# Design of Two-Stage Experiments

## with an Application to Spillovers in Tax Compliance

Guillermo Cruces, *U. of Nottingham & CEDLAS-UNLP*
Dario Tortarolo, *U. of Nottingham & IFS*
Gonzalo Vazquez-Bare, *UC Santa Barbara*
Julian Amendolaggine, *CEDLAS-UNLP*
Juan Luis Schiavoni, *CEDLAS-UNLP*

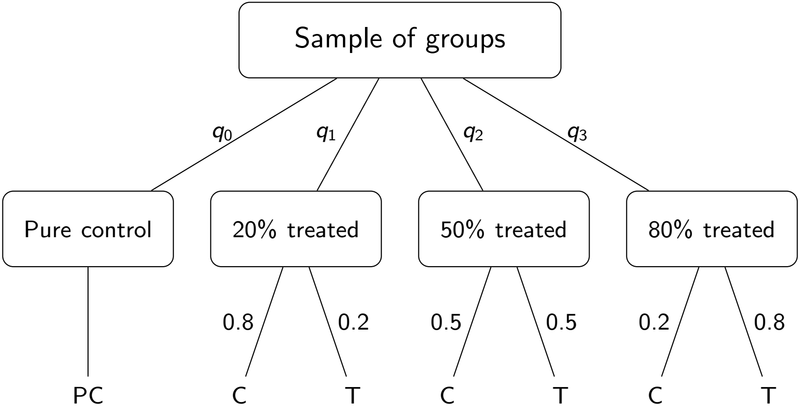
UCSB Applied Micro Economics Lunch

February 11, 2022

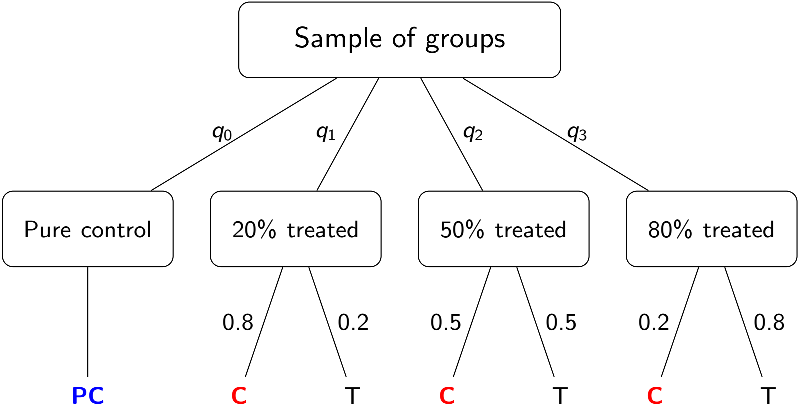
# Design of Partial Population Experiments

- Goal: estimate within-group spillovers
  - Households in villages
  - Employees in firms
  - Students in schools
- Two-step design:
  - Groups randomly divided into treatment "intensities" (saturations)
  - Units within each group randomly assigned to treatment and control
- Compare units across groups with different treatment intensities

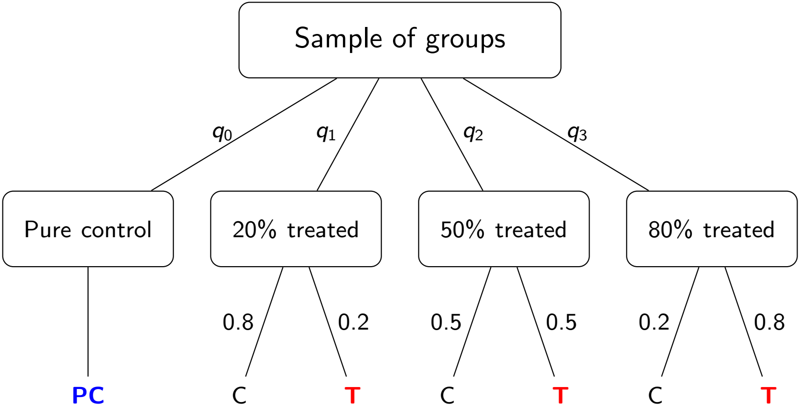
# Experimental design: example



# Experimental design: example



# Experimental design: example



# Designing PP Experiments

- Key choices:
  - Number of saturations and within-group probabilities
  - Probability of each saturation $q_0, q_1, q_2, \ldots$ (**this talk**)
  - Within-group assignment mechanism (**this talk**)
- Key inputs:
  - Parameters (outcome variances, intracluster correlations,...)
  - Variance of estimators (**this talk**)
  - Power function to calculate power, MDE (**this talk**)

