Causal Strategic Learning under Competitive Selection

Kiet Q. H. Vo, Muneeb Aadil, Siu Lun Chau, Krikamol Muandet
CISPA Helmholtz Center for Information Security.

1. Causal Strategic Learning (CSL)

Example of causal strategic learning (CSL):

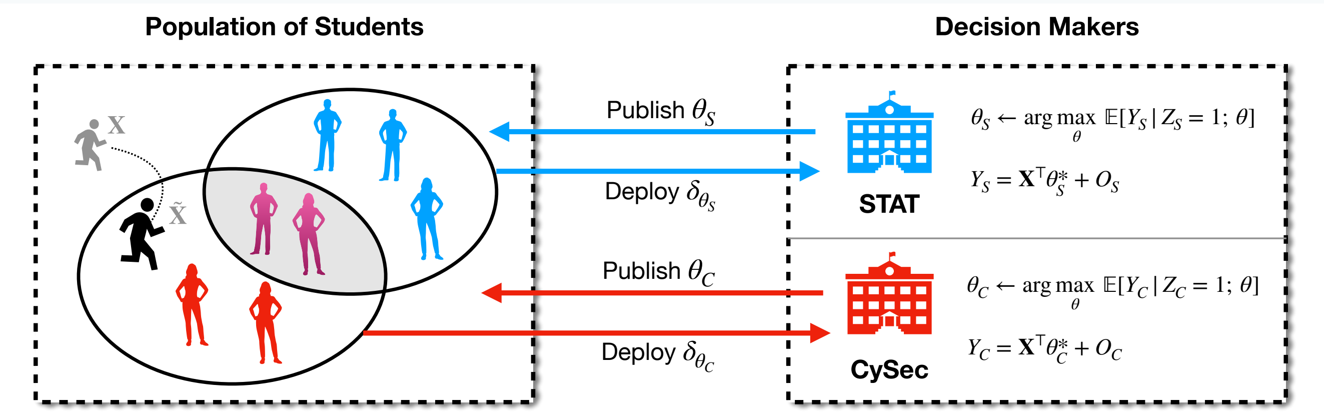
$$\text{College GPA} := \begin{bmatrix}\text{HS GPA}\\ \text{SAT}\end{bmatrix}^\top \begin{bmatrix}0.9\\ 0\end{bmatrix} + \text{Noise}$$

$$\text{Predicted GPA} := \begin{bmatrix}\text{HS GPA}\\ \text{SAT}\end{bmatrix}^\top \begin{bmatrix}0.7\\ 0.001\end{bmatrix}$$

- ***Gaming***: Changing SAT affects the prediction but not the true outcome,
- **Genuine Improvement**: Changing HS GPA also affects the true outcome.

→ Learning the true causal mechanism for better decision making,
(and there are complex interactions between decision makers and strategic agents).

2. CSL under Competitive Selection (An Example)



- 2 departments select from the same pool of students,
- They deploy their own selection rules δ_S and δ_C, parameterised by $\boldsymbol{\theta}_S$ and $\boldsymbol{\theta}_C$,
- Each wants to maximise the future GPA of their **own** students,
- Students modify $\mathbf{X}$ to $\tilde{\mathbf{X}}$ to maximise admission chance.

3. Impact of Competitive Selection

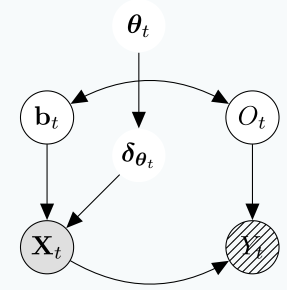


Figure 1: Prior work

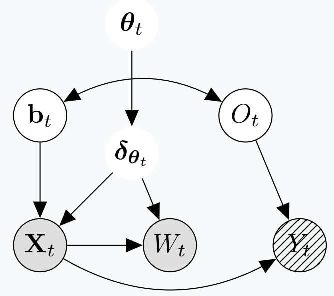


Figure 2: Ours with Selection

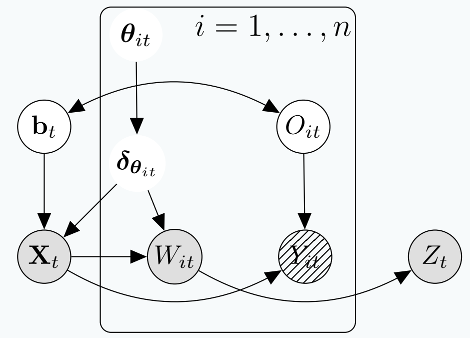


Figure 3: Ours with Competitive Selection

In Fig. (2),
- cond. dependencies, $\boldsymbol{\theta}_t \not\perp \{\mathbf{b}_t, O_t\} \mid W_t$ → harder to solve $\max_{\boldsymbol{\theta}_t} \mathbb{E}[Y_t \mid W_t = 1; \boldsymbol{\theta}_t]$;
- biased data, $\{\boldsymbol{\theta}_t, \mathbf{x}_t, \{y_t | w_t = 1\}\}_{t=1}^T$ → harder to infer causation.

