GR 6307
Public Economics and Development

2. Anti-Poverty Programs: Reaching the Poor

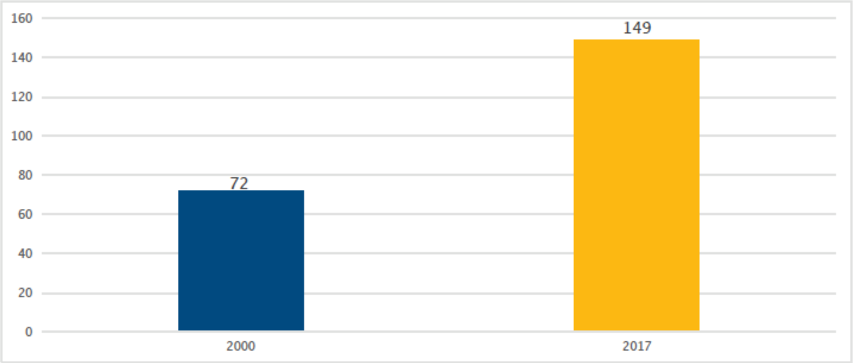
Michael Best

Fall 2019

Outline

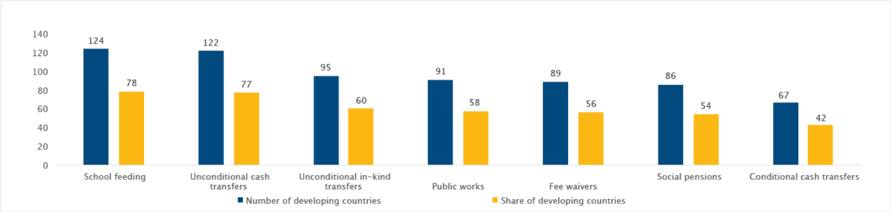
Trends in Social Programs over time and across countries

Figure 1. Number of developing countries with SSN programs



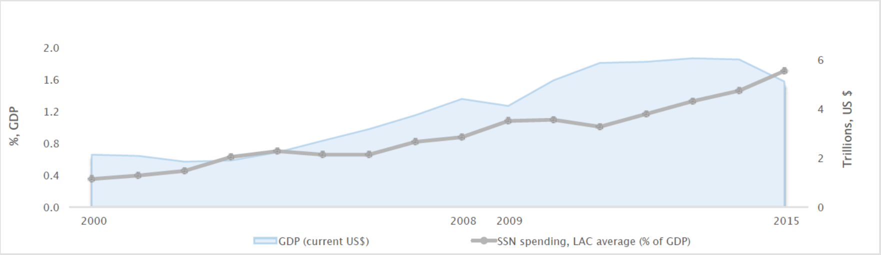
Trends in Social Programs over time and across countries

Figure 2. Number and share of developing countries with SSN instrument



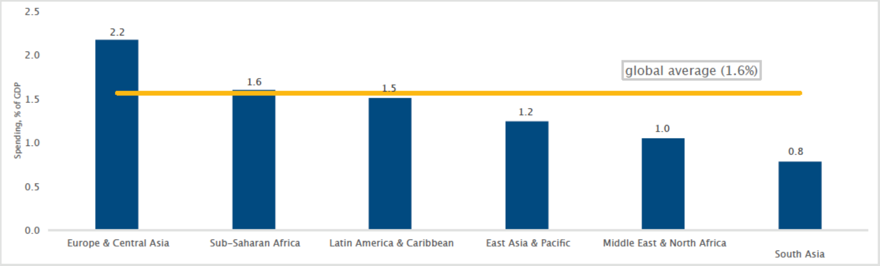
Trends in Social Programs over time and across countries

Figure 3. Spending on SSN programs in LAC over time, % of GDP



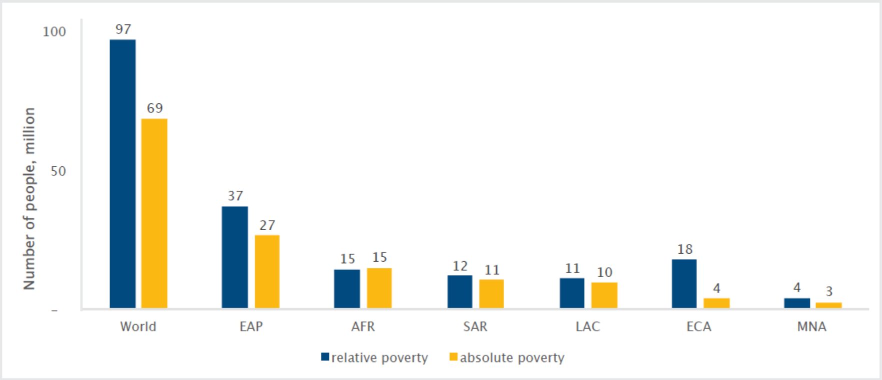
Trends in Social Programs over time and across countries

Figure 4. Spending on SSN programs across the regions, % of GDP



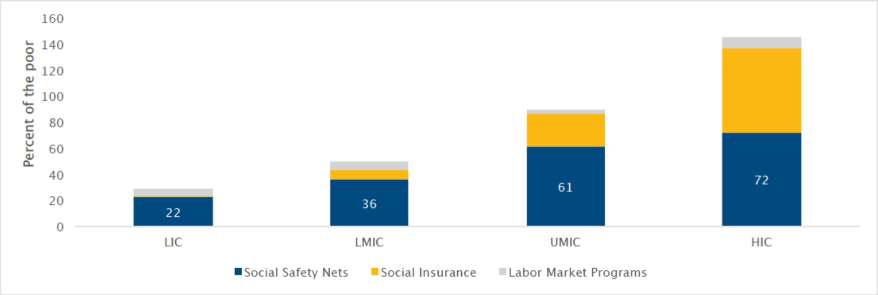
Trends in Social Programs over time and across countries

Figure 6. Estimated number of people escaping poverty because of SSN, millions



Trends in Social Programs over time and across countries

Figure 7. Coverage of the poor (bottom 20%) by SSN programs



Outline

Outline

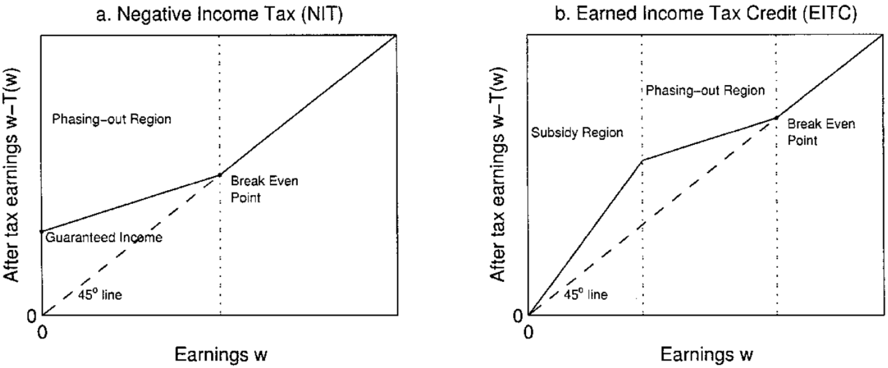
Theory

- Saez (QJE 2002) *Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses*
- Akerlof (AER 1978) *The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning*
- Nichols & Zeckhauser (AER 1982) *Targeting Transfers Through Restrictions on Recipients*
- Besley & Coate (AER 1992) *Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs*
- Niehaus, Atanassova, Bertrand & Mullainathan (AEJ:Pol 2013) *Targeting with Agents*

Saez (2002): Overview

- Most rich countries provide *lots* of income support programs
- 2 Key margins along which people can respond:
  1. *Intensive* margin: Work, but earn less pre-tax/transfer income
  2. *Extensive* margin: Stop working
- This paper takes an optimal-tax approach to the design of income support programs.
- Key results
  1. Simple theory with both intensive- and extensive- margins
  2. Apply to 2 types of policies: Negative income taxes and Earned Income Tax Credit
  3. Conditions for EITC $\succ$ NIT: Extensive margin responses stronger than intensive-
  4. Calibration to US suggests EITC $\succ$ NIT

Saez (2002): NIT vs EITC



Saez (2002): Model Setup

- There are $I + 1$ occupations. Occupation 0 is unemployment.
- Salaries: $w_0 = 0$, $w_1 < \ldots < w_i < \ldots w_I$
- Government observes income and charges net taxes $T_i \rightarrow c_i = w_i - T_i$
- Total population normalized to one. $h_i =$proportion of individuals in occupation i. $\sum_{i=0}^{I} h_i = 1$
- Individuals choose which occupation to work in. In principle, depending on consumption in every occupation:

$$h_i = h_i \left(c_0, c_1, \ldots, c_I \right)$$

- The h_is embody all the behavioral responses

Saez (2002): Model Setup

- The government sets taxes to maximize welfare.
- Subject to resource constraint: Has to finance H of per-capita government spending

$$\sum_{i=0}^{I} h_i T_i = H$$

- Government attaches welfare weight g_i to people in each occupation
- NB not clear if $g_0 \lesseqgtr g_1$: e.g. “Lazy Poor” $\rightarrow g_0 < g_1$
- Weights are endogenous, depend on the tax schedule. Without income effects

$$\sum_{i=0}^{I} h_i g_i = 1$$

Saez (2002): Only Extensive Margin

- Suppose every individual is endowed with a skill $i \in \{0, 1, \ldots, I\}$.
- Only *extensive margin* choice: work as i or be unemployed.
- Without income effects, participation depends only on $c_i - c_0$
- Clearly $c_i \geq c_0 \forall i$. Define extensive elasticities

$$\eta_i = \frac{c_i - c_0}{h_i} \frac{\partial h_i}{\partial \left(c_i - c_0\right)}$$

Saez (2002): Only Extensive Margin

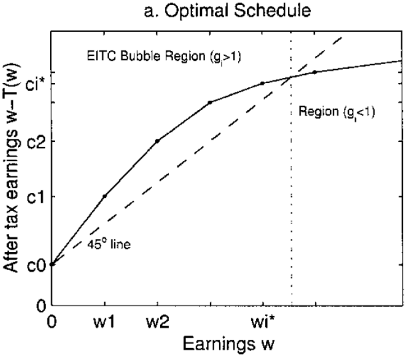
- *Proposition 1*: The optimal tax schedule satisfies

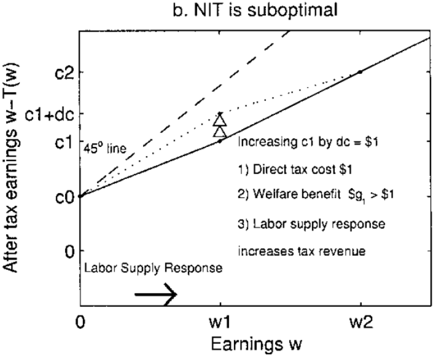
$$\frac{T_i - T_0}{c_i - c_0} = \frac{1}{\eta_i}\left(1 - g_i\right)$$

- Heuristic proof: Consider small increase dT_i in T_i. 2 effects
  - Mechanical Effect: $dM = \left(1 - g_i\right) h_i dT_i$
  - Behavioral Effect: $dB = \left(T_i - T_0\right) dh_i = -\left(T_i - T_0\right) h_i \eta_i \frac{dT_i}{(c_i - c_0)}$
  - Optimality: $dM + dB = 0$

Saez (2002): Extensive Margin: Implications for NIT vs EITC

- In the model with only extensive responses, NIT is always suboptimal, and EITC is usually optimal





Saez (2002): Only Intensive Margin

- Suppose now that people are only choosing how much to work.
- Really, people in occupation i are choosing whether to switch to $i-1$ or $i+1$.
- Then (again, w/out income effects) we can write the employment shares $h_i\left(c_{i+1}-c_i, c_i-c_{i-1}\right)$
- Now we can define the intensive elasticities

$$\zeta_i = \frac{c_i - c_{i-1}}{h_i}\frac{\partial h_i}{\partial\left(c_i - c_{i-1}\right)}$$

Saez (2002): Only Intensive Margin

- Proposition 2: The optimal tax schedule satisifies

$$\frac{T_i - T_{i-1}}{c_i - c_{i-1}} = \frac{1}{\zeta_i} \frac{(1-g_i)\, h_i + (1-g_{i+1})\, h_{i+1} + \ldots + (1-g_I)\, h_I}{h_i}$$

- Heuristic proof: Consider small increase in tax for all jobs above i: $dT_i = dT_{i+1} = \ldots = dT_I = dT$. 2 Effects:
  - Mechanical Effect: People above i pay dT more. Valued at

$$dM = (1-g_i)\, h_i + (1-g_{i+1})\, h_{i+1} + \ldots + (1-g_I)\, h_I$$

  - Behavioral effects:
    $dB = (T_i - T_{i-1})\, dh_i = -(T_i - T_{i-1})\, h_i \zeta_i \frac{dT}{c_i - c_{i-1}}$
  - Optimality: $dM + dB = 0$

Saez (2002): Intensive Margin: Implications for NIT vs EITC

- Look at optimality condition for $i = 1$:

$$\frac{T_1 - T_0}{c_1 - c_0} = \frac{1}{\zeta_1}\left[\frac{(g_0 - 1)\, h_0}{h_1}\right]$$

- Positive marginal tax (Negative Income Tax) iff $g_0 > 1$

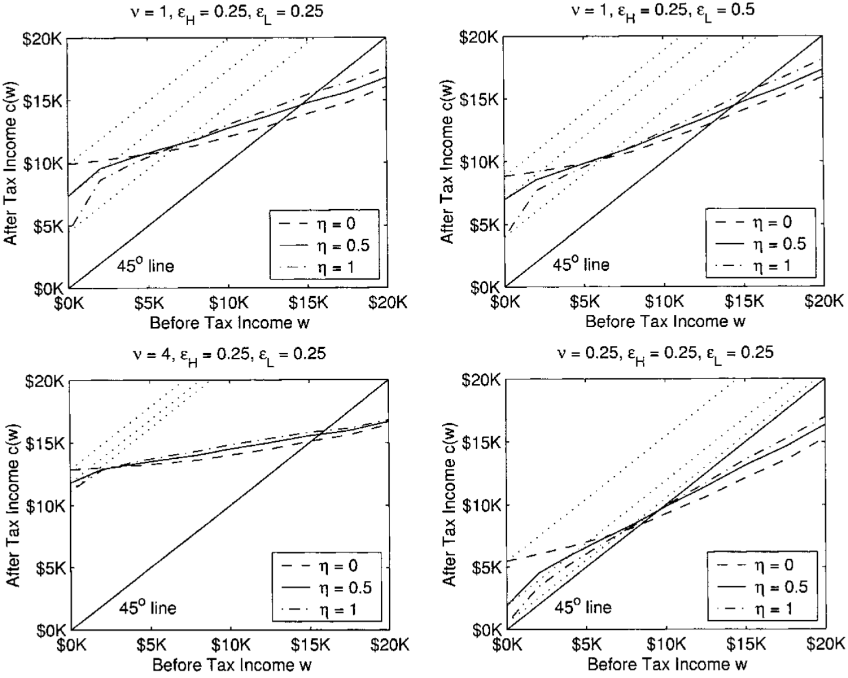
- More generally, $g_i > g_{i+1} \rightarrow$ increasing marginal tax rates, and no negative marginal tax rates (no EITC)

Saez (2002): Both Margins

- In the model with both margins, the optimal tax schedule satisfies

$$\frac{T_i - T_{i-1}}{c_i - c_{i-1}} = \frac{1}{\zeta_i h_i} \sum_{j=i}^{I} h_j \left[1 - g_j - \eta_j \frac{T_j - T_0}{c_j - c_0}\right]$$

- Ambiguous implications for optimal transfers: Depends on relative sizes of ηs and ζs.
- Calibration:
  - Extensive elasticity: $\eta \in \{0, 0.5, 1\}$ below \$20K, 0 above \$20K
  - Intensive elasticity ζ: Use traditional estimates of $\varepsilon = d \log y / d \log (1 - \tau)$. $\varepsilon_L \in \{0, 0.25, 0.5\}$ below \$20K, $\varepsilon_H \in \{0.25, 0.5\}$ above \$20K
  - No income effects.
  - Welfare: $g(c) = 1/(p \cdot c^v)$, p=marginal value of public funds, higher v = more redistributive prefrerences
  - $H = \$5K$
  - Income distribution calibrated using 1997 CPS and current tax schedule


ν = 1, ε_H = 0.25, ε_L = 0.25
ν = 1, ε_H = 0.25, ε_L = 0.5
ν = 4, ε_H = 0.25, ε_L = 0.25
ν = 0.25, ε_H = 0.25, ε_L = 0.25
After Tax Income c(w)
Before Tax Income w
$0K
$5K
$10K
$15K
$20K
45° line
η = 0
η = 0.5
η = 1

Outline

Theory

- Saez (QJE 2002) *Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses*
- Akerlof (AER 1978) *The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning*
- Nichols & Zeckhauser (AER 1982) *Targeting Transfers Through Restrictions on Recipients*
- Besley & Coate (AER 1992) *Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs*
- Niehaus, Atanassova, Bertrand & Mullainathan (AEJ:Pol 2013) *Targeting with Agents*

Akerlof (1978): Overview

- Why not use the tax system to target transfers to the poor?
- Contrast two extreme cases:

1. Negative Income Tax: $T = -\alpha\bar{Y} + tY$. $\bar{Y}$ is per capita income, t is marginal tax rate.

$$\sum_i T = -\alpha n\bar{Y} + t\sum_i Y_i \Rightarrow t = \alpha + g \quad g \equiv \sum_i T_i / \sum_i Y_i$$

   Giving the poor a bigger transfer (α) requires a one-for-one increase in t

2. Perfect tag: Imagine we can identify a group of size βn that contains all the poor people and we give the amount α to everyone in this group. Now

$$t = \beta\alpha + g$$

   Giving more to the poor only costs $\beta < 1$

Akerlof (1978): Simple Optimal Tax Model

- 2 types of workers, skilled & unskilled. Each 50% of the population
- 2 types of jobs. Easy and difficult.
  - Skilled worker's output in difficult job is $q_D > q_E =$ skilled worker's output in easy job
  - Unskilled workers only work in the easy job. Produce output q_E
- Workers in difficult job pay tax t_D. Utility of skilled workers in difficult job is $u(q_D - t_D) - \delta$
- Utility of both types in easy job is $u(q_E + t_E)$
- Assume $u(q_D) - \delta > u(q_E)$ (otherwise everyone always takes easy job).

