Identifying Causal Effects in Experiments with Social Interactions and Non-compliance

Francis J. DiTraglia[1] Camilo García-Jimeno[2] Rossa O'Keeffe-O'Donovan[1]
Alejandro Sanchez[3]

[1]University of Oxford

[2]Federal Reserve Bank of Chicago

[3]University of Pennsylvania

ESWC 2020

Empirical Example with Potential for Indirect Treatment Effects

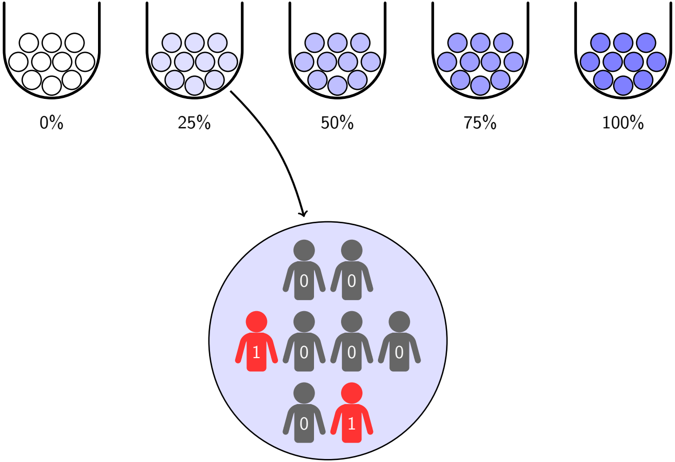
Crepon et al. (2013; QJE)

- Large-scale job-seeker assistance program in France.
- Randomized offers of intensive job placement services.

Displacement Effects of Labor Market Policies

"Job seekers who benefit from counseling may be more likely to get a job, but at the expense of other unemployed workers with whom they compete in the labor market. This may be particularly true in the short run, during which vacancies do not adjust: the unemployed who do not benefit from the program could be partially crowded out."

Studying Social Interactions Without Network Data



Partial Interference

Spillovers within but not between groups.

Randomized Saturation

Two-stage experimental design.

This Paper: Non-compliance in Randomized Saturation Experiments

Identification

Beyond ITT: Average direct/indirect causal effects under 1-sided non-compliance.

Estimation

Simple, asymptotically normal estimator under large/many-group asymptotics.

Application

French labor market experiment: Crepon et al. (2013; QJE)

Overview of Assumptions

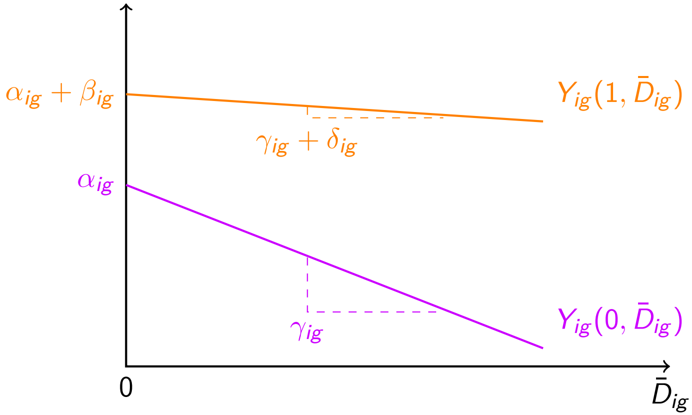
(i) Experimental Design: Randomized Saturation ✓

(ii) Exclusion Restriction for (Z_{ig}, S_g)

(iii) Treatment Take-up: $\mathbf{1}(\text{Take Treatment}) = \mathbf{1}(\text{Offered}) \times \mathbf{1}(\text{Complier})$

(iv) Potential Outcomes: Correlated Random Coefficients Model

$$Y_{ig}(D_{ig}, \bar{D}_{ig}) = \alpha_{ig} + \beta_{ig} D_{ig} + \gamma_{ig} \bar{D}_{ig} + \delta_{ig} D_{ig} \bar{D}_{ig}$$



