim k. *There exists $\hat{\delta}^{(k)} \in (0,1)$ such that for all $\delta_1, \delta_2 > \hat{\delta}^{(k)}$ and all $k_1, k_2 \geq k$, at a history $h_{i,t} \in H_{i,t}$ at which the latest action profile is (a^{k_1}, a^{k_2}), player i plays $a^{\max(k_1,k_2)}$ under any PBE of the ordered-action model.*

In Appendix B.6.2, we show that for any $k \in \{1, \cdots, K-1\}$, Claim k holds if Claim l holds for every $l \in \{k+1, \cdots, K\}$.

Since Claim K holds by the assumption on the evolution of action sets, by induction, Claim k holds for every k. □

To see that the assumptions on payoffs and the evolution of action sets are essential, Appendix B.5 provides two examples in which the assumptions are violated and the conclusion of Theorem 3 does not hold. The first example features a violation of the assumption on payoffs and has two Pareto-unranked Nash equilibria in the stage game without the original action (a^1). We show that there is an equilibrium in which each player wants to play a certain action other than a^1 as soon as possible to settle into the Nash equilibrium that she prefers before the opponent does so. The game in the second example violates the assumption on the evolution of action sets and multiple actions are available after an action other than a^1 is taken. In such a case, there are multiple equilibria in the "subgame" after an action other than a^1 is taken, just as in the reversible case in Section 4.1. This enables us to construct an equilibrium in which each player wants to play a certain action other than a^1 as soon as possible to obtain a better payoff in the subgame.

B.3 Possibility of No Private Learning of a Precedent

In our main analysis in Section 3, each player privately learns action N with positive probability p in each period (if she has not learned it yet). In reality, there may be cases in which private learning does not occur at all with some probability. This subsection shows that the conclusion of our main result (Theorem 1) is robust to such possibilities.

Formally, we consider the following two variants of our main model. In the first model, there are two states, which realize with positive probability q before the start of the game. In the first state, action N does not exist. That is, no player privately learns action N. In the second state, action N exists and the players learn N as in our main model. That is, whether there is a possibility of private learning is common across players. Call this model the *common-possibility model*. In this model, the characterization of the players' incentives is the same as our main model. This is because, in order to consider a player's incentive to take action N, we only need to consider histories at which the player has learned N and both players have been taking O. Conditional on such histories, the player knows that action N exists, and our analysis reduces to the main one in Section 3. Thus, our main result (Theorem 1) carries over to this model.

In the second model, for each player, there exists a probability-q event, which occurs independently across the players, under which the player cannot privately learn action N on her own (i.e., she learns N only when the opponent plays it). That is, whether there is a possibility of private learning is independently determined. Call this model the *independent-possibility model*. We argue that, in this second model, there exist $\overline{\delta}_1, \overline{\delta}_2 < 1$ such that if $\delta_i > \overline{\delta}_i$ for each $i \in \{1, 2\}$, then for any $p_1, p_2 \in [0, 1]$, σ^{G} is a unique PBE. In order to consider player i's incentive to take action N, it is sufficient to consider histories at which player i has learned N and both players have been taking O. Suppose that player i's strategy assigns positive probability to playing N at such a history at some finite t. Then, i's continuation payoff from such a strategy is $(1 - \delta_i)x_i$. In contrast, if she obeys strategy σ_i^{G}, then her continuation payoff is at least

$$q \cdot 1 + (1 - q)\left((1 - \delta_i) \cdot 1 + \delta_i(1 - \delta_i)(-\ell_i)\right).$$

As δ_i converges to 1, the former and the latter converge to 0 and $q > 0$, respectively. Thus, there exists $\overline{\delta}_i \in (0, 1)$ such that if $\delta_i > \overline{\delta}_i$ then σ_i^{G} is a unique equilibrium strategy for player i.

We summarize the findings of this subsection as follows.

Proposition 7. *In both the common-possibility model and the independent-possibility model, the conclusion of Theorem 1 holds.*

B.4 Full Analysis for the 2×2 Case

Our main theorem pertains to the case when p_i is small and δ_i is high. In this subsection, for any 2×2 games, we characterize the region of p_i and δ_i where we obtain uniqueness and provide comparative statics. The key takeaways from this subsection are: (i) for any (p_1, p_2), we can find a nonempty set of (δ_1, δ_2) such that σ^{G} is a unique PBE, and (ii) the set of discount factors such that the uniqueness obtains weakly expands as the probability of a player's private learning decreases.

The following lemma provides basic properties of best responses.

Lemma 9. *Fix $p_1, p_2 \in [0,1]$. For each $i \in \{1,2\}$, the following hold.*

1. *The strategy σ_i^N is a unique best response against σ_{-i}^N if and only if $x_i > \pi_i(\delta_i)$.*
2. *The strategy σ_i^{G} is a unique best response against σ_{-i}^N if and only if $x_i < \pi_i(\delta_i)$.*
3. *If $\delta_i \in (0, \hat{\delta}_i)$, then player i's strategy must be σ_i^N in any PBE.*

Using this lemma, we obtain the following theorem.

Theorem 4. *Fix $p_1, p_2 \in [0,1]$. Let $\delta_1, \delta_2 \in (0,1)$. Fix $i \in \{1,2\}$.*

1. *If $\delta_i > \hat{\delta}_i$ and $x_{-i} < \pi_{-i}(\delta_{-i})$, then σ^{G} is a unique PBE.*
2. *If $\delta_i < \hat{\delta}_i$ and $x_{-i} < \pi_{-i}(\delta_{-i})$, then $(\sigma_i^N, \sigma_{-i}^{\mathrm{G}})$ is a unique PBE.*
3. *If $\delta_j > \hat{\delta}_j$ and $x_j > \pi_j(\delta_j)$ for each $j \in \{1,2\}$, then both σ^{G} and σ^N are PBE.*
4. *If $\delta_i < \hat{\delta}_i$ and $x_{-i} > \pi_{-i}(\delta_{-i})$, then σ^N is a unique PBE.*

Figure 6 illustrates Theorem 4. Note that Lemma 5 implies that $\delta_i < \hat{\delta}_i$ and $x_i < \pi_i(\delta_i)$ cannot hold simultaneously, and thus, Theorem 4 exhausts all (non-knife-edge) values of discount factors.[1]

[1] By the concavity of the π_i function, $x_i < \pi_i(\delta_i)$ holds when either $\delta_i < \underline{\delta}_i$ or $\overline{\delta}_i < \delta_i$ holds, where $\underline{\delta}_i$ is defined later in this subsection and $\overline{\delta}_i < \delta_i$ follows from the definition of $\overline{\delta}_i$. It is easy to see that $\underline{\delta}_i > 0$ implies $\hat{\delta}_i = 0$. Also, if $\overline{\delta}_i < \delta_i$, then $\delta_i < \hat{\delta}_i$ contradicts Lemma 5 which shows $\hat{\delta}_i \leq \overline{\delta}_i$. Although Theorem 4 does not cover the cases of some knife-edge parameter values, such cases can be easily covered by using the upper-hemicontinuity of the set of PBE. For example, if $\delta_1 = \hat{\delta}_1$ and $x_2 < \pi_2(\delta_2)$, parts 1 and 2 of Theorem 4 imply that both σ^{G} and $(\sigma_1^N, \sigma_2^{\mathrm{G}})$ are PBE.

<table>
<tr><td colspan="2"></td><td colspan="2">$\delta_2 > \hat{\delta}_2$</td><td>$\delta_2 < \hat{\delta}_2$</td></tr>
<tr><td colspan="2"></td><td>$x_2 < \pi_2(\delta_2)$</td><td colspan="2">$x_2 > \pi_2(\delta_2)$</td></tr>
<tr><td rowspan="2">$\delta_1 > \hat{\delta}_1$</td><td>$x_1 <$ $\pi_1(\delta_1)$</td><td rowspan="2">Unique σ^{G} (Part 1)</td><td></td><td>Unique $(\sigma_1^{\mathrm{G}}, \sigma_2^N)$ (Part 2)</td></tr>
<tr><td rowspan="2">$x_1 >$ $\pi_1(\delta_1)$</td><td>σ^{G} and σ^N (Part 3)</td><td></td></tr>
<tr><td>$\delta_1 < \hat{\delta}_1$</td><td>Unique $(\sigma_1^N, \sigma_2^{\mathrm{G}})$ (Part 2)</td><td colspan="2">Unique σ^N (Part 4)</td></tr>
</table>

Figure 6: Illustration of Theorem 4. For example, "Unique σ^{G}" means that σ^{G} is a unique PBE when the parameters satisfy the conditions in the corresponding region. Also, "σ^{G} and σ^N" means that both σ^{G} and σ^N are PBE when the parameters satisfy the conditions in the corresponding region.

Given Theorem 4, we now provide an alternative characterization of PBE and comparative-statics results. In our comparative statics, we show that the set of the profiles of discount factors (δ_1, δ_2) under which σ^{G} is a unique PBE is weakly decreasing in p_i, x_i and ℓ_i (in the set-inclusion sense). To that end, we denote by

$$S = S(p_1, p_2, x_1, x_2, \ell_1, \ell_2)$$

the set of profiles of discount factors $(\delta_1, \delta_2) \in (0,1)^2$ such that σ^{G} is a unique PBE for given parameters $(p_1, p_2, x_1, x_2, \ell_1, \ell_2)$.

The results are stated for each of the following three cases that are exhaustive besides knife-edge cases: (i) $x_i > 1$ for each $i \in \{1,2\}$; (ii) $x_i < 1$ for each $i \in \{1,2\}$; and (iii) $x_i > 1 > x_{-i}$ for some $i \in \{1,2\}$.

First, we start with the case in which $x_i > 1$ holds for each $i \in \{1,2\}$.

Proposition 8. *Fix $p_1, p_2 \in [0,1]$. Let $\delta_1, \delta_2 \in (0,1)$. Suppose $x_i > 1$ for each $i \in \{1,2\}$. Then, $0 < \hat{\delta}_i < \overline{\delta}_i(p_{-i})$ for each $i \in \{1,2\}$. Moreover, the following hold.*

1. *If there is $i \in \{1,2\}$ such that $\delta_i > \hat{\delta}_i$ and $\delta_{-i} > \overline{\delta}_{-i}(p_i)$, then σ^{G} is a unique PBE.*
2. *If $\delta_i \in (\hat{\delta}_i, \overline{\delta}_i(p_{-i}))$ for each $i \in \{1,2\}$, then both σ^{G} and σ^N are PBE.*
3. *If there is $i \in \{1,2\}$ such that $\delta_i \in (0, \hat{\delta}_i)$ and $\delta_{-i} \in (\hat{\delta}_{-i}, \overline{\delta}_{-i}(p_i))$, then σ^N is a unique PBE.*
4. *If there is $i \in \{1,2\}$ such that $\delta_i \in (0, \hat{\delta}_i)$ and $\delta_{-i} \in (\overline{\delta}_{-i}, 1)$, then $(\sigma_i^N, \sigma_{-i}^{\mathrm{G}})$ is a unique PBE.*

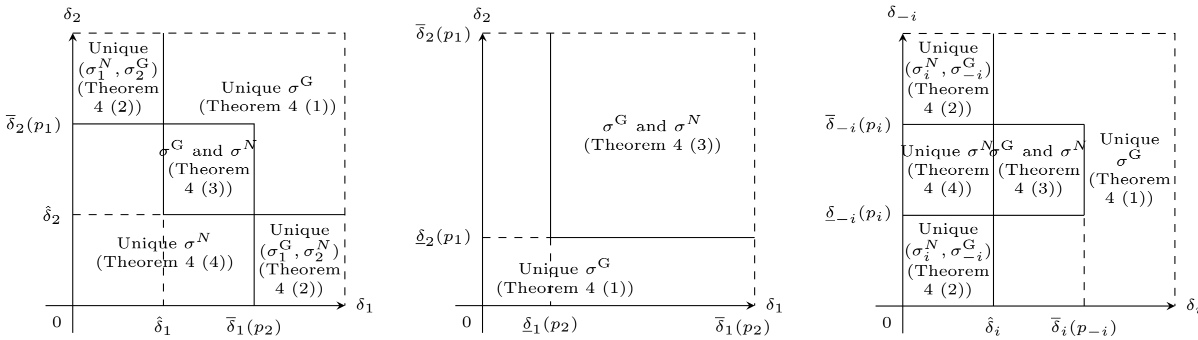

Figure 7: The alternative characterization of PBE. For example, "Unique σ^{G}" means that σ^{G} is a unique PBE in the interior of the corresponding region. Also, "σ^{G} and σ^N" means that both σ^{G} and σ^N are PBE when the parameters satisfy the conditions in the corresponding region. The same comment applies to the subsequent figures. *Left*: The case with $x_i > 1$ for each $i \in \{1, 2\}$. *Center*: The case with $x_i < 1$ for each $i \in \{1, 2\}$. The graph depicts the case in which $\underline{\delta}_i(p_{-i}) \leq \overline{\delta}_i(p_{-i})$ holds (otherwise, we have $\underline{\delta}_i(p_{-i}) = 1$ and $\overline{\delta}_i(p_{-i}) = 0$, and thus σ^{G} is a unique PBE in the entire region; this happens when p_{-i} is low) for each $i \in \{1, 2\}$. In this case, $\overline{\delta}_i(p_{-i}) = 1$ holds for each $i \in \{1, 2\}$. *Right*: The case with $x_i > 1 > x_{-i}$. The graph depicts the case in which $\underline{\delta}_{-i}(p_i) \leq \overline{\delta}_{-i}(p_i)$ holds (otherwise, $\underline{\delta}_{-i}(p_i) = 1$ and $\overline{\delta}_{-i}(p_i) = 0$, and thus only σ_{-i}^{G} is played in the entire region).