Akerlof (1978): Simple Optimal Tax Model

- Government chooses t_D, t_E to maximize expected utility

$$U = \frac{1}{2}\max\left\{u\left(q_D - t_D\right) - \delta, u\left(q_E + t_E\right)\right\} + \frac{1}{2}u\left(q_E + t_E\right)$$

- Budget balance requires

$$\begin{aligned} t_D = t_E \quad & \text{if} \quad u\left(q_D - t_D\right) - \delta \geq u\left(q_E + t_E\right) \\ t_E = 0 \quad & \text{if} \quad u\left(q_D - t_D\right) - \delta < u\left(q_E + t_E\right) \end{aligned}$$

- Optimum:

$$\begin{aligned} t_D^* = t_E^* \quad & \text{(budget balance)} \\ u\left(q_D - t_D^*\right) - \delta = u\left(q_E + t_E^*\right) \quad & \text{(binding IC)} \end{aligned}$$

Akerlof (1978): Tagging in Simple Model

- Introduce tagging: A proportion β of the poor have an observable tag.
- Difficult job taxed T_D. Untagged in easy job get T_E. Tagged get τ
- Government problem is now

$$\max_{T_D,T_E,\tau} U^{Tag} = \frac{1}{2}\max\left\{u\left(q_D - T_D\right) - \delta, u\left(q_E + T_E\right)\right\} + \frac{1}{2}\left(1-\beta\right)u\left(q_E + T_E\right) + \frac{1}{2}\beta u\left(q_E + \tau\right)$$

- Subject to

$$T_D = \left(1-\beta\right)T_E + \beta\tau \text{ if } u\left(q_D - T_D\right) - \delta \geq u\left(q_E + T_E\right)$$
$$\left(2-\beta\right)T_E + \beta\tau = 0 \text{ if } u\left(q_D - T_D\right) - \delta < u\left(q_E + T_E\right)$$

Akerlof (1978): Tagging Optimum

- Optimum features binding IC constraint

$$u\left(q_D - T_D^*\right) - \delta = u\left(q_E + T_E^*\right)$$

- More importantly, tagging allows more redistribution: $\tau^* > t_E^*$.
  - Rough proof by contradiction: Suppose that $\tau^* \leq T_E^*$.
  - Consider reducing T_E by ε and increasing τ by ε:

  1. Nobody is worse off (envelope theorem)
  2. This perturbation raises revenue
  3. This revenue can be used to redistribute

  $\Rightarrow$ τ^* or T_E^* is not optimally set

Akerlof (1978): Tags

- What makes a good tag?

1. Easily Observable. Low admin costs of using the tag
2. Correlated with need. Good at discriminating between needy and not (high β)
3. Immutable. People can't endogenously aquire the tag

- Examples?

Outline

Theory

- Saez (QJE 2002) *Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses*
- Akerlof (AER 1978) *The Economics of “Tagging” as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning*
- Nichols & Zeckhauser (AER 1982) *Targeting Transfers Through Restrictions on Recipients*
- Besley & Coate (AER 1992) *Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs*
- Niehaus, Atanassova, Bertrand & Mullainathan (AEJ:Pol 2013) *Targeting with Agents*

Nichols & Zeckhauser (1982): Overview

- What instruments should the government use to redistribute?
- Atkinson-Stiglitz (1976) $\rightarrow$ only income taxes/cash transfers.
- Akerlof (1978) $\rightarrow$ income taxes/cash transfers, but possibly dependent on *tags.*
- What about in-kind benefits? Commodity taxes/subsidies? Ordeals?
- This paper shows under what conditions to use these other instruments and how.

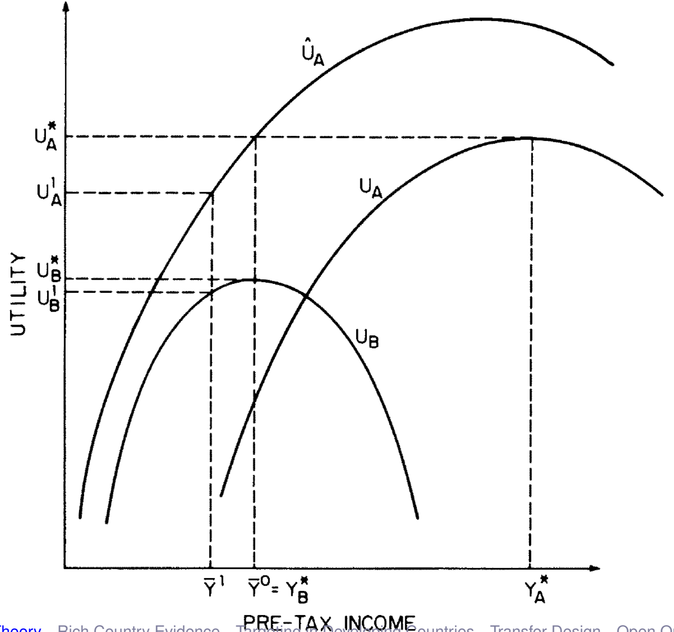
Nichols & Zeckhauser (1982): Income Tax Benchmark

- Suppose there are two individuals. Beneficiary B and high-wage earner A. Want A to finance a transfer to B. What is the most efficient way to do it?
- Both individuals have the same utility $U(C,E)$, $C=Y+T$ consumption, E effort.
- Government can't see wages or hours, only income.
- Optimal policy takes the form: If $Y>\bar{Y}$ pay tax T; if $Y\leq\bar{Y}$ receive transfer T
- IC constraint that the high type doesn't masquerade is

$$U\left(Y_A^*-T, Y_A^*/W_A\right)>U\left(\bar{Y}+T,\bar{Y}/W_A\right)$$

- Result: $\bar{Y}<Y_B^*$. It's optimal to distort B's choice

Nichols & Zeckhauser (1982): Income Constraint



Nichols & Zeckhauser (1982): In-Kind Transfers

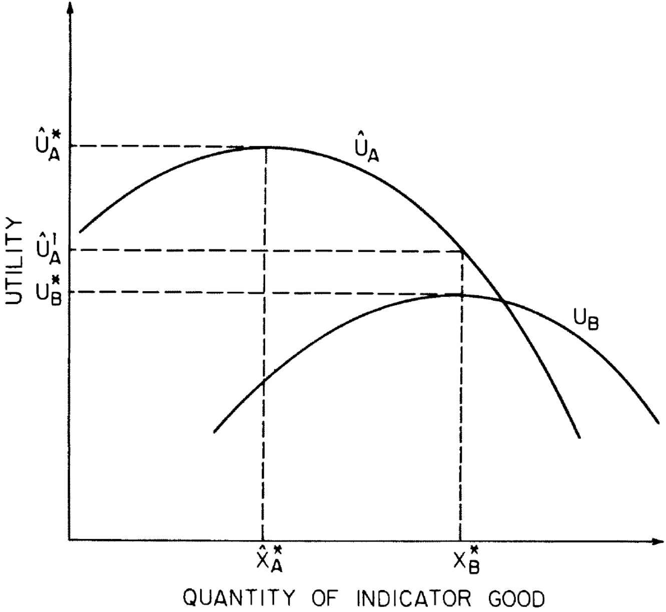
- Can it be optimal to give some of the transfer through an in-kind transfer instead of its cash equivalent?

- Usually, we think people must be at least as well off with the cash, so no.

- This is the case when demand depends only on its price and income (i.e. it's independent of leisure, ability, other observables). When it doesn't, it might act as a tag. N&Z call these *Indicator goods*

- Return to the simple income tax model. Out of total income $\bar{Y} + T$, B buys X_B^* of good X (at price 1). If A masquerades, she would only buy $\hat{X}_A^*$. $\Rightarrow X$ is an indicator good.

Nichols & Zeckhauser (1982): In-Kind Transfers



Nichols & Zeckhauser (1982): In-Kind Transfers

- Consider changing the transfer. Receive $\bar{X}$ in kind, remaining $T - \bar{X}$ in cash.
- If $\bar{X} < \hat{X}_A^*$, no effect
- If $\hat{X}_A^* < \bar{X} \leq X_B^*$ B suffers no loss, but makes mimicry more costly for A
- In general it is optimal to have $\bar{X} > X_B^*$ for deterrence.
- Similar argument can be applied to subsidies.

Outline

Theory

Saez (QJE 2002) *Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses*

Akerlof (AER 1978) *The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning*

Nichols & Zeckhauser (AER 1982) *Targeting Transfers Through Restrictions on Recipients*

Besley & Coate (AER 1992) *Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs*

Niehaus, Atanassova, Bertrand & Mullainathan (AEJ:Pol 2013) *Targeting with Agents*

Besley & Coate (1992): Overview (ET)

- Should recipients of relief be required to work in exchange for benefits?
- Two reasons why you might want to make transfers conditional on work:
  1. *Screening*. Want to direct poor support towards the truly needy. When we can't observe earnings/ability/welfare, the work requirement $\rightarrow$ self-targeting
     - Idea comes from from Nichols and Zechauser (1982) 'ordeals' on welfare claimants may improve targeting. "Say one welfare eligible would receive 100 utiles from a particular transfer [marginal utility is high], yet another would receive only 10. Then an ordeal that imposes ajn 11 utile loss in order to qualify for the transfer will be an effective sorting device"
  2. *Deterrence.* Encourage poverty-reducing investments: 'Are individuals poor just because they have experienced bad luck or because of choices made earlier in life?' If the latter is true, then public assistance may lead individuals to make choices that increase the likelihood that they will have to draw on such support in the futrue
- LDCs: not always possible to observe earnings. But in both it is not possible to observe earning opportunities

Besley & Coate (1992): Model (ET)

- n individuals. Fraction γ of low types have different 'income-generating ability' $a_L < a_H$, with quasilinear preferences over y and work l. $u(y, l) = y - h(l)$
- Poverty Alleviation Program (PAP) is a pair of benefit packages $\{b_i, c_i\}_{i=L,H}$. Transfer b_i and requirement of c_i hours of work. Assumes that public sector work is unproductive.
- Cost of program to government is $n[\gamma b_L + (1-\gamma) b_H]$, objective to minimize cost subject to constraint that z is achieved. i.e. this is a utilitarian objective: "How to give the greatest amount of needful help, with the smallest encouragement to undue reliance on it" - John Stuart Mill
- Individuals a_i choose to claim benefit package that is intended for them $\{b_{i,} c_i\}$.

Besley & Coate (1992): Model (ET)

1. Individuals who accept $\{b_i, c_i\}$ can also provide private-sector labor:

$$l\left(b, c, a_i\right) = \begin{cases} \hat{l}\left(a_i\right) - c & \text{if } c \leq \hat{l}\left(a_i\right) \\ 0 & \text{otherwise} \end{cases}$$

   where $\hat{l}$ is optimal labor supply without PAP $\left(h'\left(\hat{l}\right) = a_i\right)$ (marginal utility from labor = a_i)
2. Labor doesn't depend on the benefit, and is 0 if the work amount is higher than what they would optimally provide.
3. Yields private-sector earnings

$$y\left(c, a_i\right) = \begin{cases} a_i\left[\hat{l}\left(a_i\right) - c\right] & \text{if } c \leq \hat{l}\left(a_i\right) \\ 0 & \text{otherwise} \end{cases}$$

4. and indirect utility

$$v\left(b, c, a_i\right) = b + y\left(c, a_i\right) - h\left(l\left(c, a_i\right) + c\right)$$

Besley & Coate (1992): Benchmark Case (ET)

- Assume observable type - design to make sure that individual receives benefit package designed for him.
- Policy maker's problem - choose a PAP that minimizes the cost of poor relief as well as
  1. Participation constraints: $v\left(b_i, c_i, a_i\right) \geq v\left(0, 0, a_i\right),\ i = L, H$
  2. L types must escape poverty: $b_L + y\left(c_L, a_L\right) \geq z$
- Prop 1: observable types imply that the cost minimizing PAP is a welfare program (no work).
  - cash transfer for low types to get them out of poverty $z - y(0, a_L)$.

Besley & Coate (1992): Screening/Unobservable Case (ET)

- policymaker has no information on type but knows fraction/distribution $\gamma$of types
- Incentive Compatibility (IC): high and low types prefer the package offered to them to the others.
  - DCs: Individual's private sector earnings can't be observed. Masquerading doesn't mean individual has to reduce earnings. (can still work and also receive benefit)
  - ACs: Can be observed. Masquerading must also reduce earnings (not working)
- Unobservable Private Sector/DC context: $v(b_{L,}c_L, a_L) \geq v(b_{H,}c_H, a_L)$ and $v(b_{H,}c_H, a_H) \geq v(b_{L,}c_L, a_H)$.
  - MInimize costs subject to voluntary participation, poverty alleviation, and IC
  - impose a cost that the rich wouldn't want to take up (work/hours) 'separating work requirement' c_L^s that when coupled with a benefit suffiecient to get the poor to the poverty line, makes high types indifferent between claiming to be of low ability and receiving no benefit at all.

$$v(0, 0, a_H) \geq v(z - y(c_L^s, a_L), c_L^s, a_H)$$

Besley & Coate (1992): Screening/Observable Case (ET)

- Because we observe types,
  - Case 1: benchmark PAP is implementable if $v(0, 0, a_H) \geq z - h(y(0, a_L)/a_H)$ or if the high type prefers their income over the income they'd get from masquerading as a low type and working less.
  - Case 2: if condition is not satisfied, the IC requires that the benefit + income from being a low type less the labor implied by working as a low type.

$$v(b_{H,} c_H, a_H) \geq v(b_{L,} c_L, a_H) = b_L + y(c_{L,} a_L) - h(y(c_{L,} a_L)/a_H + c_L$$

  - Again, define the separating work requirement that satisfies

$$v\left(0, 0, a_H\right) = z - h\left(\frac{y\left(\hat{c}_L^s, a_L\right)}{a_H} + \hat{c}_L^s\right)$$

  - PROPOSITION 3: *If income-generating abilities are unobservable, individuals' incomes are observable and the benchmark PAP is not implementable, one of the following two programs is cost-minimizing: (i) (welfare) impose no work requirements and offer self-categorized high-ability individuals a*

Besley & Coate (1992): Deterrence (ET)

- If Poverty depends not only on luck but alos on choices made earlier in life, we allow each individual to make an ex-ante choice which influences his future earning ability.
- suppose $\pi(e)$function which is the probability of becoming a high ability person based on e effort (strictly concave nd increasing).
- Maximize their utility of future earnings vs the disutility of effort, and the expected cost is now:
- $$\pi\left(e\right)v\left(b_H, c_H, a_H\right) + \left[1 - \pi\left(e\right)\right]v\left(b_L, c_L, a_L\right) - e$$
- Now endogenous. Benchmark is problematic because it reduces expost utility difference - will create an undue reliance on state support
- Exists a maximal work requirement c_L^m if coupled with a transfer sufficient to get the poor to the poverty line, it would make them just indifferent between the status quo and participating in the program, (pays as much for the same work, but must be higher than the no-policy labor since low types

Besley & Coate (1992): Overview

- Should people who receive transfers be required to work in exchange?
- Common policy, but is it efficient? 2 possible reasons

1. *Screening*. When we can't (easily) observe earnings (ability), work requirement $\rightarrow$ self-targeting
2. *Deterrence.* Encourage poverty-reducing investments

- B&C formalize this, provide sufficient
- n individuals. Fraction γ of low types have ability $a_L < a_H =$ ability of high types
- $u(y, l) = y - h(l)$
- Poverty-alleviation program: $\{b_i, c_i\}_{i=L,H}$. Transfer b_i and requirement of c_i hours of work.
- PAP costs $n[\gamma b_L + (1-\gamma) b_H]$. Government objective is to minimize cost of guaranteeing everyone income z

Besley & Coate (1992): Model

- Individuals who accept $\{b_i, c_i\}$ can also provide private-sector labor:
$$l(b, c, a_i) = \begin{cases} \hat{l}(a_i) - c & \text{if } c \leq \hat{l}(a_i) \\ 0 & \text{otherwise} \end{cases}$$
where $\hat{l}$ is optimal labor supply without PAP $\left(h'\left(\hat{l}\right) = a_i\right)$
- Yields private-sector earnings
$$y(c, a_i) = \begin{cases} a_i\left[\hat{l}(a_i) - c\right] & \text{if } c \leq \hat{l}(a_i) \\ 0 & \text{otherwise} \end{cases}$$
- and indirect utility
$$v(b, c, a_i) = b + y(c, a_i) - h(l(c, a_i) + c)$$

Besley & Coate (1992): Model

- Individuals will take the package intended for them iff

$$v\left(b_i, c_i, a_i\right) \geq v\left(0, 0, a_i\right)$$

- Assume that only the L types are poor without government intervention

$$y\left(0, a_H\right) > z > y\left(0, a_L\right)$$

Besley & Coate (1992): First-best benchmark

- Suppose policymakers observe abilities a_L, a_H.

