LONG-TERM CAUSAL EFFECTS OF INTERVENTIONS IN MULTIAGENT ECONOMIC MECHANISMS

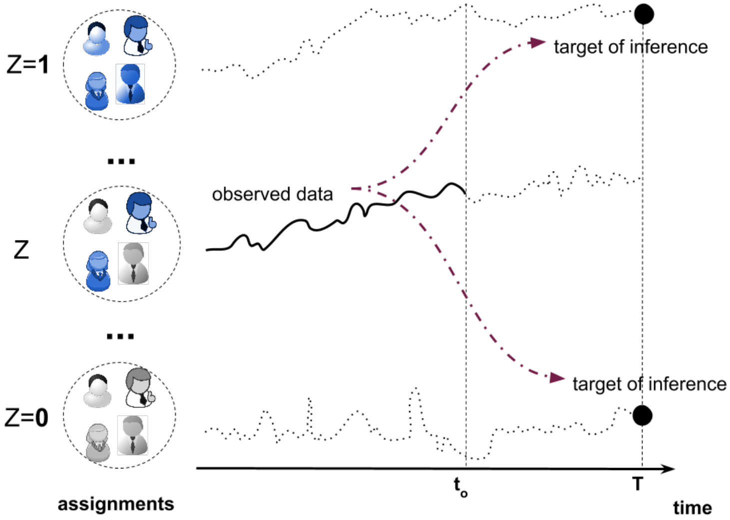
Panos Toulis and David C. Parkes

October 16, 2015

Harvard University

- We study *causal effects* of interventions in economic mechanisms.
- Effect is *causal* if comparing *counterfactuals*.
- Effects of interventions fluctuate until new equilibrium; causal effects measured in equilibrium are *long-term* causal effects;
- Long-term causal effects are more representative of the value of interventions.

(Examples: increase in reserve price, change in matching mechanism, etc.)


Z=1
Z
Z=0
observed data
target of inference
target of inference
assignments
t_o
T
time

Two main challenges:

- *Temporal dynamic actions*: statistical estimation relies on data before new equilibrium.
- *Strategic interference*: Agent actions depend on assignment of other agents.

Two main challenges:

- *Temporal dynamic actions*: statistical estimation relies on data before new equilibrium.
- *Strategic interference*: Agent actions depend on assignment of other agents.

Our contributions in this paper:

(1) Formalize problem using potential outcomes framework of causal inference (Neyman-Rubin).
(2) Develop a method to estimate long-term causal effects.
(3) State sufficient assumptions for identification.
(4) Illustrate on real-world data.

Experiment:

- Agents $\mathcal{I}$, games $\mathcal{G}$, actions $\mathcal{A}$.
- $Z =$ assignment vector (every agent assigned to one game);
- e.g., $Z_i = j \in \mathcal{G}$ assignment of agent i to game j.

Agent actions:

- $A_{it}(Z) \in \mathcal{A} =$ action of agent i at time t under assignment Z.
- $\alpha_{j,t}(Z) \in \Delta^{|\mathcal{A}|} =$ frequency of $\{A_{it}(Z) : Z_i = j\}$. (Δ is the simplex).
- **Observed data**: $\alpha_{j,0}(Z), \alpha_{j,1}(Z), \dots, \alpha_{j,t_o-1}(Z)$.

Objective:

- Experimentally select best game from $\mathcal{G}$ according to objective R.
  (e.g., $R =$revenue.)
- $R_{j,t}(Z) = h(\alpha_{j,t}(Z))$, objective value in game j, time t, assignment Z, for an appropriate function h.

Definition

The causal effect at time t on objective R of game j over game j' is equal to the quantity

$$\tau(j, j'; t) = \underbrace{R_{j,t}(j\mathbf{1})}_{\text{all agents play in } j} - \underbrace{R_{j',t}(j'\mathbf{1})}_{\text{all agents play in } j'} = h(\alpha_{j,t}(j\mathbf{1})) - h(\alpha_{j',t}(j'\mathbf{1})).$$

Long-term average causal effect (LACE) at appropriate time $t = T$.

