Information Acquisition and Reputation Dynamics

Qingmin LIU

University of Pennsylvania

August 18, 2022

Introduction

- Many tourist destinations: increasing popularity $\rightarrow$ quality and reputation decline $\rightarrow$ rejuvenation.
- In online trading: an agent behaves honestly for some time and then tries to milk his reputation by cheating.
- The information of past transaction is costly to the customers.

Model

- A long-lived firm (Player 1) meets a sequence of short-lived customers (Player 2).
- In each period, Player 1 chooses from two possible product or service qualities, high (H) and low (L).
- Player 2, without knowing the quality, chooses whether to take a trusting action h or a distrusting action l.
- If the quality is unknown, (L, l) is the unique static Nash equilibrium.
- If the quality is observable, (H, h) is the Stackelberg outcome.

Assumptions

- For an action profile $a \in \{H, L\} \times \{h, l\}$, let $u_1(a)$ and $u_2(a)$ be Player 1's and Player 2's stage-game payoffs.
- Since action L is Player 1 's stage-game dominant action, we have

$$u_1(L, h) > u_1(H, h) \quad \text{and} \quad u_1(L, l) > u_1(H, l).$$

- Moreover, since the profile (L, l) is the stage-game Nash equilibrium, we have

$$u_2(L, l) > u_2(L, h)$$

- Since the profile (H, h) is the Stackelberg outcome, we need

$$u_1(H, h) > u_1(L, l).$$

- It is assumed that a firm's short-term cheating benefit is higher when customers acquire a large quantity.

$$u_1(L, h) - u_1(H, h) > u_1(L, l) - u_1(H, l)$$

- This condition implies a growing tension between reputation building and exploitation: as Player 2's trust increases, Player 1's temptation to exploit its reputation becomes stronger.

Types

- Two types of Player 1: the good type and opportunistic type.
- Good type Player 1 always chooses high quality H.
- Opportunistic type Player 1 maximizes the expected pay-off with a discount factor $\delta \in (0, 1)$.
- The common prior probability on the good type is $\mu_0 > 0$.
- To focus on the interesting cases, assume

$$\mu_0 < \bar{\mu} = \frac{u_2(L, l) - u_2(L, h)}{u_2(L, l) - u_2(L, h) + u_2(H, h) - u_2(H, l)}.$$

- Here, $\bar{\mu}$ is such that Player 2 is indifferent between h and l in a one-shot interaction.
- If $\mu_0 > \bar{\mu}$, Player 2 will play h even though he knows that the opportunistic type plays L with probability 1.

$$\bar{\mu} \cdot u_2(H, h) + (1 - \bar{\mu}) \cdot u_2(L, h) = \bar{\mu} \cdot u_2(H, l) + (1 - \bar{\mu}) \cdot u_2(L, l).$$

Information Acquisition

- The short-lived players, upon entering the game, observe neither the previous outcomes nor the number of transactions before them: they are symmetric ex ante.
- A short-lived Player 2 can pay a cost, $C(n)$ to observe Player 1's actions in the previous n periods, $n \in \{0, 1, ...\}$.

Assumption (1)

(1) $C(n)$ is weakly increasing. (2) $C(0) = 0$ and there exists $N_C > 0$ such that

$$C(n) > \max\left\{u_2(H, h), u_2(L, l)\right\} \text{ for any } n > N_C.$$

- The state space can be defined as a finite set consisting of finite histories.
- Examples: linear cost $C(n) = cn$ or "wholesale of information".

State Space and Stationary Strategies

- For a given cost function C , individual rationality of Player 2 implies that he never buys more than N_C periods of information.
- Fix any $N \geq N_C$. Let us define the state space $S = \{H, L\}^N$ as the set of Player 1 's feasible plays in the last N periods.
- For a state/history $s = (s_N, s_{N-1}, \ldots, s_1) \in S$, s_1 is the most recent outcome and s_N the oldest.
- Denote Player 2's information acquisition strategy by a probability measure

$$\alpha \in \Delta\{1, 2, \ldots, N\}.$$

- With probability $\alpha(n)$, Player 2 acquires n periods of information, and his information is represented by a partition $\mathcal{P}^n$ on S.
- The partition element containing s, $\mathcal{P}^n(s)$, is the set of finite histories having the same most recent n entries as s.
- In particular, $\mathcal{P}^0(s) = S$ for each $s \in S$. Note that since $\mathcal{P}^n$ becomes finer as n increases, more information is more informative ex post.