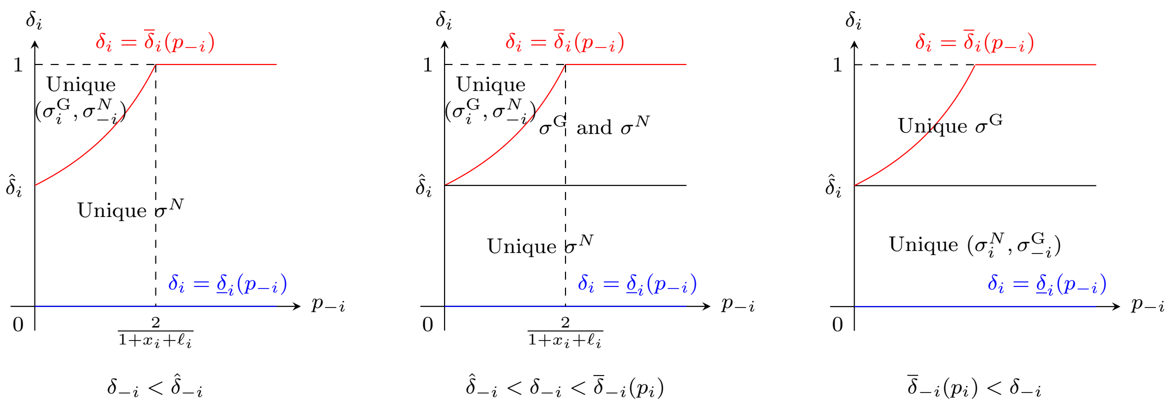

Figure 8: Illustration of Proposition 8 for different values of (p_{-i}, δ_i). *Left*: The case with $\delta_{-i} < \hat{\delta}_{-i}$. *Center*: The case with $\hat{\delta}_{-i} < \delta_{-i} < \overline{\delta}_{-i}(p_i)$. *Right*: The case with $\overline{\delta}_{-i}(p_i) < \delta_{-i}$.

This result is illustrated in the left panel of Figure 7, which illustrates PBE for different values of (δ_1, δ_2). The figure demonstrates, in particular, that for any (p_1, p_2), we can find a nonempty set of (δ_1, δ_2) such that σ^{G} is a unique PBE. Figure 8 illustrates the alternative characterization of PBE for different values of (p_{-i}, δ_i). For the symmetric cases, the left panel of Figure 9 illustrates the characterization. The specific shapes and cutoff values in these graphs can be obtained by algebra (the same comment applies to the other figures to be introduced in this subsection).

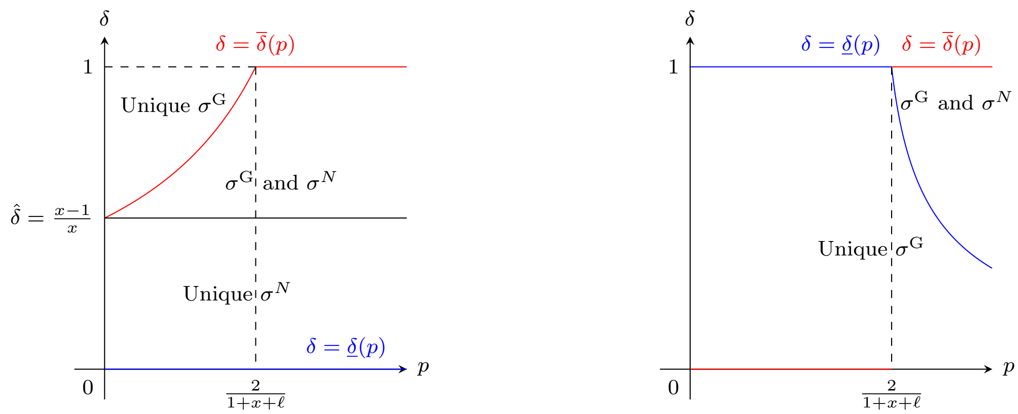

Figure 9: The alternative characterization of PBE for the symmetric cases ($p := p_1 = p_2$, $\delta := \delta_1 = \delta_2$, $x := x_1 = x_2$, and $\ell := \ell_1 = \ell_2$). *Left*: The case with $x > 1$. *Right*: The case with $x < 1$. We let $\overline{\delta}(p) = \overline{\delta}_1(p) = \overline{\delta}_2(p)$, $\underline{\delta}(p) = \underline{\delta}_1(p) = \underline{\delta}_2(p)$, and $\hat{\delta} = \hat{\delta}_1 = \hat{\delta}_2$.

Proposition 8 implies that, when $x_i > 1$ for each $i \in \{1, 2\}$, we have

$$S = \{(\delta_1, \delta_2) \in (0,1)^2 \mid \delta_i > \hat{\delta}_i \text{ and } \delta_{-i} > \overline{\delta}_{-i}(p_i) \text{ for some } i \in \{1,2\}\}. \tag{B.3}$$

Now, we provide the comparative-statics result.

Remark 2. Suppose $x_j > 1$ for each $j \in \{1, 2\}$. Fix $i \in \{1, 2\}$.

1. If $p_i > p_i'$, then $S(p_i, p_{-i}, x_i, x_{-i}, \ell_i, \ell_{-i}) \subseteq S(p_i', p_{-i}, x_i, x_{-i}, \ell_i, \ell_{-i})$.
2. If $x_i > x_i' > 1$, then $S(p_i, p_{-i}, x_i, x_{-i}, \ell_i, \ell_{-i}) \subseteq S(p_i, p_{-i}, x_i', x_{-i}, \ell_i, \ell_{-i})$.
3. If $\ell_i > \ell_i'$, then $S(p_i, p_{-i}, x_i, x_{-i}, \ell_i, \ell_{-i}) \subseteq S(p_i, p_{-i}, x_i, x_{-i}, \ell_i', \ell_{-i})$.

That is, the uniqueness of PBE holds for a weakly wider range of discount factors when private learning is less likely, the instantaneous gain from deviation is smaller, or the loss from the opponent deviating is smaller. These conditions are intuitive.

Second, we consider the case in which $x_i < 1$ for each $i \in \{1, 2\}$. To understand this case, we need to define a new variable:

$$\underline{\delta}_i(p_{-i}) := \max\{\delta_i' \in [0,1] \mid x_i < \pi_i(\delta_i) \text{ for all } \delta_i \in [0, \delta_i')\},$$

where the dependence on p_{-i} is sometimes omitted as for $\overline{\delta}_i$. Recall that $\overline{\delta}_i(p_{-i})$ is the minimum discount factor above which a one-period deviation from σ^N is always profitable for player i. The new variable $\underline{\delta}_i(p_{-i})$ is similar, but it is rather the maximum discount factor below which such one-period deviation is always profitable for player i. When the discount

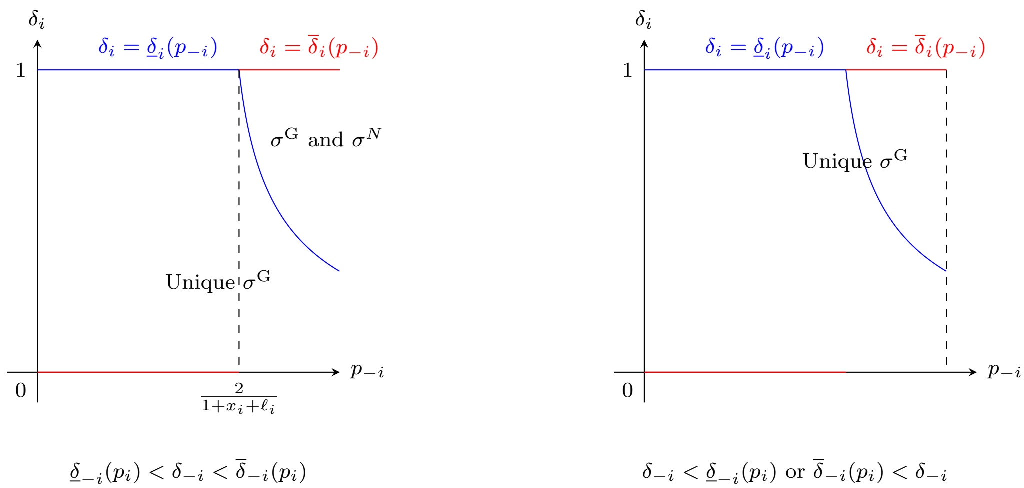

Figure 10: Illustration of Proposition 9 for different values of (p_{-i}, δ_i). *Left*: The case with $\underline{\delta}_{-i}(p_i) < \delta_{-i} < \overline{\delta}_{-i}(p_i)$. *Right*: The case with $\delta_{-i} < \underline{\delta}_{-i}(p_i)$ or $\overline{\delta}_{-i}(p_i) < \delta_{-i}$.

factor is smaller than this value, the player is so impatient that the impact of the one-period loss from playing N is too large ($0 < \underline{\delta}_{-i}$ holds only when $x_{-i} < 1$).

Proposition 9. *Fix $p_1, p_2 \in [0,1]$. Let $\delta_1, \delta_2 \in (0,1)$. Suppose $x_i < 1$ for each $i \in \{1,2\}$.*

1. *If $\delta_i \in (\underline{\delta}_i(p_{-i}), \overline{\delta}_i(p_{-i}))$ for each $i \in \{1,2\}$, then both σ^{G} and σ^N are PBE.*

2. *If $\delta_i \notin [\underline{\delta}_i(p_{-i}), \overline{\delta}_i(p_{-i})]$ for some $i \in \{1,2\}$, then σ^{G} is a unique PBE.*

The central panel of Figure 7 depicts this result when p_i satisfies $\underline{\delta}_{-i}(p_i) \leq \overline{\delta}_{-i}(p_i)$ for each $i \in \{1,2\}$.[2] Figure 10 illustrates Proposition 9 for different values of (p_{-i}, δ_i). For the symmetric cases, the right panel of Figure 9 depicts the characterization for different values of (p, δ).

Proposition 9 implies that, when $x_i < 1$ for each $i \in \{1,2\}$, we have

$$S = \{(\delta_1, \delta_2) \in (0,1)^2 \mid \delta_i \notin [\underline{\delta}_i(p_{-i}), \overline{\delta}_i(p_{-i})] \text{ for some } i \in \{1,2\}\}. \tag{B.4}$$

We obtain the same comparative statics as before, as follows:

Remark 3. Suppose $x_j < 1$ for each $j \in \{1,2\}$. Fix $i \in \{1,2\}$.

1. If $p_i > p_i'$, then $S(p_i, p_{-i}, x_i, x_{-i}, \ell_i, \ell_{-i}) \subseteq S(p_i', p_{-i}, x_i, x_{-i}, \ell_i, \ell_{-i})$.

2. If $x_i > x_i'$, then $S(p_i, p_{-i}, x_i, x_{-i}, \ell_i, \ell_{-i}) \subseteq S(p_i, p_{-i}, x_i', x_{-i}, \ell_i, \ell_{-i})$.

3. If $\ell_i > \ell_i'$, then $S(p_i, p_{-i}, x_i, x_{-i}, \ell_i, \ell_{-i}) \subseteq S(p_i, p_{-i}, x_i, x_{-i}, \ell_i', \ell_{-i})$.

[2] If $\overline{\delta}_{-i}(p_i) < \underline{\delta}_{-i}(p_i)$, then $\overline{\delta}_{-i}(p_i) = 0$ and $\underline{\delta}_{-i}(p_i) = 1$. Thus, σ^{G} is a unique PBE for any (δ_1, δ_2).

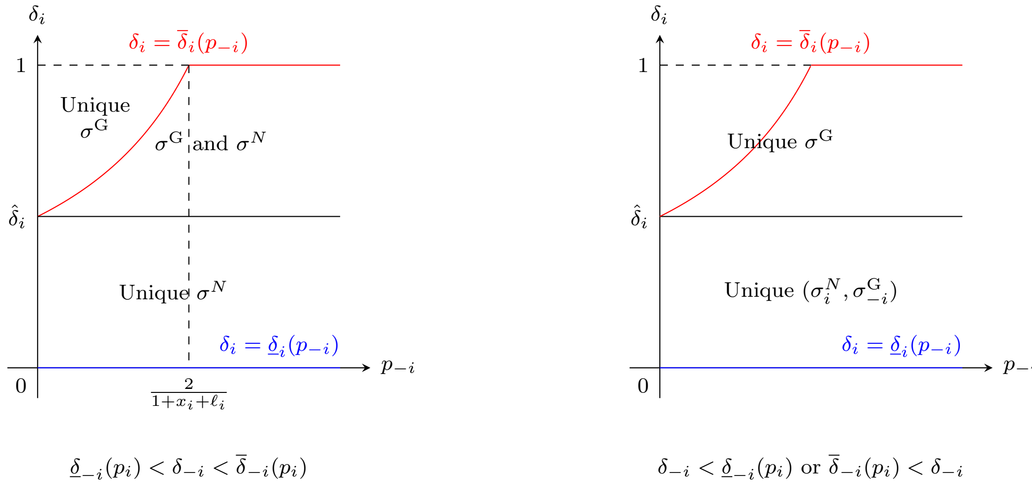

Figure 11: Illustration of Proposition 10 for different values of (p_{-i}, δ_i). *Left*: The case with $\underline{\delta}_{-i}(p_i) < \delta_{-i} < \overline{\delta}_{-i}(p_i)$. *Right*: The case with $\delta_{-i} < \underline{\delta}_{-i}(p_i)$ or $\overline{\delta}_{-i}(p_i) < \delta_{-i}$.