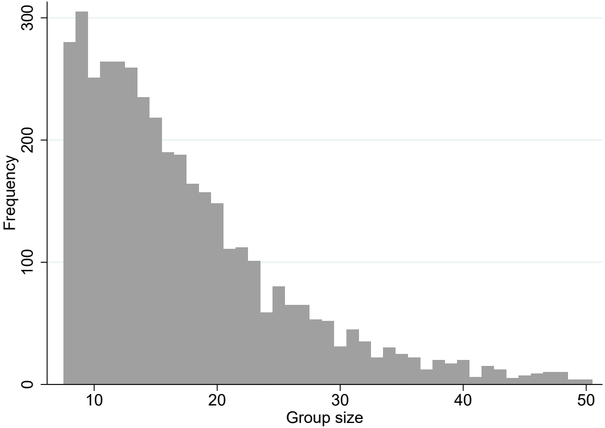
# Challenges for Designing PP Experiments

- Two-stage design
- Multiple treatments
  - Compare units exposed to different saturations
- Within-group correlations (clustering)
- Heterogeneity in group sizes
  - Group sizes tend to vary widely in practice

# Existing tools for designing PP Experiments

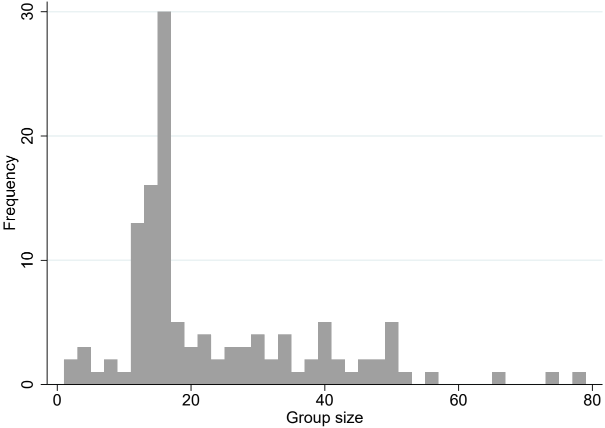
- Hirano and Hahn (2010), Baird et al (2018)
  - Homoskedasticity, random effects structure
  - Ignore group size heterogeneity
- Software (e.g. `Stata`'s `power` command) makes restrictive assumptions about group size distribution
  - Equally-sized groups, $N_T$ proportional to $N_C$,...