State Space and Stationary Strategies

- Player 2's strategy after acquiring n periods of information is a $\mathcal{P}^n$-measurable function $\sigma_2^n : S \to [0, 1]$ which specifies a probability of playing the trusting action h for each $s \in S$. Let us write $\sigma_2 = \left(\sigma_2^n\right)_{0 \leq n \leq N}$.
- Since information acquisition activity is private, Player 1 might not know what Player 2 actually observes.
- Player 1 will respond to Player 2's expected strategy weighted by the information acquisition probability α.
- Write Player 2's expected strategy as $\bar{\sigma}_2(s) = \sum_n \alpha(n)\sigma_2^n(s)$.
- Consequently, Player 1 always has a stationary best response that only depends on the finite state space S.
- Denote $\sigma_1 : S \to [0, 1]$ as Player 1's stationary strategy that specifies a probability of playing H for each $s \in S$.

Belief Updating

- From Player 2's point of view, nature chooses one of the two types of Player 1 and hence one of the following two stochastic processes on S:
  - The good type plays H constantly and the process is degenerate. Let λ^* be the steady-state distribution of this process, $\lambda^*(\underbrace{HH\ldots H}_{N}) = 1$.
  - The opportunistic type's strategy σ_1 induces a Markov process on the directed graph whose state space is S. Let $\lambda \in \Delta(S)$ be a steady-state distribution of this process.

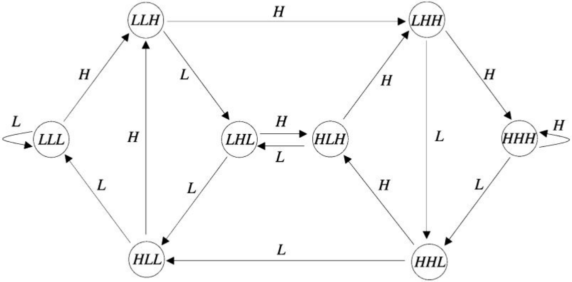

FIGURE 1
Transition on $S = \{H, L\}^3$ with eight states

Belief Updating

- After learning his information set $\mathcal{P}^n(s)$, Player 2 updates his belief about the first process (the good type) according to Bayes' rule whenever possible:

$$\frac{\mu_0 \lambda^* \left(\mathcal{P}^n(s)\right)}{\mu_0 \lambda^* \left(\mathcal{P}^n(s)\right) + (1 - \mu_0)\, \lambda \left(\mathcal{P}^n(s)\right)}.$$

- Consistent with this formulation, Player 2 holds the prior belief μ_0 when he does not acquire information (i.e. $n = 0$), and he assigns probability 0 to the good type if he sees an L.

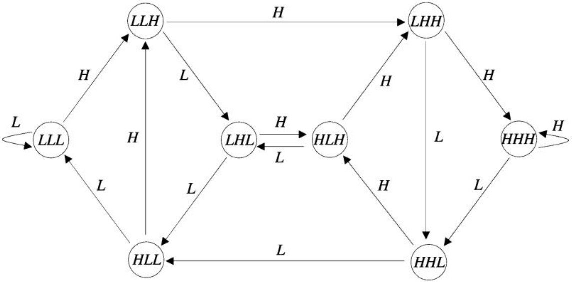

FIGURE 1
Transition on $S = \{H, L\}^3$ with eight states

Equilibria and Preliminary Results

Definition (1)

A quintuple $(S, \sigma_1, \alpha, \sigma_2, \lambda)$ is a perfect Bayesian equilibrium under the cost function C if (1) $S = \{H, L\}^N, N \geq N_C$, (2) Player 2 updates his belief using prior and λ, and σ_1 and (α, σ_2) are best responses given belief, and (3) λ is a steady-state distribution on S consistent with strategy σ_1

Lemma (1)

An equilibrium exists.

Theorem (1)

In the complete information game ($\mu_0 = 0$), the infinite repetition of the static Nash equilibrium (L, l) is the unique perfect equilibrium outcome for any discount factor $\delta \in (0, 1)$. Consequently, Player 2 never acquires information if $C(1) > 0$.

Equilibria and Preliminary Results

Definition (1)

A quintuple $(S, \sigma_1, \alpha, \sigma_2, \lambda)$ is a perfect Bayesian equilibrium under the cost function C if (1) $S = \{H, L\}^N, N \geq N_C$, (2) Player 2 updates his belief using prior and λ, and σ_1 and (α, σ_2) are best responses given belief, and (3) λ is a steady-state distribution on S consistent with strategy σ_1

Lemma (1)

An equilibrium exists.