Finally, we consider the case in which $x_i > 1 > x_{-i}$ for some $i \in \{1, 2\}$.

Proposition 10. *Fix $p_1, p_2 \in [0, 1]$. Let $\delta_1, \delta_2 \in (0, 1)$. Suppose $x_i > 1 > x_{-i}$ for some $i \in \{1, 2\}$.*

1. *Suppose $\delta_i > \hat{\delta}_i$.*
   (a) *If $\delta_{-i} \in (\underline{\delta}_{-i}(p_i), \overline{\delta}_{-i}(p_i))$ and $\delta_i < \overline{\delta}_i(p_{-i})$, then both σ^{G} and σ^N are PBE.*
   (b) *If $\delta_{-i} \notin [\underline{\delta}_{-i}(p_i), \overline{\delta}_{-i}(p_i)]$ or $\delta_i > \overline{\delta}_i(p_{-i})$, then σ^{G} is a unique PBE.*
2. *Suppose $\delta_i < \hat{\delta}_i$.*
   (a) *If $\delta_{-i} \in (\underline{\delta}_{-i}(p_i), \overline{\delta}_{-i}(p_i))$, then σ^N is a unique PBE.*
   (b) *If $\delta_{-i} \notin [\underline{\delta}_{-i}(p_i), \overline{\delta}_{-i}(p_i)]$, then $(\sigma_i^N, \sigma_{-i}^{\mathrm{G}})$ is a unique PBE.*

The right panel of Figure 7 illustrates this proposition. The figure depicts the case in which $\underline{\delta}_{-i}(p_i) \leq \overline{\delta}_{-i}(p_i)$. Figure 11 illustrates PBE for different values of (p_{-i}, δ_i).

Proposition 10 implies that, fixing $i \in \{1, 2\}$ with $x_i > 1 > x_{-i}$, we have

$$S = \{(\delta_1, \delta_2) \in (0, 1)^2 \mid \delta_i > \hat{\delta}_i \text{ and } (\delta_{-i} \notin [\underline{\delta}_{-i}(p_i), \overline{\delta}_{-i}(p_i)] \text{ or } \delta_i > \overline{\delta}_i(p_{-i}))\}. \tag{B.5}$$

Again, we obtain the following comparative statics:

Remark 4. Suppose there exists $j \in \{1, 2\}$ such that $x_j > 1 > x_{-j}$. For each $i \in \{1, 2\}$, the following hold.

1. If $p_i > p'_i$, then $S(p_i, p_{-i}, x_i, x_{-i}, \ell_i, \ell_{-i}) \subseteq S(p'_i, p_{-i}, x_i, x_{-i}, \ell_i, \ell_{-i})$.

2. If $x_i > x'_i > 1$ or $1 > x_i > x'_i$, then $S(p_i, p_{-i}, x_i, x_{-i}, \ell_i, \ell_{-i}) \subseteq S(p_i, p_{-i}, x'_i, x_{-i}, \ell_i, \ell_{-i})$.

3. If $\ell_i > \ell'_i$, then $S(p_i, p_{-i}, x_i, x_{-i}, \ell_i, \ell_{-i}) \subseteq S(p_i, p_{-i}, x_i, x_{-i}, \ell'_i, \ell_{-i})$.

We note that, if p_{-j} is small, then we have $\overline{\delta}_j(p_{-j}) < 1$ in the left and right panels of Figure 7 and $\underline{\delta}_j(p_{-j}) = 1$ in the central panel of the same figure. Thus, σ^{G} is indeed a unique PBE when the discount factors are high as shown in Theorem 1.

B.5 Supplementary Discussions for Appendix B.2

We provide two examples in which the assumptions made in Appendix B.2 are violated and the conclusion of Theorem 3 does not hold. In the first example, the assumption on payoffs is violated. The second example pertains to the assumption on the evolution of action sets.

Example 1. Consider the symmetric normal-form game as depicted in Table 3. Suppose $\delta_1 = \delta_2 =: \delta$.

	A	B	C
A	$2, 2$	$-2, 3$	$-2, 3$
B	$3, -2$	$-1, -1$	$0, 1$
C	$3, -2$	$1, 0$	$-2, -2$

Table 3: The payoff matrix of the stage game in Example 1

Note that this game does not satisfy our conditions on the payoffs: If it were to satisfy the conditions, then, since $u_i(A, A) > u_i(B, B) > u_i(C, C)$ for each i, we must have $(a^1, a^2, a^3) = (A, B, C)$ by condition (A). However, we have $u_i(C, B) \geq u_i(C, C)$, violating condition (B).

Suppose, in contrast, that the evolution of action sets satisfies the assumption specified in this subsection, where the actions are ordered as $(a^1, a^2, a^3) = (A, B, C)$. We assume there is $p \in (0, 1)$ such that $p^i_{k,k'} = p$ for each i, k, and k'.

The following strategy profile, in which the players play A forever on the path of play, is a PBE when $\delta \geq \frac{1}{2}$ for any p: (i) player i plays A as long as only A has been taken; (ii) i plays C if i has privately learned C and $-i$'s current action is B; (iii) otherwise, i plays B. See Appendix B.6.4 for the proof of this claim.

The following strategy profile σ^* does not play A forever on the path of play and constitutes a PBE when δ is large and p is small.

- At a history at which player i has taken C, i plays C (this is her only choice).

- At a history at which player i has not taken C:

  - if the opponent has taken C, then player i takes B.
  - if the opponent has not taken C either, then:
    * if action C is available to player i, then i takes C.
    * if action C is not available to player i, then:
      · if the opponent has played B, then player i takes B.
      · if the opponent has not played B, then player i takes A.

Under this strategy profile, the players play C as soon as it becomes available. The reason why this constitutes an equilibrium is that, in the small game that excludes action A, there are two Pareto-unranked Nash equilibria: (B, C) and (C, B), where each player prefers the one in which she plays C. Given that the opponent $-i$ will play C as soon as possible, it is i's best response to play C as soon as possible. This sort of construction depends on the multiple Pareto-unranked equilibria in the small game, and our assumption on the payoff functions excludes such a possibility. Appendix B.6.4 formally shows that σ^* is a PBE. $\square$

Example 2. Consider the symmetric normal-form game as depicted in Table 4. Suppose $\delta_1 = \delta_2 =: \delta$.

	A	B	C
A	$2, 2$	$-2, 3$	$-3, 4$
B	$3, -2$	$1, 1$	$-1, 2$
C	$4, -3$	$2, -1$	$0, 0$

Table 4: The payoff matrix of the stage game in Example 2

Note that this game satisfies the conditions on the payoffs, where the actions are ordered as $(a^1, a^2, a^3) = (A, B, C)$.

Suppose, however, that the evolution of action sets does not satisfy our condition but is given as follows.

1. Each player's action set is $\{A\}$ at the beginning.
2. At each period, with probability $p \in (0, 1)$, each player's action set $\{A\}$ becomes $\{A, B, C\}$.
3. Once i plays B or C, her action set becomes $\{B, C\}$ forever.
4. If i has been always playing A and $-i$ has played B or C in the past, then i's action set is $\{A, B, C\}$.

The following strategy profile, in which the players play A forever on the path of play, is a PBE when $\delta \geq \frac{1}{2}$ for any p: Play A as long as only A has been taken; otherwise, play C. See Appendix B.6.4 for the proof of this claim.

We argue that the following strategy profile σ^* does not play A forever on the path of play and, for any p, it is a PBE when the players are sufficiently patient. For each i, construct a PBE strategy profile $\sigma^{(i)}$ and denote by $u_j^{(i)}$ the payoff this strategy profile yields to player j. The standard folk-theorem argument shows that, if the players are sufficiently patient, we can construct $(\sigma^{(i)})_{i\in\{1,2\}}$ satisfying $u_1^{(1)} > u_1^{(2)}$ and $u_2^{(2)} > u_2^{(1)}$. Now we define σ^* as follows:

- Player i plays A if it is the only available action.
- Player i plays C if C is available to i and no player has taken C in the past.
- Once some player has taken C, the players play $\sigma^{(j)}$ forever, where j is the first player who has taken C.

We show in Appendix B.6.4 that, for any p, the strategy profile σ^* is a PBE when the players are sufficiently patient. The reason why this constitutes an equilibrium is analogous to the analysis of the reversible model in Section 4.1 because the "subgame" after each player chooses an action other than A is that of a reversible model. Recall that the reversible model has multiple equilibria due to the flexibility of action switches. Our assumption on the evolution of action sets shuts down such flexibility of action switches. □

B.6 Proofs for the Results in the Online Appendix

B.6.1 Proofs for Appendix B.1

Proof of Proposition 4. Suppose $\overline{q} = 1$. The proof of Proposition 2 goes through up to (6) under $\overline{q} = 1$. Now, (6) reduces to

$$x_i \leq \frac{1-\delta_i^s}{1-\delta_i} + \delta_i^s \min\{x_i, 0\},$$

which is equivalent to

$$\max\{(1-\delta_i^s)x_i, x_i\} \leq \frac{1-\delta_i^s}{1-\delta_i}. \tag{B.6}$$

Suppose that $x_i \leq 1$ for each $i \in \{1,2\}$. Then, the left-hand side of (B.6) is no greater than 1. Also, the right-hand side of (B.6) is no less than $\frac{1-\delta_i^1}{1-\delta_i} = 1$ because it is increasing in s. Hence, (B.6) holds. This implies that the continuation payoff from playing O is no less than the one from playing N. □

Proof of Proposition 5. We consider two cases.

Case 1. Consider any history $h_{i,t} \in \overline{H}_{i,t}$ such that $t = n{\cdot}T$ for some $n \in \mathbb{N}$. The continuation payoff from N is $(1-p_{-i})^T(1-\delta_i)x_i$. The continuation payoff from O is

$$(1-(1-p_{-i})^T)(1-\delta_i)(-\ell_i) + (1-p_{-i})^T((1-\delta_i^T)\cdot 1 + \delta_i^T(1-\delta_i)(1-p_{-i})^T x_i).$$

Thus, the continuation payoff from N is no less than the one from O if and only if

$$(1-p_{-i})^T(1-\delta_i)x_i \geq (1-(1-p_{-i})^T)(1-\delta_i)(-\ell_i)+(1-p_{-i})^T((1-\delta_i^T)\cdot 1+\delta_i^T(1-\delta_i)(1-p_{-i})^T x_i),$$

or

$$(1-p_{-i})^T x_i \geq (1-(1-p_{-i})^T)(-\ell_i) + (1-p_{-i})^T\left(\frac{1-\delta_i^T}{1-\delta_i} + \delta_i^T(1-p_{-i})^T x_i\right).$$

There exists $\delta'' < 1$ such that this holds for all $\delta_i > \delta''$ if and only if:

$$(1-p_{-i})^T x_i \geq (1-(1-p_{-i})^T)(-\ell_i) + (1-p_{-i})^T\left(T + (1-p_{-i})^T x_i\right),$$

which we obtain by letting $\delta_i \to 1$. This is condition (B.1).

Case 2. Consider any history $h_{i,t} \in \overline{H}_{i,t}$ such that $t \neq n \cdot T$ for any $n \in \mathbb{N}$. Let $s \in \{1,\dots,T-1\}$ be such that the current period is $t = nT - s$ for some $n \in \mathbb{N}$. If i plays N at $h_{i,t}$, then her continuation payoff is $(1-\delta_i)x_i$. If she plays O instead, then her continuation payoff is

$$(1-\delta_i^s)\cdot 1 + \delta_i^s(1-p_{-i})^T(1-\delta_i)x_i. \tag{B.7}$$

Notice that this is a convex combination of 1 and $(1-p_{-i})^T(1-\delta_i)x_i$.

Case 2-1. If $1 \leq (1-p_{-i})^T(1-\delta_i)x_i$, then (B.7) is no greater than $(1-p_{-i})^T(1-\delta_i)x_i$. Thus, the continuation payoff from O is no less than the one from N only if

$$(1-p_{-i})^T(1-\delta_i)x_i \geq (1-\delta_i)x_i,$$

which is equivalent to $(1-p_{-i})^T \geq 1$, a contradiction. Hence, we need $1 > (1-p_{-i})^T(1-\delta_i)x_i$ for $\sigma^{(T,1)}$ to be a PBE.

Case 2-2. If $1 > (1-p_{-i})^T(1-\delta_i)x_i$, then (B.7) is minimized at $s = 1$, and the minimized value is

$$(1-\delta_i)\cdot 1+\delta_i(1-p_{-i})^T(1-\delta_i)x_i.$$

Thus, at any history $h_{i,t}$ that we consider in Case 2, the continuation payoff from O is no less than the one from N if and only if

$$(1-\delta_i)x_i \leq (1-\delta_i)\cdot 1+\delta_i(1-p_{-i})^T(1-\delta_i)x_i,$$

which is equivalent to

$$x_i \leq 1+\delta_i(1-p_{-i})^T x_i,$$

or

$$(1-p_{-i})^T \geq \frac{x_i-1}{\delta_i x_i}.$$

Combining with $1>(1-p_{-i})^T(1-\delta_i)x_i$, we need

$$\frac{x_i-1}{\delta_i x_i} \leq (1-p_{-i})^T < \frac{1}{(1-\delta_i)x_i}.$$

There exists $\delta''<1$ such that this holds for all $\delta_i>\delta''$ if and only if:

$$\frac{x_i-1}{x_i} < (1-p_{-i})^T,$$

which we obtain by letting $\delta_i \to 1$. This is condition (B.2).