- However, **only one** assignment Z observed $\Rightarrow$ only $A_{it}(Z)$ observed. All other outcomes will be **missing**.
- Challenge: **Predict** missing outcomes that are important.
- **Every** method makes assumptions on that prediction; assumptions usually not made explicit.

- A *behavior* is a distribution over actions; finite set $\mathcal{B} = \{1, 2, \ldots, |\mathcal{B}|\}$; (e.g., “aggressive”, “passive”, etc.)
- $B_{it}(Z) \in \mathcal{B}$ = behavior agent i adopts at time t, assignment Z.
- $\beta_{j,t}(Z) \in \Delta^{|\mathcal{B}|}$ = frequency of behaviors $\{B_{it}(Z) : Z_i = j\}$, game j.

FIRST IDEA: AUGMENT WITH LATENT BEHAVIORS

- A *behavior* is a distribution over actions; finite set $\mathcal{B} = \{1, 2, \ldots, |\mathcal{B}|\}$; (e.g., "aggressive", "passive", etc.)
- $B_{it}(Z) \in \mathcal{B}$ = behavior agent i adopts at time t, assignment Z.
- $\beta_{j,t}(Z) \in \Delta^{|\mathcal{B}|}$ = frequency of behaviors $\{B_{it}(Z) : Z_i = j\}$, game j.

Why behaviors?

(1) Given $\beta_{j,t}(Z)$ we have distribution on actions $\alpha_{j,t}(Z)$.
(2) Leverage behavioral game theory.
(3) Identification assumptions (next slides) more natural on behavioral space.

Assumption (#1, Temporal model of behaviors)

Let $\mathcal{F}_t$ be the filtration for $\beta_{j,t}(Z)$. Under assignment Z, for a known prior π and observation model f, there exist parameters $\theta_z = (\phi_z, \psi_z)$, such that

$$\beta_{j,0}(Z) \sim \pi(\cdot|\phi_z)$$
$$\beta_{j,t}(Z) \mid \mathcal{F}_{t-1} \sim f(\cdot|\psi_z, \mathcal{F}_{t-1}),$$

Assumption (#1, Temporal model of behaviors)

Let $\mathcal{F}_t$ be the filtration for $\beta_{j,t}(Z)$. Under assignment Z, for a known prior π and observation model f, there exist parameters $\theta_z = (\phi_z, \psi_z)$, such that

$$\beta_{j,0}(Z) \sim \pi(\cdot|\phi_z)$$
$$\beta_{j,t}(Z) \mid \mathcal{F}_{t-1} \sim f(\cdot|\psi_z, \mathcal{F}_{t-1}),$$

- The model is known but its parameters $\theta_z = (\phi_z, \psi_z)$ are unknown.
- Parameters may depend on assignment Z as well as game j.
- xx TODO: *Should depend on game as well.* xx

Assumption (#2, Initial behaviors)

Every agent picks one fixed but possibly unknown behavior at $t = -1$*. Thus, for every assignment* Z*,*

$$\beta_{j,-1}(Z) = \beta^{(0)}.$$

Assumption (#2, Initial behaviors)

Every agent picks one fixed but possibly unknown behavior at $t = -1$. *Thus, for every assignment* Z,

$$\beta_{j,-1}(Z) = \beta^{(0)}.$$

- Precludes that agents change their initial behavior at $t = 0$ on anticipation of assignment Z.
- More relaxed assumption possible. (e.g., agents sampling i.i.d. from $\beta^{(0)}$.)

Assumption (#3, Behavioral ignorability)

For every assignment Z,

$$Z \perp\!\!\!\perp \beta_{j,t}(Z) \quad | \quad \mathcal{F}_{t-1}, Game_j. \tag{1}$$

- Assignment Z does not add information about behaviors at t given the behaviors up to $t-1$. (assignment mechanism is *ignorable*)
- Precludes interference between games, dependence on identity of agents, or on number of agents, etc.

Theorem (Estimation of long-term effects)

Suppose that Assumptions #1, #2, #3 hold. Then, LACE can be identified if parameters $\theta = (\phi, \psi)$ *of the temporal model can be identified as* $t_o \to \infty$.