Theorem (1)

In the complete information game ($\mu_0 = 0$), the infinite repetition of the static Nash equilibrium (L, l) is the unique perfect equilibrium outcome for any discount factor $\delta \in (0, 1)$. Consequently, Player 2 never acquires information if $C(1) > 0$.

An Example

- Consider the following cost structure: $C(1) = 0$ and $C(2) = 5$.
- We consider the state space $S = \{H, L\}$ and assume that Player 2 always **observes Player 1's previous-period play**, i.e. $\alpha(1) = 1$.

	h	l
H	2,3	0,2
L	4,0	1,1

FIGURE 2
A version of the product choice game

An Example: $\mu_0 = 0$

- Theorem 1 predicts that repetition of (L, l) is the unique perfect equilibrium outcome.
- IC condition for Player 1 to play action L at history L is

$$(1-\delta)\cdot\big(4\cdot\sigma_2(L)+1\cdot(1-\sigma_2(L))\big)+\delta U_2(L) \geq (1-\delta)\cdot\big(2\cdot\sigma_2(L)+0\cdot(1-\sigma_2(L))\big)+\delta U_2(H).$$

  or equivalently,

$$\delta\left[U_1(H) - U_1(L)\right] \leq (1-\delta)\left[2\sigma_2(L) + 1 - \sigma_2(L)\right].$$

- By assumption $\sigma_2(L) < \sigma_2(H)$, we have

$$(1-\delta)\left[2\sigma_2(L) + 1 - \sigma_2(L)\right] \leq (1-\delta)\left[2\sigma_2(H) + 1 - \sigma_2(H)\right].$$

- This ensures that the IC condition for exploitation to hold for Player 1 on history H.

$$\delta\left[U_1(H) - U_1(L)\right] < (1-\delta)\left[2\sigma_2(H) + 1 - \sigma_2(H)\right].$$

- This implies that Player 2 should not play h upon this history, i.e. $\sigma_2(H) = 0$, a contradiction.
- The driving force of this result is that in a reputation game Player 1's incentive to play a cheating action is higher if Player 2 trusts him more.

An Example: $\mu_0 > 0$

- The intertemporal incentive will not fully collapse as $\mu = 0$.
- To see it, if Player 1 plays L repeatedly in a perfect equilibrium, Player 1's discounted average pay-off is
$$(1-\delta)\cdot\left(1+\delta+\delta^2+\ldots+\delta^n+\ldots\right)=1.$$
- However, if Player 1 defects to playing H forever, future Player 2 will assign probability 1 to the good type and his pay-off will be at least
$$(1-\delta)\cdot\left(0+2\delta+2\delta^2+\ldots+2\delta^n+\ldots\right)>2\delta.$$
- Hence, the deviation will be profitable for $\delta > \frac{1}{2}$.
- However, the Stackelberg outcome (H, h) cannot be enforced.

An Example: $\mu_0 > 0$

- The Player 1's payoff from Stackelberg outcome (H, h) is

$$(1-\delta)\cdot\left(2+2\delta+2\delta^2+\ldots+2\delta^n+\ldots\right)=2.$$

- Following a history H, Player 1 has a profitable deviation of alternating between L and H, which results in a sequence of pay-offs of at least 4,0,4,0,... with a discounted average larger than 2.

$$(1-\delta)\cdot\left(4+0\delta+4\delta^2+\ldots+4\delta^{2n}+\ldots\right)=4\cdot\frac{1-\delta}{1-\delta^2}>2.$$

- A grim trigger strategy argument does not work by punishing the deviation with worst payoff.
- Reason is that the Player 1 can recover from the deviation!

An Example: $\mu_0 > 0$

- In this equilibrium, the opportunistic type plays a completely mixed strategy in state L.
- When H is observed, Player 2 updates his belief on the good type to such an extent that he is willing to play the trusting action h, while the opportunistic Player 1 exploits this trust by playing L.
- State L is a reputation-building state and state H is a reputation-exploitation state.
- We can verify this is an equilibrium:
  - Let us verify that the prescribed strategy profile is indeed an equilibrium. First, in state L, Player 2 knows Player 1 's type; we need $\sigma_1(L) = \frac{1}{2}$ to make Player 2 indifferent. Second, in state H, the posterior belief on the good type is given by $\mu_H = \frac{\mu_0}{\mu_0 + (1-\mu_0)\frac{\sigma_1(L)}{1+\sigma_1(L)}} = \frac{\mu_0}{\mu_0 + \frac{1}{3}(1-\mu_0)}$, where $\frac{\sigma_1(L)}{1+\sigma_1(L)}$ in the denominator is Player 2 's steady-state belief on state H; to induce Player 2 to play the trusting action h, we need $\mu_H \geq \frac{1}{2}$, i.e. $\mu_0 \geq \frac{1}{4}$.
  - We still need to check Player 1 's incentives upon the two histories. For Player 1 to be indifferent in state L, we need $\sigma_2(L) = \frac{3\delta-1}{3\delta+1}$, which requires $\delta \geq \frac{1}{3}$. It is straightforward to check that action L is Player 1's best response in state H.