□

Proof of Proposition 6. Fix any $p_1,p_2 \in (0,1]$ and $\overline{q}\in(0,1)$. We have shown in the proof of Proposition 2 that there exist $\delta' \in (0,1)$ and $T'<\infty$ such that $\sigma^{(T,\overline{q})}$ is a PBE of the synchronous model if $\delta_i \in (\delta',1)$ for each $i\in\{1,2\}$ and $T>T'$.

Let H_i^1 be the set of histories at period $t=n\cdot T+1$ for some $n=0,1,\dots$, where the realization of private learning or the public randomization device has not occurred yet.[3] For any $h_{i,t}\in H_i^1$, denoting by $V_i(h_{i,t})$ player i's continuation payoff at $h_{i,t}$ under $\sigma^{(T,\overline{q})}$, we have:

$$V_i(h_{i,t}) \leq (1-\delta_i^{T-1})\cdot 1+\delta_i^{T-1}\Bigg(\overline{q}(1-(1-p_1)^T(1-p_2)^T)(1-\delta_i)\max\{x_i,0\} + (\overline{q}(1-p_1)^T(1-p_2)^T+1-\overline{q})((1-\delta_i)+\delta_i \sup_{h'_{i,t}\in H_i^1} V_i(h'_{i,t}))\Bigg).$$

[3]Strictly speaking, these are not "histories" as defined in Section 2, and thus this involves an abuse of terminology. Nonetheless, the payoff is still well defined.

Since this inequality must hold for every $h_{i,t} \in H_i^1$, it follows that $\sup_{h'_{i,t} \in H_i^1} V_i(h'_{i,t})$ is no greater than the right-hand side of the above inequality. Rearranging, we obtain:

$$(1-\delta_i^{T-1}(\overline{q}(1-p_1)^T(1-p_2)^T+1-\overline{q})\delta_i) \sup_{h'_{i,t}\in H_i^1} V_i(h'_{i,t}) \leq (1-\delta_i^{T-1}) \cdot 1$$
$$+\,\delta_i^{T-1}\left(\overline{q}(1-(1-p_1)^T(1-p_2)^T)(1-\delta_i)\max\{x_i,0\}+(\overline{q}(1-p_1)^T(1-p_2)^T+1-\overline{q})(1-\delta_i)\right). \tag{B.8}$$

For any fixed $T<\infty$, notice that the limit superior of the left-hand side of (B.8) as $\delta_i \to 1$ is

$$(1-(\overline{q}(1-p_1)^T(1-p_2)^T+1-\overline{q})) \cdot \limsup_{\delta_i \to 1} \sup_{h'_{i,t}\in H_i^1} V_i(h_{i,t}).$$

Notice that $1-(\overline{q}(1-p_1)^T(1-p_2)^T+1-\overline{q})>0$ because $\overline{q}>0$, $p_1, p_2>0$, and $T>0$. On the other hand, the right-hand side of (B.8) goes to 0 as $\delta_i \to 1$. This implies that

$$\limsup_{\delta_i \to 1} \sup_{h'_{i,t}\in H_i^1} V_i(h'_{i,t}) \leq 0.$$

Finally, notice that we must have

$$\liminf_{\delta_i \to 1} \inf_{h_{i,t}\in H_i^1} V_i(h_{i,t}) \geq 0$$

because the limit payoff of σ_i^N as $\delta_i \to 1$ is no less than 0 for any strategy of the opponent. Hence, for any $h_{i,t} \in H_i^1$, i's expected payoff under $\sigma^{(T,\overline{q})}$ at that history converges to 0 as $\delta_i \to 1$. By setting $n=0$, we conclude that i's expected payoff of the game under $\sigma^{(T,\overline{q})}$ converges to 0 as $\delta_i \to 1$. □

B.6.2 Proofs for Appendix B.2

Proof of Theorem 3. First, Claim K holds by the assumption on the evolution of action sets.

Next, fix $k \leq K-1$. Suppose, as an induction hypothesis, that Claim l holds for every $l=k+1,\cdots,K$. We show that Claim k holds. For this purpose, assume $\delta > \hat{\delta}^{(k+1)}$ and fix a PBE σ. Fix i and t. The proof consists of three steps.

For the first step, consider a history $h_{i,t} \in H_{i,t}$ at which player i's latest action is a^k and player $-i$'s latest action is $a^{k'}$, where $k'>k$. The induction hypothesis implies the following.

1. If i plays a^l such that $l > k'$, then her payoff is

$$(1-\delta_i)u_i(a^l, a^{k'}) + \delta_i u_i(a^l, a^l).$$

   Due to condition (A), there is $\hat{\delta}_i < 1$ such that for all $\delta_i \in (\hat{\delta}_i, 1)$, this is strictly less than $u_i(a^{k'}, a^{k'})$.

2. If i plays a^l such that $l = k'$, then her payoff is

$$u_i(a^{k'}, a^{k'}).$$

3. If i plays a^l such that $k < l < k'$, then her payoff is at most

$$(1-\delta_i^2)u_i(a^l, a^{k'}) + \delta_i^2 u_i(a^{k'}, a^{k'}).$$

   Due to condition (B), this is strictly less than $u_i(a^{k'}, a^{k'})$.

4. If i plays a^l such that $l = k$, then cases 1–3 above imply that her payoff is a convex combination of $u_i(a^k, a^{k'})$ and at most $u_i(a^{k'}, a^{k'})$ when $\delta_i \in (\hat{\delta}_i, 1)$, with a strictly positive weight on the former payoff. Due to condition (B), such a convex combination is strictly less than $u_i(a^{k'}, a^{k'})$.

Overall, we find that it is i's unique best response to play $a^{k'}$ if $k' > k$, provided $\delta_i > \max(\hat{\delta}^{(k+1)}, \hat{\delta}_i) < 1$.

For the second step, consider a history $h_{i,t} \in H_{i,t}$ at which player i's latest action is $a^{k'}$ and player $-i$'s latest action is a^k, where $k' > k$. Then, if i plays a^l such that $l \geq k'$, we know from the first step that $-i$ plays a^l at $t+1$, and by the induction hypothesis, the action profile (a^l, a^l) will be played thereafter at every period. Hence, the payoff is

$$(1-\delta_i)u_i(a^l, a^k) + \delta_i u_i(a^l, a^l).$$

Due to condition (A), there is $\hat{\delta}'_i < 1$ such that for all $\delta \in (\hat{\delta}'_i, 1)$, this is maximized when $l = k'$. Hence, we find that it is i's unique best response to play $a^{k'}$ if $k' > k$, provided $\delta_i > \max(\hat{\delta}^{(k+1)}, \hat{\delta}_{-i}, \hat{\delta}'_i)$. Note that $\max(\hat{\delta}^{(k+1)}, \hat{\delta}_{-i}, \hat{\delta}'_i) < 1$ holds.

For the third step, consider a history $h_{i,t} \in H_{i,t}$ at which player i's latest action is a^k and player $-i$'s latest action is also a^k.

We start with showing that for all time $t' \geq t$, for all history $h_{j,t'}$ where j is the moving player at t' such that the current action profile is (a^k, a^k), j does not play a^l with $k+1 < l$.

To see this, note that our first step implies that the payoff from playing $a^{l'}$ with $k < l'$ is

$$(1-\delta_j)u_j(a^{l'}, a^k) + \delta_j u_j(a^{l'}, a^{l'}).$$

Due to condition (A), there is $\hat{\delta}''_j < 1$ such that for all $\delta_j \in (\hat{\delta}''_j, 1)$, this is maximized when $l' = k+1$ among $l' \in \{k+1, \ldots, K\}$. Since action a^{k+1} is available to a player when action $a^{l'}$ with $k+1 < l'$ is available to that player, this implies that when the current action profile is (a^k, a^k), the moving player j will never take any action a^l with $k+1 < l$ if $\delta > \hat{\delta}''_j$.

Note that the first step, the second step, and the induction hypothesis pin down the players' actions under σ when the current action profile is (a^k, a^{k+1}),(a^{k+1}, a^k), and (a^{k+1}, a^{k+1}), respectively, and they imply that the chosen action is a^{k+1} with probability 1 in either case. This implies that at every time $t' \geq t$ such that the current action profile is (a^k, a^k), the moving player's optimization problem is the same as the one for 2×2 games. Thus, by letting $\hat{\delta}^{(k)} := \max(\hat{\delta}^{(k+1)}, \hat{\delta}_1, \hat{\delta}_2, \hat{\delta}'_1, \hat{\delta}'_2, \hat{\delta}''_1, \hat{\delta}''_2) < 1$, Theorem 1 implies that it is i's unique best response to play a^k if the current action profile is (a^k, a^k), provided $\delta_i > \hat{\delta}^{(k)}$. This concludes the third step.

Overall, we find that Claim k holds. Since Claim k holds if Claim l holds for all $l \in \{k+1, \ldots, K\}$, by induction, Claim k holds for every k. Thus, we conclude that σ^{G} is a unique candidate for a PBE if $\delta > \max_k \hat{\delta}^{(k)} (= \hat{\delta}^{(1)})$. The same argument shows that σ^{G} is indeed a PBE under the same condition. The proof is complete. □

B.6.3 Proofs for Appendix B.4

Proof of Lemma 9. Fix $i \in \{1, 2\}$.

1. First, notice that σ_i^N designates a unique best response at any history such that player i's opponent has already chosen N in the past.

   So, consider a history in $\overline{H}_{i,t}$. If i takes action N, her continuation payoff is $(1-\delta_i)x_i$. If, instead, she takes action O, then her continuation payoff is

   $$(1-\delta_i)\cdot 1 + \delta_i \left(p_{-i}((1-\delta_i)(-\ell_i) + \delta_i \cdot 0) + (1-p_{-i})((1-\delta_i)\cdot 1 + \delta_i(1-\delta_i)x_i)\right),$$

   which is $(1-\delta_i)\pi_i(\delta_i)$. Then, we have that the former is strictly greater than the latter if and only if $x_i > \pi_i(\delta_i)$.[4] Thus, σ_i^N is a unique best response against σ_{-i}^N whenever $x_i > \pi_i(\delta_i)$.

[4]Note that this calculation is essentially the same as in the proof of Lemma 6.

2. First, notice that σ_i^{G} designates a unique best response at any history such that player i's opponent has already chosen N in the past.

   So, consider a history in $\overline{H}_{i,t}$. If i takes action N, her continuation payoff is $(1-\delta_i)x_i$. If, instead, she takes action O, then her continuation payoff is

   $$(1-\delta_i)\cdot 1+\delta_i\left(p_{-i}((1-\delta_i)(-\ell_i)+\delta_i\cdot 0)+(1-p_{-i})((1-\delta_i)\cdot 1+\delta_i(1-\delta_i)x_i)\right),$$

   which is $(1-\delta_i)\pi_i(\delta_i)$. Then, we have the latter being strictly greater than the former if and only if $x_i<\pi_i(\delta_i)$.[5] Thus, σ_i^{G} is a unique best response against σ_{-i}^N whenever $x_i<\pi_i(\delta_i)$.

3. Fix any PBE. Note first that, if player i chooses N at time t, then $-i$ chooses N at time $t+1$. Consider a history $h_{i,t}\in\overline{H}_{i,t}$ at time t in which player i moves. Notice that if i plays N at t, her payoff is $(1-\delta_i)\cdot x_i+\delta_i\cdot 0=(1-\delta_i)x_i$. If she plays O at t but $-i$ plays N at time $t+1$, then i's payoff is $(1-\delta_i)\cdot 1+\delta_i[(1-\delta_i)\cdot(-\ell_i)+\delta_i\cdot 0]=(1-\delta_i)(1-\delta_i\ell_i)$. If no one plays N at any time, i's continuation payoff is 1. Hence, i's continuation payoff at $h_{i,t}$ is in the following set:

   $$\mathrm{co}\left(\{\delta_i^{2n}(1-\delta_i)x_i|n=0,1,\dots\}\cup\{\delta_i^{2n+1}(1-\delta_i)(1-\delta_i\ell_i)|n=0,1,\dots\}\cup\{1\}\right),$$

   where $\mathrm{co}(S)$ denotes the convex hull of the set S. If $\delta_i<\frac{x_i-1}{x_i}$, the maximum element in this set is $(1-\delta_i)x_i$ because $(1-\delta_i)x_i>1$ and $(1-\delta_i)(1-\delta_i\ell_i)<1$. Hence, any action that induces the continuation payoff of $(1-\delta_i)x_i$ is a best response at $h_{i,t}$. Hence, choosing action N is a best response at $h_{i,t}$. Since there is an action that induces the payoff of $(1-\delta_i)x_i$ at $h_{i,t}$, no other action is a best response. Hence, at $h_{i,t}$, choosing action N is a unique best response and thus, i chooses N at $h_{i,t}$ under the PBE we fixed.

   Since this was true for any choice of $h_{i,t}$ from $\overline{H}_{i,t}$, this shows that if $\delta_i<\frac{x_i-1}{x_i}$, in any PBE, i chooses action N at any history in $\overline{H}_{i,t}$. This completes the proof.

$\square$

To prove Theorem 4, we generalize Lemmas 6 and 7, respectively, as follows.

Lemma 6'. *Fix $i\in\{1,2\}$. Suppose $\delta_i>\hat{\delta}_i$ and $\delta_{-i}\in(0,\underline{\delta}_{-i})\cup(\overline{\delta}_{-i},1)$. Then, for any $k\in\mathbb{N}$ such that i moves at t_k, $\sigma_i^*(h_{i,t_k})(N)<1$ holds for some $h_{i,t_k}\in\overline{H}_{i,t_k}$.*

[5] Again, this calculation is essentially the same as in the proof of Lemma 6.