Indirect Effects

Treated: $\gamma_{ig} + \delta_{ig}$

Untreated: γ_{ig}

Direct Effects

$\beta_{ig} + \delta_{ig} \bar{D}_{ig}$

Näive IV Does Not Identify The Effects of Interest

Unoffered Individuals

$$\begin{aligned} Y_{ig} &= \alpha_{ig} + \cancel{\beta_{ig} D_{ig}} + \gamma_{ig} \bar{D}_{ig} + \cancel{\delta_{ig} D_{ig} \bar{D}_{ig}} \\ &= \mathbb{E}[\alpha_{ig}] + \mathbb{E}[\gamma_{ig}]\bar{D}_{ig} + (\alpha_{ig} - \mathbb{E}[\alpha_{ig}]) + (\gamma_{ig} - \mathbb{E}[\gamma_{ig}])\,\bar{D}_{ig} \\ &= \alpha + \gamma \bar{D}_{ig} + \varepsilon_{ig} \end{aligned}$$

IV Estimand

$$\gamma_{IV} = \frac{\text{Cov}(Y_{ig}, S_g)}{\text{Cov}(\bar{D}_{ig}, S_g)} = \gamma + \frac{\text{Cov}(\varepsilon_{ig}, S_g)}{\text{Cov}(\bar{D}_{ig}, S_g)} = \ldots = \gamma + \frac{\text{Cov}(\gamma_{ig}, \bar{C}_{ig})}{\mathbb{E}(\bar{C}_{ig})}$$

Identification – Unoffered Individuals: $Y_{ig}(0, \bar{D}_{ig}) = \alpha_{ig} + \gamma_{ig}\bar{D}_{ig}$

Theorem

$(Z_{ig}, S_g, \bar{D}_{ig})$ are independent of the unobserved heterogeneity given $(\bar{C}_{ig}, N_g)$.

$$
\begin{aligned}
\mathbb{E}\left\{ \begin{bmatrix} 1 \\ \bar{D}_{ig} \end{bmatrix} Y_{ig} \middle| (Z_{ig} = 0, \bar{C}_{ig}, N_g) \right\} &= \mathbb{E}\left\{ \begin{bmatrix} 1 & \bar{D}_{ig} \\ \bar{D}_{ig} & \bar{D}_{ig}^2 \end{bmatrix} \begin{bmatrix} \alpha_{ig} \\ \gamma_{ig} \end{bmatrix} \middle| (Z_{ig} = 0, \bar{C}_{ig}, N_g) \right\} \\
&= \mathbb{E}\left\{ \begin{bmatrix} 1 & \bar{D}_{ig} \\ \bar{D}_{ig} & \bar{D}_{ig}^2 \end{bmatrix} \middle| (Z_{ig} = 0, \bar{C}_{ig}, N_g) \right\} \mathbb{E}\left\{ \begin{bmatrix} \alpha_{ig} \\ \gamma_{ig} \end{bmatrix} \middle| (\bar{C}_{ig}, N_g) \right\}
\end{aligned}
$$

"Localized" Average Coefficients

$$
\mathbb{E}\left\{ \begin{bmatrix} \alpha_{ig} \\ \gamma_{ig} \end{bmatrix} \middle| (\bar{C}_{ig}, N_g) \right\} = \mathbb{E}\left\{ \begin{bmatrix} 1 & \bar{D}_{ig} \\ \bar{D}_{ig} & \bar{D}_{ig}^2 \end{bmatrix} \middle| (Z_{ig} = 0, \bar{C}_{ig}, N_g) \right\}^{-1} \mathbb{E}\left\{ \begin{bmatrix} 1 \\ \bar{D}_{ig} \end{bmatrix} Y_{ig} \middle| (Z_{ig} = 0, \bar{C}_{ig}, N_g) \right\}
$$

Identification – Unoffered Individuals: $Y_{ig}(0, \bar{D}_{ig}) = \alpha_{ig} + \gamma_{ig}\bar{D}_{ig}$

Iterated Expectations over $(\bar{C}_{ig}, N_g)$

$$\mathbb{E}\begin{bmatrix}\alpha_{ig}\\ \gamma_{ig}\end{bmatrix} = \mathbb{E}\left\{\left[\mathbf{Q}_0(\bar{C}_{ig}, N_g)\right]^{-1}\begin{bmatrix}1\\ \bar{D}_{ig}\end{bmatrix} Y_{ig}\middle| Z_{ig} = 0\right\}, \quad \mathbf{Q}_0(\bar{C}_{ig}, N_g) \equiv \mathbb{E}\left\{\begin{bmatrix}1 & \bar{D}_{ig}\\ \bar{D}_{ig} & \bar{D}_{ig}^2\end{bmatrix}\middle|(Z_{ig} = 0, \bar{C}_{ig}, N_g)\right\}$$

$\mathbf{Q}_0$ is *Known*

Distribution of $\bar{D}_{ig}|(\bar{C}_{ig}, N_g)$ determined by experimental design.

