Slow Observational Learning and Reputation Failures

HARRY PEI
Department of Economics, Northwestern University

Jan 6th, 2020
AEA Meeting, San Diego, CA
Model

- Time: $t = 0, 1, 2, ...$

- Long-lived P1 (e.g., seller), chooses $a_t \in A$, discount $\delta \in (0, 1)$. Short-lived P2s (e.g., buyers), choose $b_t \in B$, with $A$ and $B$ finite.

- Stage game payoffs: $u_1(a_t, b_t)$ and $u_2(a_t, b_t)$.

- Seller has two possible types:
  1. with prob $\pi_0 \in (0, 1)$, mechanically plays pure Stackelberg action,
  2. with prob $1 - \pi_0$, strategic type that maximizes payoff.
Model: Reputation Building Through Social Learning

Period $t$ buyer observes:

1. buyers’ actions from 0 to $t - 1$, namely, $b_0, b_1, ..., b_{t-1}$.
2. and a bounded (possibly stochastic) subset of seller’s past actions.

Most of this talk: Period $t$ buyer observes:

- $b_0, ..., b_{t-1}$,
- and $a_{t-K}, ..., a_{t-1}$, with $K \in \mathbb{N}$ a parameter.

By the end: Stochastic network monitoring.

- private monitoring of P1’s actions, private learning of P1’s type.
Motivation & Takeaway

Heterogenous accessibility of different types of information:

• buyer can skim through online reviews and observe how frequent each product was purchased and the time trend;

• buyer needs to read reviews carefully to figure out seller’s action, and she has limited capacity to process such detailed info.

Effectiveness of reputation building through social learning:

• info about seller’s actions is dispersed among buyers.

Result: Exist equilibria s.t. patient seller receives low payoff.

• Contrasts to Fudenberg and Levine (89,92) in which patient seller guarantees high payoff.

Why?

• Learning cannot stop, buyers cannot herd on bad actions.

• The speed of observational learning vanishes to 0 as $\delta \to 1$. 
Assumption on Stage-Game Payoffs

Assumption 1

$u_1$ and $u_2$ satisfy:

1. P1 has a unique pure Stackelberg action, denoted by $a^* \in A$.
2. P2 has a unique best reply against $a^*$, denoted by $b^* \in B$.
3. There exists a pure strategy Nash Equilibrium in the stage-game.

Interesting case: P1 can strictly benefit from committing to $a^*$.

<table>
<thead>
<tr>
<th></th>
<th>$T$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H$</td>
<td>2,1</td>
<td>$-1,0$</td>
</tr>
<tr>
<td>$L$</td>
<td>3,$-1$</td>
<td>0,0</td>
</tr>
</tbody>
</table>
Result: Reputation Failure

Let $v_1$ be P1’s worst pure stage-game NE payoff, and $\delta \in (0, 1)$ is a cutoff discount factor that depends only on $u_1$ and $u_2$.

**Theorem 1**

*If $u_1$ and $u_2$ satisfy Assumption 1,
then for every $K \in \mathbb{N}$, there exists $\pi_0 \in (0, 1)$,
such that for every $\pi_0 \in (0, \pi_0)$ and $\delta > \delta$,*

$\exists$ a sequential equilibrium s.t. strategic P1 receives payoff $v_1$.

Recall: In Fudenberg and Levine (1989, 1992) and Gossner (2011),

- Fix $\pi_0$ and let $\delta \to 1$,

  P1’s payoff in all equilibria is no less than $u_1(a^*, b^*)$. 
Remark: No Bad Herd

Proposition 1

At every on-path history $h^t$ of every Bayes Nash equilibrium, if $P_2$ attaches positive probability to $P_1$ being committed at $h^t$, then $P_2$s cannot herd on any action that is not $b^*$ at $h^t$. 
Proof Sketch of Theorem 1

Focus on Product Choice Game with Public Randomization

<table>
<thead>
<tr>
<th></th>
<th>$T$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H$</td>
<td>2, 1</td>
<td>−1, 0</td>
</tr>
<tr>
<td>$L$</td>
<td>3, −1</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

I construct a three-phase equilibrium:

1. Reputation-building phase.
   
   *Play starts from here, P1’s payoff is $v_1$, P2 slowly learns.*

2. Reputation-maintenance phase.
   
   *Play eventually moves here, P1’s payoff is $u_1(a^*, b^*)$.*
   
   *Learning stops on-path.*

3. Punishment phase.
   
   *Only reached off-path, P1’s payoff is $v_1$. Learning stops.*
Reputation-Building Phase

Play starts from a reputation-building phase, in which:

- P2 plays $N$.
- Strategic P1 mixes between $H$ and $L$ s.t. P2 believes that $H$ is played with prob $1/2$ (more sophisticated construction under private learning).