- Now the government has to satisfy two constraints:

1. Participation constraints: $v\left(b_i, c_i, a_i\right) \geq v\left(0, 0, a_i\right), \; i = L, H$

2. L types must escape poverty: $b_L + y\left(c_L, a_L\right) \geq z$

- PROPOSITION 1: *Cost-minimizing PSP is a welfare program. Low-ability individuals are offered a cash transfer that is just high enough to get them out of poverty, $z - y\left(0, a_L\right)$, and high-ability individuals are offered no benefits.*

Besley & Coate (1992): Screening

- Now suppose, govt only knows γ, the proportion of L types.
- Now the government also needs to respect incentive compatibility constraints:

$$v\left(b_L, c_L, a_L\right) \geq v\left(b_H, c_H, a_L\right) \quad v\left(b_H, c_H, a_H\right) \geq v\left(b_L, c_L, a_H\right)$$

- Note, without work requirements, these imply $b_L = b_H$
- The work requirement will allow the government to screen because H types have a higher opportunity cost of government work.
- But, the work requirement reduces the L types' income, so requires a bigger transfer to get them back to $y_L = z$.
- Tradeoff of increasing work requirement: lower transfers to H to respect IC, and higher transfers to L to achieve $y_L = z$

Besley & Coate (1992): Screening

- Define the *separating work requirement* c_L^s that gets the poor out of poverty and prevents masquerading

$$v(0, 0, a_H) = v\left(z - y\left(c_L^s, a_L\right), c_L^s, a_H\right)$$

- PROPOSITION 2: *If both income-generating abilities and incomes are unobservable, one of the following two PAPs is cost-minimizing: (i) (welfare) impose no work requirements and offer both ability groups a transfer of* $z = y(0, a_L)$*; (ii) (workfare) offer self-categorized high-ability individuals no benefits and offer self-categorized low-ability individuals a transfer of* $z - y(c_L^s, a)$ *in exchange for a work requirement of* c_L^s*. A sufficient condition for the workfare solution to be cost-minimizing is that* $a_L < (1 - \gamma) a_H$*.*

Besley & Coate (1992): Screening

- In the previous case we assumed the government can't observe earnings (so transfers can't depend on earnings).
- When the government *can* observe earnings, to mimic the low type, the high type must reduce her earnings to $y\left(c_L, a_L\right)$, requiring labor supply $y\left(c_L, a_L\right)/a_H$.
- We might even be able to implement the first-best benchmark if

$$v\left(0,0,a_H\right) \geq z - h\left(\frac{y\left(0,a_L\right)}{a_H}\right)$$

(H prefers no benefit to earning $y\left(0,a_L\right)$ and consuming z)

- If this condition isn't met, we have to satisfy the IC constraint

$$v\left(b_H, c_H, a_H\right) \geq b_L + y\left(c_L, a_L\right) - h\left(\frac{y\left(c_L, a_L\right)}{a_H} + c_L\right)$$

Besley & Coate (1992): Screening

- Again, define the separating work requirement that satisfies

$$v\left(0,0,a_H\right) = z - h\left(\frac{y\left(\hat{c}_L^s, a_L\right)}{a_H} + \hat{c}_L^s\right)$$

- PROPOSITION 3: *If income-generating abilities are unobservable, individuals' incomes are observable and the benchmark PAP is not implementable, one of the following two programs is cost-minimizing: (i) (welfare) impose no work requirements and offer self-categorized high-ability individuals a transfer of* $z - h\left(y\left(0,a_L\right)/a_H\right) - v\left(0,0,a_H\right)$ *and offer self-categorized low-ability individuals a transfer of* $z - y\left(0,a_L\right)$*; (ii) (workfare) offer self-categorized high-ability individuals no benefits and offer self-categorized low-ability individuals a transfer of* $z - y\left(\hat{c}_L^s, a_L\right)$ *in exchange for a work requirement of* $\hat{c}_L^s$*. A sufficient condition for the workfare solution to be cost-minimizing is that* $\gamma a_L < \left(1-\gamma\right) h'\left(\frac{y(0,a_L)}{a_H}\right)\left(1 - \frac{a_L}{a_H}\right)$.

Besley & Coate (1992): Deterrence

- Assume again that the government observes ability. However ability depends on effort earlier in life.
- Probability of being high ability is $\pi(e)$ (strictly concave), costs e
- Individuals will choose e to maximize

$$\pi(e)\, v(b_H, c_H, a_H) + [1 - \pi(e)]\, v(b_L, c_L, a_L) - e$$

- e^* increasing in ex-post difference between high- and low-ability

$$e^*(\chi(\cdot)) \quad \chi(b_L, c_L, b_H, c_H) \equiv v(b_H, c_H, a_H) - v(b_L, c_L, a_L)$$

Besley & Coate (1992): Deterrence

- Define the *maximal work requirement* c_L^m that satisfies

$$v\left(z-y\left(c_L^m,a_L\right),c_L^m,a_L\right)=v\left(0,0,a_L\right)$$

- Note that $c_L^m > \hat{l}\left(a_L\right)$ and so the low-ability individuals do not work.

- PROPOSITION 4: *If income-generating abilities are observable but depend partly on choices made earlier in life, the cost-minimizing PAP either imposes no work requirements and offers low-ability individuals a transfer of* $z-y\left(0,a_L\right)$*, or imposes the maximal work requirement* c_L^m *on low-ability individuals and offers them a transfer of* z

Outline

Theory

- Saez (QJE 2002) *Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses*
- Akerlof (AER 1978) *The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning*
- Nichols & Zeckhauser (AER 1982) *Targeting Transfers Through Restrictions on Recipients*
- Besley & Coate (AER 1992) *Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs*
- Niehaus, Atanassova, Bertrand & Mullainathan (AEJ:Pol 2013) *Targeting with Agents*

Niehaus et al (2013): Overview

- How should targeting rules be designed when they must be implemented by corruptible agents?
- Consider a proxy means test (PMT) that must be implemented by an official who may be corrupt.
- Show that having more indicators (which makes identifying the poor easier statistically) can backfire if enforcement is weak, by making it easier to pretend that ineligible households are eligible and hence easier to receive bribes.
- Test the theory with data on Below Poverty Line (BPL) cards in Karnataka, India.

Niehaus et al (2013): Households

- The principal wants to allocate slots among a set of households.
- Household i has
  - Income $y_i \in \{\underline{y}, \overline{y}\}$
  - other characteristics $\mathbf{x}_i \in \mathbf{X}$
  - Values the slot at $v_i \sim G(v_i)$, exponential with rate $1/\eta$
- Joint distribution of attributes is $F(y_i, \mathbf{x}_i)$

Niehaus et al (2013): Official

- Principal would like to use y_i for targeting, but it's unobserved.
- Instead, use easier to observe characteristics $\mathbf{x}$
- A targeting rule is a subset $R \subseteq \mathbf{X}$: A household is eligible iff $\mathbf{x}_i \in R$
- The official (agent)
  - Implements R.
  - Observes $y_i, \mathbf{x}_i$, but not v_i
  - Cares about his income Y and the allocation.

$$U\left(Y, \{a_i\}\right) = Y + \underline{\alpha} \int_{y_i = \underline{y}} a_i di + \overline{\alpha} \int_{y_i = \overline{y}} a_i di$$

  where $a_i \in \{0, 1\}$ indicates whether household i gets a slot; $(\underline{\alpha}, \overline{\alpha})$ summarize the official's distributive preferences

Niehaus et al (2013): Official

- If official violates R, he is caught with probability $\pi(a_i, \mathbf{x}_i, R)$
  - $\pi(a, \mathbf{x}, R) = 0$ if $a = \mathbf{1}\{\mathbf{x} \in R\}$
  - $\pi(a, \mathbf{x}, R) > 0$ if $a \neq \mathbf{1}\{\mathbf{x} \in R\}$
  - Punishment is a fine f

- Official allocates slots by establishing a menu of prices $p(y_i, \mathbf{x}_i) \geq 0$

$$\max_{\{p_i\}} \int [1 - G(p_i)] [p_i - c(y_i, \mathbf{x}_i)] \, dF(y_i, \mathbf{x}_i)$$

where the implicit marginal cost $c(y_i, \mathbf{x}_i)$ is

$$c(y_i, \mathbf{x}_i) = f[\pi(1, \mathbf{x}_i, R) - \pi(0, \mathbf{x}_i, R)] - \underline{\alpha}\mathbf{1}\{y_i = \underline{y}\} - \overline{\alpha}\mathbf{1}\{y_i = \overline{y}\}$$

Niehaus et al (2013): Official

- The official's problem looks just like a monopolist's problem.
- The solution will satisfy MR = MC, markups follow the rule $(p - c)/p = -1/\epsilon$

$$p^*(y_i, \mathbf{x}_i) = \max\{0, c(y_i, \mathbf{x}_i) + \eta\}$$

- The probability that household i gets a slot is then

$$\Pr(a_i = 1 | \mathbf{x}_i, y_i) = 1 - G(\max\{0, c(y_i, \mathbf{x}_i) + \eta\})$$

- Comparative statics:
    - prices increase in income iff $\underline{\alpha} > \overline{\alpha}$
    - prices decrease in eligibility (strictly if $f > 0$)
    - If f is sufficiently large, all eligible households get a slot at price 0. But might require arbitrarily harsh penalties, which seems unlikely.

Niehaus et al (2013): Principal

- Principal values the poor's surplus at $\underline{\omega}$. Similarly for the rich at $\overline{\omega} < 1/\eta < \underline{\omega}$
- Normalize cost of slots to 1.

$$V(\{p_i\}) = \int_{y_i=\underline{y}} \mathbf{1}\{v_i > p_i\}\left[\underline{\omega}(v_i - p_i) - 1\right] dG(v_i)\, dF(y_i, \mathbf{x}_i)$$
$$+ \int_{y_i=\overline{y}} \mathbf{1}\{v_i > p_i\}\left[\overline{\omega}(v_i - p_i) - 1\right] dG(v_i)\, dF(y_i, \mathbf{x}_i)$$

- Using the exponential distribution $G(v_i) = e^{-v_i/\eta}/\eta$

$$V(\{p_i\}) = (\underline{\omega}\eta - 1)\int_{y_i=\underline{y}} e^{-p_i/\eta} dF(y_i, \mathbf{x}_i)$$
$$+ (\overline{\omega}\eta - 1)\int_{y_i=\overline{y}} e^{-p_i/\eta} dF(y_i, \mathbf{x}_i)$$

Niehaus et al (2013): When agency doesn't matter

- Contrast with the case where the agent is completely honest, $p_i = 0$ if $\mathbf{x}_i \in R$, $p_i = \infty$ if $\mathbf{x}_i \notin R$. Now, principal solves

$$\max_{R \subseteq \mathbf{X}} (\underline{\omega}\eta - 1) \int_{y_i = \underline{y}} \mathbf{1}\{\mathbf{x}_i \in R\} \, dF(y_i, \mathbf{x}_i)$$
$$+ (\overline{\omega}\eta - 1) \int_{y_i = \overline{y}} \mathbf{1}\{\mathbf{x}_i \in R\} \, dF(y_i, \mathbf{x}_i)$$

- PROPOSITION 1: *Let R^* be statistically optimal (solves above equation). Then*
  1. *As $f \to \infty$ the payoff from R^* approaches the constrained optimal payoff.*
  2. *As $\underline{\alpha} \to \infty$ while $\overline{\alpha} \to -\infty$ the payoff from R^* approaches the constrained optimal payoff.*
  3. *If $\underline{\alpha} = \overline{\alpha}$ and there exists $\tilde{\pi}$ such that $\pi(a_i, \mathbf{x}_i, R) = \tilde{\pi} \cdot \mathbf{1}\{a_i \neq \mathbf{1}\{\mathbf{x}_i \in R\}\}$ then rule R^* yields at least as high a payoff as any other nontrivial rule.*

Niehaus et al (2013): When agency doesn't matter

- In any of these cases agency issues are unimportant: Principal wants to use the statistically optimal rule.
- Third case is the interesting one: What does $\pi(a_i, \mathbf{x}_i, R) = \tilde{\pi} \cdot \mathbf{1}\{a_i \neq \mathbf{1}\{\mathbf{x}_i \in R\}\}$ mean?
    - Principal detects deviations from R with pr $\tilde{\pi}$, *regardless of what R is.* (e.g. random audits which fully verify $\mathbf{x}_i \in / \notin R$)
    - That means the choice of R doesn't affect how likely it is that the principal can detect deviations.
    - That means the choice of R doesn't affect the official's incentives to deviate from R. Every household is equally risky.
    - Then changing R to make one household eligible won't affect the probability that any other households get a slot.
    - Now question of what R should be is only question of whether R includes eligible households, the statistical question.
- Full verification of R seems like a stretch though. So what happens with more realistic information structures?

Niehaus et al (2013): Means testing

- Let's apply this framework to pure means testing $\mathbf{X} = \{\underline{y}, \overline{y}\}$
- Assume $\underline{\alpha} = \overline{\alpha} = 0$. Official only cares about profit
- w/pr π_e the principal observes the existence of a household
- w/pr $\pi_t \leq \pi_e$ the principal observe y_i
- Consider 2 policies:
  1. Universal eligibility: $R = \mathbf{X}$
  2. Means testing by official: $R = \underline{y}$
- Let's work through the math of which one is better

Niehaus et al (2013): Means testing

- Start with Universal Eligibility:
- The poor: $c\left(\underline{y}\right) = f\left[0 - \pi_e\right] \rightarrow p\left(\underline{y}\right) = \eta - f\pi_e$
- The rich: $c\left(\overline{y}\right) = f\left[0 - \pi_e\right] \rightarrow p\left(\overline{y}\right) = \eta - f\pi_e$
- Principal's value:

$$V\left(UE\right) = \left(\underline{\omega}\eta - 1\right) e^{\left(\frac{f\pi_e}{\eta} - 1\right)} F\left(\underline{y}\right) + \left(\overline{\omega}\eta - 1\right) e^{\left(\frac{f\pi_e}{\eta} - 1\right)} \left[1 - F\left(\underline{y}\right)\right]$$

Niehaus et al (2013): Means testing

- What about Means Testing?
- The poor: $c\left(\underline{y}\right) = f\left[0 - \pi_t\right] \rightarrow p\left(\underline{y}\right) = \eta - f\pi_t$
- The rich: $c\left(\overline{y}\right) = f\left[\pi_t - 0\right] \rightarrow p\left(\overline{y}\right) = f\pi_t + \eta$
- Principal's value:

$$V\left(MT\right) = \left(\underline{\omega}\eta - 1\right) e^{\left(\frac{f\pi_t}{\eta} - 1\right)} F\left(\underline{y}\right) + \left(\overline{\omega}\eta - 1\right) e^{\left(-\frac{f\pi_t}{\eta} - 1\right)} \left[1 - F\left(\underline{y}\right)\right]$$

Niehaus et al (2013): Means testing

- Comparing the two:

$$V(MT) - V(UE) = (1 - \overline{\omega}\eta) \underbrace{\left[e^{f\pi_e/\eta} - e^{-f\pi_t/\eta} \right]}_{\text{fewer rich recipients}} \frac{F(\underline{y})}{e^{\eta}}$$

$$- (\underline{\omega}\eta - 1) \underbrace{\left[e^{f\pi_e/\eta} - e^{f\pi_t/\eta} \right]}_{\text{fewer poor recipients if } \pi_e > \pi_t} \frac{1 - F(\underline{y})}{e^{\eta}}$$

$\Rightarrow$ With perfect enforcement or $\pi_t = \pi_e$, targeting is optimal.

$\Rightarrow$ If exclusion errors are sufficiently costly (high $\underline{\omega}$, $F(\underline{y})$ relative to $\overline{\omega}$, $1 - F(\overline{y})$) the constrained optimal policy is universal eligibility.

Niehaus et al (2013): Proxy Means Testing

- Now imagine that the principal wants to use land x_1 and jewelry x_2 as proxies for poverty.
- Should she use both? Or just land?

- The principal considers anyone whose total assets $x_i^1 + x_i^2 \leq y^*$ as poor. Optimal rule statistically is simply

$$R_{12} \equiv \left\{ \mathbf{x} : x^1 + x^2 \leq y^* \right\}$$

  achieving perfect targeting when no agency concerns.
- Consider using just land:

$$R_1 \equiv \left\{ \mathbf{x} : x^1 \leq x^{1*} \right\}$$

- NB these are both scoring rules of the type $\sum h_n \left(x^n \right) < 0$
- Suppose the principal learns $x_j, \ j \in \{1, 2\}$ for household i with independent probability ϕ_j
- If the principal learns enough to determine the household is incorrectly classified, she fines the official f

Niehaus et al (2013): Proxy Means Testing

- LEMMA 1: *Fix any $\phi_1 > 0$ and let x^{1*} satisfy*

$$\mathbb{P}\left(x^1 + x^2 \leq y^* | x^1 = x^{1*}\right) \underline{\omega} \\ + \left[1 - \mathbb{P}\left(x^1 + x^2 \leq y^* | x^1 = x^{1*}\right)\right] \overline{\omega} = 1/\eta,$$

or $x^{1*} = 0$ if that equation has no solution. Then the rule R_1 defined by threshold x^{1*} is uniquely optimal within the class of rules that condition only on x^1.

- Equate marginal benefits ($\mathbb{P}$ more poor people and $1 - \mathbb{P}$ more rich people get the transfer) with the marginal cost $1/\eta$.
- ..Even when there are agency issues.

Niehaus et al (2013): Proxy Means Testing

- Is there a rule that uses x^2 that's preferable to R_1?
- PROPOSITION 2: *Given a fixed rule R that conditions nontrivially on x^2, there exists $\phi_2^* (R) > 0$ such that if $\phi_2 < \phi_2^* (R)$, then rule R_1 yields a strictly higher payoff than R*
- Intuition:
  - Using x^2 improves targeting.
  - Using x^2 is harder to enforce for poor"ish" people: need to observe both x^1 and x^2 to be sure that someone is ineligible
  - Using x^2 is actually easier to enforce for very rich people: Observing *either* x^1 or x^2 sufficient to determine ineligibility.
  - Balance of effects depends on ϕ_2. As $\phi_2 \to 0$ R_1 becomes easier to enforce.

Niehaus et al (2013): Proxy Means Testing

- Is it always the case that having stronger enforcement increases welfare? No!
- PROPOSITION 3: *Let the probability of detecting a violation be constant ($\pi(a_i, \mathbf{x}_i, R) = \pi > 0$ whenever $a_i \neq \mathbf{1}\{\mathbf{x}_i \in R\}$), If R perfectly targets the poor, then $\partial V/\partial f \geq 0$. If R does not perfectly target the poor, so that there are some ineligible poor and some eligible rich, then there exist a scalar f^* and $\underline{\alpha}^*(f)$ and $\overline{\alpha}^*(f)$ such that if $f > f^*$, $\underline{\alpha} > \underline{\alpha}^*(f)$, and $\overline{\alpha} < \overline{\alpha}^*(f)$ then $\partial V/\partial f < 0$.*
- Intuition:
  - If you have a perfect rule, you want to force the official to use it
  - If you have an imperfect rule, then how much you want to force the official to use it depends on how much you disagree about who is deserving.
  - If sufficiently aligned: $\underline{\alpha}$ is large, and $\overline{\alpha}$ is small, then let the official use his discretion to violate your rule sometimes.

Outline

Motivating Facts

Theory

Evidence from Rich Countries

Targeting in Developing Countries: Who gets the Benefit?

Transfer Design: What is the Benefit?