Feasible Estimation

Need $\widehat{C}_{ig} \to_p \bar{C}_{ig}$. Hence: large/many–group asymptotics.

We Identify these Average Causal Effects:

Spillover

$\bar{D}_{ig} \rightarrow Y_{ig}$ for the population, holding $D_{ig} = 0$.

Treatment on the Treated

$D_{ig} \rightarrow Y_{ig}$ for compliers as a function of $\bar{D}_{ig}$.

Spillover on the Treated

$\bar{D}_{ig} \rightarrow Y_{ig}$ for compliers holding $D_{ig} = 0$ or $D_{ig} = 1$.

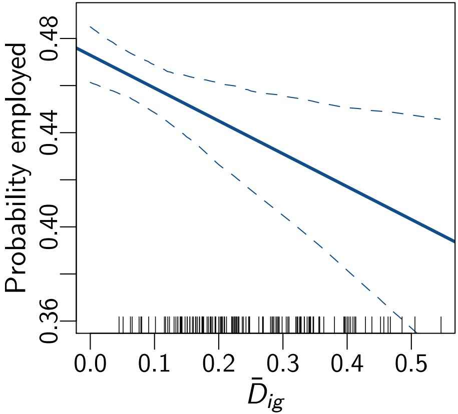
Spillover on the Untreated

$\bar{D}_{ig} \rightarrow Y_{ig}$ for never-takers holding $D_{ig} = 0$.

Average Spillover to Long-term Employment: $Y_{ig}(0, \bar{D}_{ig}) = \alpha_{ig} + \gamma_{ig}\bar{D}_{ig}$

Data from Crepon et al. (2013; QJE)

	$\mathbb{E}(\alpha_{ig})$	$\mathbb{E}(\gamma_{ig})$
Our estimator	0.47	-0.14
	(0.01)	(0.07)
Naïve IV	0.47	-0.06
	(0.01)	(0.06)



Conclusion

Identification

Go beyond ITTs to identify average direct and indirect effects in randomized saturation experiments with 1-sided non-compliance.

Estimation

Simple asymptotically normal estimator under large/many-group asymptotics.

Application

Substantial labor market spillovers in Crepon et al. (2013; QJE) experiment.

Appendix – Notation

Sample Size and Indexing

- Groups: $g = 1, \ldots, G$
- Individuals in g: $i = 1, \ldots, N_g$

Observables

- Z_{ig} = binary treatment offer to (i, g)
- D_{ig} = binary treatment take-up of (i, g)
- Y_{ig} = outcome of (i, g)
- S_g = saturation of group g
- $\bar{D}_{ig}$ = take-up fraction in g excluding (i, g)

Appendix – Correlated Random Coefficients Model

$$Y_{ig}(\boldsymbol{D}) = Y_{ig}(\boldsymbol{D}_g) = Y_{ig}(D_{ig}, \bar{D}_{ig}) = \mathbf{f}(\bar{D}_{ig})' \left[(1 - D_{ig})\boldsymbol{\theta}_{ig} + D_{ig}\boldsymbol{\psi}_{ig}\right]$$

- $\mathbf{f}$ is a vector of known functions, bounded on $[0, 1]$
- $\boldsymbol{\theta}_{ig}$ and $\boldsymbol{\psi}_{ig}$ are RVs that may be dependent on $(D_{ig}, \bar{D}_{ig})$.

This Talk

$$Y_{ig}(D_{ig}, \bar{D}_{ig}) = \alpha_{ig} + \beta_{ig} D_{ig} + \gamma_{ig} \bar{D}_{ig} + \delta_{ig} D_{ig} \bar{D}_{ig}$$

Appendix – Assumptions on Treatment Take-up

One-sided Non-compliance

Only those offered treatment can take it up.