# Cruces et al (2022)



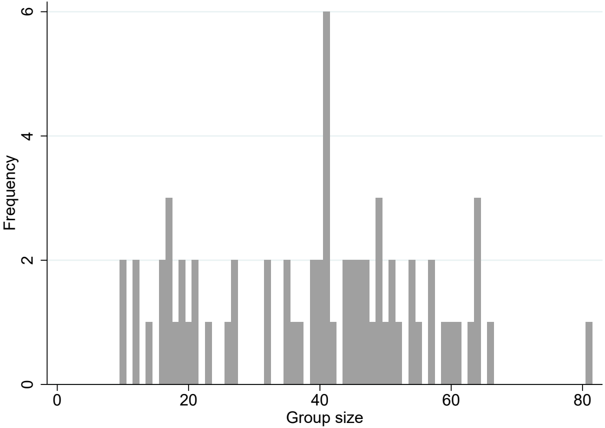

Distribution of group sizes

# Haushofer and Shapiro (2016, QJE)



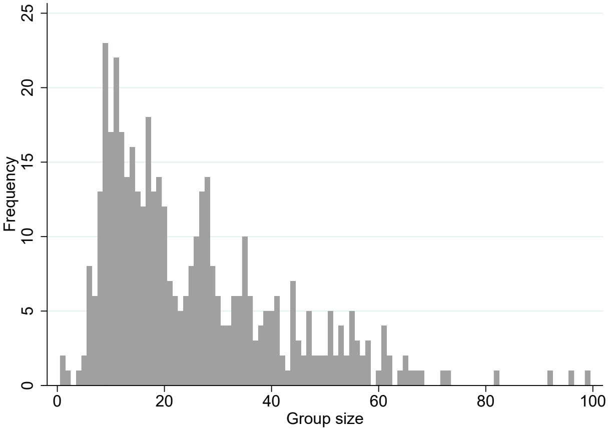

Distribution of group sizes

# Giné and Mansuri (2018, AEJ Applied)




Distribution of group sizes

# Imai et al (2020, JASA)




Distribution of group sizes

# Why is group size heterogeneity important?

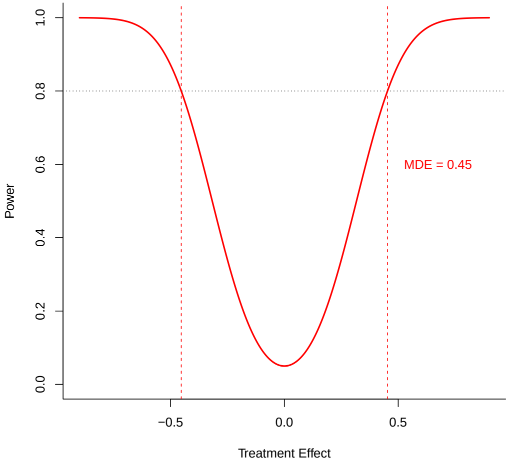
- It affects the variance of estimators

$$\mathbb{V}[\hat{\beta}] \approx \sigma^2 \left[1 + \rho(ICC, \bar{n}, Var(n_g))\right]$$

  - Ignoring $Var(n_g)$ underestimates $\mathbb{V}[\hat{\beta}]$ $\rightarrow$ overestimates power

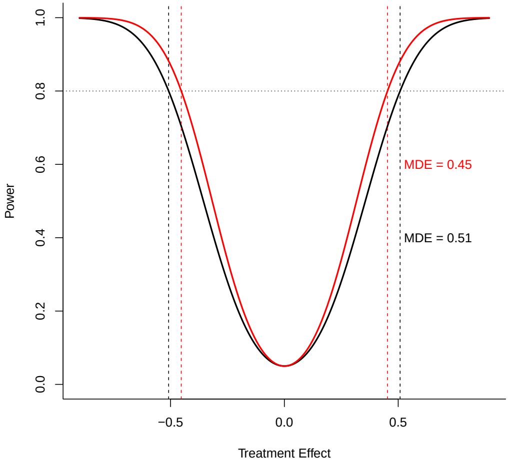
- It affects inference and power calculations
  - Normal approx may be inaccurate if groups are "too heterogeneous"
  - Carter et al (2017), Djogbenou et al (2019), Hansen and Lee (2019)