Open Questions

Outline

Evidence from Rich Countries

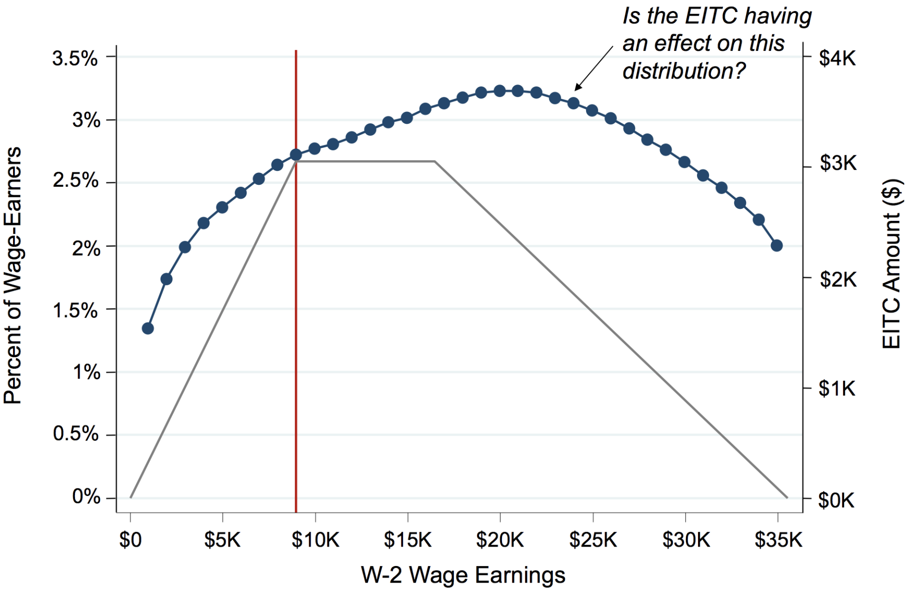
- Chetty, Friedman & Saez (AER 2013): *Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings*
- Deshpande & Li (2017) *Who is Screened Out? Application Costs and the Targeting of Disability Programs*
- Deshpande & Li (2017) [ET] *Who is Screened Out? Application Costs and the Targeting of Disability Programs*
- Finkelstein & Notowidigdo (2019) *Take-up and Targeting: Experimental Evidence from SNAP*

Chetty et al (2013): Overview

- How do income support programs affect workers' earnings?
- Direct effect on the earnings of the beneficiaries
- Indirect effect on earnings of all workers in the same labor market
- Very hard to estimate these market-level impacts.
- Use variation across zip-codes in how many people are aware of EITC to estimate impact of EITC on the whole earnings distribution

Chetty et al (2013): EITC

Chetty et al (2013): EITC



Chetty et al (2013): A Simple Model

- Individuals make two choices:
  - labor supply l_i
  - tax evasion e_i
  - $z_i = wl_i$ is true earnings. Report $\hat{z}_i = wl_i - e_i$
- Taxes:
  - When $\hat{z}_i < K$ tax rate is $\tau_1 < 0$ (a work subsidy)
  - When $\hat{z}_i > K$ tax rate is $\tau_2 > 0$. $\boldsymbol{\tau} = (\tau_1, \tau_2)$
- Compliance behavior
  - Non-compliers always report $\hat{z}_i = K$ to maximize refund.
  - Compliers always report the truth $\hat{z}_i = z_i$
- Utility $U(C_i, l_i, \alpha_i) = C_i - h(l_i, \alpha_i)$
- Heterogeneity in α_i gives rise to earnings distribution $F(z)$

Chetty et al (2013): A Simple Model

- There are N cities of equal size $c = 1, \ldots, N$
- In city c, fraction λ_c of the workers are aware of taxes.
- Remaining $1 - \lambda_c$ optimize as if $\boldsymbol{\tau} = \mathbf{0}$
- Cities differ in skill distributions $G_c(\alpha_i)$ and the fraction of non-compliers θ_c
- Each city has an earnings distribution $F_c(z|\boldsymbol{\tau})$
- The goal is to estimate

$$F_c(z|\boldsymbol{\tau} \neq \mathbf{0}) - F_c(z|\boldsymbol{\tau} = \mathbf{0})$$

Chetty et al (2013): Empirical Strategy

- Basic strategy is to use cities where $\lambda_c = 0$ to estimate the counterfactual:

$$F_c\left(z|\boldsymbol{\tau} \neq \mathbf{0}, \lambda_c = 0\right) = F_c\left(z|\boldsymbol{\tau} = \mathbf{0}, \lambda_c = 0\right)$$

- How to measure λ_c? Use the degree of sharp bunching at K. Denote the fraction of individuals reporting $\hat{z}_i = K$ by ϕ_c. $\phi_c = \theta_c \lambda_c$ and so
- ASSUMPTION 1 (Tax Knowledge): *Individuals in neighborhoods with no sharp bunching at the kink have no knowledge of the policy's marginal incentives and perceive $\boldsymbol{\tau} = 0$:* $\phi_c = 0 \Rightarrow \lambda_c = 0$.
- Under assumption 1 cities with no bunching reveal the distribution when there are no taxes:

$$F_c\left(z|\boldsymbol{\tau} \neq \mathbf{0}, \phi_c = 0\right) = F_c\left(z|\boldsymbol{\tau} = \mathbf{0}, \phi_c = 0\right)$$

Chetty et al (2013): Empirical Strategy

- Can we use cities with no bunching as a counterfactual for cities with bunching?

- ASSUMPTION 2A (Cross-Sectional Identification): *Individuals' skills do not vary across cities with different levels of knowledge about the tax credit:*

$$G\left(\alpha_i|\lambda_c\right) = G\left(\alpha_i\right)\ \forall\lambda_c$$

- Now we can compare cities with and without bunching:

$$\hat{\Delta F} = F\left(z|\boldsymbol{\tau}\right) - F\left(z|\boldsymbol{\tau}, \phi_c = 0\right)$$

Chetty et al (2013): Empirical Strategy

- Assumption 2A is strong (why?)
- We can relax it by studying earnings changes when people become eligible for EITC. Here, use birth of first child which makes people eligible for EITC.
- ASSUMPTION 2B (Panel Identification): *Changes in skills when an individual becomes eligible for the credit do not vary across cities with different levels of knowledge about the tax credit:*

$$G_t\left(\alpha_i|\lambda_c\right) - G_{t-1}\left(\alpha_i|\lambda_c\right) = G_t\left(\alpha_i\right) - G_{t-1}\left(\alpha_i\right) \; \forall \lambda_c$$

- Under this assumption we can construct a difference in differences estimator:

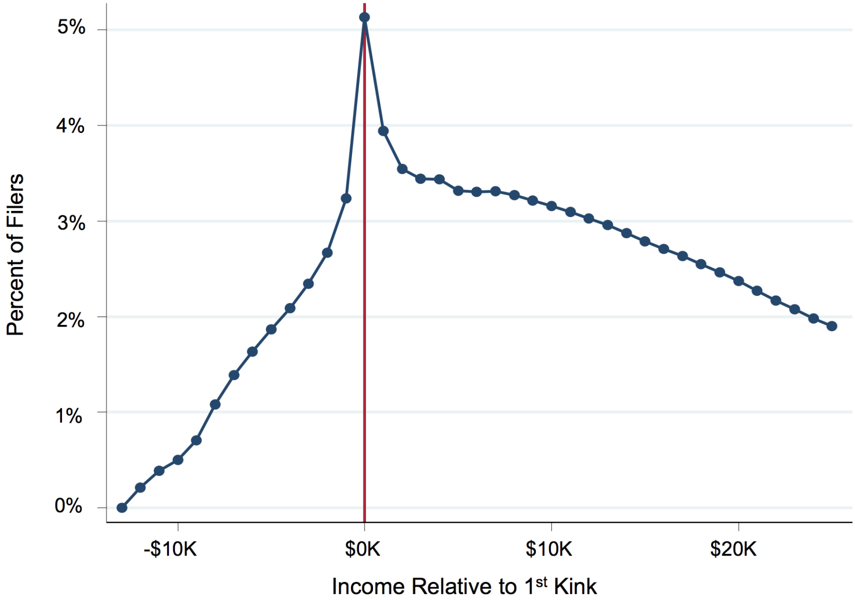
$$\begin{aligned} \hat{\Delta F}_{DD} = & \left[F_t\left(z|\boldsymbol{\tau}\right) - F_t\left(z|\boldsymbol{\tau}, \phi_c = 0\right)\right] \\ & - \left[F_{t-1}\left(z|\boldsymbol{\tau}\right) - F_{t-1}\left(z|\boldsymbol{\tau}, \phi_c = 0\right)\right] \end{aligned}$$

Chetty et al (2013): Data

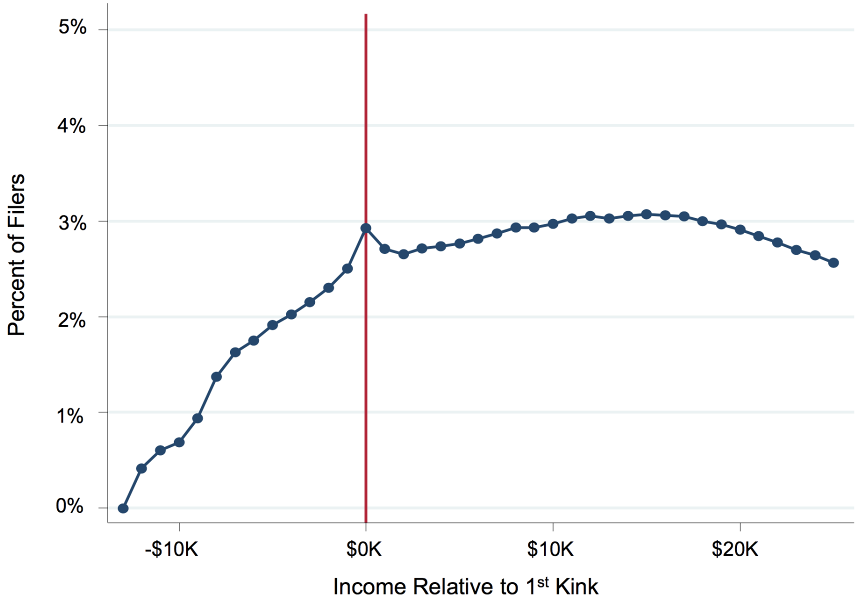
- Universe of US tax returns 1996-2009. Restrict to ppl with a dependent, income<$50K $\rightarrow$ 77.6 mn taxpayers, 1 bn obs.

Variable	Mean (1)	Std. Dev. (2)
Income Measures		
Total Earnings	$20,091	$10,784
Wage Earnings	$18,308	$12,537
Self-Employment Income	$1,770	$6,074
Non-Zero Self-Emp. Income	19.6%	39.7%
Tax Credits		
EITC Refund Amount	$2,543	$1,454
Claimed EITC	88.9%	31.4%
Professionally Prepared Return	69.6%	46.0%
Demographics		
Age	37	13
Number of Children	1.7	0.8
Married	30.3%	45.9%
Female (for single filers)	73.0%	44.4%
Number of Observations	219,742,011	

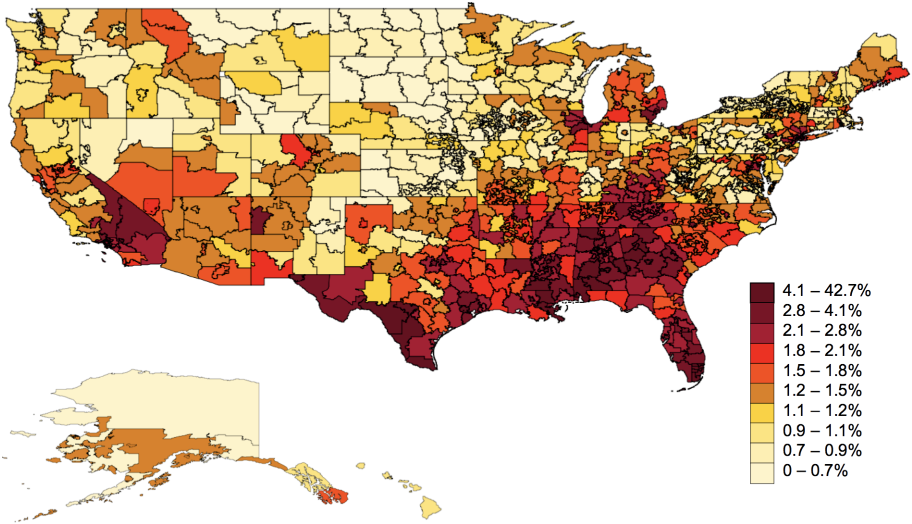
Earnings Distribution in Texas



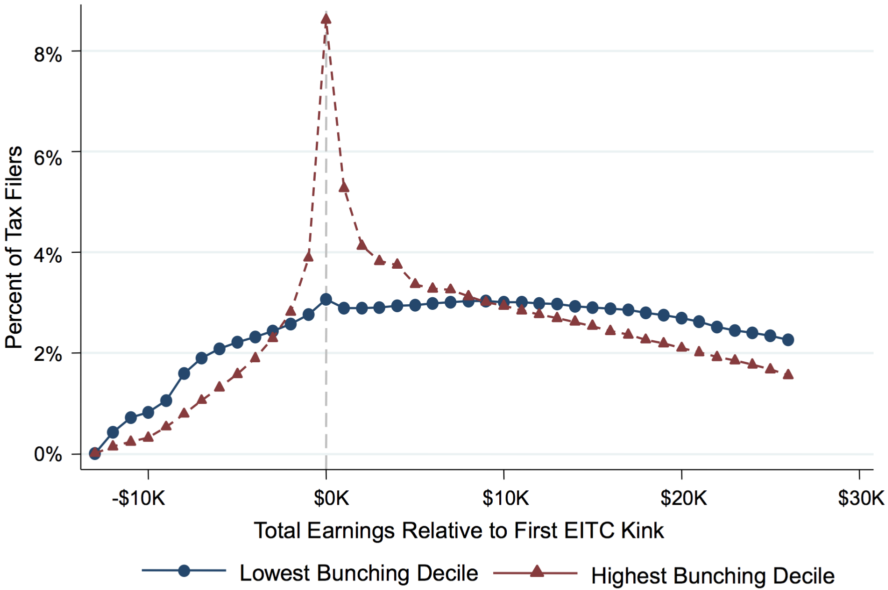
Earnings Distribution in Kansas

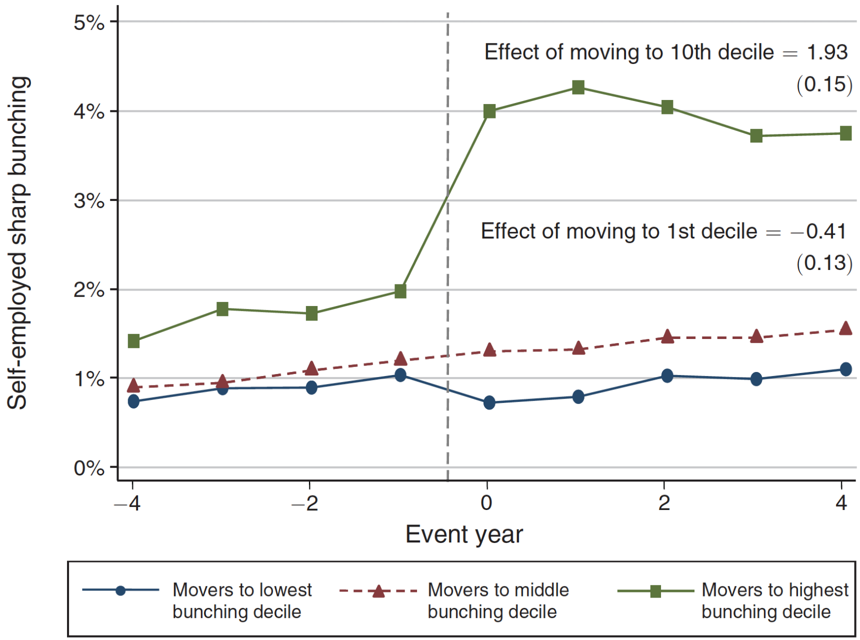


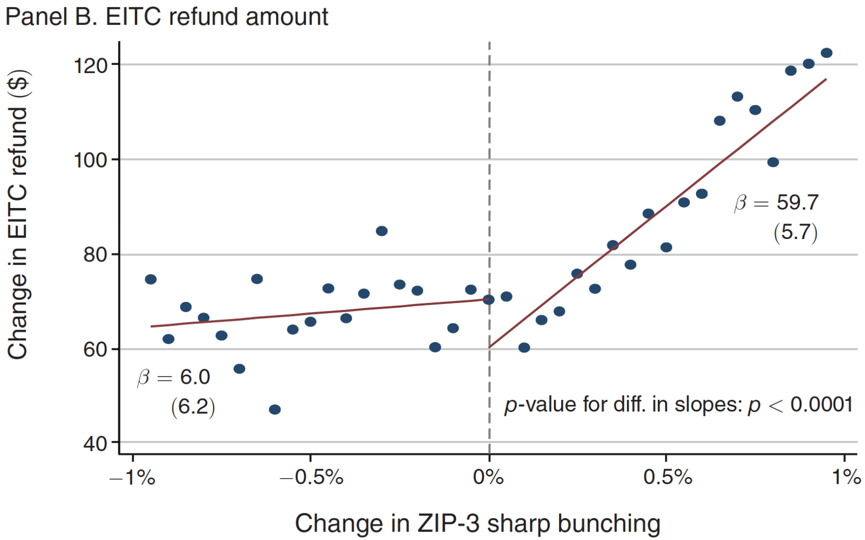
Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 2008



Earnings Distributions in Lowest and Highest Bunching Deciles

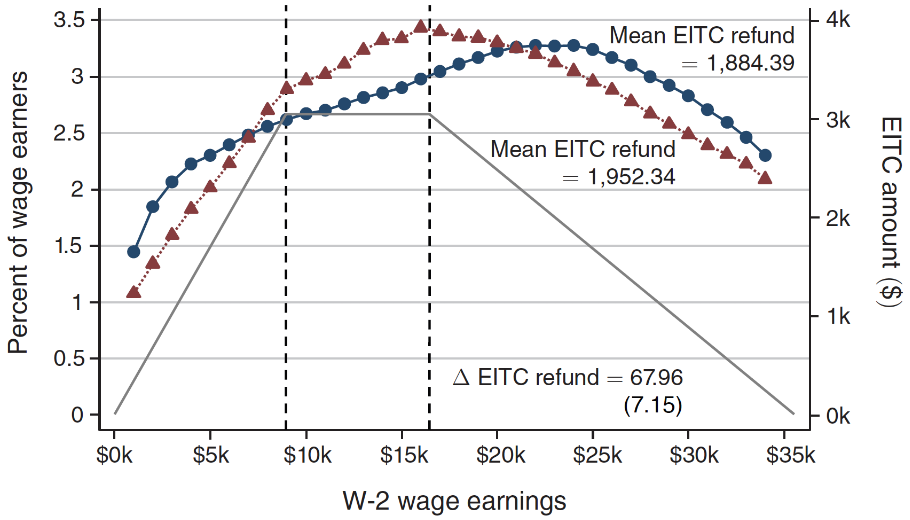



Effect of moving to 10th decile = 1.93
(0.15)
Effect of moving to 1st decile = −0.41
(0.13)
Self-employed sharp bunching
5%
4%
3%
2%
1%
0%
−4
−2
0
2
4
Event year
Movers to lowest bunching decile
Movers to middle bunching decile
Movers to highest bunching decile