Individualistic Offer Response (IOR)

$$D_{ig}(\boldsymbol{Z}) = D_{ig}(\boldsymbol{Z}_g) = D_{ig}(Z_{ig}, \bar{Z}_{ig}) = D_{ig}(Z_{ig})$$

Notation

$C_{ig} = 1$ iff (i, g) is a complier; $\bar{C}_{ig} \equiv$ share of compliers among (i, g)'s neighbors.

$$\boxed{\text{(IOR)} + \text{(1-Sided)} \Rightarrow D_{ig} = C_{ig} \times Z_{ig}}$$

Appendix – Exclusion Restriction

Notation

- $\mathbf{B}_g$ = random coefficients for everyone in group g.
- $\boldsymbol{C}_g$ = complier indicators for everyone in group g
- $\boldsymbol{Z}_g$ = treatment offers for everyone in group g

Exclusion Restriction

(i) $S_g \perp\!\!\!\perp (\boldsymbol{C}_g, \mathbf{B}_g, N_g)$

(ii) $\boldsymbol{Z}_g \perp\!\!\!\perp (\boldsymbol{C}_g, \mathbf{B}_g) | (S_g, N_g)$

Theorem 1

Under the randomized saturation design, IOR assumption, and exclusion restriction, $(Z_{ig}, \bar{D}_{ig}, S_g) \perp\!\!\!\perp (\mathbf{B}_{ig}, C_{ig}) | (\bar{C}_{ig}, N_g)$.

Appendix - Testable Implications of IOR in Crepon et al. (2013; QJE)

Recall: IOR Assumption

Person (i, g)'s take-up D_{ig} depends only on her *own* treatment offer Z_{ig}.

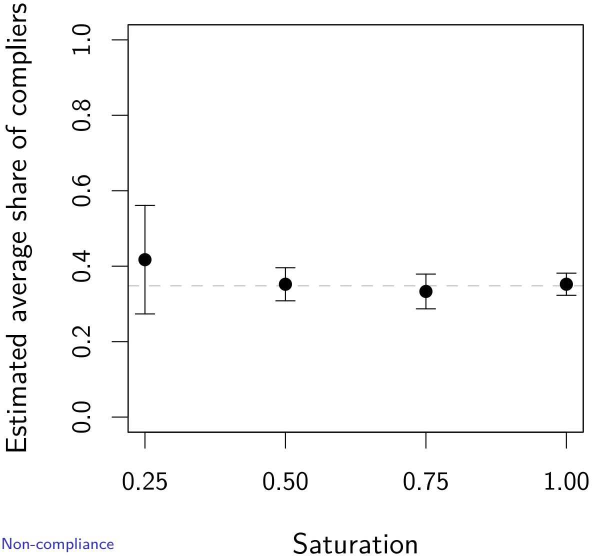
A "Regression-based" Test

$\mathbb{E}[D_{ig}|Z_{ig} = 1, S_g = s]$ should not depend on s.

Implementation

- Regress D_{ig} on an intercept and saturation dummies for those offered treatment.
- Test the joint null that the coefficients on all saturation dummies are zero.
- (Equivalently test that share of compliers is the same across saturation "bins")
- Crepon Experiment: p-value of 0.62

Appendix – Testable Implications of IOR



Appendix – Identification: Rank Condition

$Y_{ig}(D_{ig}, \bar{D}_{ig}) = \alpha_{ig} + \beta_{ig} D_{ig} + \gamma_{ig} \bar{D}_{ig} + \delta_{ig} D_{ig} \bar{D}_{ig}$

Notation

$$\mathbf{Q}_z(\bar{c}, n) \equiv \mathbb{E}\left\{ \left. \begin{bmatrix} 1 & \bar{D}_{ig} \\ \bar{D}_{ig} & \bar{D}_{ig}^2 \end{bmatrix} \right| (Z_{ig} = z, \bar{C}_{ig} = \bar{c}, N_g = n) \right\}$$

Rank Condition

$\mathbf{Q}_0(\bar{c}, n)$ and $\mathbf{Q}_1(\bar{c}, n)$ are invertible for all $(\bar{c}, n)$ in the support of $(\bar{C}_{ig}, N_g)$.

Bernoulli Offers

$$\mathbf{Q}_0(\bar{c}, n) = \frac{1}{\mathbb{E}(1 - S_g)} \begin{bmatrix} \mathbb{E}\{1 - S_g\} & \bar{c}\, \mathbb{E}\{S_g(1 - S_g)\} \\ \bar{c}\, \mathbb{E}\{S_g(1 - S_g)\} & \bar{c}^2\, \mathbb{E}\left\{S_g^2(1 - S_g)\right\} + \frac{\bar{c}}{n-1} \mathbb{E}\left\{S_g(1 - S_g)^2\right\} \end{bmatrix}$$