Lemma 7'. *Suppose $p_j \in [0,1)$ for each $j \in \{1,2\}$ and suppose $\delta_i > \hat{\delta}_i$ and $\delta_{-i} \in (0, \underline{\delta}_{-i}) \cup (\overline{\delta}_{-i}, 1)$. Then, there exist $j \in \{1,2\}$ and $\alpha_j^* > 1$ such that, for any $t \geq t_1$ such that j moves, we have $1 - \mu_{t+2} = \alpha_j^*(1 - \mu_t)$.*[6]

The proofs for these results are similar to those of Lemmas 6 and 7, respectively, and are omitted.[7]

Proof of Theorem 4.

1. This part follows from Lemmas 1-5, 6' and 7'.
2. By part 3 of Lemma 9, in any PBE, player i's equilibrium strategy is σ_i^N. Then, by part 2 of Lemma 9, player $-i$'s unique best response is σ_{-i}^{G}. Thus, the result obtains.
3. Lemma 1 shows that σ^{G} is a PBE. Also, by part 2 of Lemma 9, σ^N is a PBE.
4. By part 3 of Lemma 9, in any PBE, player i's equilibrium strategy is σ_i^N. Then, by part 1 of Lemma 9, player $-i$'s unique best response is σ_{-i}^N. Thus, the result obtains.

□

Proof of Proposition 8. For each $i \in \{1,2\}$, when $x_i > 1$, it follows from the proof of Lemma 5 that $\hat{\delta}_i = \frac{x_i - 1}{x_i} > 0$ and $\hat{\delta}_i < \overline{\delta}_i(p_{-i})$. Also, for each $i \in \{1,2\}$, since $\pi_i(0) = 1 < x_i$, we have $\underline{\delta}_i(p_{-i}) = 0$. Hence, $x_i < \pi_i(\delta_i)$ if $\delta_i > \overline{\delta}_i(p_{-i})$, and $x_i > \pi_i(\delta_i)$ if $\delta_i < \overline{\delta}_i(p_{-i})$.

Then, the proposition follows from Theorem 4. □

Proof of Proposition 9. Recall that, for each $i \in \{1,2\}$, if $x_i < 1$ then $\hat{\delta}_i = 0$. Also, for each $i \in \{1,2\}$, since $\pi_i(0) = 1 > x_i$, we have $\underline{\delta}_i(p_{-i}) > 0$.

Since $\hat{\delta}_i = 0$ for each $i \in \{1,2\}$, it follows from parts 1 and 3 of Theorem 4 that σ^{G} is a unique PBE if $x_i < \pi_i(\delta_i)$ for some $i \in \{1,2\}$; and that both σ^{G} and σ^N are PBE if $x_i > \pi_i(\delta_i)$ for all $i \in \{1,2\}$.

Thus, it suffices to show the following two assertions for each $i \in \{1,2\}$. First, if $\delta_i \in (\underline{\delta}_i(p_{-i}), \overline{\delta}_i(p_{-i}))$, then $x_i > \pi_i(\delta_i)$. Second, if $\delta_i \notin [\underline{\delta}_i(p_{-i}), \overline{\delta}_i(p_{-i})]$, then $x_i < \pi_i(\delta_i)$.

Fix $i \in \{1,2\}$. We consider the following three exhaustive cases. As the first case, suppose $\underline{\delta}_i(p_{-i}) = 1$. Then, the first assertion vacuously follows because $(\underline{\delta}_i(p_{-i}), \overline{\delta}_i(p_{-i})) = \emptyset$. For the second assertion, it follows from the definition of $\underline{\delta}_i(p_{-i})$ that $x_i < \pi_i(\delta_i)$ for all $\delta_i \in (0,1)$. Then, we have $x_i < \pi_i(\delta_i)$ for all $\delta_i \notin [\underline{\delta}_i(p_{-i}), \overline{\delta}_i(p_{-i})]$, as desired.

[6]Note that, in Lemma 7, we are able to replace $\delta_i > \hat{\delta}_i$ in Lemma 7' with $\delta_i > \overline{\delta}_i$ because Lemma 5 implies $\overline{\delta}_i \geq \hat{\delta}_i$.

[7]The interested reader may refer to an old working-paper version of the present paper (available on our websites) which provides the proofs of Lemmas 6' and 7'.

As the second case, suppose that $\underline{\delta}_i(p_{-i}) \neq 1$ and $\overline{\delta}_i(p_{-i}) = 1$. Then, $\underline{\delta}_i(p_{-i}) \in (0,1)$. Also, $x_i \geq \pi_i(1)$ (otherwise, $\overline{\delta}_i(p_{-i}) < 1$). This means that the quadratic equation $x_i = \pi_i(\delta_i)$ has a unique interior solution in $(0,1)$, which is $\underline{\delta}_i(p_{-i})$. Thus, $x_i < \pi_i(\delta_i)$ if $\delta_i \in (0, \underline{\delta}_i(p_{-i}))$, that is, $\delta_i \notin [\underline{\delta}_i(p_{-i}), \overline{\delta}_i(p_{-i})]$. Also, $x_i > \pi_i(\delta_i)$ if $\delta_i \in (\underline{\delta}_i(p_{-i}), 1)$, that is, $\delta_i \in (\underline{\delta}_i(p_{-i}), \overline{\delta}_i(p_{-i}))$. Hence, the two assertions hold.

As the third case, suppose that $\underline{\delta}_i(p_{-i}) \neq 1$ and $\overline{\delta}_i(p_{-i}) \neq 1$. Then, $\underline{\delta}_i(p_{-i}) \in (0,1)$. Also, $x_i < \pi_i(1)$ (otherwise, $\overline{\delta}_i(p_{-i}) = 1$). This means that the quadratic equation $x_i = \pi_i(\delta_i)$ has either (i) two interior solutions in $(0,1)$, which are $\underline{\delta}_i(p_{-i})$ and $\overline{\delta}_i(p_{-i})$ with $\underline{\delta}_i(p_{-i}) < \overline{\delta}_i(p_{-i})$, or (ii) a unique solution in $(0,1)$, which is $\underline{\delta}_i(p_{-i}) = \overline{\delta}_i(p_{-i})$. Thus, $x_i < \pi_i(\delta_i)$ if $\delta_i \in (0, \underline{\delta}_i(p_{-i})) \cup (\overline{\delta}_i(p_{-i}), 1)$, that is, $\delta_i \notin [\underline{\delta}_i(p_{-i}), \overline{\delta}_i(p_{-i})]$. Also, $x_i > \pi_i(\delta_i)$ if $\delta_i \in (\underline{\delta}_i(p_{-i}), \overline{\delta}_i(p_{-i}))$. Hence, the two assertions hold.

In sum, for each of the three cases, the statement of the proposition holds. $\square$

Proof of Proposition 10. Suppose that $x_i > 1 > x_{-i}$ for some $i \in \{1,2\}$. Then, it follows from the proof of Lemma 5 that $\overline{\delta}_i(p_{-i}) > \hat{\delta}_i = \frac{x_i - 1}{x_i} > 0$. Also, $\hat{\delta}_{-i} = 0$. For player i, as in the proof of Proposition 8, $x_i < \pi_i(\delta_i)$ if $\delta_i > \overline{\delta}_i(p_{-i})$, and $x_i > \pi_i(\delta_i)$ if $\delta_i < \overline{\delta}_i(p_{-i})$. For player $-i$, similarly to the proof of Proposition 9, $x_{-i} < \pi_{-i}(\delta_{-i})$ if $\delta_{-i} \notin [\underline{\delta}_{-i}(p_i), \overline{\delta}_{-i}(p_i)]$, and $x_{-i} > \pi_{-i}(\delta_{-i})$ if $\delta_{-i} \in (\underline{\delta}_{-i}(p_i), \overline{\delta}_{-i}(p_i))$. Given these conclusions, the statement of the proposition follows from Theorem 4. $\square$

To prove Remarks 2 to 4, we provide the following auxiliary result.

Lemma 10. *Fix $i \in \{1,2\}$.*

1. *The threshold discount factor $\overline{\delta}_i$ is non-decreasing in p_{-i}, x_i, and ℓ_i.*

2. *The threshold discount factor $\underline{\delta}_i$ is non-increasing in p_{-i}, x_i, and ℓ_i.*

Proof of Lemma 10. We prove both parts at once. We define

$$D(p_{-i}, x_i, \ell_i) := \{\delta_i \in [0,1] \mid x_i < \pi_i(\delta_i; p_{-i}, x_i, \ell_i)\},$$

where we made explicit the dependence of π_i on the parameters.

(i) Comparative statics with respect to p_{-i}: For any x_i and ℓ_i, we define

$$A(\delta) := 1 + \delta_i(-\ell_i) + \delta_i^2 \cdot 0 \quad \text{and} \quad B(\delta) := 1 + \delta_i \cdot 1 + \delta_i^2 x_i.$$

We consider the following two cases.

Case 1. Suppose $A(1) \geq B(1)$ and $x_i < 0$. Since $B(1) = 2 + x_i$, we have $x_i < \pi_i(1; p_{-i}, x_i, \ell_i)$. Since π_i is concave in δ_i and $\pi_i(1; p_{-i}, x_i, \ell_i) = 0$, it follows that $x_i < \pi_i(\delta_i; p_{-i}, x_i, \ell_i)$ for all $\delta_i \in [0, 1]$. Thus, we have $\overline{\delta}_i(p_{-i}) = 0$ and $\underline{\delta}_i(p_{-i}) = 1$ for all p_{-i}.

Case 2. Suppose $A(1) < B(1)$ or $x_i \geq 0$. We first show $A(\delta_i) \leq B(\delta_i)$ for all $\delta_i \in [0, 1]$. We consider the following two subcases.

Case 2-1. Suppose $x_i \geq 0$. Then, we have $B(\delta) - A(\delta) = x_i\delta_i^2 + (1 + \ell_i)\delta_i \geq 0$ for all $\delta_i \in [0, 1]$.

Case 2-2. Suppose $x_i < 0$ and $A(1) < B(1)$. Since B is concave in δ_i and $A(0) = B(0)$, we have $A(\delta_i) \leq B(\delta_i)$ for all $\delta_i \in [0, 1]$.

Since $\pi_i(\delta_i; p_{-i}, x_i, \ell_i) = p_{-i}A(\delta_i) + (1 - p_{-i})B(\delta_i)$, we have

$$\pi_i(\delta_i; p''_{-i}, x_i, \ell_i) - \pi_i(\delta_i; p'_{-i}, x_i, \ell_i) = (p''_{-i} - p'_{-i})(B(\delta_i) - A(\delta_i)).$$

If $p'_{-i} < p''_{-i}$, then this value is non-negative. Thus, $D(p'_{-i}, x_i, \ell_i) \supseteq D(p''_{-i}, x_i, \ell_i)$. Hence, by the definition of $\overline{\delta}_i$, it follows that $\overline{\delta}_i$ is no greater under p'_{-i} than under p''_{-i}. Similarly, by the definition of $\underline{\delta}_i$, it follows that $\underline{\delta}_i$ is no smaller under p'_{-i} than under p''_{-i}.

(ii) Comparative statics with respect to x_i: For any δ_i, p_{-i}, and ℓ_i, we have

$$\pi_i(\delta_i; p_{-i}, x''_i, \ell_i) - \pi_i(\delta_i; p_{-i}, x'_i, \ell_i) = (1 - p_{-i})\delta_i^2(x''_i - x'_i) \leq x''_i - x'_i.$$

Thus,

$$\pi_i(\delta_i; p_{-i}, x'_i, \ell_i) - x'_i \geq \pi_i(\delta_i; p_{-i}, x''_i, \ell_i) - x''_i,$$

which implies $D(p_{-i}, x'_i, \ell_i) \supseteq D(p_{-i}, x''_i, \ell_i)$. Hence, by the definition of $\overline{\delta}_i$, it follows that $\overline{\delta}_i$ is no greater under x'_i than under x''_i. Similarly, by the definition of $\underline{\delta}_i$, it follows that $\underline{\delta}_i$ is no smaller under x'_i than under x''_i.

(iii) Comparative statics with respect to ℓ_i: For any δ_i, p_{-i}, and x_i, we have $\pi_i(\delta_i; p_{-i}, x_i, \ell'_i) \geq \pi_i(\delta_i; p_{-i}, x_i, \ell''_i)$ if $\ell'_i \leq \ell''_i$. Thus, $D(p_{-i}, x_i, \ell'_i) \supseteq D(p_{-i}, x_i, \ell''_i)$. Hence, by the definition of $\overline{\delta}_i$, it follows that $\overline{\delta}_i$ is no greater under ℓ'_i than under ℓ''_i. Similarly, by the definition of $\underline{\delta}_i$, it follows that $\underline{\delta}_i$ is no smaller under ℓ'_i than under ℓ''_i.