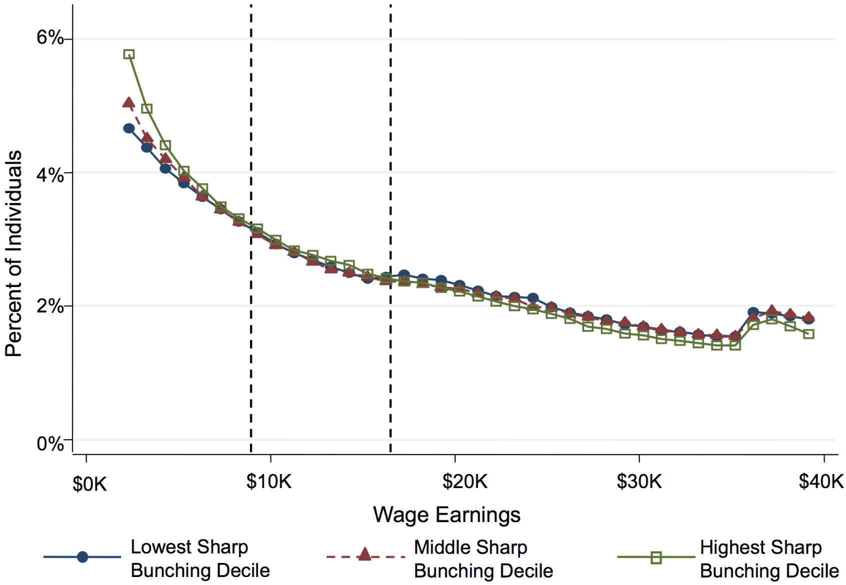
Panel B. EITC refund amount
Change in EITC refund ($)
120
100
80
60
40
−1%
−0.5%
0%
0.5%
1%
Change in ZIP-3 sharp bunching
$\beta = 6.0$
(6.2)
$\beta = 59.7$
(5.7)
p-value for diff. in slopes: $p < 0.0001$

Panel A. Wage earners with one child

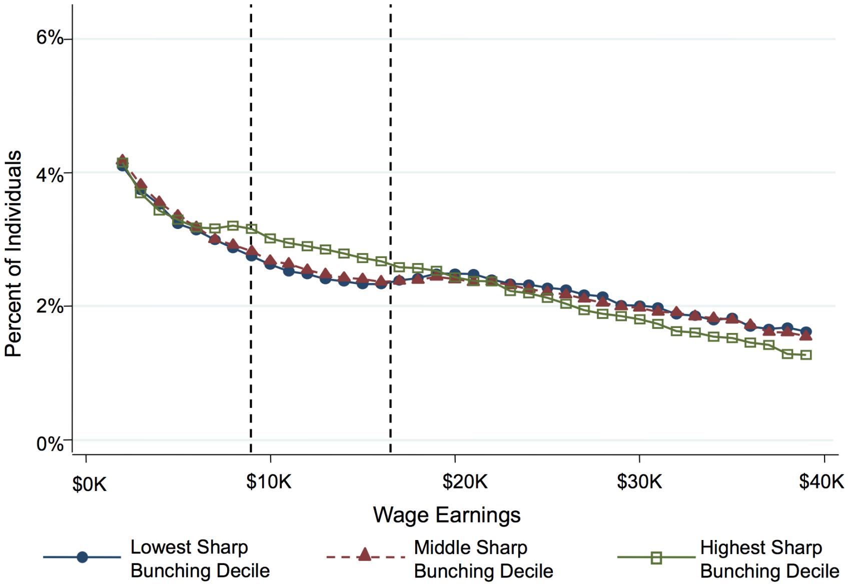


Difference in Wage Earnings Distributions Between Top and Bunching Decile Wage Earners with One Child

Earnings Distribution in the Year Before First Child Birth for Wage Earners



Earnings Distribution in the Year of First Child Birth for Wage Earners



Simulated EITC Credit Amount for Wage Earners Around First Child Birth

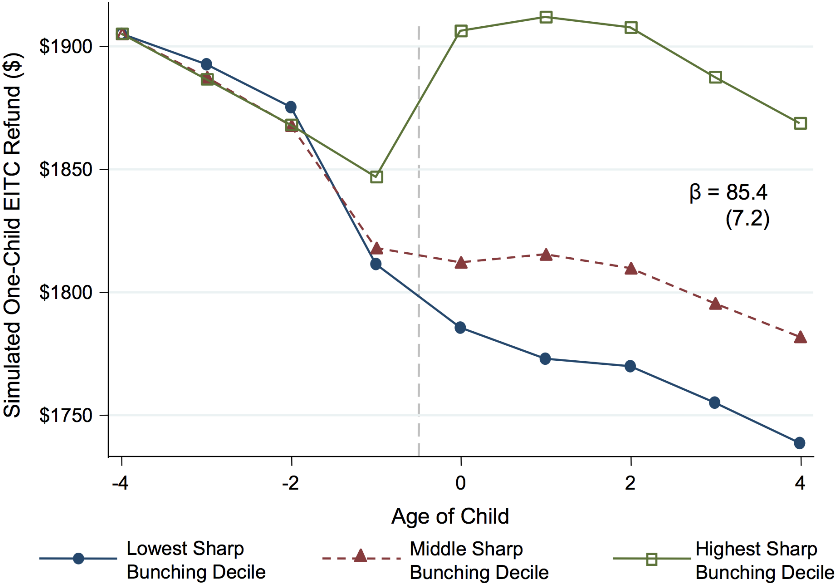


TABLE 3—ELASTICITY ESTIMATES BASED ON CHANGE IN EITC REFUNDS AROUND BIRTH OF FIRST CHILD

	Mean elasticity (1)	Phase-in elasticity (2)	Phase-out elasticity (3)	Extensive elasticity (4)
Panel A. Wage earnings				
Elasticity in US 2000–2005	0.21	0.31	0.14	0.19
	(0.012)	(0.018)	(0.015)	(0.019)
Elasticity in top decile ZIP-3s	0.55	0.84	0.29	0.60
	(0.020)	(0.031)	(0.020)	(0.034)
Panel B. Total earnings				
Elasticity in US 2000–2005	0.36	0.65	0.11	0.36
	(0.017)	(0.030)	(0.006)	(0.019)
Elasticity in top decile ZIP-3s	1.06	1.70	0.31	1.06
	(0.029)	(0.047)	(0.010)	(0.040)

TABLE 4—IMPACT OF EITC ON WAGE EARNINGS DISTRIBUTION OF EITC-ELIGIBLE HOUSEHOLDS

	Percent of EITC-eligible households below threshold			
	50 percent of poverty line (1)	100 percent of poverty line (2)	150 percent of poverty line (3)	200 percent of poverty line (4)
No EITC counterfactual	13.15	31.31	53.81	77.06
EITC with no behavioral response	8.92	21.37	41.56	70.82
EITC with avg. behavioral response in United States	8.16	21.00	41.97	71.29
EITC with top decile behavioral response	6.73	20.24	42.56	72.08

Outline

Evidence from Rich Countries

- Chetty, Friedman & Saez (AER 2013): *Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings*
- Deshpande & Li (2017) *Who is Screened Out? Application Costs and the Targeting of Disability Programs*
- Deshpande & Li (2017) [ET] *Who is Screened Out? Application Costs and the Targeting of Disability Programs*
- Finkelstein & Notowidigdo (2019) *Take-up and Targeting: Experimental Evidence from SNAP*

Deshpande & Li (2017): Overview

- When eligibility is hard to observe/verify, application costs can act as a screening device (ordeal).
- Study this in the context of disability insurance.
- Use closings of Social Security Administration offices to get variation in application costs.
- Find large impacts,
  - Fewer people apply
  - Targeting worsens
- Suggests social costs of closing SSA offices outweigh social benefits 5 to 1

Deshpande & Li (2017): A Simple Framework

- Consider an increase in application costs from η to $\eta' > \eta$.
- Assume adjudicators don't change their standards. Then targeting efficiency increases iff

$$\mathbb{P}\left(R|A,\eta'\right) > \mathbb{P}\left(R|A,\eta\right)$$

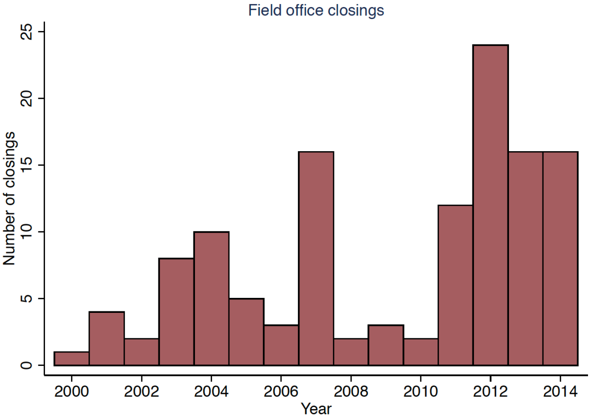
  where $\mathbb{P}\left(R|A,\eta\right)$ is the probability of receiving benefits conditional on applying.
- Rewrite: Targeting improves iff

$$1 < \frac{\mathbb{P}\left(R|A,\eta'\right)}{\mathbb{P}\left(R|A,\eta\right)} = \frac{\frac{\mathbb{P}(R|\eta')}{\mathbb{P}(A|\eta')}}{\frac{\mathbb{P}(R|\eta)}{\mathbb{P}(A|\eta)}} = \frac{\frac{\mathbb{P}(R|\eta')}{\mathbb{P}(R|\eta)}}{\frac{\mathbb{P}(A|\eta')}{\mathbb{P}(A|\eta)}} = \frac{\Delta_R + 1}{\Delta_A = 1}$$

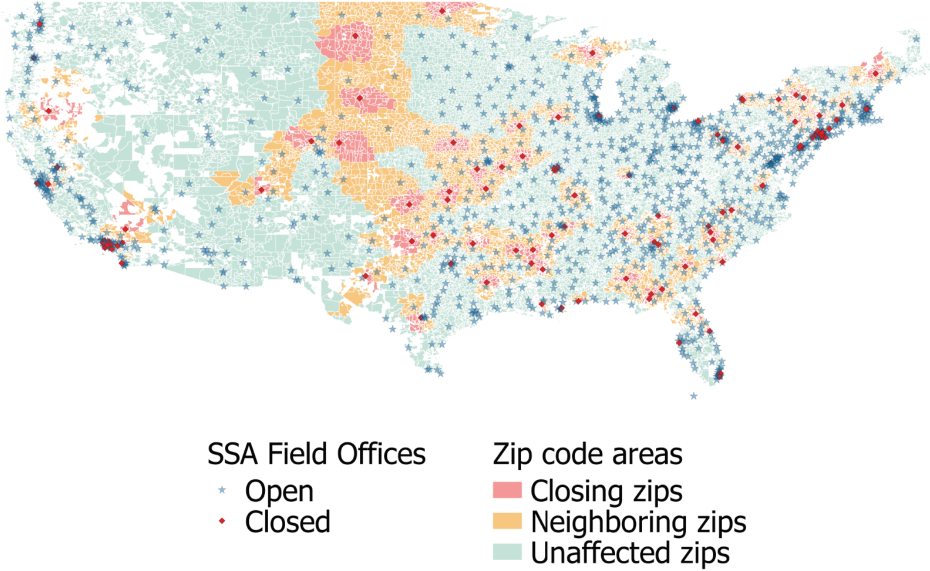
  where $\Delta_R = \frac{\mathbb{P}(R|\eta') - \mathbb{P}(R|\eta)}{\mathbb{P}(R|\eta)}$ and $\Delta_A = \frac{\mathbb{P}(A|\eta') - \mathbb{P}(A|\eta)}{\mathbb{P}(A|\eta)}$

Deshpande & Li (2017): Context

- To get Social Security Disability Insurance (SSDI) and Supplemental Security Income (SSI), you have to apply to the Social Security Administration.
- They review medical/work history, determine eligibility.
- 1,230 Social Security Field Offices provide in-person help.



Deshpande & Li (2017): SSA Office Closures



Deshpande & Li (2017): Data

- Data from the Social Security Administration
  - All field offices with addresses & closing dates
  - walk-in wait times
  - Staff numbers at each office
  - volume of calls to 800 numbers
  - application data
  - decision data
  - ZIP code of applicants
- Collapse by ZIP and assign each ZIP code a nearest, second nearest, and third nearest SSA office.

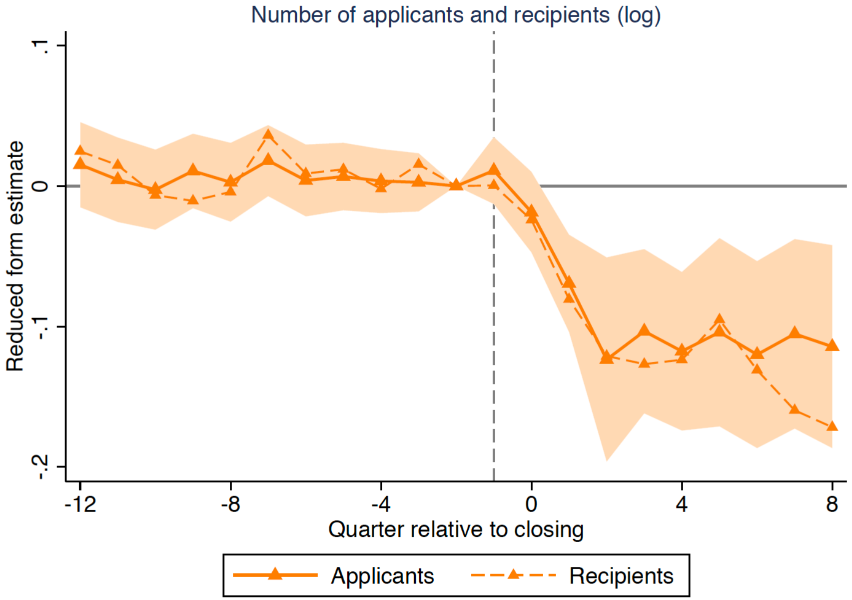
Deshpande & Li (2017): Empirical Approach

- Compare applications and receipts in the ZIP codes where the office closed to ZIP codes where the office hasn't closed yet, but will do.
- For each closure, take the ZIPs that experience the closure as treated, and ZIPS that experience a closure more than 2 years in the future are control
- Event Study estimates:

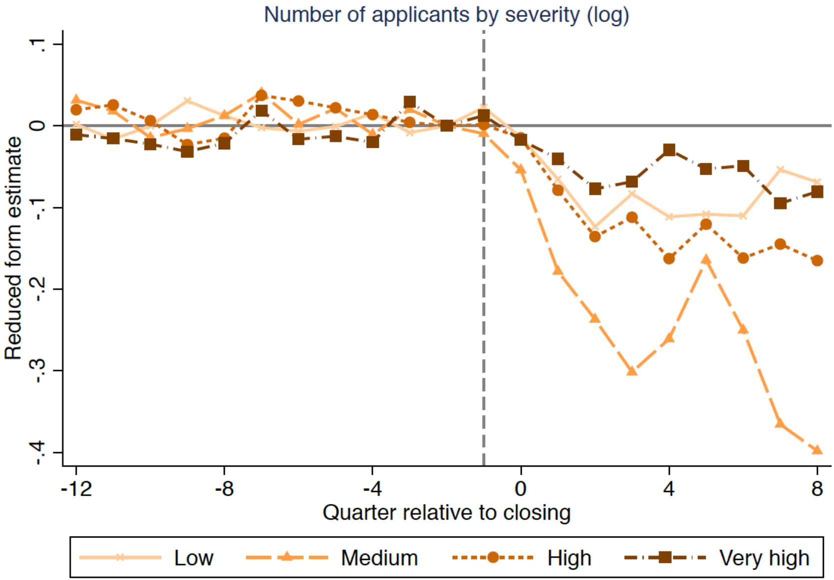
$$Y_{isct} = \alpha_i + \gamma_{st} + \sum_{\tau} D_{ct}^{\tau} + \sum \delta_{\tau} \left(Treated_{ic} \times D_{ct}^{\tau} \right) + \epsilon_{isct}$$

where Y_{isct} is outcome for ZIP i in state s for closure c in quarter t; α_i are ZIP FEs; γ_{st} are state$\times$quarter FEs; $Treated_{ic}$ indicates treatment; D_{ct}^{τ} are indicators for observations τ quarters before/after the closure.