# Why is group size heterogeneity important?



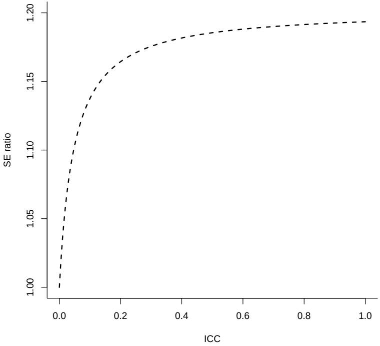

$G = 95,\ \bar{n} = 23.3,\ sd(n_g) = 0,\ \sigma^2_Y = 1,\ ICC = 0.2$

# Why is group size heterogeneity important?

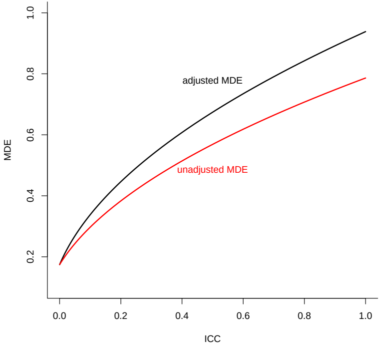



$$G = 95,\ \bar{n} = 23.3,\ sd(n_g) = 15.2,\ \sigma_Y^2 = 1,\ ICC = 0.2$$

# Why is group size heterogeneity important?




(a) Adjusted SE / Unadjusted SE




(b) Adjusted and unadjusted MDEs

# This paper

- We derive asymptotic variance approximations allowing for:
  - Multiple treatments
  - General intracluster correlation and heteroskedasticity
  - Group size heterogeneity
  - Varying probabilities across groups
- Calculate power and MDEs
- Our formulas can be applied in a wide range of designs
  - Two-stage, PP, clustered, stratified experiments...
- We conduct a field experiment on tax compliance in Argentina

# Setup

- Random sample of groups $g = 1, \ldots, G$ with units $i = 1, \ldots, n_g$
- Total sample size $n = \sum_g n_g$
- First stage: randomly divide groups into categories:

$$T_g \in \{0, 1, 2, \ldots, M\}, \quad \mathbb{P}[T_g] = q_t$$

- Within each group, assign binary individual-level treatment:

$$D_{ig} \in \{0, 1\}, \quad \mathbb{P}_g[D_{ig} = d | T_g = t] = p_g(d, t)$$

# Setup

- Estimands:

$$\beta_{dt} = \mathbb{E}[Y_{ig}|D_{ig} = d, T_g = t] - \mathbb{E}[Y_{ig}|T_g = 0]$$

  - Direct effects $= \beta_{1t}$
  - Spillover effects $= \beta_{0t}$

- Second moments:

$$\sigma^2_{dt} = \mathbb{V}[Y_{ig}|D_{ig} = d, T_g = t]$$
$$\rho_{dt} = cor(Y_{ig}, Y_{jg}|D_{ig} = d, D_{jg} = d, T_g = t)$$

# Setup

- Estimation strategy:

$$Y_{ig} = \alpha + \sum_{t=1}^{M} \beta_{0t}(1 - D_{ig})\mathbb{1}(T_g = t) + \sum_{t=1}^{M} \beta_{1t} D_{ig} \mathbb{1}(T_g = t) + \varepsilon_{ig}$$

- Equivalent to:

$$\hat{\beta}_{dt} = \bar{Y}_{dt} - \bar{Y}_{00}$$

- Allow for correlated errors within groups

# Main Result

## Asymptotic Approximation

Under regularity conditions, if

$$\max_{g \leq G} \frac{n_g^2}{n} \to 0, \quad \frac{\sum_{g=1}^{G} n_g^4}{n^2} \leq C < \infty,$$

then:

$$\hat{\beta}_{dt} \overset{a}{\sim} \mathcal{N}(\beta_{dt}, V_{dt})$$

where:

$$V_{dt} = \frac{\sigma_{dt}^2}{q_t \sum_g n_g p_g(d,t)} \left\{ 1 + \rho_{dt} \frac{\sum_g n_g(n_g - 1)\mathbb{P}_g[D_{ig} = d, D_{jg} = d | T_g = t]}{\sum_g n_g p_g(d,t)} \right\}$$
$$+ \frac{\sigma_{00}^2}{q_0 n} \left\{ 1 + \rho_{00} \left( \frac{\sum_g n_g^2}{n} - 1 \right) \right\}$$

# Main result: intuition

- Variance: $\mathbb{V}[\hat{\beta}_{dt}] = \mathbb{V}[\bar{Y}_{dt}] + \mathbb{V}[\bar{Y}_{00}]$ allowing for:
  - Heteroskedasticity: $\sigma^2_{dt} \neq \sigma^2_{d't'}$
  - Intracluster correlation: $\rho_{dt} \neq 0$
  - Unequal probabilities between groups: $p_g(d,t) \neq p_{g'}(d,t)$
  - Group size heterogeneity: $Var(n_g) \neq 0$

# Main Result: Intuition

- Condition:
$$\max_{g \leq G} \frac{n_g^2}{n} \to 0$$
restricts the relative size of the largest group
  - Ensures that no group "dominates" the sample

- Condition:
$$\frac{\sum_{g=1}^{G} n_g^4}{n^2} \leq C < \infty$$
bounds the fourth moment of the distribution
  - Rules out fat tails (outliers)