Phase transition: By the end of period $t$,

- If $a_t = L$, then remains in the reputation-building phase in period $t + 1$.
- If $a_t = H$, then transits to the reputation-maintenance phase in period $t + 1$ with probability:
  
  $p(\delta) \equiv \frac{1 - \delta}{2\delta}$,

  determined by public randomization in the beginning of $t + 1$.

- This transition prob makes P1 indifferent between $H$ and $L$, which vanishes to 0 as $\delta \to 1$. 
Reputation-Maintenance Phase & Punishment Phase

After play transits to reputation-maintenance phase.

- P1 plays $H$ and P2 plays $T$ on the equilibrium path.

Phase transition: In period $t + 1$,

- Play remains in the reputation-maintenance phase if $(a_t, b_t) = (H, T)$.
- Otherwise, play transits to the punishment phase.

Punishment phase is absorbing, in which P1 plays $L$ and P2 plays $N$.

- Future P2 knew play is in the punishment phase when $N$ occurs after $T$.

In the $t \to \infty$ limit:

- Play reaches the reputation maintenance phase with probability 1.
  But the number of periods it takes goes to infinity as $\delta \to 1$. 
How to Square this with Gossner (2011)?

Gossner’s upper bound on the sum of P2s’ *1-step-ahead prediction errors*:

$$
\mathbb{E}^{a^*} \left[ \sum_{t=0}^{\infty} d\left(y_t(\cdot|a^*)\bigg|y_t\right) \right] \leq -\log \pi_0
$$

The above inequality implies a payoff lower bound for P1 if

- whenever P2 does not have strict incentive to play $b^*$,
  
  $$d(y_t(\cdot|a^*)\bigg|y_t)$$
  
  is bounded from below by a positive number.

- This implies at most a bounded number of bad periods.

- As $\delta \to 1$, the payoff consequence of bad periods vanishes.
How to Square this with Gossner (2011)?

Gossner’s upper bound on the sum of P2s’ *1-step-ahead prediction errors*:

\[
\mathbb{E}^{a^*} \left[ \sum_{t=0}^{\infty} d \left( y_t (\cdot | a^*) \big | y_t \right) \right] \leq - \log \pi_0
\]

My model applying to the product choice game (or any MSM game):

- If P1 plays \( a^* \) in every period, then either \( d \left( y_t (\cdot | a^*) \big | y_t \right) > 0 \) or \( b_t = b^* \) or \( b_{t+i} = b^* \) for all \( i \in \{1, 2, \ldots, K\} \).
- As \( \delta \to 1 \), \( d \left( y_t (\cdot | a^*) \big | y_t \right) \) goes to 0, and expected number of bad periods explodes.
- As \( \delta \to 1 \), the payoff consequence of bad periods is not negligible.
Remark: Low Consumer Welfare

Suppose a social planner discounts future consumers’ payoffs by $\delta$.

- $v_2$ is P2’s worst pure stage-game NE payoff.

Proposition 2

*For every $K \in \mathbb{N}$ and $\varepsilon > 0$,*

there exist $\pi_0 \in (0, 1)$ and $\delta \in (0, 1)$,

such that for every $\pi_0 \in (0, \pi_0)$ and $\delta \geq \delta$,

$\exists$ a sequential equilibrium s.t. P2’s welfare is less than $v_2 + \varepsilon$.

In product choice game, exists equilibrium s.t. both players’ payoffs are close to their minmax payoff.
Extension to Stochastic Monitoring

Stochastic network among buyers: \( \mathcal{N} \equiv \{ \mathcal{N}_t \}_{t=1}^{\infty} \), with

\[ \mathcal{N}_t \in \Delta \left( 2^{\{0,1,\ldots,t-1\}} \right), \quad \text{with } N_t \text{ the realization of } \mathcal{N}_t. \]

Buyer in period \( t \) observes:

- \( b_0, b_1, \ldots, b_{t-1} \).
- Realization of \( \mathcal{N}_t \) and \( \{a_j\}_{j \in \mathcal{N}_t} \).