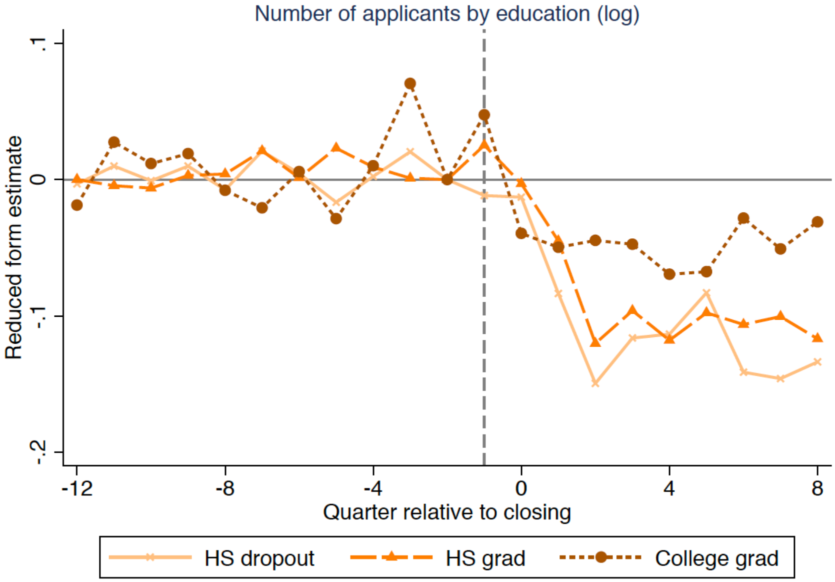
Deshpande & Li (2017): Results



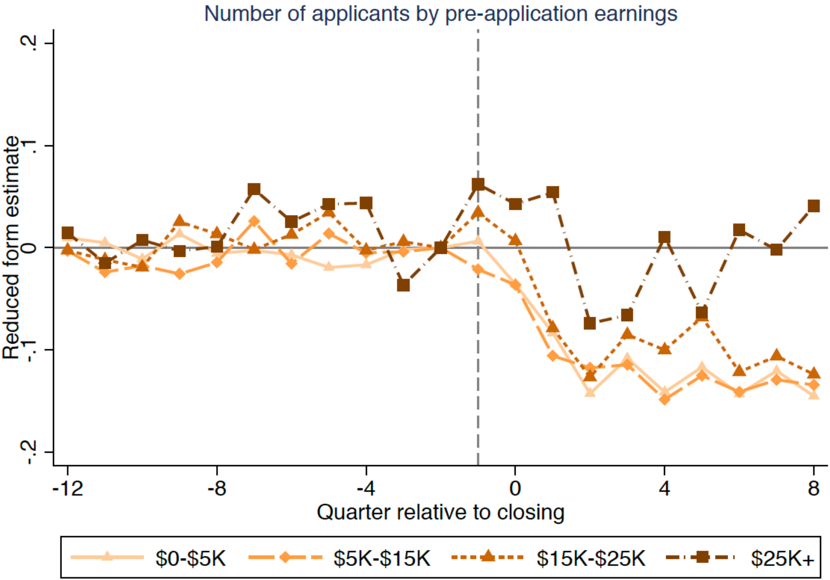
Deshpande & Li (2017): Subgroups



Deshpande & Li (2017): Subgroups



Deshpande & Li (2017): Subgroups



Deshpande & Li (2017): Spillovers

Table 4: Estimates of the Effect of Closings on Types of Application Costs

	Closing ZIP			Neighboring ZIP		
	Pt. Est.	Std. Err.	Mean	Pt. Est.	Std. Err.	Mean
Applications (log)	-0.110***	(0.0300)	39.7	-0.0539***	(0.0176)	42.5
Recipients (log)	-0.133***	(0.0312)	21.7	-0.0904***	(0.0181)	22.6
Congestion measures						
FO processing time	3.032***	(1.094)	28.8	2.804***	(0.731)	28.4
Walk-in wait times	4.352***	(1.412)	13.6	3.472***	(1.126)	16.3
Travel cost measures						
Driving time	9.974***	(1.636)	23.5			
Driving distance	11.98***	(1.338)	24.3			
Transit time	35.76***	(6.426)	89.4			

Deshpande & Li (2017): Mechanisms

- How do office closures affect applications?
  1. Congestion at neighboring offices
  2. Travel time

$$Y_{isct} = \alpha_i + \gamma_{st} + \beta Congestion_{ict} + \kappa Distance_{ict} + \delta NewOffice_{ict}$$

- To overcome endogeneity (why?) use closures and difference with next-closest office as instruments:

$$\begin{aligned} Congestion_{isct} &= \alpha_i + \gamma_{st} + \beta_1 \left(Treated_{ic} \times Post_t \times WDiff_{ic} \right) \\ &\quad + \beta_2 \left(TreatedNbr_{ic} \times Post_t \right) + \nu_{isct} \\ Distance_{isct} &= \alpha_i + \gamma_{st} + \kappa_1 \left(Treated_{ic} \times Post_t \times DDiff_{ic} \right) + \xi_{isct} \end{aligned}$$

where $WDiff_{isc}$ is difference in walk-in times with now-closest office; $TreatedNbr_{ic}$ indicates closing neighbor; $DDiff_{ic}$ is difference in driving distance with the now-closest office

Deshpande & Li (2017): Mechanisms

Table 5: IV Estimates of the Effect of Different Application Costs on Disability Applications

	First Stage		Red. Form	OLS	IV	Inc after	OLS Δ in	IV Δ in
	Driving dist	Wait time	Log(app)	Log(app)	Log(app)	closing	log(app)	log(app)
TrtXPostXDDiff	0.996***		-0.00320					
	(0.00442)		(0.00250)					
TrtXPostXWDiff		0.452***	-0.000454					
		(0.0981)	(0.000588)					
NbrXPost		4.564***	-0.0781***					
		(1.168)	(0.0107)					
New Office (TrtXPost)			-0.148***	-0.0830***	-0.0229	1	-0.083	-0.023
			(0.0348)	(0.0246)	(0.0221)			
Driving distance (km)				-0.000483	-0.000253	12	-0.006	-0.003
				(0.000431)	(0.000408)			
Wait time (min)				-0.00273***	-0.0158***	4.3	-0.012	-0.068
				(0.000647)	(0.00327)			
N	101,008	80,779	98,557	77,786	76,280			

Deshpande & Li (2017): Welfare

- Let the benefits of approving a disability application be

$$b_r(n) \equiv b_1 - c_1 - c_2(n)$$

  where b_1 is net social benefit of providing benefits, c_1 is cost of reviewing application and c_2 is the cost to the applicant

- Let the costs of rejecting an application be

$$b_n(n) \equiv c_1 + c_2(n)$$

- And let F be the cost of running an office. Then social benefit of n offices is

$$\begin{aligned} W(n) &= b_r(n)\, r(n) - b_n(n)\left[a(n) - r(n)\right] - Fn \\ &= b_1 r(n) - \left[c_1 + c_2(n)\right] a(n) - Fn \end{aligned}$$

Deshpande & Li (2017): Welfare

- And so change in welfare from closing one office is

$$
\begin{aligned}
W(n-1) - W(n) = & \underbrace{b_1 \left[r(n-1) - r(n)\right]}_{\text{benefits to fewer recipients}} \\
& - \underbrace{\left[c_2(n-1) - c_2(n)\right] a(n-1)}_{\text{higher costs for applicants}} \\
& - \underbrace{c_1 \left[a(n-1) - a(n)\right]}_{\text{review fewer applications}} \\
& - \underbrace{c_2(n) \left[a(n-1) - a(n)\right]}_{\text{fewer applicants}} + F
\end{aligned}
$$

Deshpande & Li (2017): Welfare

Table 6: Costs and Benefits of Field Office Closings

Costs of closing (thousands)	
Lower receipt in areas surrounding closed office	\$3,100
Lower receipt in areas surrounding neighboring office	\$13,000
Higher applicant time and earnings decay	\$3,200
Total	\$19,300
Benefits of closing (thousands)	
Administrative savings from processing fewer applications	\$2,600
Administrative savings from closing field office	\$500
Application cost savings from discouraged applicants	\$1,000
Total	\$4,100
Ratio of costs to benefits	5

Outline

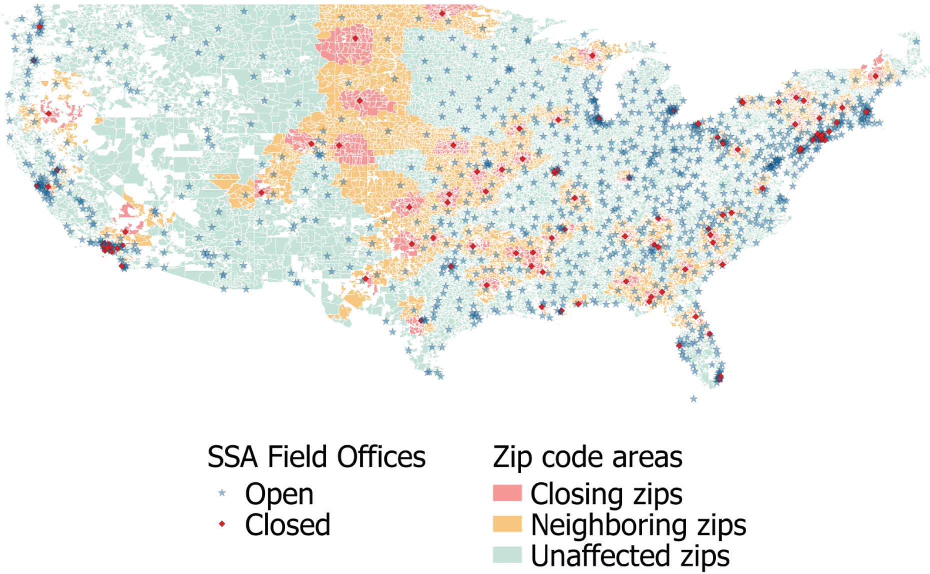
Evidence from Rich Countries

- Chetty, Friedman & Saez (AER 2013): *Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings*
- Deshpande & Li (2017) *Who is Screened Out? Application Costs and the Targeting of Disability Programs*
- Deshpande & Li (2017) [ET] *Who is Screened Out? Application Costs and the Targeting of Disability Programs*
- Finkelstein & Notowidigdo (2019) *Take-up and Targeting: Experimental Evidence from SNAP*

Deshpande & Li (2017): Overview [ET]

- Nichols and Zeckhauser (1982): Application costs can improve targeting by screening out high-ability individuals with a high opportunity cost of time. Productive efficiency loss (lost time) is more than offset by the gain in targeting efficiency (less wastage).
- vs. Bertrand, Mullainathan and Shafir (2004): Hassles may discourage those most in need. Inverse U curve? Most likely depends on type of hassle and characteristics of the marginal population.
- Empirically test for this debate in Disability Programs using the timing of closings of SSA field offices that provide assistance with filing disability applications.
- Estimate the effect of increase in application costs on the number and composition of disability applicants and recipients using diff-in-diff.
- Find that disability recipients declines, reduces targeting efficiency...

Deshpande & Li (2017): SSA Office Closures [ET]



Deshpande & Li (2017): Model Framework [ET]

- Application costs increase from η to $\eta' > \eta$.
- Assuming disability inspectors don't change their standards, targeting efficiency increases iff

$$\mathbb{P}\left(R|A,\eta'\right) > \mathbb{P}\left(R|A,\eta\right)$$

- $\mathbb{P}\left(R|A,\eta\right)$: probability of receiving benefits conditional on applying.
- Rewrite: Targeting improves iff

$$1 < \frac{\mathbb{P}\left(R|A,\eta'\right)}{\mathbb{P}\left(R|A,\eta\right)} = \frac{\frac{\mathbb{P}(R|\eta')}{\mathbb{P}(A|\eta')}}{\frac{\mathbb{P}(R|\eta)}{\mathbb{P}(A|\eta)}} = \frac{\frac{\mathbb{P}(R|\eta')}{\mathbb{P}(R|\eta)}}{\frac{\mathbb{P}(A|\eta')}{\mathbb{P}(A|\eta)}} = \frac{\Delta_R + 1}{\Delta_A = 1}$$

where $\Delta_R = \frac{\mathbb{P}(R|\eta') - \mathbb{P}(R|\eta)}{\mathbb{P}(R|\eta)}$ and $\Delta_A = \frac{\mathbb{P}(A|\eta') - \mathbb{P}(A|\eta)}{\mathbb{P}(A|\eta)}$ = % change in recipients/applicants resulting frrom the closing.

Deshpande & Li (2017): Context [ET]

- SSDI (Social Security Disability Insurance) and SSI (Supplemental Security Income)
- SSDI (Online/Phone/SSA Field Office), processed by field office assigned to applicant zip code.
- Verified by disability examiner at DDS (disability determination services office)
- Disability claims take the most time of SSA staff.

Deshpande & Li (2017): Idea and Data [ET]

- Use field office closings to study the effect of application costs on selection into disability programs.
- BUT SSA might be closing offices in areas where disability applications are already falling or where composition of disability applicants is already changing. –>
  - Treatment: Areas that expeirience a closing today.
  - Control: Areas that experience a closing in the future
  - (assumes exact timing of the closing is uncorrelated with changes in the number and type of disability applicants.
- Data: SSA Application data + disability diecision, incl ZIP, + SSA data field office data, walk-in wait times, number of staff members, volume of calls
- Constructed: great-circle + Google Map Driving distance to assign nearest, second nearest, and third-nearest field offices.

Deshpande & Li (2017): Empirical Strategy [ET]

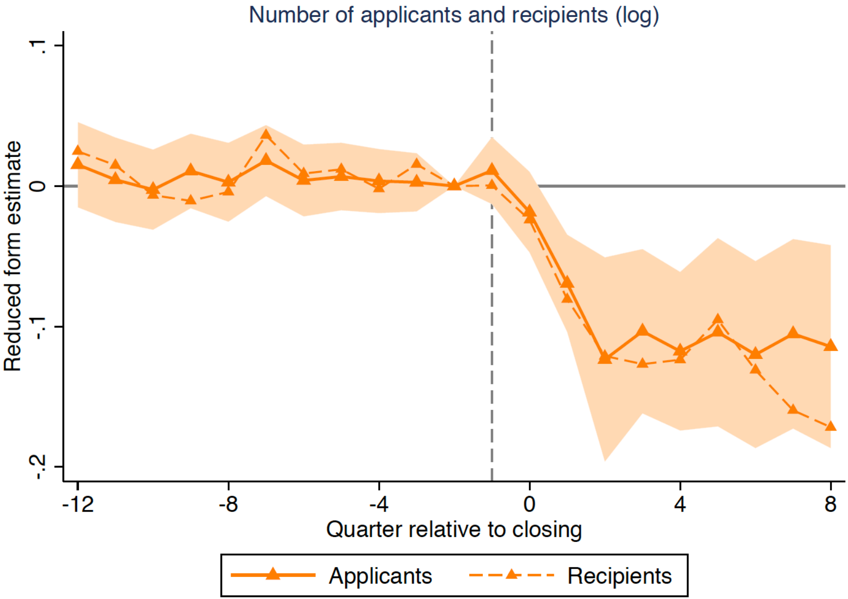
$$Y_{isct} = \alpha_i + \gamma_{st} + \sum_{\tau} D_{ct}^{\tau} + \sum \delta_{\tau} \left(Treated_{ic} \times D_{ct}^{\tau} \right) + \epsilon_{isct}$$

- zip i in state s for closing c in quarter t, D_{ct}^{τ} indicator variables before and after the closing.
- Treated indicator for whether it is the zip that actually closed for that c.

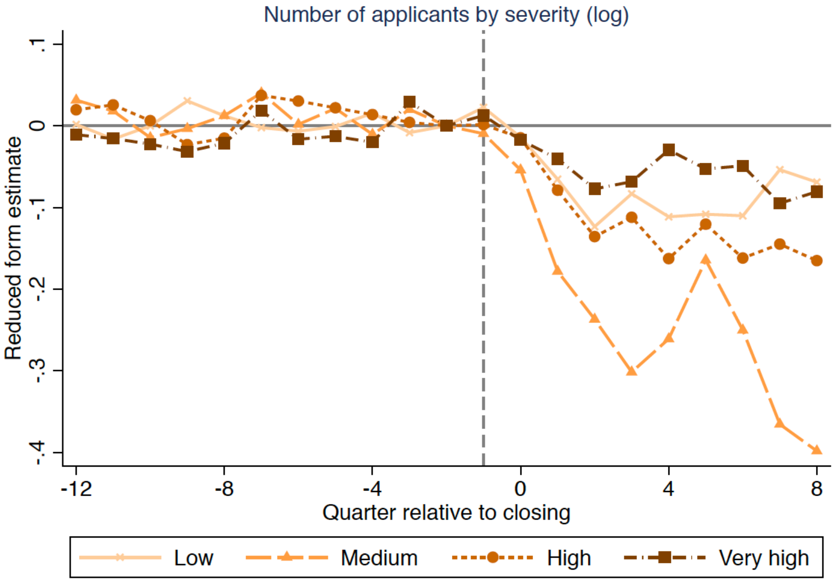
$$Y_{isct} = \alpha_i + \gamma_{st} + \sum_{\tau} D_{ct}^{\tau} + \beta(Treated_{ic} \times Post_{ct}) + \kappa \left(Treated_{ic} \times Zero_{ct} \right) + \epsilon_{is}$$

- Post is indicator if quarter t is after the closing and Zero is an indicator equal to 1 if quarter t is the quarter of the closing.