# Power and MDE calculations

- Based on the normal approximation, the power function is

$$\Gamma(\beta_{dt}) \approx 1 - \Phi\left(\frac{\beta_{dt}}{\sqrt{V_{dt}}} + z_{1-\alpha/2}\right) + \Phi\left(\frac{\beta_{dt}}{\sqrt{V_{dt}}} - z_{1-\alpha/2}\right)$$

- Depends on:
  - Treatment effect $\beta_{dt}$
  - Group sizes $\{n_g\}_{g=1}^{G}$ and total sample size $n$
  - Assig mech: $\{q_t\}_t$, $\{p_g(d,t)\}_{t,g}$, $\{\mathbb{P}_g[D_{ig} = d, D_{jg} = d | T_g = t]\}_{t,g}$
  - Outcome moments $\{\sigma^2_{dt},\ \rho_{dt}\}_t$

# Choice of $\{q_t\}_t$

- Optimal choice requires defining an optimality criterion
  - How to combine variances of multiple estimators
- Optimal design literature has proposed several alternatives
- We discuss two scenarios:
  - Unconstrained designs: minimize the average of all estimator variances (*A-optimality*)
  - Constrained designs

# Choice of $\{q_t\}_t$: unconstrained optimization

## A-optimal design

The solution to the optimal design problem:

$$\min_{q_0, q_1, \ldots, q_M} \sum_{t=1}^{M} \left\{ \mathbb{V}[\hat{\beta}_{0t}] + \mathbb{V}[\hat{\beta}_{1t}] \right\}, \quad q_t > 0, \quad \sum_{t=0}^{M} q_t = 1$$

is:

$$q_0^* = \frac{\sqrt{2MB_0}}{\sqrt{2MB_0} + \sum_{t>0} \sqrt{B_t}}, \quad q_t^* = \frac{\sqrt{B_t}}{\sqrt{2MB_0} + \sum_{t>0} \sqrt{B_t}}, \quad t > 0,$$

where $\{B_t\}_t$ are constants depending on $\{n_g\}_g$, $\{p_g(d,t)\}_{d,t,g}$ and $\{\mathbb{P}_g[D_{ig} = d, D_{jg} = d | T_g = t]\}_{t,g}$, $\{\sigma^2_{dt}, \rho_{dt}\}_t$

# Choice of $\{q_t\}_t$: incorporating constraints

- Researchers may need to incorporate constraints in choice of $q_t$
  - Logistical, administrative, etc
- We provide an example in our field experiment
  - "Minimax-like" approach with fixed number of treated

# Within-group treatment assignment

- We want to assign exactly $n_g p_t$ units to treatment
- But $n_g p_t$ may not be an integer (e.g. $p_t = 0.5$, $n_g = 11$)
- Let $\xi_g \in \{0, 1\}$ be a random adjustment factor and let

$$N_g^1 = \lfloor n_g p_t \rfloor + \xi_g \mathbb{1}(n_g p_t \notin \mathbb{N})$$

be the (random) number of treated in group $g$ with $T_g = t$

- Setting $\mathbb{P}_g[\xi_g = 1 | T_g = t] = (n_g p_t - \lfloor n_g p_t \rfloor)\mathbb{1}(n_g p_t \notin \mathbb{N})$ gives:

$$\mathbb{E}[N_g^1 | T_g = t] = n_g p_t, \quad \mathbb{P}_g[D_{ig} = 1 | T_g = t] = p_t$$

# Direct and spillover effects in tax compliance

- We teamed up with a large municipality in Greater Buenos Aires
- Neighbors are required to pay a monthly bill on their real estate
- Information campaign with personalized letters
  - One-page letter informing of new electronic billing option
  - Instructions on how to sign up and pay online
  - Information on current billing period and past due debt
- Are there spillovers between neighbors from the same block?