Theorem (Estimation of long-term effects)

Suppose that Assumptions #1, #2, #3 hold. Then, LACE can be identified if parameters $\theta = (\phi, \psi)$ of the temporal model can be identified as $t_o \to \infty$.

Proof sketch:

- Under Assumptions #1 and #2, $\theta_z = (\phi_z, \psi_z) \equiv (\phi, \psi_z)$. (Prior π of initial behavior $\beta_{j,0}(Z)$ completely defined for any Z by **randomization**.)
- Under Assumptions #1 and #3, $\theta_z = (\phi_z, \psi_z) \equiv (\phi_z, \psi)$. (Temporal dynamics of behaviors governed by same parameters across assignments Z.)
- Then, if we can learn (ϕ, ψ), then we can predict counterfactuals to any assignment Z and time T.
- xx TODO: *Dependence on game. Can use behavioral game theory to predict between games.* xx

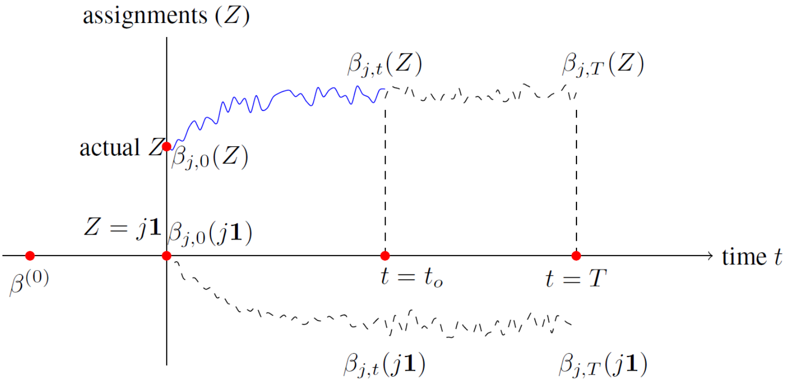
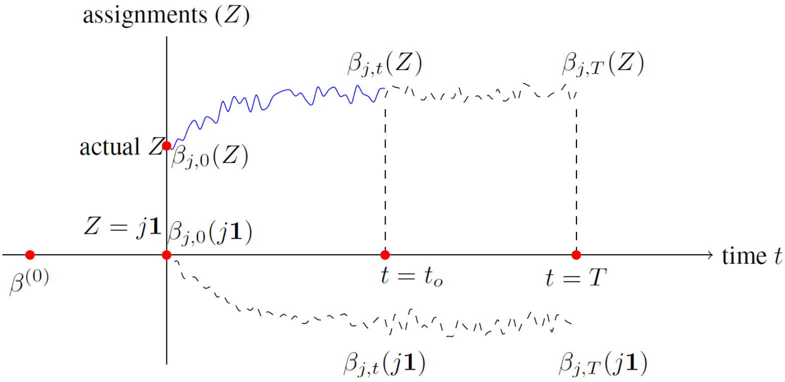


Figure 1: **Step (1)**: learn ψ of $\beta_{j,t}(Z)$ under Z using data (blue line); **Step (2)**: learn ϕ of $\beta_{j,0}(Z)$ under Z; **Step (3)**: Under randomization, use ϕ to estimate $\beta^{(0)}$; **Step (4)**: Use ψ to estimate $\beta_{j,T}(j1)$ starting from $\beta_{j,0}(j1) = \beta^{(0)}$.

ARE ASSUMPTIONS #1-#3 NECESSARY?



- **Difference-in-differences** (DID). DID compares difference in treatment vs. control; requires strong additive modeling assumptions.
- **Structural approach**, e.g., Athey et. al., (2008): Estimate bidder valuations from observed data in one auction and predict counterfactual bids in other auction, assuming equilibrium play. ignorability assumption; no long-term effects.
- **Directed acyclical graphs** (DAGs), e.g., Bottou et. al. (2012): Create full DAG and predict counterfactuals. Crucial assumption: underlying DAG remains stable under treatment assignment (form of ignorability); problematic in equilbrium systems (Dash, 2011).