An Example

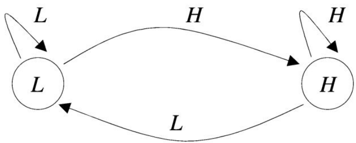

FIGURE 3

In state H: (L, h) is played, and hence the transition back to state H is off the equilibrium path. In state L: Player 1 plays H with probability $\sigma_1(L) = \frac{1}{2}$; and Player 2 plays h with probability $\sigma_2(L) = \frac{3\delta-1}{3\delta+1}$

Reputation Dynamics

- The state space $S = \{H, L\}^N$ grows exponentially with N, and *a* priori the direct graph on the 2^N states lacks an obvious, tractable transition.
- Note that a history reveals Player 1's type as long as it contains a single cheating action L and the number of H's since the most recent L measures how far Player 1 has gone towards a clean record with straight H's. This observation motivates a natural index which cuts through the graph on S.

Definition (2)

The reputation index of s, $I(s)$, is the number of good actions H since the most recent cheating action, L, in state s. Formally,

$$I(s) = \begin{cases} N & \text{if } \forall i \in \{1, \ldots, N\}, s_i = H \\ \min\{i : s_i = L\} - 1 & \text{otherwise} \end{cases}$$

- For example, $I(\cdots LHH) = 2, I(\cdots L) = 0$, and $I(H \cdots HH) = N$.
- By definition, an additional H will increase the index from i to $\min\{i + 1, N\}$ and an L will reduce the index down to 0.

Reputation Dynamics

Lemma (2)

Not acquiring information is an equilibrium if and only if $C(1) \geq \mu_0 [u_2(H, h) - u_2(H, l)]$. In this equilibrium, Player 2 plays l and the opportunistic Player 1 plays L.

- If Player 2 never acquires information, the opportunistic type of Player 1 plays L.
- Player 2 is tempted to acquire one period of information to check Player 1's type.
- The benefit of such information is $\mu_0 [u_2(H, h) - u_2(H, l)]$ because Player 2 should play h instead of l if he finds out that Player 1 is the good type. The cost is $C(1)$. Information acquisition is beneficial if $C(1) < \mu_0 [u_2(H, h) - u_2(H, l)]$.
- We focus on $C(1) < \mu_0 [u_2(H, h) - u_2(H, l)]$ in the later.

Reputation Dynamics

Lemma (2)

Not acquiring information is an equilibrium if and only if $C(1) \geq \mu_0 [u_2(H, h) - u_2(H, l)]$. In this equilibrium, Player 2 plays l and the opportunistic Player 1 plays L.

- If Player 2 never acquires information, the opportunistic type of Player 1 plays L.
- Player 2 is tempted to acquire one period of information to check Player 1's type.
- The benefit of such information is $\mu_0 [u_2(H, h) - u_2(H, l)]$ because Player 2 should play h instead of l if he finds out that Player 1 is the good type. The cost is $C(1)$. Information acquisition is beneficial if $C(1) < \mu_0 [u_2(H, h) - u_2(H, l)]$.

Long-run Reputation Dynamics

Theorem (2)

If $\delta > \bar{\delta}$, then every equilibrium $(\mathrm{I}, \sigma_1, \alpha, \sigma_2, \lambda)$ exhibits a stochastic reputation cycle: there exists an integer $n^*, 0 < n^* \leq N$, such that the set $\{0, 1, \ldots, n^*\}$ consists of all states on the equilibrium path and it can be broken down into the following two phases:

- the set $\{0, 1, \ldots, n^* - 1\}$ forms a reputation-building phase in which Player 1 plays strict mixed strategies: $0 < \sigma_1(i) \leq \bar{\mu} < 1$ for each state i in this set.
- the state n^* is a reputation-exploitation phase where Player 1 plays L with probability 1.