# Example of the intervention letter

**Tus impuestos municipales ahora vienen en la BOLETA DIGITAL**

3F MUNICIPALIDAD

ID: XXXXX

TITULAR:

DIRECCIÓN: CAP. MADARIAGA N°

LOCALIDAD: 11 de Septiembre

C.P.: 1657

PARTIDA: XXXXXX/7

Te queremos contar que ahora en Tres de Febrero tu boleta municipal de la Tasa por Servicios Generales (TSG) es 100% digital. O sea, ya no se usa más el papel. Podés acceder a ella y pagarla desde el celular o la computadora. De esta manera, nos cuidamos entre todos al reducir la circulación y también cuidamos el medio ambiente. Es una situación difícil y te agradecemos el esfuerzo que estás haciendo para estar al día con tus impuestos, porque eso se transforma directamente en obras y servicios que no paran en tu barrio. Te informamos el estado de tu cuenta y te mostramos lo fácil que es:

| PARTIDA: XXXXX/7 | |
|---|---|
| Cuota 10 vencimiento 10 de octubre 2020: | 347,29 |
| Deuda año en curso*: | 1.702,58 |
| Deuda años anteriores*: | 289,54 |

* Al 15/09/2020

**¿CÓMO PAGAR?**

Ingresando a **tasas.tresdefebrero.gov.ar** completá los datos:

**DESCARGÁ O PAGÁ TU BOLETA**

TASA O IMPUESTO A PAGAR

Servicios generales

IDENTIFICADOR

Ingresar / Ingresar

BUSCAR BOLETA

RECIBIR LA BOLETA POR MAIL

CLICKEÁ ESTE BOTÓN y recibila todos los meses en tu mail.

También podés entrar a miboleta.tresdefebrero.gov.ar

**1)** Podés pagar **ONLINE** con

mercado pago → En el momento desde nuestra web.

PagoMisCuentas, LINK → Obteniendo el **código de pago electrónico** para pagar desde **la plataforma de tu banco o cajero automático.**

**2)** Podés pagar en **EFECTIVO** en

Pago Fácil, rapipago → **DESCARGALA** o llevá tu **NÚMERO DE PARTIDA.**

Por dudas comunicate con nosotros a **reclamos.mistasas@tresdefebrero.gov.ar**
Si esta carta llegó por error a tu domicilio, informanos en ese mismo correo electrónico

**¡Muchas gracias!**

# PP Experiment: Design

- We randomly divide blocks into four categories:
  - $T_g = 0$: pure controls with prob $q_0$
  - $T_g = 1$: 20% treated with prob $q_1$
  - $T_g = 2$: 50% treated with prob $q_2$
  - $T_g = 3$: 80% treated with prob $q_3$
- We set up a system of eqs to incorporate constraints on $\{q_t\}_t$

# Constrained choice of $\{q_t\}_t$

- Choose $q_1, q_2, q_3$, with $q_0 = 1 - q_1 - q_2 - q_3$
- The total number of letters sent ($L$) should equal the expected number of treated:

$$L = n(0.2q_1 + 0.5q_2 + 0.8q_3)$$

- Categories $T_g = 1$ and $T_g = 3$ are symmetric, so $q_1 = q_3$
- This leaves two probabilities to be determined: $q_2$ and $q_3$
- Idea: balance variances across assignments

# Constrained choice of $\{q_t\}_t$

- The "hardest" effects (smallest cells) to estimate are $\beta_{03}$ and $\beta_{11}$
  - Spillover effect in 80% groups and direct effect in 20% groups
- We choose $q_2$ and $q_3$ by setting:

$$\mathbb{V}[\hat{\beta}_{03}] = \mathbb{V}[\hat{\beta}_{02}]$$

  based on our variance approximation
- We assume $\sigma^2 = 0.25$ (upper bound for binary outcomes) and $\rho \approx 0.1$ (based on baseline data)