TOY EXAMPLE: BEHAVIORAL EXPERIMENT OF RAPOPORT & BOEBEL (1992)

	B_1	B_2	B_3	B_4	B_5
A_1	W	L	L	L	L
A_2	L	L	W	W	W
A_3	L	W	L	L	W
A_4	L	W	L	W	L
A_5	L	W	W	L	L

- Game consisted of two players A (=row) and B(=column), each having 5 actions.
- Numbers W, L indicate *payoffs* for row-player; W= win, L=loss.
- Two different *game designs*: $(W, L) = (\$10, -\$6)$ and $(W, L) = (\$10, -\$1)$.

Game	Period	A_1	A_2	A_3	A_4	B_1	B_2	B_3	B_4
1	1	0.308	0.307	0.113	0.120	0.350	0.218	0.202	0.092
1	2	0.293	0.272	0.162	0.100	0.333	0.177	0.190	01.40
1	3	0.273	0.350	0.103	0.123	0.353	0.133	0.258	0.102
1	4	0.295	0.292	0.113	0.135	0.372	0.192	0.222	0.063
2	1	0.258	0.367	0.105	0.143	0.332	0.115	0.245	0.140
2	2	0.290	0.347	0.118	0.110	0.355	0.198	0.208	0.108
2	3	0.355	0.313	0.082	0.100	0.355	0.215	0.187	0.110
2	4	0.323	0.270	0.093	0.105	0.343	0.243	0.168	0.107

Figure 2: Columns: Game=design, A_k=frequency of action A_k from players A. B_k=frequency of action B_k from players B.

- Randomized 40 players in one of two designs; $\mathcal{I} = \{1, 2, \ldots, 40\}$.
- $\mathcal{G} = \{1, 2\}$; $\mathcal{A} = \{1, 2, 3, 4, 5\}$ (row/column); each player played both as row & column vs. two different opponents in matchups.
- (In Table) Proportions are aggregate actions $\alpha_{j,t}(Z)$, $t = 1, 2, 3, 4$.

Game	Period	A_1	A_2	A_3	A_4	B_1	B_2	B_3	B_4
1	1	0.308	0.307	0.113	0.120	0.350	0.218	0.202	0.092
1	2	0.293	0.272	0.162	0.100	0.333	0.177	0.190	01.40
1	3	0.273	0.350	0.103	0.123	0.353	0.133	0.258	0.102
1	4	0.295	0.292	0.113	0.135	0.372	0.192	0.222	0.063
2	1	0.258	0.367	0.105	0.143	0.332	0.115	0.245	0.140
2	2	0.290	0.347	0.118	0.110	0.355	0.198	0.208	0.108
2	3	0.355	0.313	0.082	0.100	0.355	0.215	0.187	0.110
2	4	0.323	0.270	0.093	0.105	0.343	0.243	0.168	0.107

- Assume objective $R_{j,t}(Z) = c'\alpha_{j,t}(Z)$, linear function of aggregate action.
- “What is the long-term causal effects of design #1 over #2 on objective?”.
  - game designs = different options for payoffs (W, L).
  - Assume short-term $t = 1, 2, 3$; long-term = 4 (held-out).

STEP 1. GAME-THEORETIC MODEL, BEHAVIORS $\rightarrow$ ACTIONS

- We use the Quantal Level-k (Stahl and Wilson, 1994).
- Behaviors $\mathcal{B} = \{b_0, b_1, b_2\}$; increased level of sophistication; parametrized by $\Lambda = (\lambda_1, \lambda_{(1)2}, \lambda_2)$ (precision parameters).
- Model gives distribution over aggregate action $\alpha_{j,t}(Z)$:

$$g(\alpha_{j,t}(Z) | \beta_{j,t}(Z), \Lambda, G_j),$$

  where G_j = game design, $\beta_{j,t}(Z)$ = aggregate behavior.
- Other modeling approaches are possible; (QRE, McKelvey and Palfrey, 1995), cognitive hierarchy (Camerer et al., 2004).