□

Proof of Remark 2. Recall that the set S is given by equation (B.3). All the statements of this remark follow if, for each $j \in \{1, 2\}$, the threshold discount factors $\hat{\delta}_j$ and $\overline{\delta}_j$ are non-decreasing in x_j, p_{-j}, and ℓ_j. Since $x_j > 1$, $\hat{\delta}_j = \frac{x_j - 1}{x_j}$ is indeed non-decreasing in x_j, p_{-j}, and ℓ_j. For $\overline{\delta}_j$, it follows from part 1 of Lemma 10 that $\overline{\delta}_j$ is non-decreasing in x_j, p_{-j}, and ℓ_j. □

Proof of Remark 3. Recall that the set S is given by equation (B.4). All the statements of this remark follow if, for each $i \in \{1, 2\}$, (i) the threshold discount factor $\overline{\delta}_i$ is non-decreasing in x_i, p_{-i}, and ℓ_i; and (ii) the threshold discount factor $\underline{\delta}_i$ is non-increasing in x_i, p_{-i}, and ℓ_i. Statement (i) follows from part 1 of Lemma 10, and statement (ii) follows from part 2 of Lemma 10. □

Proof of Remark 4. Take $i \in \{1, 2\}$ such that $x_i > 1 > x_{-i}$. Recall that the set S is given by equation (B.5). All the statements of this remark follow if the following hold: (i) $\hat{\delta}_i$ is non-decreasing in x_i, p_{-i}, and ℓ_i; (ii) $\overline{\delta}_i$ is non-decreasing in x_i, p_{-i}, and ℓ_i; (iii) $\overline{\delta}_{-i}$ is non-decreasing in x_{-i}, p_i, and ℓ_{-i}; and (iv) $\underline{\delta}_{-i}$ is non-increasing in x_{-i}, p_i, and ℓ_{-i}. Statement (i) follows because, as shown in Remark 2, $\hat{\delta}_i = \frac{x_i - 1}{x_i}$ is non-decreasing in x_i, p_{-i}, and ℓ_i. Statements (ii) and (iii) follow from part 1 of Lemma 10. Statement (iv) follows from part 2 of Lemma 10. □

B.6.4 Proofs for Appendix B.5

Proof for Example 1. For the first strategy profile, we show that this constitutes a PBE if $\delta \geq \frac{1}{2}$ for any p. We consider player i's incentives. If $-i$ has been taking A and A is available to i, then i's following the given strategy yields a payoff of 2. If i deviates, her payoff is at most $(1 - \delta) \cdot 3 + \delta \cdot 1$. This is no greater than 2 if $\delta \geq \frac{1}{2}$. After any history in which some player has taken B or C, following the given strategy is weakly better than following any other strategy at any period for any realization of private learning against the opponent's given strategy.

Now we show that the strategy profile σ^* is a PBE when δ is large and p is small. First, at a history at which player i has taken C, the only available action to her is C. Thus, below we consider a history at which player i has not taken C.

Second, suppose that player $-i$ has taken C. Since player $-i$ will keep taking C in the future, it is player i's best response to take B. For this reason, below we assume that no player has taken C.

Third, suppose that action C is available to player i. If her opponent $-i$ has taken B (note that this implies that player $-i$ has taken B in the last period), then it is player i's best response to follow σ_i^* to obtain the maximum continuation payoff (after the opponent takes action B) of 1. Thus, assume that player $-i$ has taken only A. Now, if player i takes C, then her continuation payoff, which we denote by V_C, is $V_C = (1 - \delta) \cdot 3 + \delta \cdot 1$. If player

i takes A (this means that both players have been taking A), then her continuation payoff, which we denote by V_A, is

$$V_A = (1-\delta)\cdot 2 + \delta\left(\tilde{p}\left((1-\delta)(-2) + \delta\cdot 0\right) + (1-\tilde{p})\left((1-\delta)\cdot 2 + \delta\cdot V_C\right)\right),$$

where $\tilde{p}$ is the probability that player $-i$ takes C in the next period, which is a convex combination of $p^{-i}_{A,C}$ and $p^{-i}_{B,C}$ and thus is equal to p. Hence, in the limit as $\delta \to 1$, the right-hand side becomes $(1-p)V_C$. Thus, there is $\delta' \in (0,1)$ such that if $\delta > \delta'$, then $V_C \geq V_A$ so that the deviation is not profitable.

If player i takes B instead, then her continuation payoff, which we denote by V_B, is

$$V_B = (1-\delta)\cdot 3 + \delta\left(p\cdot 0 + (1-p)((1-\delta)(-1) + \delta\cdot 1)\right).$$

We have $V_B \leq V_C$ because V_B is a convex combination of 3 and a term which is less than 1, V_C is a convex combination of 3 and 1, and the weights on the convex combinations are the same.

Fourth, assume that action C is not available to player i and that the opponent has played B. If player i follows σ_i^*, then letting $\tilde{q}$ be the probability that player $-i$ takes C in the next period, her continuation payoff, which we denote by W_B, satisfies:

$$W_B = (1-\delta)(-1) + \delta\left(\tilde{q}\cdot 0 + (1-\tilde{q})(p\cdot 1 + (1-p)W_B)\right).$$

If A is available to player i and she takes A, then her continuation payoff is

$$(1-\delta)(-2) + \delta\left(\tilde{q}((1-\delta)(-2) + \delta\cdot 0) + (1-\tilde{q})W_B\right).$$

To see that such a deviation is not profitable, for any $\tilde{q}$, it is enough to show that $p\cdot 1 + (1-p)W_B \geq W_B$, that is, $1 \geq W_B$. This inequality follows because player i's maximum possible continuation payoff (after the opponent takes action B) is 1.

Fifth, suppose that action C is not available to player i and player $-i$ has been taking A. If the only available action to player i is A, then she takes A. Suppose action B is available to player i. Note that the probability that player $-i$ takes C in the next period is p. If player i follows σ_i^*, then her continuation payoff, which we denote by X_A, satisfies:

$$\begin{aligned}X_A &= (1-\delta)2 + \delta\left(p((1-\delta)(-2) + \delta\cdot 0) + (1-p)((1-\delta)2 + \delta\left(p\cdot V_C + (1-p)X_A\right))\right)\\ &= (1-\delta)2 + \delta\left(p(1-\delta)(-2) + (1-p)((1-\delta)2 + \delta\left(p\cdot((1-\delta)3 + \delta\cdot 1) + (1-p)X_A\right))\right).\end{aligned}$$

The limit of X_A as $\delta \to 1$, which we denote by X_A^*, satisfies

$$X_A^* = (1-p)p + (1-p)^2 X_A^*.$$

Thus,

$$X_A^* = \frac{(1-p)p}{1-(1-p)^2}.$$

If player i instead takes B, then her continuation payoff, which we denote by X_B, is

$$X_B = (1-\delta)\cdot 3 + \delta Y_B,$$

where Y_B is the continuation payoff when the latest action profile is (B,A) or (B,B), it is $-i$'s turn to move, and i has not privately learned C:

$$Y_B = p \cdot 0 + (1-p)\left((1-\delta)(-1) + \delta\left(p\cdot 1 + (1-p)\left((1-\delta)(-1) + \delta Y_B\right)\right)\right).$$

The limit of Y_B as $\delta \to 1$, which we denote by Y_B^*, satisfies

$$Y_B^* = (1-p)p + (1-p)^2 Y_B^*.$$

Thus,

$$Y_B^* = \frac{(1-p)p}{1-(1-p)^2}.$$

Hence, the limit of X_B as $\delta \to 1$, which we denote by X_B^*, satisfies

$$X_B^* = Y_B^*.$$

Thus, we have $X_A^* = X_B^*$. Now, by algebra, we have that

$$\left.\frac{d(X_A - X_B)}{d\delta}\right|_{\delta=1} = 2 + \frac{4}{2-p} - \frac{3}{p}.$$

Note that this is increasing in p and is negative for a sufficiently small p. Hence, there exist $\delta'' \in (0,1)$ and $p' \in (0,1)$ such that $X_A > X_B$ for all $\delta > \delta''$ and $p < p'$.

In sum, there exist $\overline{\delta} \in (0,1)$ and $\overline{p} \in (0,1)$ such that σ^* constitutes a PBE for all $\delta \in (\overline{\delta}, 1)$ and $p \in (0, \overline{p})$. □

Proof for Example 2. For the first strategy profile, we show that it constitutes a PBE if $\delta \geq \frac{1}{2}$ for any p. We consider player i's incentives. If player i can choose her action from $\{A, B, C\}$ because she has privately learned B and C and the opponent has been taking A, then she receives a payoff of 2 as long as she follows the strategy. If she deviates by taking C, then she receives a payoff of $(1-\delta)4$ (if she deviates by taking B, her payoff is bounded by $(1-\delta)3$, which is less than $(1-\delta)4$). Thus, she plays A if

$$(1-\delta)\cdot 4 \leq 2, \text{ that is, } \delta \geq \frac{1}{2}.$$

If player i can choose her action from $\{A, B, C\}$ because the opponent has taken B or C in the past, then it is player i's best response to always choose C.

If player i's available action set is $\{B, C\}$, then it is player i's best response to always choose C.

Second, we show that, for any p, the second strategy profile σ^* is a PBE when the players are sufficiently patient. We consider two cases. Suppose first that both players have been taking only A. Player i plays C when she privately learns it if

$$(1-\delta)4+\delta u_i^{(i)} \geq (1-\delta)2+\delta\left(p\left\{(1-\delta)(-3)+\delta u_i^{(-i)}\right\}+(1-p)\left\{(1-\delta)2+\delta\left((1-\delta)4+\delta u_i^{(i)}\right)\right\}\right),$$

or,

$$(1-\delta^2(1-p))\left((1-\delta)4+\delta u_i^{(i)}\right) \geq (2-p)(1-\delta)2+\delta p\left\{(1-\delta)(-3)+\delta u_i^{(-i)}\right\}.$$

When $\delta = 1$, since $p > 0$, the above inequality holds because it reduces to $u_i^{(i)} > u_i^{(-i)}$. Thus, there exists $\overline{\delta} \in (0,1)$ such that the above inequality holds when $\delta \in (\overline{\delta}, 1)$.

Next, suppose that some player has taken B in the past and that C has never been taken. Player i plays C when she privately learns it if

$$(1-\delta)2+\delta u_i^{(i)} \geq (1-\delta)2+\delta\left((1-\delta)(-1)+\delta u_i^{(-i)}\right),$$

where the left-hand side is a lower bound of the payoff from playing C and the right-hand side is an upper bound of the payoff from playing A or B. This inequality holds for any δ because $u_i^{(i)} > u_i^{(-i)} \geq -1$: The strict inequality follows by assumption, and the weak inequality follows because $u_{-i}^{(-i)}$ must be feasible in the game with actions B and C while -1 is the worst payoff in such a game.

In sum, σ^* is a PBE when $\delta \in (\overline{\delta}, 1)$. $\square$

C Supplementary Discussions for Section 5

C.1 0-Waiting Strategy Profile

In this subsection, we consider the breakthrough model and solve for the payoff region such that action N is chosen immediately in equilibrium:

Proposition 11. *There exists $\delta' < 1$ such that the 0-waiting strategy profile is an equilibrium of the breakthrough model with $(K, 0)$ for all $\delta \in (\delta', 1)$ if and only if $\frac{K^2}{2(K-1)} - 1 \leq x_i + \ell_i$ for each $i \in \{1, 2\}$.*

That is, the condition for the existence of 0-waiting strategy profile is that the one-time gain from deviating to N or the one-time loss of being deviated is large.

Proof of Proposition 11. The existence of this equilibrium for sufficiently high δ is equivalent to (10) in the proof of Proposition 3. □

C.2 Comparison with the Bubble Models

This subsection compares our result in the breakthrough model with the ones in the literature on "greater fool" bubbles that aims to explain why rational investors ride bubbles (Abreu and Brunnermeier, 2003; Conlon, 2008; Doblas-Madrid, 2012). On the one hand, those explanations feature asymmetric information akin to our breakthrough model: The discrepancy between the asset price and fundamentals happens at a random unobserved time t^O, and each investor receives a signal of such information at a random time between t^O and $t^O + \eta$, where η is a constant. Thus, given any realization of t^O, for a sufficiently large t, it is almost common knowledge in the Rubinstein (1989) sense that the asset price has departed from fundamentals, but it is never common knowledge. The investors nonetheless do not sell their assets meanwhile, which resembles our result that action O keeps being taken.

On the other hand, the predicted outcomes are different. In the bubble-riding equilibria in the literature, the investors sell their assets after riding the bubble for a while. What corresponds to this in our model would be that action N is chosen at some point on the equilibrium path of play, which we do not predict. To understand why this difference arises, let us compare the incentive compatibility conditions. First, in Abreu and Brunnermeier (2003)'s continuous-time model, suppose that all other investors use the τ-waiting strategy and that investor i learns at time t_i.[8] The condition for not selling to be a best response at

[8] We use Abreu and Brunnermeier (2003)'s model for the comparison, but analogous comparison can be made with other models of bubbles in the literature that are based on asymmetric information.

time $t = t_i + \tau$ can be written as:[9]

$$\underbrace{0}_{\text{flow payoff today}} + \underbrace{(1 - \Delta \cdot h(t|t_i))}_{\text{Prob(no burst in }\Delta)} \cdot \underbrace{p(t + \Delta)}_{\text{increased price}} + \underbrace{\Delta \cdot h(t|t_i)}_{\text{Prob(burst in }\Delta)} \cdot \underbrace{(p(t) - L(t))}_{\text{reduced price}} \geq \underbrace{p(t)}_{\text{current price}}, \tag{B.9}$$

where $h(t|t_i)$ is the hazard rate that the bubble bursts at time t and $L(t)$ is the magnitude of the loss incurred when the bubble bursts.[10] Next, our incentive compatibility condition at time $t^* + \tau + 2$ (i.e., the period at which the player is supposed to choose N if she is adopting the τ-waiting strategy) when the opponent uses the τ-waiting strategy can be written as:

$$\underbrace{(1-\delta)}_{\text{flow payoff today}} + \underbrace{(1-\hat{q})}_{\text{Prob(no }N\text{ tomorrow)}} \underbrace{\delta(1-\delta)(1+\delta x_i)}_{\text{increased payoff}} + \underbrace{\hat{q}}_{\text{Prob(}N\text{ tomorrow)}} \underbrace{\delta(1-\delta)(-\ell_i)}_{\text{reduced payoff}} \geq \underbrace{(1-\delta)x_i}_{\text{payoff from }N\text{ today}}, \tag{B.10}$$

where $\hat{q}$ is the probability that the opponent chooses N in the next period given his τ-waiting strategy. This probability goes to 0 as K tends to infinity because, although player $-i$'s learning could have occurred in any of the $\frac{K}{2}$ periods after the period in which i learned, the only case in which $-i$ plays N in the next period is when he learned in the period right before i learned.[11] As is apparent, the two conditions are similar to each other, except that the flow payoff is positive in our condition (which is not important for the discussion that follows). Why does Abreu and Brunnermeier (2003)'s model allow for an equilibrium in which the investors sell at some point, while our model does not, despite this similarity? To see this, rewrite (B.9) as follows:[12]

$$\Delta \cdot h(t_i + \tau|t_i) \cdot L(t_i + \tau) \leq (1 - \Delta \cdot h(t_i + \tau|t_i)) \cdot (p(t_i + \tau + \Delta) - p(t_i + \tau)). \tag{B.11}$$

The left-hand side and the right-hand side of (B.11) represent, respectively, the benefits and costs of selling. Notice that the benefits are associated with the event when the bubble burst in the next Δ time interval and the magnitude of the loss upon bursting, while the costs are associated with the complementary event and the growth of the price during the Δ time interval (i.e., capital gain) which the selling investor would miss.

