# Causal Interaction

Kosuke Imai

Princeton University

Joint work with Naoki Egami (University of Tokyo)

Experiments in Governance and Politics Conference

University of California, Los Angeles

February 20, 2015

# Interaction and Causal Heterogeneity

- Heterogenous treatment effects:

1. Moderation
   - How do treatment effects vary across individuals?
   - Who benefits from (or is harmed by) the treatment?
   - Interaction between treatment and pre-treatment covariates
2. Causal interaction
   - What aspects of a treatment are responsible for causal effects?
   - What combination of treatments is most efficacious?
   - Interaction between treatment variables
3. Individualized treatment regimes
   - What combination of treatments is optimal for a given individual?

- The focus of this talk: causal interaction

# Two Interpretations of Causal Interaction

1. Conditional effect interpretation:
   - Does the effect of one treatment change as we vary the value of another treatment?
   - Does the effect of being black change depending on whether an applicant is male or female?
   - Useful for testing moderation among treatments

2. Interactive effect interpretation:
   - Does a combination of treatments induce an *additional effect* beyond the sum of separate effects attributable to each treatment?
   - Does being a black female induce an additional effect beyond the effect of being black and that of being female?
   - Useful for finding efficacious treatment combinations in high dimension

# An Illustration in the $2 \times 2$ Case

- Two binary treatments: $A$ and $B$
- Potential outcomes: $Y(a, b)$ where $a, b \in \{0, 1\}$
- Conditional effect interpretation:

$$\underbrace{[Y(1,1) - Y(0,1)]}_{\text{effect of } A \text{ when } B=1} - \underbrace{[Y(1,0) - Y(0,0)]}_{\text{effect of } A \text{ when } B=0}$$

- Interactive effect interpretation:

$$\underbrace{[Y(1,1) - Y(0,0)]}_{\text{effect of } A \text{ and } B} - \underbrace{[Y(1,0) - Y(0,0)]}_{\text{effect of } A \text{ when } B=0} - \underbrace{[Y(0,1) - Y(0,0)]}_{\text{effect of } B \text{ when } A=0}$$

- The same quantity but two different interpretations
- The interactive interpretation requires the specification of the baseline condition: $(A, B) = (0, 0)$ in this example

# Causal Interaction in High Dimension

- In the $2 \times 2$ case, computing all four average potential outcomes gives a complete picture
- The dimensionality rapidly increases as the number of levels and treatments increase:
  - 3 trichotomous treatments: $3^3 = 27$
  - 4 treatments with each having 4 levels: $4^4 = 256$
- A motivating example: Conjoint analysis (Hainmueller *et al.* 2014 )
  - survey experiments to measure immigration preferences
  - a representative sample of 1,396 American adults
  - `gender`$^2$, `education`$^7$, `origin`$^{10}$, `experience`$^4$, `plan`$^4$, `language`$^4$, `profession`$^{11}$, `application reason`$^3$, `prior trips`$^5$
  - Over 1 million treatment combinations
  - What combinations of profiles characterize (un)preferred immigrants?
- We focus on the interactive interpretation in high dimension