STEP 2. TEMPORAL BEHAVIORAL MODEL, BEHAVIOR$_{t-1} \rightarrow$ BEHAVIOR$_t$

- We use a simple VAR(1) model

$$\omega_{j,t} = \psi_0 + \psi_1 \omega_{j,t-1} + \psi_2 \varepsilon_t, \tag{2}$$

where $\omega_{j,t} \stackrel{\text{def}}{=} \mathrm{L}(\beta_{j,t}(Z))$; $L(x) = (0, \log(x_1/x_0), \log(x_2/x_0), \ldots)$ (because $\beta_{j,t}(Z)$ is on the simplex.).

- Other modeling approaches are possible; e.g., time-series on compositional data (Grunwald et. al., 1993), Brownian motion on simplex (Evans, 2003).

The posterior distribution is

$$p(\Lambda, \phi, \psi | \alpha_{j,1:t_o}(Z)) \propto \int_{\beta_{j,1:t_o}(Z)} \underbrace{g(\alpha_{j,0:t_o}(Z) | \Lambda, \beta_{j,1:t_o}(Z))}_{\text{game-theoretic model}} \times$$

$$\times \underbrace{f(\beta_{j,1:t_o}(Z) | \psi, \beta_{j,0}(Z))}_{\text{temporal behavioral model}} \times$$

$$\times \quad \underbrace{\pi(\beta_{j,0}(Z) | \phi)}_{\text{prior for behavior } t\text{=0}}$$

$$\times \underbrace{\pi_0(\Lambda, \phi, \psi)}_{\text{other priors}}.$$

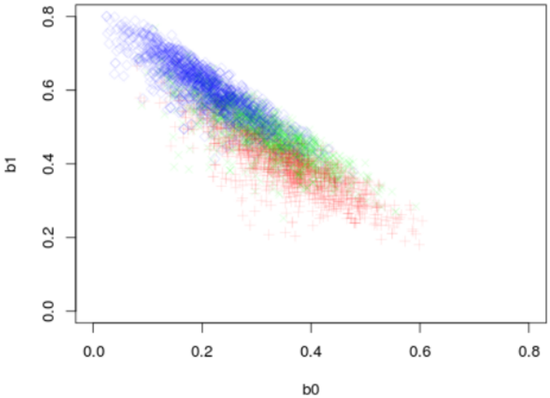


Figure 3: Proportion of (b_0, b_1) behaviors of QL_3 for $t = 1$ (red), $t = 2$ (green), $t = 3$ (blue). A slight trend to more sophisticated behaviors (decreasing b_0) is observed over time, which provides statistical evidence of agents learning the game.

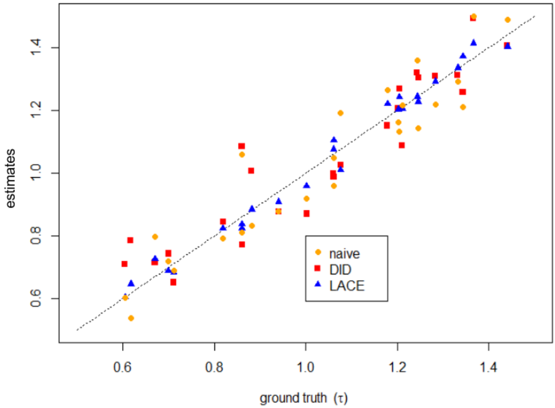


Figure 4: Random objectives. x-axis= true long-term effect. LACE=long-term effect estimation; DID=difference-in-differences; naive=use $t-1$ for estimation.

CONCLUSION

- Problem of estimating long-term causal effects of interventions in multiagent systems.
- Challenges: (1) strategic interference among competing agents; (2) temporal/adaptive actions by agents introduces short-term effects
- Contributions

(a) Explicate sufficient assumptions for identification.
(b) Provide algorithm that identifies long-term causal effects under said assumptions.
(c) Methods uses latent behavioral space; can leverage behavioral game theory to make more informed statistical predictions of counterfactuals.

- Open issues: (1) strategic interference *between* games; (2) necessary assumptions for identification.