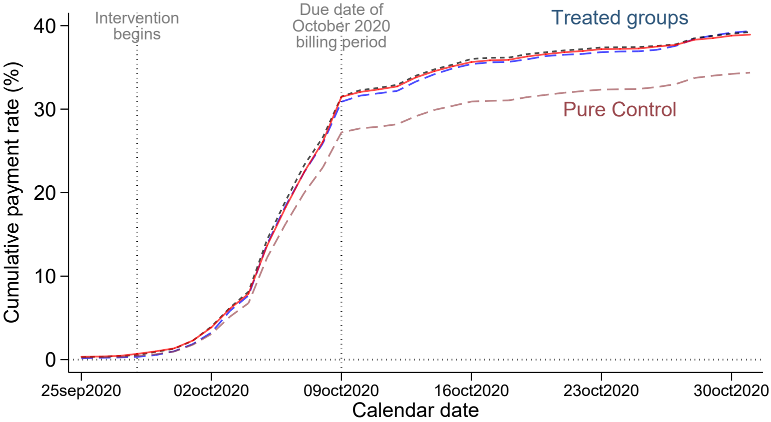
# Sample sizes

| | Blocks | Control Obs | Treated Obs |
|---|---|---|---|
| $T_g = 0$ | $1,102$ | $19,105$ | $0$ |
| $T_g = 1$ | $1,100$ | $15,049$ | $\mathbf{3,864}$ |
| $T_g = 2$ | $\mathbf{680}$ | $\mathbf{5,898}$ | $\mathbf{5,904}$ |
| $T_g = 3$ | $1,100$ | $\mathbf{3,707}$ | $15,281$ |
| Total | $3,982$ | $43,759$ | $25,049$ |

MDEs range from 2.6 to 3.3 p.p.

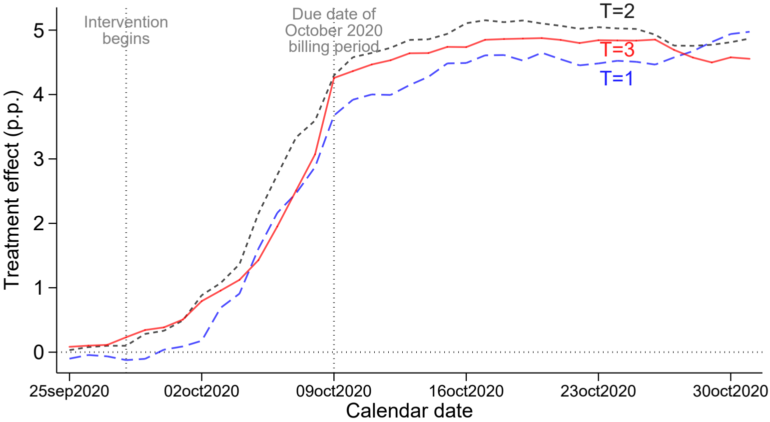
# Treated groups: Payment rate (Oct'20 bill)

Figure: Payment rates in levels



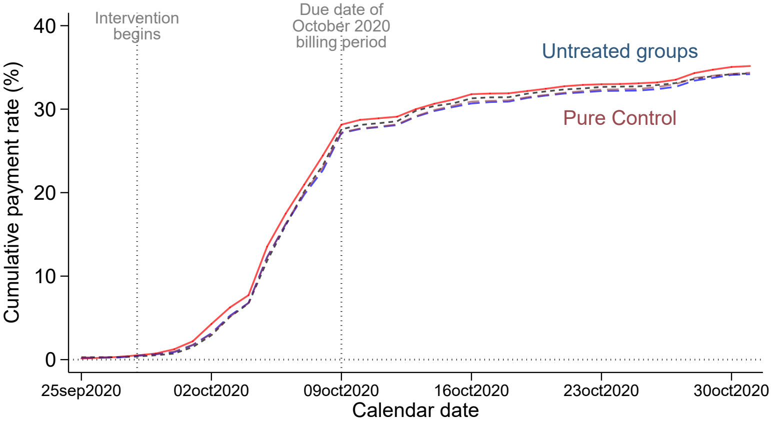
# Treated groups: Payment rate (Oct'20 bill)