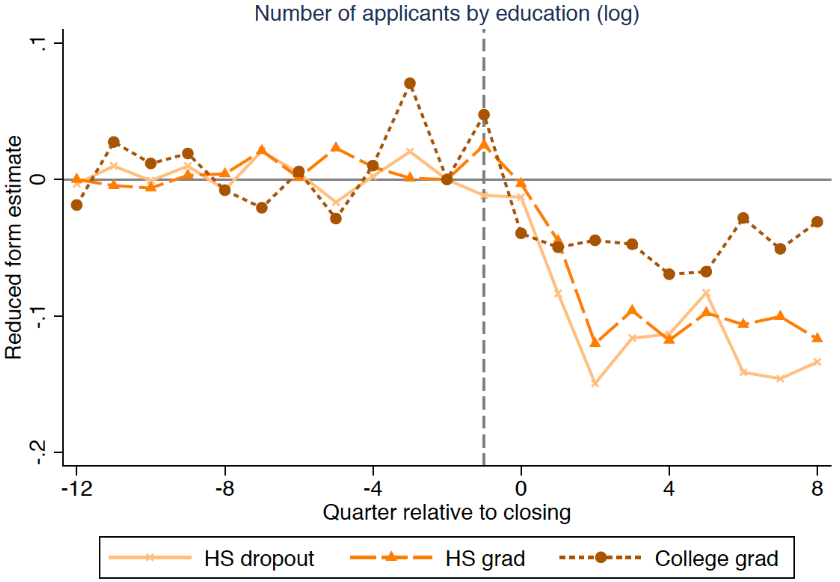
Deshpande & Li (2017): Results [ET]



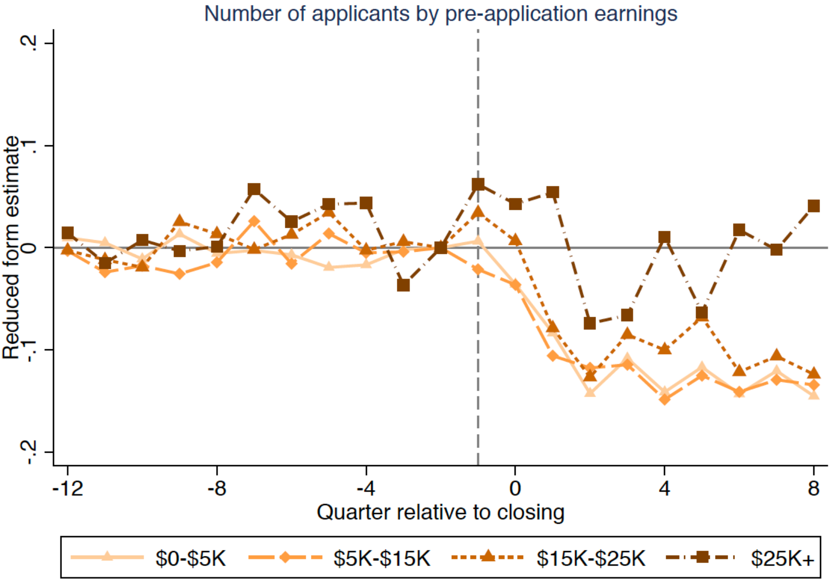
Deshpande & Li (2017): Subgroups [ET]



Deshpande & Li (2017): Subgroups [ET]



Deshpande & Li (2017): Subgroups [ET]



Deshpande & Li (2017): Spillovers [ET]

Table 4: Estimates of the Effect of Closings on Types of Application Costs

	Closing ZIP			Neighboring ZIP		
	Pt. Est.	Std. Err.	Mean	Pt. Est.	Std. Err.	Mean
Applications (log)	-0.110***	(0.0300)	39.7	-0.0539***	(0.0176)	42.5
Recipients (log)	-0.133***	(0.0312)	21.7	-0.0904***	(0.0181)	22.6
Congestion measures						
FO processing time	3.032***	(1.094)	28.8	2.804***	(0.731)	28.4
Walk-in wait times	4.352***	(1.412)	13.6	3.472***	(1.126)	16.3
Travel cost measures						
Driving time	9.974***	(1.636)	23.5			
Driving distance	11.98***	(1.338)	24.3			
Transit time	35.76***	(6.426)	89.4			

Outline

Evidence from Rich Countries

- Chetty, Friedman & Saez (AER 2013): *Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings*
- Deshpande & Li (2017) *Who is Screened Out? Application Costs and the Targeting of Disability Programs*
- Deshpande & Li (2017) [ET] *Who is Screened Out? Application Costs and the Targeting of Disability Programs*
- Finkelstein & Notowidigdo (2019) *Take-up and Targeting: Experimental Evidence from SNAP*

Finkelstein & Notowidigdo (2019): Overview

- Incomplete take-up of safety net programs is pervasive.
- 2 typical explanations:
  1. Lack of information about eligibility
  2. Transaction costs of enrollment (including stigma)
- How does incomplete takeup affect social welfare?
  - "Ordeals" may serve as userful screens (Nichols & Zeckhauser 1982, Besley & Coate 1992)
  - Behavioral work suggesting ordeals have higher psychic costs for the poor $\Rightarrow$ Worsening targeting effectiveness.
- Ultimately empirical question

Finkelstein & Notowidigdo (2019): Overview

- This paper:
- framework for welfare effects of ordeal changes: How many people enrol? Which types of people?
- Apply framework to RCT of interventions aimed at elderly non-participants in SNAP
- 30K subjects
- 3 treatment arms

1. Information about eligibility
2. Information plus assistance with application
3. control

Finkelstein & Notowidigdo (2019): Framework

- 2 types of individuals: $j \in \{L, H\}$ with unobserved wages θ_j, $\theta_H > \theta_L$. Unit mass of each type.
- Individuals choose hours h_j and earn pre-tax income $\theta_j h_j$, incurring taxes $\tau(\theta_j h_j)$, leading to net of tax earning $y_j \equiv \theta_j h_j - \tau(\theta_j h_j)$.
- The program grants benefits B if income is below r^*.
- Individuals misperceive the benefit by ϵ_j so their *perceived* benefit of applying is $(1 + \epsilon_j) B$.
    - $\epsilon_j < 0$: underestimate benefits of applying
    - $\epsilon_j = 0$ for $j \in \{L, H\}$ is the neoclassical benchmark.

Finkelstein & Notowidigdo (2019): Framework

- Individuals have common utility
  - $u(x_j) - v(h_j)$ if they do not apply
  - $u(x_j) - v(h_j) - \left(\bar{\Lambda}\kappa_j + c\right)$ if they apply
- Disutility from applying:
  - c: individual-specific utility cost of applying. has distribution $f_j(c)$
  - $\bar{\Lambda}$: parameter affecting utility cost
  - κ_j: how this utility cost varies across types. e.g. ordeals can impose greater utility cost on H types $(\kappa_H > \kappa_L)$ or low types $(\kappa_H < \kappa_L)$

Finkelstein & Notowidigdo (2019): Framework

- Individuals make application and labor supply choices.
- Type js work h_j^A hours if they apply, h_j^{-A} if they don't.
- Low types: assume $y_j^A, y_j^{-A} < r^*$
- High types: assume $y_j^{-A} > r^*$ and to get the benefit, if apply $h_H^A = r^*/\theta_H$
- NB $h_j^A \leq h_j^{-A}$ for both types
- Individuals with $c < c_j^*$ will apply. Private welfare is therefore

$$
\begin{aligned}
V_j &= \mathbb{P}\left(\text{apply}\right)\mathbb{E}\left[U\left(\right)|\text{apply}\right] + \mathbb{P}\left(\cancel{\text{apply}}\right)\mathbb{E}\left[U\left(\right)|\ \cancel{\text{apply}}\right] \\
&= \int_0^{c_j^*} \left(u\left(y_j^A + B\right) - v\left(h_j^A\right) - \bar{\Lambda}\kappa_j - c\right) dF_j(c) \\
&\quad + \int_{c_j^*}^{\infty} \left(u\left(y_j^{-A}\right) - v\left(h_j^{-A}\right)\right) dF_j\left(c\right)
\end{aligned}
$$

Finkelstein & Notowidigdo (2019): Framework

- Individuals' decisions create a fiscal externality: Their hours decisions change tax revenue and this has a social cost that they don't internalize. For people who apply, the fiscal externality is $G_j^A = \tau\left(h_j^A \theta_j\right)$ while for those who don't it's $G_j^{-A} = \tau\left(h_j^{-A} \theta_j\right)$
- We can write total social welfare as

$$W = \underbrace{V_L + V_H}_{\text{Private Welfare}} - \underbrace{B(A_L + A_H)}_{\text{Program Cost}}$$
$$+ \underbrace{A_L G_L^A + (1 - A_L) G_L^{-A} + A_H G_H^A + (1 - A_H) G_H^{-A}}_{\text{Fiscal Externality}}$$

  where $A_j = F_j(c^*)$, the expected number of applicants of type j.
- Planner chooses tax system $\tau()$ and transfer program (including ordeal $\bar{\Lambda}$) to maximize W.

Finkelstein & Notowidigdo (2019): Framework

- Characterize the marginal effects on welfare of the two treatments.
- Information treatment: Increase perceived benefits by $d\epsilon$
- Information plus assistance: increase benefits $d\epsilon$ and reduce cost of applying, $-d\bar{\Lambda}$
- Define:

$$\mu_j \equiv u\left(y_j^A + B\right) - u\left(y_j^A + (1+\epsilon_j)\,B\right) \quad \text{and} \quad \xi_j \equiv u'\left(y_j^A + B\right)$$

Finkelstein & Notowidigdo (2019): Framework

- We can write the effect of the information treatment on welfare as

$$
\begin{aligned}
\frac{dW}{dT}^{\text{Information only}} = & \underbrace{\mu_L \frac{dA_L}{dT} + \mu_H \frac{dA_H}{dT}}_{\text{Change in Private Welfare}} \\
& - \underbrace{B\left(\frac{dA_L}{dT} + \frac{dA_H}{dT}\right)}_{\text{Change in Mechanical Program Costs}} \\
& + \underbrace{\left(G_L^A - G_L^{-A}\right)\frac{dA_L}{dT} + \left(G_H^A - G_H^{-A}\right)\frac{dA_H}{dT}}_{\text{Fiscal Externality}}
\end{aligned}
$$

Finkelstein & Notowidigdo (2019): Framework

- We can write the effect of the information plus assistance treatment as

$$
\begin{aligned}
\frac{dW}{dT}^{\text{Information + Assistance}} &= \underbrace{\mu_L \frac{dA_L}{dT} + \mu_H \frac{dA_H}{dT} + \kappa_L A_L + \kappa_H A_H}_{\text{Change in Private Welfare}} \\
&- \underbrace{B\left(\frac{dA_L}{dT} + \frac{dA_H}{dT}\right)}_{\text{Change in Mechanical Program Costs}} \\
&+ \underbrace{\left(G_L^A - G_L^{-A}\right)\frac{dA_L}{dT} + \left(G_H^A - G_H^{-A}\right)\frac{dA_H}{dT}}_{\text{Fiscal Externality}}
\end{aligned}
$$

Finkelstein & Notowidigdo (2019): Framework

- Define targeting as $e = E_L / (E_L + E_H)$, the share of enrollees who are L types.
- Then a treatment T increases targeting if $de/dT > 0$
- *Holding constant the change in applications due to an intervention, the change in social welfare in response to an improvement in targeting* $(de/dT > 0)$ *from an Information Only (or Information Plus Assistance) treatment is given by the following expression:*

$$\frac{\partial}{\partial (de/dT)} \left(\frac{dW}{dT} \right) \Bigg|_{\frac{dA}{dT}} = \left[(\mu_L - \mu_H) + \left(G_L^A - G_L^{-A} \right) - \left(G_H^A - G_H^{-A} \right) \right] \times (E_H + E_L)$$

- Note this is increasing in $\mu_L - \mu_H$. Sufficient condition for better targeting to increase private welfare is $\epsilon_L \leq \epsilon_H \leq 0$ with at least 1 strict inequality, or $\epsilon_H = \epsilon_L < 0$

Finkelstein & Notowidigdo (2019): Empirical Setting

- Work with households aged 60 or over in Pennsylvania in 2016
- To enroll in SNAP, individuals must complete an application, provide documents verifying circumstances, and do an interview (in person or by phone). Provide info on income, various household expenses (medical expenses, rent, utilities etc).
- Once enrolled, eligible for 36 months of benefits. Benefits are a decreasing function of net income.
- SNAP benefits average 15% of annual income for eligible households.

Finkelstein & Notowidigdo (2019): Interventions

- Partnered with Benefits Data Trust (BDT), a national non-profit working to help people access benefits.
- State of Pennsylvania provided BDT admin data on individuals aged >60 and on Medicaid but not SNAP. Likely to be eligible for SNAP.
- Randomized the 30K individuals into equally-sized groups.
- Information Treatment got outreach materials informing them of likely SNAP eligibility and benefits they might receive, and info on how to call DHS to apply.
- Assistance component: Info Treatment contained BDT number. If individuals call BDT, they help with the application. Advise on documents, can populate and submit application on their behalf etc.
- BDT submitted $\sim 70\%$ of applications in this treatment arm, spent on average 47 mins on phone with applicants, 30 with those who don't end up applying.

Finkelstein & Notowidigdo (2019)

Table 1: Description of Study Population

	Original Outreach List	List, After Exclusions	After Exclusions Receving SNAP	After Exclusions Not Receiving SNAP	Study Population
	(1)	(2)	(3)	(4)	(5)
Observations (N)	229,584	143,923	84,038	59,885	31,888
Panel A - Demographics					
Age (as of October 31, 2015)	72.91	70.45	69.77	71.42	68.83
Share Age above Median = 65	0.72	0.66	0.66	0.66	0.50
Share Age 80+	0.27	0.18	0.15	0.23	0.16
Male	0.35	0.36	0.36	0.36	0.38
Share White[a]	0.71	0.79	0.79	0.79	0.75
Share Black[a]	0.17	0.10	0.11	0.07	0.08
Share Primary Language not English	0.04	0.03	0.03	0.03	0.04
Share Living in Philadelphia	0.18	0.00	0.00	0.00	0.00
Share Living in Pittsburgh	0.05	0.07	0.07	0.06	0.06
Share Last Medicaid Spell Starting before 2011	0.45	0.47	0.55	0.36	0.33
Share Enrolled in Medicaid for 2015 Full Year	0.83	0.84	0.89	0.77	0.73
Panel B - (Annual) Health Care Measures, 2015					
Total Health Care Spending ($)[b]	18,347	7,683	6,036	9,995	11,838
Number of Hospital Days	5.41	1.51	1.24	1.88	2.16
Number of ER Visits	0.41	0.41	0.41	0.40	0.50
Number of Doctor Visits	6.25	5.87	5.97	5.74	7.11
Number of SNF Days	66.23	1.57	0.85	2.58	2.67
Number of Chronic Conditions	6.50	4.93	5.08	4.70	5.45

Finkelstein & Notowidigdo (2019)

Table 2: Behavioral Responses to "Information Only" and "Information Plus Assistance"

	Control	Information Only	Information Plus Assistance	P Value of Difference (Column 2 vs 3)
	(1)	(2)	(3)	(4)
SNAP Enrollees	0.058	0.105	0.176	
		[0.000]	[0.000]	[0.000]
SNAP Applicants	0.077	0.147	0.238	
		[0.000]	[0.000]	[0.000]
SNAP Rejections among Applicants	0.233	0.266	0.255	
		[0.119]	[0.202]	[0.557]
Callers	0.000	0.267	0.301	
		[0.000]	[0.000]	[0.000]
Adjusted Callers	0.000	0.289	0.301	
		[0.000]	[0.000]	[0.156]

Finkelstein & Notowidigdo (2019)

SNAP Applicants among Non-Callers	0.077	0.086	0.081	
		[0.063]	[0.324]	[0.363]
SNAP Applicants among Callers	0.000	0.313	0.602	
		[0.000]	[0.000]	[0.000]
SNAP Enrollees among Non-Callers	0.058	0.061	0.059	
		[0.442]	[0.713]	[0.688]
SNAP Enrollees among Callers	0.000	0.226	0.450	
		[0.000]	[0.000]	[0.000]
Observations (N)	10,630	5,314	10,629	

Notes: Columns 1 through 3 shows means by intervention arm with the p-value (relative to the control arm) in [square brackets]. Column 1 shows the control. Column 2 shows the Information Only arm (for the two equally-sized pooled sub-treatments). Column 3 shows the Information Plus Assistance arms (weighted so that the two pooled sub-treatments received equal weight). Column 4 reports the p-value of the difference between the Information Plus Assistance and Information Only treatment arms. All outcomes are binary rates measured during the nine months from the initial mail date. All p-values are based on heteroskedasticity-robust standard errors. Callers are measured for the relevant call number and are therefore mechanically zero for the control; see text for a description of the adjusted caller rate.

Finkelstein & Notowidigdo (2019)

Table 4: Enrollee Monthly Benefits and Predicted Benefits

	Control	Information Only	Information Plus Assistance	P Value of Difference (Column 2 vs 3)
	(1)	(2)	(3)	(4)
Benefit Amount	145.94	115.38	101.32	
		[0.000]	[0.000]	[0.013]
Share $16 Benefit	0.192	0.312	0.367	
		[0.000]	[0.000]	[0.021]
Share $194 Benefit	0.206	0.164	0.147	
		[0.076]	[0.003]	[0.352]
Share $357 Benefit	0.060	0.052	0.040	
		[0.587]	[0.077]	[0.259]
Share Missing Benefit	0.073	0.043	0.028	
		[0.025]	[0.000]	[0.139]

Finkelstein & Notowidigdo (2019)

Predicted Benefit for Enrollees w/ Actual Benefit	140.20	112.49	102.93	
		[0.000]	[0.000]	[0.086]
Predicted Benefit for All Enrollees	138.65	114.01	104.03	
		[0.000]	[0.000]	[0.068]
Share of Enrollees in Household Size of 1	0.657	0.714	0.760	
		[0.038]	[0.000]	[0.036]
Benefit Amount for Enrollees in Household Size of 1	116.97	93.35	85.82	
		[0.000]	[0.000]	[0.134]
Observations (N)	613	559	1,861	

Notes: Sample is individuals who enrolled in the 9 months after their initial mailing. Columns 1 through 3 shows means by intervention arm with the p-value (relative to the control arm) in [square brackets] for SNAP enrollees. Column 1 shows the control. Column 2 shows the Information Only arm (with the two equally-sized sub-treatments pooled). Column 3 shows the Information Plus Assistance arms (weighted so that the two pooled sub-treatments received equal weight). Column 4 reports the p-value of the difference between the Information Plus Assistance and Information Only treatment arms. See text for a description of the predicted benefits. All p-values are based on heteroskedasticity-robust standard errors. N reports the sample size of enrollees.