Seller does not observe the realization of \( \mathcal{N}_t \).

In MSM games (e.g., product choice game), my result generalizes when:

Assumption 2

For every \( t \neq s \), \( \mathcal{N}_t \) and \( \mathcal{N}_s \) are independent random variables.

There exist \( K \in \mathbb{N} \) and \( \gamma \in (0,1) \) such that for every \( t \geq 1 \),

\[ \Pr \left( \left| \mathcal{N}_t \right| \leq K \right) = 1 \text{ and } \Pr \left( t-1 \in \mathcal{N}_t \right) \geq \gamma. \]
Challenges

Period $t$ player 2 observes:

$$h^t_2 \equiv \left\{ N_t, b_0, b_1, \ldots, b_{t-1}, (a_s)_{s \in N_t} \right\}.$$

Player 1 observes:

$$h^t_1 \equiv \left\{ b_0, b_1, \ldots, b_{t-1}, a_0, a_1, \ldots, a_{t-1} \right\}$$

Two challenges in constructing equilibrium:

1. **Private monitoring** of player 1’s past actions.
2. Player 2s’ **private learning** about player 1’s type.

Proof uses a combination of **belief-free approach** and **belief-based approach**.
Conclusion

Reputation model in which short-run player observes:

- all his predecessors’ actions,
- a bounded subset of long-run player’s past actions.

In a large class of games,

- reputation fails since the speed of learning vanishes as \( \delta \to 1 \).

Novel questions on social learning:

- Social learning about endogenous actions rather than exogenous state.
- Speed of social learning rather than asymptotic beliefs.
- Discounted payoff rather than long-run outcomes.
1. Social learning: Banerjee (92), Bikhchandani, Hirshleifer, and Welch (92), Smith and Sørensen (00).
   Difference: Speed and welfare consequences instead of \( t \to +\infty \).

2. Efficiency of social learning: Rosenberg and Vieille (19).
   Difference: My efficiency standard takes discounting into account.

3. Reputation effects: Fudenberg and Levine (89,92), Gossner (11).
   Difference: Players’ endogenous actions as public signals.

4. Reputation with limited memory: Liu (11), Liu and Skrzypacz (14).
   Difference: Their models deliberately shut down social learning.

5. Bad reputation: Ely and Valimaki (03), Ely, Fudenberg and Levine (08)
   Difference: P2’s action can statistically identify P1’s past actions.

6. Logina, Lukyanov and Shamruk (19)
   Difference: P2 observes current P1’s action versus P1’s past actions.
   P1 can strictly benefit from commitment or not.
Construction without Public Randomization

Reputation Building Phase:

1. P2 has never played $T$ before & $a_{t-1} = L$,
   - P1 mixes between $H$ and $L$ s.t. overall prob of $H$ is $1/2$.
   - P2 plays $N$ with prob 1.

2. P2 has never played $T$ before & $a_{t-1} = H$,
   - P1 mixes between $H$ and $L$ s.t. overall prob of $H$ is $1/2$.
   - P2 plays $T$ with prob $\frac{1-\delta}{2\delta}$. 
Construction without Public Randomization

Reputation Maintenance Phase:

1. P2 plays $T$ for the first time in period $t-1 \& a_{t-1} = L$,
   P1 plays $H$ for sure.
   P2 plays $T$ with prob $\frac{4\delta - \delta^2 - 1}{3 - \delta}$.

2. P2 plays $T$ for the first time in period $t-1 \& a_{t-1} = H$,
   P1 plays $H$ for sure $\&$ P2 plays $T$ for sure.

3. $N$ has never occurred after $T$, $T$ occurs at least twice $\& a_{t-1} = H$,
   P1 plays $H$ for sure $\&$ P2 plays $T$ for sure.
Construction without Public Randomization

Punishment Phase:

1. \( N \) has never occurred after \( T \), \( T \) occurs at least twice & \( a_{t-1} = L \),
   
P1 plays \( L \) for sure & P2 plays \( N \) for sure.

2. \( N \) has occurred after \( T \),
   
P1 plays \( L \) for sure & P2 plays \( N \) for sure.