[9] We set the risk-free interest rate in their model to be zero here to facilitate the comparison because our model does not feature what would correspond to a risk-free interest rate (the payoff from taking N does not vary over time).

[10] Abreu and Brunnermeier (2003) write $L(t) = [p(t)\beta(t - T^{*-1}(t))]$, where $p(t)$ is the asset price at time t, $\beta(s)$ denotes the fraction of the "bubble component" of the price when time s has elapsed since the first time the asset price departs from fundamentals, and $T^*(t^O)$ denotes the bursting time of the bubble when t^O is the departing time.

[11] The proof in the Appendix explains that this probability is $\frac{K-1}{(\frac{K}{2})^2}$.

[12] This is the expression used in Abreu and Brunnermeier (2003), where we set the (risk-free) interest rate to be 0.

Two things are important. First, both sides of (B.11) are of the order of Δ, so they are comparable. This comes, in particular, from the fact that the size of the loss conditional on bursting, L, is independent of the period length Δ. Thus, we are comparing a large loss that happens with a small probability (left-hand side) with a small gain (price difference between times t and $t+\Delta$) that happens with a large probability. Second, inequality (B.11) will be violated at a large τ because either $h(t|t_i)$ is increasing in t (when the environment is such that the bubble is determined to burst for an exogenous reason after a certain period of time after the departure) or L increases faster than p (in the other case, where a later burst means more time passing from the departure, implying greater losses).

Now, manipulating (B.10) in a similar way yields:

$$\hat{q}(x_i + \delta\ell_i) \leq (1-\hat{q})(\delta(1+\delta x_i) - x_i) + 1. \tag{B.12}$$

First, the "loss" on the left-hand side of (B.12) is not large because it is just a single-period payoff. Consequently, as $\hat{q} \to 0$ (corresponding to $\Delta \to 0$), the left-hand side vanishes, ensuring the inequality to hold. Second, even if we do not consider the limit $\hat{q} \to 0$, all parameters are independent of τ, and thus if a condition is satisfied for a particular τ, it holds for any τ.

The crucial difference is that the bubble is supposed to burst at some point in the Abreu and Brunnermeier (2003) model or the loss from bursting is assumed to grow faster than the asset price, which makes the benefits of selling increase as time passes and eventually exceed the costs, and one can use this as a foothold to show the uniqueness of an equilibrium. Our breakthrough model, however, does not have such a foothold, and because of this, (i) τ-waiting strategy profile is not an equilibrium for a fixed payoff structure when $\hat{q}$ is close to 0, and (ii) $x_j + \ell_j$ must be large when K is large for the τ-waiting strategy profile to be an equilibrium.

Note that (i) and (ii) are based on the fact that $\hat{q}$ goes to 0 as $K \to \infty$. For (i), the magnitude of the "loss" is fixed, so the left-hand side of (B.12) goes to zero as $\hat{q}$ goes to 0. For (ii), for the left-hand side to be large enough, the magnitude of the "loss" needs to be high when $\hat{q}$ is small.

Third, although our model does not have any "foothold" as was used in the Abreu and Brunnermeier (2003) model, we can still use an iterative argument as we described to show the uniqueness of an equilibrium. The difficulty is that when showing the uniqueness in our model, $\hat{q}$ must be replaced with the maximum probability across all equilibria (not just the τ-waiting strategy profile) that the opponent plays N in the next period, and such a probability does *not* go to zero because we must consider the possibility that the opponent might plan to take N at a particular calendar date irrespective of when he learns it before that date. A priori, such a probability can be taken to be arbitrarily close to 1 by considering a large t for any realization of the learning time, and this is why we rely on the novel iterative argument to prove uniqueness.

D Detailed Discussions on Real-World Examples

We provide a detailed account of each example we presented in Section 6. As explained there, the examples we provide have a feature that a certain action had been unprecedented for a long time, but it was then taken due to some unforeseen contingency or its existence was revealed by later research.

In explaining each example, we aim to explain why we think our assumptions are met. Given that a model ought to be a simplification of reality, one may argue that the assumptions do not perfectly fit with the model. Even so, we think that understanding what our model predicts under our assumptions is important in deepening our understanding of the phenomena we see in the examples. For instance, regarding asynchronicity, although we do view it as a natural assumption as we will see, our stance is that it is at best not obvious which of synchronicity and asynchronicity (or perhaps some mixture of them) would be applicable, just like any dynamic models. Thus, it is important to understand what each model predicts and why there are differences, if any.

D.1 *Predator: The Secret Scandal of J-Pop*

Our first example is about a series of sexual abuse allegations against Johnny Kitagawa, the founder of Johnny & Associates, which is by far the largest production agency for numerous popular boy bands in Japan. Kitagawa was the president of this agency from its foundation until his death in 2019.[13] The first allegation dates back to 1964 when there was a trial regarding his sexual abuse towards a talent in his production. Several books have been published between 1988 and 2005, accusing him of sexual abuse targeting various talents in the production. There was even a court trial in which some accusations were proved true.[14] However, no TV stations had reported on these matters. In March 2023, however, the BBC aired a documentary titled *Predator: The Secret Scandal of J-Pop* accusing Kitagawa (BBC

[13] Various pieces of information presented in this subsection are, unless otherwise specified, based on the investigation report prepared by the "Special Recurrence Prevention Team" composed of external experts established by Johnny & Associates. The report (in Japanese) is available at `https://www.smile-up.inc/s/su/group/detail/info-711?&link=ROBO004` (Last Access: June 10, 2026).

[14] Kitagawa was not found guilty because it was a civil trial regarding a magazine article accusing the abuse.

World News Japan aired the documentary in Japan about 10 days later[15]). This broadcast triggered a cascade of accusations by victims. Kauan Okamoto is one of the first accusers in public, and he claimed that there would be at least 100-200 victims even if we only consider the time when he was in the production. Following these accusations, virtually all TV stations started reporting the problem in May 2023.

The situation can be modeled as a game played by TV stations.[16] For simplicity, consider the case of two TV stations, stations A and B, where each station's actions are "report the problem" or "not report the problem." Each period, each station privately learns direct evidence of the abuse with some probability, while they can only choose "not report" before private learning occurs. After a station privately learns such evidence, it can choose whether to report or not. Thus, the set of feasible actions changes after private learning, which we view as reasonable in our context given Japan's generally cautious culture and their Article 230 of the Penal Code (defamation), which provides that anyone who publicly defames another by alleging facts may be guilty of a crime, *regardless of* whether those facts are true or false. Once a TV station reports a story, another station can also report it by investigating the same source of information (e.g., interviewing the same victim) or simply by relying on the first station's reporting. Thus, a precedent also changes the set of feasible actions. Once reporting is done, it cannot go back to not reporting—a report cannot be erased; thus, the action change is irreversible. Once a TV station makes a report, there are two effects. First, since Johnny & Associates had such a strong impact in the showbiz world, they could stop sending their talents to the TV station that made a report, which is detrimental to the station: virtually all TV stations have heavily relied on the talents from Johnny & Associates. Second, given the strong impact, which is widely recognized by Japanese viewers, reporting enhances viewership as they associate it with high-quality journalism from that TV station. If station A is the only station that made a report, then the second effect is large and station A's payoff increases.[17] Station B would lose viewers so the payoff decreases. But if both A and B report, then the number of viewers at each station would be unchanged while the relationship with Johnny & Associates deteriorates, so they are in a worse situation than the case when neither has reported. Since TV stations typically air their news programs at different times of a day and they can condition their report on what has been reported, asynchronous moves would be a natural assumption on

[15]See `https://www.youtube.com/watch?v=zaTV5D3kvqE` (Last Access: June 10, 2026).

[16]This is a case that is still evolving at the time of this writing. We describe the payoffs that the players (the TV stations) would have had in mind when deciding to make a report or not, to the best of our knowledge.

[17]This makes the stage game a prisoner's dilemma. If the second effect is not large enough, then A's payoff may decrease. In such a case, the stage game is a coordination game, for which our result still applies.

timing. In this situation, our prediction is that the sexual abuse would never be reported on TV, and that was exactly what had happened until the BBC, a foreign media outlet that does not rely on the talents from Johnny & Associates, made a move. After the BBC's report, the sexual abuse became common knowledge, so there is an equilibrium where both stations report as soon as possible.

We note that it is hard to imagine that the TV stations did not have any direct evidence of the sexual abuse at the time when the BBC made a report: It has been more than 50 years since the first allegation. Indeed, an investigation report by external experts states that many in the industry had chances to learn the evidence of Kitagawa's sexual abuses.[18] Kenichi Mikawa, a singer who has been in the showbiz since 1965 (only a year after the very first allegation), later reflected at a press conference that he "has known about sexual assault for a long time. It was a lot. It's a tacit agreement that we pretend not to see it. I think the media felt the same way. At that time, everyone felt that way. It was a time when you shouldn't touch or say anything."[19] This suggests that Mikawa believed accusing Kitagawa had been considered taboo. Our theory provides an alternative explanation for why accusation has not been done from the perspective of incentives.

Our point is based on the observation that the incident of sexual misconduct might have been known by everyone, and in some sense "close to" common knowledge, but might not have been exactly so.[20] If instead the evidence was shown to all TV stations in a way public to them, then our uniqueness argument would break down, although keeping the taboo is one possible prediction. However, given the nature of how information about this type of incidents would spread out, we believe that the incident was "almost" common knowledge at best.

D.2 The Olympus Accounting Fraud

Olympus Corporation, a Japanese manufacturer of optical and reprographic products, engaged in accounting fraud for over a decade, a practice that the *Wall Street Journal* described as "one of the biggest and longest-running loss-hiding arrangements in Japanese corporate

[18]The report states "[A]ccording to the results of the hearing, it appears that Johnny's sexual assault was widely known to those in the entertainment industry as early as the early 1970s."

[19]Part of the press conference (in Japanese), aired by All-Nippon News Network in September 2023, is available at `https://www.youtube.com/watch?v=KxAJWd2XlXg` (Last Access: June 10, 2026). The translations are made by the authors.

[20]Analogous comments apply to the other three examples in the subsequent sections, which would show the extent to which our theory is relevant, and why we view our information structure to be appropriate in the context of those examples.

history."[21] The company incurred significant losses during the Japanese asset-price bubble in the late 1980s and early 1990s, but those losses did not appear in their financial report as they used what is called "Tobashi scheme" where funds are shifted to paper companies. This practice remained confidential among a few top executives for a considerable period, until it was exposed by an article in a rather minor financial magazine: An employee at Olympus happened to read materials from the board of directors and found out the wrongdoing. Then he shared this information with an independent journalist who happened to be his acquaintance, who wrote the article at the magazine.

Michael C. Woodford, a British-born executive who had just become the company's president at the time, became aware of the article after a reader told him about it, and initiated an investigation. However, he was promptly dismissed by the board of directors. The scandal subsequently gained widespread public attention after Woodford himself became a whistleblower in the *Financial Times*.[22] The company's stock price plummeted, and its top executives were arrested in the end.

This situation can be modeled as a game played by the board directors of the company. They have a chance to learn the top executives' wrongdoing when they have chances to review financial statements in detail. Once (and only if) a director learns the wrongdoing, they have an option to whistleblow (thus, the set of feasible actions changes after learning).[23] If no whistleblowing occurs, then the scandal remains concealed, and the company remains stable. If both directors whistleblow, the company's stock prices plummet, resulting in a worse situation for both directors. If only one director blew the whistle, the other director would be perceived as complicit in the wrongdoing, receiving the worst possible payoff. The whistleblower's payoff depends on whether she is an inside director or an outside director. If

[21]See `https://www.wsj.com/articles/SB10001424052970204190704577024680506345936` (Last Access: June 10, 2026). Unless otherwise specified, various pieces of information presented in this subsection are based on the investigation report prepared by the third party committee launched by Olympus. The English translation is available at `https://www.olympus-global.com/en/common/pdf/if111206corpe_2.pdf` (Last Access: June 10, 2026).

[22]See `https://www.ft.com/content/87cbfc42-f612-11e0-bcc2-00144feab49a` (Last Access: June 10, 2026).