Figure: Difference relative to pure control group



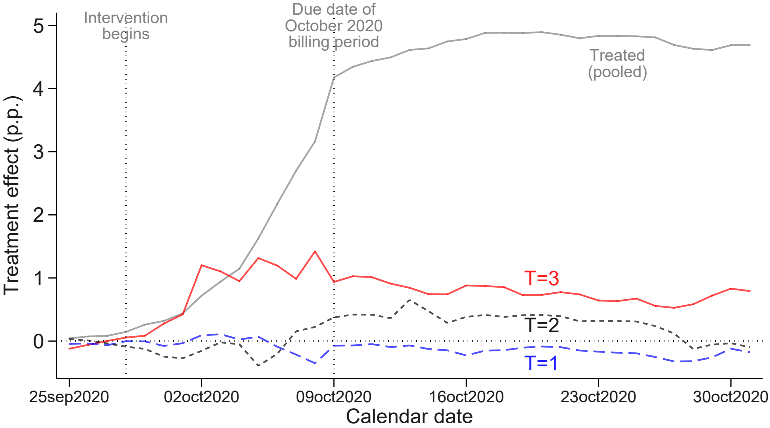
# Untreated groups: Payment rate (Oct'20 bill)

Figure: Payment rates in levels

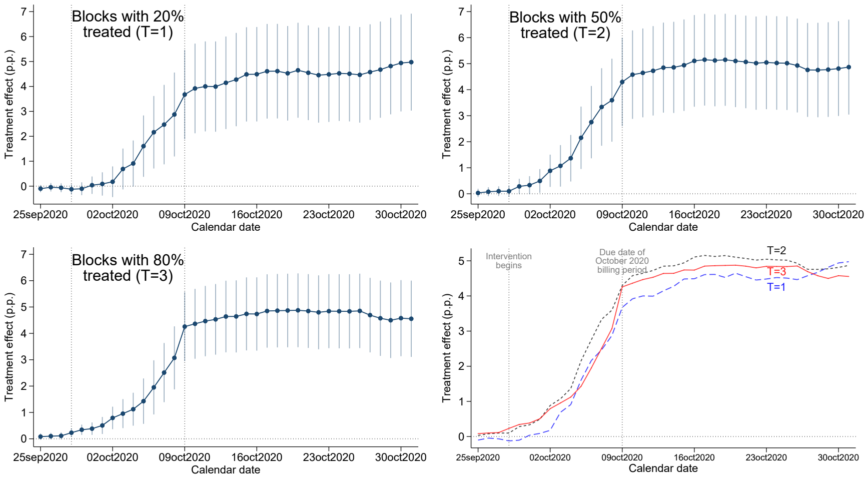


# Untreated groups: Payment rate (Oct'20 bill)

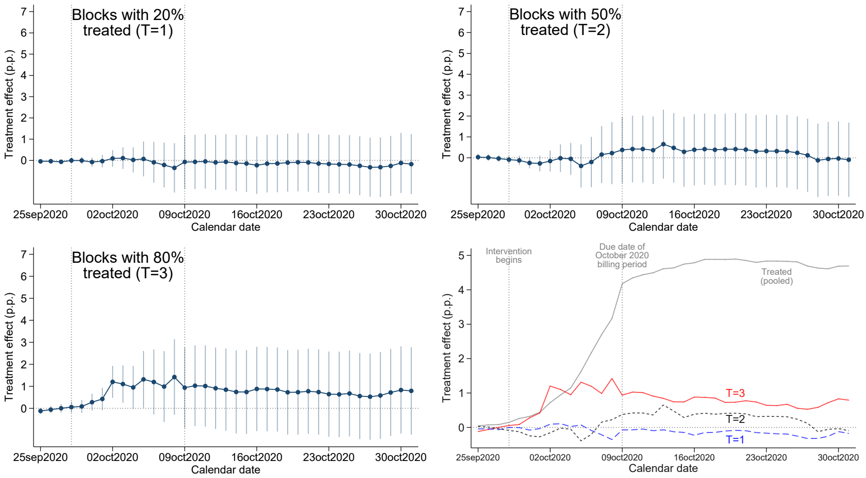
Figure: Difference relative to pure control group



# Diff treated/pure controls (Oct'20 bill)



# Diff. untreated/pure controls (Oct'20 bill)



# Summary

- Framework to calculate power and MDE in PP experiments
  - Allow for group size heterogeneity, heteroskedasticity, ICC,...
  - Derive optimal choice of group-level probabilities
- Application to tax compliance in Argentina
  - Strong and significant direct effects of the letters
  - No clear evidence of reinforcement effects between treated
  - Some evidence of within-neighbor spillovers in highest saturation

Thank you!