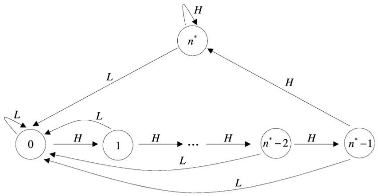

FIGURE 4

$\sigma_1(i) \in (0, \bar{\mu}]$ if $i < n^*$ and $\sigma_1(n^*) = 0$. Therefore, reputation collapses completely at n^* and the transition from n^* back to itself is off the equilibrium path

Proof Sktech

- Step 1. Unique ergodic set.
  - The Player 1's strategy induce at least one ergodic set on I.
  - An ergodic set must take the form of $\{m, m+1, ..., n\}$, where $0 \leq m \leq n \leq N$.
  - Two extreme cases, $\{N\}$ and $\{0\}$.
- Step 2. Increasing incentive.
  - Step 1 shows that Player 1 must go from state 0 to higher states in equilibrium. Therefore, Player 1 must be compensated for "climbing" towards higher states by playing the static dominated action H.
  - Player 2's expected strategies $\bar{\sigma}_2(i) = \sum_n \alpha(n)\sigma_2^n(i)$ must be non-decreasing on the equilibrium path in order to provide reputation-building incentives.
- Step 3. Uniform bound on mixing probabilities.
  - By contradiction: suppose exist i such that $\sigma_1(i) > \bar{\mu}$.
  - The Play 2 chooses h whenever observes this state.
  - By monotonicity of the Player 2's expected strategies, $\sigma_2^m(i) = 1$ for $m > i$. Hence, $\bar{\sigma}_2(i) = \bar{\sigma}_2(i+1) = \cdots = \bar{\sigma}_2(n^*)$.
  - Player 1 has no incentive to play H in state i now and L later: he prefers to play L as soon as possible as he discounts his pay-off.

Additional Properties

Lemma (3)

On the equilibrium path, the probability that Player 2 plays the trusting action $h, \bar{\sigma}_2(i) = \sum_n \alpha(n)\sigma_2^n(i)$, and Player 1's expected pay-off, $U_1(i)$, are strictly increasing in reputation index i.

Corollary (1)

If the cost structure is $C(N) = 0$ and $C(N+1) = \infty$ as in the example in Section 4 , then $n^* = N$.

Corollary (2)

Fix an equilibrium $(\mathrm{I}, \sigma_1, \alpha, \sigma_2, \lambda)$. Then $\lambda(i)$, the probability that state i is reached in equilibrium, is strictly decreasing in i and $0 < \lambda(i) < \bar{\mu}^i$ for any $i \in \{1, 2, \ldots, n^*\}$, where n^* is characterized in Theorem 2 .

The Short-lived Players

Theorem (3)

Suppose $\delta > \bar{\delta}$ and the cost function is strictly increasing. Consider any equilibrium $(\mathrm{I}, \sigma_1, \alpha, \sigma_2, \lambda)$ and let n^* be the highest reputation index in this equilibrium characterized in Theorem 2. Then the following properties hold:

- n^* is the maximum number of periods of information that Player 2 acquires for generic cost functions.
- Player 2 plays completely mixed strategies: $\alpha(i) > 0, i \in \{0, 1, \ldots, n^*\}$.
- Upon acquiring information, Player 2 plays pure strategy l whenever he sees an L and he plays pure strategy h whenever he sees straight H's.

Proposition (1)

If $\delta > \bar{\delta}$, then n^* and σ_1 are independent of δ.

The Short-lived Players

- Here, we demonstrate the construction for the family of linear cost functions: $C(n) = cn$, where $0 < c < \bar{c} = \mu_0 \left[u_2(H, h) - u_2(H, l) \right]$.
- We focus on the characterization of n^* and Player 2's action when he does not buy information.

Proposition (6.1)

If $C(n) = cn$, the equilibrium on I is unique for generic $c \in (0, \bar{c})$ when $\delta > \bar{\delta}$. Moreover, there exist a strictly decreasing sequence $\{c_n\}_{n=1}^{\infty}$ and a cut-off $c^* > 0$ that are independent of δ, such that

- if $c \in (c_n, c_{n+1})$, then $n^* = n$;
- if Player 2 does not acquire information, he plays h if $c \in (0, c^*)$ and l if $c \in (c^*, \bar{c})$.

Extensions

- Other information acquisition schemes.
  - One may also consider a sampling procedure where the short-lived players observe a random subset of past transactions.
- Multiple types
  - We conjecture that the nature of reputation dynamics will not change qualitatively if in addition there is a third type who always plays the cheating action L.
- Observability
  - The short-lived players' actions are unobservable to each other.
  - If, instead, the history of short-lived players is observable, then the short-lived players could coordinate their actions.

Thanks!