[23]The fact that even a low-ranking employee could discover the arrangement suggests that there was a significant chance for the directors, even those who were not the very top executives, to learn about the arrangements during the extended period in which it was operational. In fact, in one of the fraudulent M&As, Olympus paid an exorbitantly high percentage (approximately 30%) of the acquisition price to a middleman as a fee, which suggests that the board members would have been able to easily notice the fraud if they had been accurately informed about the deal or had reviewed the deal carefully. The investigation report points out, however, that accurate information was not provided by top executives in board meetings and there are no records that board members reviewed the M&A in sufficient detail, suggesting that the probability of private learning, p, was small. However, once a board director had blown the whistle, it would have been easy for other board directors to do the same, given how easy it was to find evidence. Thus, a precedent would also have changed the feasible set of actions.

she is an inside director, the substantial negative impact on stock prices would leave her worse off.[24] On the other hand, if she is an outside director, then given that she is the only one who has blown the whistle, she enhances her reputation as a functional outside director, opening up future opportunities at other companies. Consequently, her payoff would increase. The stage game would thus be a coordination game in the former case and a prisoner's dilemma in the latter case. The action change is irreversible because, once a director blows the whistle, the action cannot be undone. Asynchronicity seems reasonable given that board members, as high-ranking members of the company, tend to work quite independently of one another.

In this situation, our result predicts that the directors would refrain from whistleblowing, which was what had happened for more than a decade until the situation was altered by a chance revelation by an employee who happened to know a magazine writer and the involvement of a new foreign president.

D.3 Nerve Agents during World War II

During World War II, both Germany and the Allies actively pursued the development of nerve agents, recognizing the possible significant influence on the outcome of the war.[25] The potential power and strategic advantage offered by nerve agents made them a critical focus for military research and development on both sides. Germany, in particular, invested substantial resources and expertise into the synthesis of nerve agents, with efforts comparable to the American Manhattan Project, which resulted in the production and storage of tons of nerve agent munitions. Meanwhile, the Allies, although trailing in the chemical weapons race in fact, were also aware of the strategic importance of these weapons and developed their own countermeasures.

The discoveries related to nerve agents during the war were primarily the result of private learning. On the German side, the development of lethal nerve agents was shrouded in utmost secrecy, with knowledge of these capabilities limited to a select few within the highest echelons of the military and scientific community. This secrecy ensured that the Allies remained unaware of the full extent of Germany's advancements in nerve agents. Conversely, the Allies stumbled upon the toxic effects of nerve agents somewhat accidentally. British scientists Bernard Charles Saunders and Hamilton McCombie made a serendipitous discovery of the harmful properties of these chemicals. It has been said that, although the Allies' technology related to nerve agents was in its infancy, German high command believed the Allies had

[24] In Japan, whistleblowers are protected, but there is no monetary compensation.

[25] See Hilmas et al. (2008) for the source of various pieces of historical information presented in this subsection. We note that there often exist multiple interpretations of the historical facts. Our discussion does not necessarily encompass all possible interpretations or reflect our personal views.

developed nerve agents as well.

Indeed, the mutual perception of chemical-weapon capabilities played a crucial role in the conduct of both sides during the war. Intelligence reports and rumors about the other side's capabilities reinforced this belief. The constant fear of massive chemical attacks kept military leaders on both sides in a state of high alert and pushed them to maintain rigorous preparedness.

The situation can be modeled as a game played by Germany and the Allies, where each group's actions are "deploy nerve agents" or "not deploy nerve agents." Each period, each group privately learns about production methods of nerve agents with some probability, while they can only choose "not deploy" before private learning occurs. After a group privately learns it, it can choose whether to deploy or not (thus, the set of feasible actions changes after learning). Learning can also happen by observing a precedent: Once one side uses a nerve agent, the other side can learn about its chemical composition by analyzing environmental samples such as air and soil. Moreover, the use of the weapon would trigger extensive efforts to uncover the underlying production process. We therefore assume that once one side uses a nerve agent, the other side learns the production method. Deployment is irreversible in the sense that its effect remains for a long time on the environment and the health of people, physically and mentally. If no deployment occurs, then each group remains stable. If a group, say Germany, deploys their weapons, it causes extensive damage to the Allies and gives an advantage to Germany. However, the effect of such an advantage is smaller than the effect of the damage, and thus each group values the status quo more than the situation where both groups deploy. Asynchronous moves seem reasonable given that the two sides operate independently, and they are on different time zones.

In this situation, our prediction is that the deployment would never occur, and that was exactly what happened in World War II: Despite the development of nerve agents, neither side deployed them during the war. The delicate balance of mutual deterrence, coupled with the sporadic nature of private learning, resulted in a state of cautious behavior, and the use of nerve agents was ultimately prevented during the war.

D.4 Zombie Lending

"Zombie firms" are insolvent firms with no hope of recovery that are nevertheless kept alive by banks that support them (Hoshi, 2006; Caballero et al., 2008). They were first documented for the Japanese economy in the early 1990s during the post-bubble era (Hoshi, 2000, 2006; Caballero et al., 2008). Zombie firms are also found to have globally risen particularly since the Global Financial Crisis (see Acharya et al. (2022) for a survey) and the COVID-19 pandemic (Albuquerque and Iyer, 2024). A central puzzle in the literature is the economic rationale for why banks continue to extend credit to zombie firms. The leading explanation is that, following a large economic shock, undercapitalized banks may have incentives to continue providing credit to zombie firms to avoid recognizing loan losses (i.e., by "evergreening" loans), thereby avoiding penalties such as regulatory sanctions and the associated reputational costs (Caballero et al., 2008; Acharya et al., 2022). We show that our theory with incomplete information predicts that zombie lending occurs in a unique equilibrium, although a simple model with penalties admits an equilibrium in which banks call in no matter how large those penalties are. The theory also shows a limitation of a possible remedy of zombie lending based on the leading explanation of penalties.

We model this situation as a game played by the banks that are providing credit to an insolvent firm. For simplicity, we assume that two banks choose between "Roll over" (i.e., keep providing credit) and "Call in" (i.e., withdraw credit). Each period, each bank privately learns the evidence that the firm is a zombie with a small probability. A bank may call in its loan only after obtaining such evidence; absent such evidence, calling in is prohibited by the contractual terms unless an unforeseen contingency arises. Thus, private learning changes the set of feasible actions.[26] Once one bank calls in its loan, the other bank is likely to investigate the firm, which in turn would likely uncover evidence that the firm is a zombie. We therefore assume that a precedent also makes the action "Call in" available to the other bank.

Let A be the firm's aggregate recoverable value (i.e., the amount of loan that the firm is able to repay). In one period, the bank that calls in can collect part of that value, $B \in (0, A)$, where we assume that A is an odd multiple of B for simplicity. After one bank calls in, both banks will be able to collect the remaining value gradually (up to B each period), and we assume that the banks split the remaining value equally. By algebra, one can show that this corresponds to receiving $(1 - \delta^{\frac{A-B}{2B}})B$ each period, where δ is the common discount factor.[27] Finally, let R be the interest payment each bank receives every period.[28] This payoff structure is depicted in Table 5.[29] We assume $R \in (0, B)$. Note that, for the case

[26] Technically, they would be able to call in when the lending contract expires, but we ignore such a possibility because it happens infrequently given the typical duration of such contracts.

[27] Consider paying B each period to total $\frac{A-B}{2}$. Let C be the discounted sum of the payments. We have $C = (1 + \delta + \delta^2 + \cdots + \delta^{\frac{A-B}{2B}-1})B$, which is equal to $\frac{1-\delta^{\frac{A-B}{2B}}}{1-\delta}B$. The bank receives the same continuation payoff by receiving $(1-\delta)C$ each period, and this amount is equal to $(1 - \delta^{\frac{A-B}{2B}})B$.

[28] In practice, zombie firms may not be able to pay the interest every period. We note that R could be interpreted as an expected interest and the realization could thus be lower than R. We also note that the expectation of the interest payment may be different when a bank does not know that the firm is a zombie compared to when it does. However, such an expectation does not matter because the bank does not have an action choice when it does not know the firm's insolvency.

[29] The symmetry of the payoff structure is not essential. Our argument goes through, for example, if

	Roll over	Call in
Roll over	R, R	$0, B$
Call in	$B, 0$	$(1-\delta^{\frac{A-B}{2B}})B, (1-\delta^{\frac{A-B}{2B}})B$

Table 5: The payoff matrix

when the two banks call in, we have put the payoff as calculated above to fit the story with our framework, with an understanding that once the opponent bank calls in, the unique best response is to call in.[30] The bank's action is irreversible because calling in would jeopardize the trust relationship between the bank and the firm. Asynchronicity is a natural assumption because each bank independently conducts a site visit.

In this situation, our prediction of the base model is that when the discount factor is large enough, banks keep rolling over in a unique equilibrium.[31] This is consistent with what has been observed in the post-bubble era in Japan, after the Global Financial Crisis, and after the Covid-19 pandemic. If we employ the breakthrough model, where banks receive information about insolvency of the firm at correlated times, assuming that the discount factor is large, the uniqueness of perpetual rolling over still holds for some (K, p) when the interest payment R is relatively large, and the τ-waiting strategy profile is an equilibrium for some K as long as $R \leq B$.[32]

Policy implication: This result shows a limitation of a possible remedy of zombie lending based on the leading explanation of penalties. To see the effect of penalties, consider a model where we introduce penalties in their payoff function: When there is a bank that initiates calling in, the firm goes bankrupt and as a result, the losses are realized on the banks' balance sheets. In this event, we assume that penalties of magnitude P are imposed on each bank. This corresponds to changing the payoff profile from (Roll over, Call in) to $(B-P, -P)$. Our result shows the uniqueness of the "roll over" equilibrium irrespective of the size of P. One possible remedy in light of the leading explanation—that the banks keep lending in fear of penalties—is to lessen the size of penalties. According to our result, however, the effect of such a policy is limited at best. Given the negative effect of lessening the penalties (the risk that banks may become undercapitalized), this argument suggests that policymakers need

each bank receives the fraction of the capital (close to) proportional to the amount they were lending. Furthermore, although the per-period payoff depends on δ, one can check that all our proofs go through nonetheless.

[30] This makes the payoff matrix dependent on δ. One can check that our proofs for high δ in both our base model and the breakthrough model go through even with such dependence.

[31] Note that large δ implies that $R > (1-\delta^{\frac{A-B}{2B}})B$ holds, which is necessary for our results.

[32] These conditions for uniqueness can be obtained by normalizing the payoffs in Table 5 as in Table 1 of the main text and applying the condition on $x_i + \ell_i$.

to be cautious about adjusting the size of penalties.

Common knowledge about the zombie firm: Now, consider a simple model where the firm is commonly known to be a zombie among the banks. As above, consider a payoff matrix such that the payoff profile from $(\text{Roll over}, \text{Call in})$ is changed to $(B - P, -P)$.

The strategy profile in which the banks call in at every history is an equilibrium if

$$(1-\delta)(B-P) + \delta(1-\delta^{\frac{A-B}{2B}})B \geq (1-\delta)R + \delta((1-\delta)(-P) + \delta(1-\delta^{\frac{A-B}{2B}})B).$$

This is equivalent to:

$$\delta(1-\delta^{\frac{A-B}{2B}})B - (1-\delta)P \geq R - B.$$

Since the left-hand side approaches 0 as $\delta \to 1$ while the right-hand side is strictly negative, the inequality holds for a sufficiently large δ. Note that this argument holds for any P. The intuition is that of strategic interdependence: Given the expectation that the opponent bank will call in and thus the penalties realize, it is optimal to call in early to receive a larger share of the capital of the firm. Such a logic breaks down in our model because the evidence of the firm's insolvency is never common knowledge (whether it is our base model or the breakthrough model), which (by our theorems) leads to a unique equilibrium. Thus, our prediction provides a new explanation for zombie lending based on incomplete information, which does not rely on regulatory sanctions or reputational costs.

Overall, we applied our model to zombie lending and explained why banks keep lending to insolvent firms. An introduction of incomplete information, which is a natural assumption in our context, provides a strong explanation (a unique prediction) for why zombie lending persisted. Our explanation sheds new light on the role of regulatory sanctions toward undercapitalized banks, offering a novel policy implication.

References for the Online Appendix

Abreu, D. and M. K. Brunnermeier (2003): "Bubbles and Crashes," *Econometrica*, 71, 173–204.

Acharya, V. V., M. Crosignani, T. Eisert, and S. Steffen (2022): "Zombie Lending: Theoretical, International, and Historical Perspectives," *Annual Review of Financial Economics*, 14, 21–38.

Albuquerque, B. and R. Iyer (2024): "The Rise of the Walking Dead: Zombie Firms around the World," *Journal of International Economics*, 104019.

Caballero, R. J., T. Hoshi, and A. K. Kashyap (2008): "Zombie Lending and Depressed Restructuring in Japan," *American Economic Review*, 98, 1943–77.

Conlon, J. R. (2008): "Greater Fool Bubbles: A Survey," Working paper.

Doblas-Madrid, A. (2012): "A Robust Model of Bubbles with Multidimensional Uncertainty," *Econometrica*, 80, 1845–1893.

Hilmas, C. J., J. K. Smart, and J. Hill, Benjamin A. (2008): "History of Chemical Warfare," in *Medical Aspects of Chemical Warfare*, ed. by S. D. Tuorinsky, Office of the Surgeon General, chap. Chapter 2, 9–76.

Hoshi, T. (2000): "Naze Nihonwa Ryudosei no Wana kara Nogarerarenaika," in *Zero Kinri to Nihon Keizai*, ed. by M. Fukao and H. Yoshikawa, Nihon Keizai Shimbunsha, 233–266.

——— (2006): "Economics of the Living Dead," *Japanese Economic Review*, 57, 30–49.

Rubinstein, A. (1989): "The Electronic Mail Game: Strategic Behavior Under "Almost Common Knowledge"," *American Economic Review*, 79, 385–391.