In Fig. (3), $p(Y_{jt} \mid Z_t = j)$ depends on $p(Z_t = j \mid \mathbf{X}_t)$ which depends on $\{p(W_{it} \mid \mathbf{X}_t)\}_{i=1}^n$
→ interference from rival decision makers.

4. An Optimal Solution

Table 1: Our assumptions on agents' behaviour.

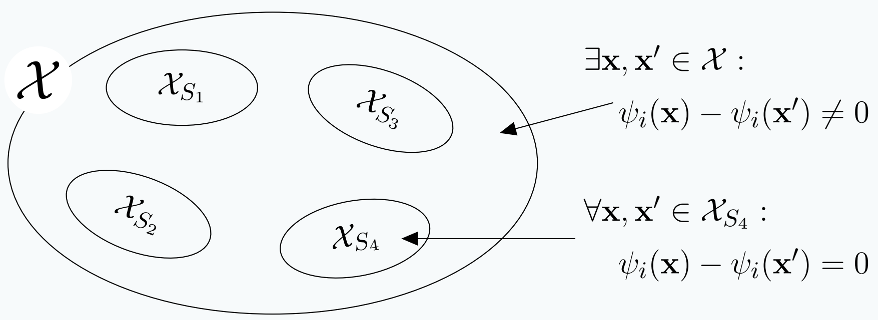
	base covariates	strategic response	preference factors	compliance model
is included	✓	✓	✓	✓
is heterogeneous	✓	✗	✗	✗

We provide:
- the optimal solution for the case where interactions result in a linear function of $\boldsymbol{\theta}_{it}$,
- conditions that can be imposed by the government to safeguard social welfare.

$$\begin{aligned}\arg\max_{\boldsymbol{\theta}_{it}} \mathbb{E}\left[Y_{it} \mid Z_t = i;\ \boldsymbol{\theta}^{\text{all}}\right] &= \arg\max_{\boldsymbol{\theta}_{it}} \left\{\mathbb{E}\left[\text{Base} \mid Z_t = i;\ \boldsymbol{\theta}^{\text{all}}\right] + \mathbb{E}\left[\text{Improvement} \mid Z_t = i; \boldsymbol{\theta}^{\text{all}}\right]\right\}\\ &= \arg\max_{\boldsymbol{\theta}_{it}} \left\{ \left(A_i^\top \boldsymbol{\theta}_{it} + B_i + h_i\left(\boldsymbol{\theta}_t^{\text{all}\setminus i}\right)\right) + \left(\boldsymbol{\theta}_{it}^\top \gamma_{it} \mathcal{E}\mathcal{E}^\top \boldsymbol{\theta}_i^* + C_i\right) \right\}\\ &= \frac{A_i + \gamma_{it}\mathcal{E}\mathcal{E}^\top \boldsymbol{\theta}_i^*}{\|A_i + \gamma_{it}\mathcal{E}\mathcal{E}^\top \boldsymbol{\theta}_i^*\|}\end{aligned}$$

Ideally,
- $A_i = k \times \mathcal{E}\mathcal{E}^\top \boldsymbol{\theta}_i^*$ for some $k > 0$,
- and more...

5. Inferring Causal Parameters



$$Y_{it} = \mathbf{X}_t^\top \boldsymbol{\theta}_i^* + \psi_i(\mathbf{X}_t) \Rightarrow \frac{dY_{it}}{d\mathbf{X}_t} = \boldsymbol{\theta}_i^* + \frac{d(\psi_i(\mathbf{X}_t))}{d\mathbf{X}_t} \quad ✗ \tag{1}$$

$$\exists \mathcal{X}_S : \Delta Y_{it} = (\Delta \mathbf{X}_t)^\top \boldsymbol{\theta}_i^* + \underbrace{\Delta\psi_i(\mathbf{X}_t)}_{=0} \Rightarrow \frac{d(\Delta Y_{it})}{d(\Delta \mathbf{X}_t)} = \boldsymbol{\theta}_i^* \quad ✓ \tag{2}$$

A cooperative protocol for all decision makers to partition $\mathbf{X}_t \mid Z_t$ correctly!

All variables here are conditioned on Z_t, but not shown here for the sake of simplicity.

6. Simulation