Finkelstein & Notowidigdo (2019)

Table 5: Demographic and Health Characteristics: Applicants and Enrollees

	Applicants				Enrollees			
	Means			P Value	Means			P Value
	Control	Info Only	Info Plus Assistance	Info Plus Assistance vs Info Only	Control	Info Only	Info Plus Assistance	Info Plus Assistance vs Info Only
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A - Predicted Benefits								
Predicted Benefits	148.26	125.65	115.36		138.65	114.01	104.03	
		[0.000]	[0.000]	[0.037]		[0.000]	[0.000]	[0.068]
Panel B - (Annual) Health Care Measures, 2015								
Total Health Care Spending ($) [a]	9,424	8,605	8,334		10,238	9,532	8,603	
		[0.517]	[0.300]	[0.781]		[0.661]	[0.208]	[0.459]
Total Number of Visits and Days	13.33	11.67	9.92		14.79	10.90	9.92	
		[0.331]	[0.018]	[0.166]		[0.058]	[0.008]	[0.467]
Weighted Total Number of Visits and Days	4,661	3,273	2,818		5,407	3,288	2,779	
		[0.128]	[0.022]	[0.442]		[0.064]	[0.011]	[0.461]
Number of Chronic Conditions	6.21	5.55	5.27		6.54	5.43	5.37	
		[0.094]	[0.006]	[0.383]		[0.019]	[0.005]	[0.875]

Finkelstein & Notowidigdo (2019)

<u>Panel C - Demographics</u>								
Share Age above Median = 65	0.41	0.46	0.46		0.39	0.43	0.46	
		[0.072]	[0.014]	[0.764]		[0.282]	[0.006]	[0.159]
Share Age 80+	0.06	0.11	0.14		0.07	0.12	0.14	
		[0.001]	[0.000]	[0.042]		[0.005]	[0.000]	[0.085]
Male	0.41	0.40	0.38		0.39	0.42	0.38	
		[0.983]	[0.232]	[0.250]		[0.446]	[0.444]	[0.104]
Share White [b]	0.67	0.73	0.74		0.71	0.78	0.78	
		[0.005]	[0.000]	[0.554]		[0.004]	[0.001]	[0.958]
Share Black [b]	0.10	0.08	0.11		0.11	0.07	0.10	
		[0.103]	[0.577]	[0.011]		[0.011]	[0.833]	[0.004]
Share Primary Language not English	0.08	0.06	0.04		0.06	0.05	0.03	
		[0.141]	[0.000]	[0.012]		[0.242]	[0.002]	[0.067]
Share Living in Pittsburgh	0.05	0.06	0.07		0.05	0.06	0.07	
		[0.385]	[0.066]	[0.459]		[0.374]	[0.028]	[0.310]
Share Last Medicaid Spell Starting before 2011	0.25	0.30	0.29		0.26	0.33	0.31	
		[0.022]	[0.017]	[0.704]		[0.009]	[0.026]	[0.348]
Observations (N)	817	781	2,519		613	559	1,861	

Finkelstein & Notowidigdo (2019)

Table 6: Demographic and Health Characteristics: Callers and non-Callers

	Callers (1)	Non-callers (2)	P Value of Difference (3)
Panel A - Predicted Benefits			
Predicted Benefits	106.99	114.68	[0.000]
Predicted Enrollment	0.05	0.05	[0.752]
Panel B - (Annual) Health Care Measures, 2015			
Total Health Care Spending ($) [a]	7,316	13,656	[0.000]
Total Number of Visits and Days	9.52	13.50	[0.000]
Weighted Total Number of Visits and Days	2,853	5,064	[0.000]
Number of Chronic Conditions	5.16	5.48	[0.024]

Finkelstein & Notowidigdo (2019)

Panel C - Demographics			
Share Age above Median = 65	0.49	0.51	[0.014]
Share Age 80+	0.16	0.17	[0.190]
Male	0.38	0.38	[0.977]
Share White [b]	0.77	0.74	[0.000]
Share Black [b]	0.09	0.07	[0.006]
Share Primary Language not English	0.03	0.05	[0.000]
Share Living in Pittsburgh	0.06	0.06	[0.658]
Share Last Medicaid Spell Starting before 2011	0.32	0.34	[0.044]
Observations (N)	4,597	11,346	

Finkelstein & Notowidigdo (2019): Interpreting Results

- How can we take these results back to the normative framework?

1. Add exogenous probability π_j that application is accepted
2. Allow for different benefit levels $\overline{B}$ for low types and B_{min} for high types with $\overline{B} > B_{min}$

- Expressions for welfare effects are as before but mechanical effect becomes

$$\pi_H B_{min} \frac{dA_H}{dT} + \pi_L \overline{B} \frac{dA_L}{dT}$$

- $B_{min} = \$16/$month, $\overline{B} = \$178/$month and $\pi_L = \pi_H = 0.75$ yields $\pi_L B_L = \$4,086$ for L types and $\$432$ for H types.
- Baseline: Fiscal externality is public cost of processing applications: $G_L^A = G_H^A = -g = -\$267$ and $G_L^{-A} = G_H^{-A} = 0$
- First-order approximation to expected utility and time cost of applying of \$75 yields $\epsilon_L = -0.98$ and $\epsilon_H = -0.83$

Finkelstein & Notowidigdo (2019): Interpreting Results

- To make statements about social welfare impact of interventions need
  - $G_j^A - G_j^{-A}$. Use $-g$
  - μ_J. First-order approximation to utility yields approximations $\mu_H \approx \xi_H \pi_H \epsilon_H B_{min}$ and $\mu_L \approx \xi_L \pi_L \epsilon_L \overline{B}$.
- Write welfare effect as ratio of marginal benefits (private welfare increase) to marginal social costs (mech and fiscal externality) and use these approximations to get

$$MVPF^{\text{Info}} = \frac{-\epsilon_L \pi_L \overline{B} \frac{dA_L}{dT} - \epsilon_H \pi_H B_{min} \frac{dA_H}{dT}}{\left(\pi_L \overline{B} + g\right) \frac{dA_L}{dT} + \left(\pi_H B_{min} + g\right) \frac{dA_H}{dT}} \approx 0.89$$

$$MVPF^{\text{I + A}} = \frac{-\epsilon_L \pi_L \overline{B} \frac{dA_L}{dT} - \epsilon_H \pi_H B_{min} \frac{dA_H}{dT} - \left(A_H - A_L + \frac{dA_H}{dT} + \frac{dA_L}{dT}\right) \frac{dc}{dT}}{\left(\pi_L \overline{B} + g\right) \frac{dA_L}{dT} + \left(\pi_H B_{min} + g\right) \frac{dA_H}{dT}}$$

$$\approx 0.93$$

Outline

- Motivating Facts
- Theory
- Evidence from Rich Countries
- Targeting in Developing Countries: Who gets the Benefit?
- Transfer Design: What is the Benefit?
- Open Questions

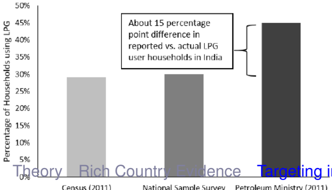
Outline

Targeting in Developing Countries: Who gets the Benefit?

- Barnwal (IGC 2017) *Curbing Leakage in Public Programs* [ET]
- Alatas, Banerjee, Hanna, Olken & Tobias (AER 2012) *Targeting the Poor: Evidence from a Field Experiment in Indonesia*
- Alatas, Banerjee, Hanna, Olken, Purnamasari & Wai-Poi (JPE 2016) *Self-Targeting: Evidence from a Field Experiment in Indonesia*
- Cohen Dupas & Schaner (AER 2015) *Price Subsidies, Diagnostic Tests, and Targeting of Malaria Treatment: Evidence from a Randomized Controlled Trial*
- Banerjee Hanna Olken & Sumarto (WP 2018) *The (lack of) Distortionary Effects of Proxy-Means Tests: Results from a Nationwide Experiment in Indonesia*

Barnwal (2017): Overview (ET)

- Policy change: Direct Benefit Transfer (DBT) to verified enrollees of a fuel subsidy benefit enforced and then removed. Direct Benefit Transfer for LPG (DBTL) available subsidized to households, LPG for commercial use is taxed.
- 'Ghost beneficiaries' who deliver subsidized fuel to black market - so enforcement increases prices in black market. Black market prices going up induce firms to buy formally. Easy to counterfeit. But DBT has biometrics-based unique ID (fingerprints, iris scan)
- DBT an 'enforcement' mechanism: minimizes role of local officials in disbursing subsidies and increase the cost of carrying out transfers in the name of ghost and duplicate beneficiaries



Barnwal (2017): Overview (ET)

- Dataset with 23 million transactions and black-market fuel prices, 3.79 million households in 509 districts, LPG monthly data from 3000 distributors.
- Simulated Customers to survey 'gas hawkers' (performing dpoorstep deliveries (in 89 districts, 11 states).
- 11-14% less 'domestic' LPG purchases during DBT period, which then converge to controls after.
- Black market $\downarrow$13-19% prices after DBT removal, $\downarrow$9% fuel purchases in formal sector by firms.
- Beneficiaries that drew the most subsisdies didn't comply (i.e. are ghosts), genuine households tend not to resell. Simple governance reform that take away monitoring and authentication tasks from intermediate agents can contain fraud and corruption.
- Barriers/ Frictions for real housholds to resell: Visually different containers and illegal redistribution laws. Cap on subsidized fuel, and search costs. But only delivery men hold permits to carry both types of containers.

Barnwal (2017): Model (ET)

We can add and subtract ghost demand to represent transfers from households to firms.

$$D_{total} = D_{household} + D_{ghost} + D_{firm} - D_{ghost}$$

Black market supply and demand is conditional on enforcement C. DBTL increases C_{ghost}.

$$S_{bm}(P - C_{ghost}) = D_{bm}(P + C_{firm})$$

Firms have an outside option and so will purchase fuel in the formal sector when Black market prices are too high:

$$D_{firm} : D_f[min(p + t, P + C_{firm})]$$

So before the DBTL policy, the price 'ceiling and floor' are determined by subsidy diversion opportunities:

$$P \in (p - s + C_{ghost}, p + t - C_{firm})$$

Barnwal (2017): Empirical Model (ET)

Diff-in-diff

$$Y_{idm} = \alpha + \beta Post_m * DBTL_d + \mu_i + \pi_m + \epsilon_{idm}$$

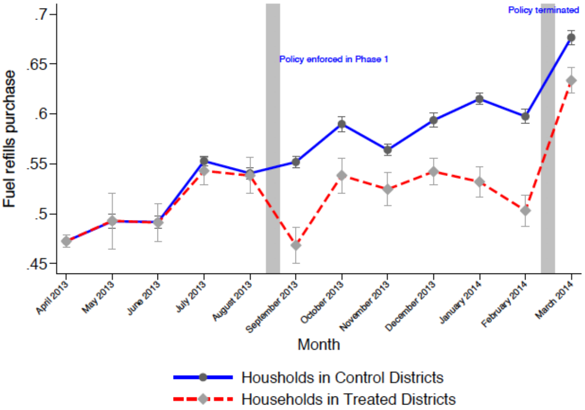
- Household i in district d in month m.
- $Post_m$ is the dummy for treatment period: 1 after DBTL enforced and 0 otherwise
- $DBTL_d$ dummy for districts reflecting treatment status: some districts did not have DBTL roll out
- clustered standard errors

Dynamic Response

$$Y_{idm} = \alpha + \beta . DBTL_d \theta_m + \theta_m + \mu_i + \pi_m + \epsilon_{idm}$$

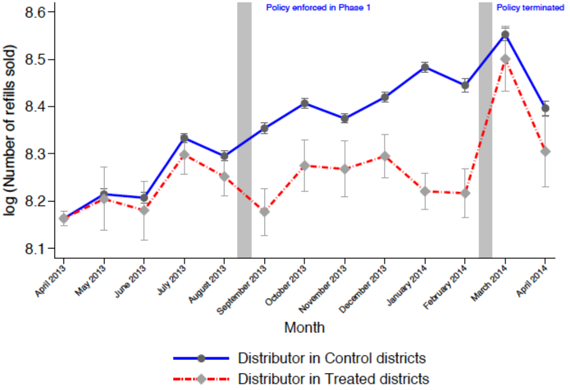
Barnwal (2017): Results (ET)

Figure 5: Impact of DBTL on domestic fuel purchase by beneficiaries



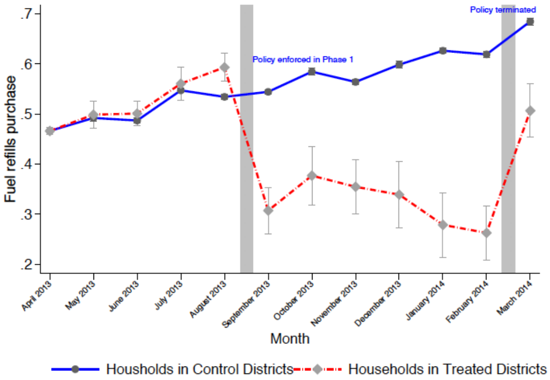
Barnwal (2017): Results (ET)

Figure 6: Impact of DBTL on domestic fuel sales by distributors



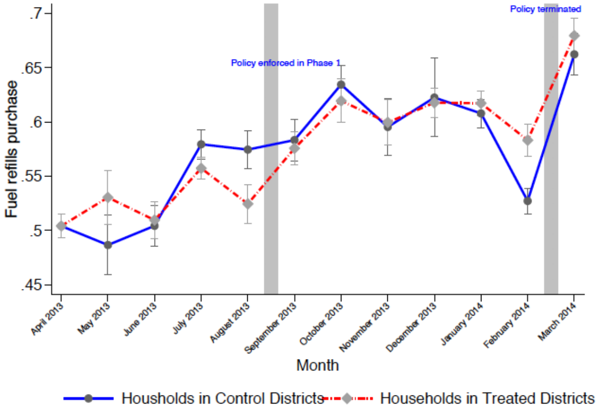
Barnwal (2017): Results (ET)

Figure 9: Fuel purchase by non-compliant beneficiaries



Barnwal (2017): Results (ET)

Figure 10: Fuel purchase by compliant beneficiary households



Barnwal (2017): Results (ET)

Table 3: Impact of DBTL on domestic fuel purchase by beneficiary households

	(1)	(2)	(3)
Outcome variable: Household monthly LPG refill purchase			
DBTL X Post	-0.0664***	-0.0621***	-0.0769***
	(0.00375)	(0.00401)	(0.00466)
Constant	0.484***	0.485***	0.475***
	(0.00319)	(0.00378)	(0.00396)
Observations	37,408,250	27,389,714	13,064,788
Household	3,400,750	2,489,974	1,187,708
Mean of outcome var	0.561	0.556	0.556
Control group	Ph 3-6 & Non-policy	Ph 3-6	Non-policy
Household FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes

Note: This regression estimates the impact of DBTL policy on domestic fuel purchase using OLS. A household-month level panel is used. Difference-in-differences estimates suggest about 11% to 14% reduction in fuel purchase in domestic cooking sector (i.e. coefficient on the interaction term as the percentage of mean value). Outcome variable is – number of LPG refills purchased by a beneficiary in a given month. Household and month fixed effects are included. Treated group includes all Phase 1 districts in the sample (16 districts). Phase 2 districts are not included. Col (1) combines all upcoming phases and non-policy districts together in the control group. Col(2) and Col (3) present estimates from the same specification, but with two different sub-groups as control. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Standard errors are clustered at the district level.

Barnwal (2017): Results (ET)

Table 4: Impact of DBTL on domestic fuel sales by distributors

	(1)	(2)	(3)
Outcome variable: log(Domestic LPG refills sales)			
DBTL X Post	-0.149***	-0.134***	-0.174***
	(0.0110)	(0.0118)	(0.0128)
Constant	8.178***	8.357***	7.975***
	(0.00716)	(0.00927)	(0.00953)
Observations	31,322	19,944	13,135
District	485	236	264
Distributor	3013	1909	1269
Control	Ph3-6 & Non-policy	Ph3-6	Non-policy
Distributor FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes

Note: This regression estimates the impact of DBTL program on domestic-use fuel sales using OLS. A distributor-month level panel is used. Difference-in-differences estimates suggest about 13% to 17% reduction in domestic-use LPG purchase. Outcome variable is – log(Total domestic-use LPG refills sold to households in a given month). Distributor and month fixed effects are included. Treated group includes all Phase 1 districts in the sample (16 districts). Phase 2 districts are not included. Col (1) combines all upcoming phases and non-policy districts together in the control group. Col(2) and Col (3) present estimates from same specification, but with two different control groups. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Standard errors are clustered at the district level.

Outline

Targeting in Developing Countries: Who gets the Benefit?

- Barnwal (IGC 2017) *Curbing Leakage in Public Programs* [ET]
- Alatas, Banerjee, Hanna, Olken & Tobias (AER 2012) *Targeting the Poor: Evidence from a Field Experiment in Indonesia*
- Alatas, Banerjee, Hanna, Olken, Purnamasari & Wai-Poi (JPE 2016) *Self-Targeting: Evidence from a Field Experiment in Indonesia*
- Cohen Dupas & Schaner (AER 2015) *Price Subsidies, Diagnostic Tests, and Targeting of Malaria Treatment: Evidence from a Randomized Controlled Trial*
- Banerjee Hanna Olken & Sumarto (WP 2018) *The (lack of) Distortionary Effects of Proxy-Means Tests: Results from a Nationwide Experiment in Indonesia*

Alatas et al (2012): Overview

- How effective are proxy means tests (PMT) in practice in a developing country?
- In particular, do communities know more than the proxies and can that be leveraged to improve targeting?
- What is the tradeoff?
  - Communities have better information.
    - Assets capture permanent income. Wouldn't observe severe negative shocks.
    - Poor PMT seen as ilegitimate by community.
  - Communities may disagree with government
    - May perceive poverty differently
    - May favor friends/relatives of elites
- This paper does an experiment in Indonesia to explore this

Alatas et al (2012): Experimental Design

- Indonesia has a large targeted cash transfer: Bantuan Langsung Tunai (BLT)
  - Provides \$10/month to 19 million households
  - Seen as badly targeted: WB estimates 45% of funds go to nonpoor households, 47% of poor excluded.
- For the experiment sample
  - 640 villages in 3 provinces (North Sumatra, South Sulawesi, Central Java).
  - Within each, sample a neighborhood.
  - 30% urban, 70% rural
- In each village,
  - Government and an Indonesian NGO Mitra Samya implemented an unconditional cash transfer.
  - Each beneficiary household to receive RP. 30,000 ($\sim$\$3)
  - Overall number of beneficiaries set in advance

Alatas et al (2012): Experimental Design

T1 *PMT*. Government used 49 indicators (home attributes, assets, hh composition, education, occupation etc) and a regression to create poverty scores. Enumerators collect this data and generate poverty scores. Lowest scores got the transfer.

T2 *Community Targeting.* Facilitator holds community meeting. List all households in the village and go through them one by one to rank them by poverty. After ranking finalized, quota of beneficiaries revealed.

T3 *Hybrid.* Do the community ranking. Then "verify" 1.5 $\times$ quota using PMT scores.

- Community subtreatments:
  - Elite capture: Either whole community or only elite invited to the meeting
  - Randomize the order in which households are ranked to test for effort effects (fatigue).
  - Preferences: Some meetings in the day (more women) others in the evening (more men), others place high emphasis on poverty.

Alatas et al (2012): Randomization

- To ensure balance, create 51 strata by geography.
- Randomize 640 subvillages into the 3 main treatments. Equal proportions in each stratum (up to integer constraint)
- Randomize the community and hybrid villages into subtreatments, stratifying by stratum $\times$ treatment
- Timing
  - Nov/Dec 2008. Carry out baseline survey
  - Dec 2008/ Jan 2009: Implement treatments, create beneficiary lists
  - Feb 2009: Distribute funds, interview village heads
  - March 2009: Endline survey

Alatas et al (2012): Randomization

TABLE 1—RANDOMIZATION DESIGN

Community/hybrid subtreatments			Main treatments		
			Community	Hybrid	PMT
Elite	10 poorest first	Day	24	23	
		Night	26	32	
	No 10 poorest first	Day	29	20	
		Night	29	34	
Whole community	10 poorest first	Day	29	28	
		Night	29	23	
	No 10 poorest first	Day	28	33	
		Night	20	24	
		Total	214	217	209

Alatas et al (2012): BLT targeting

Table 2—Summary Statistics

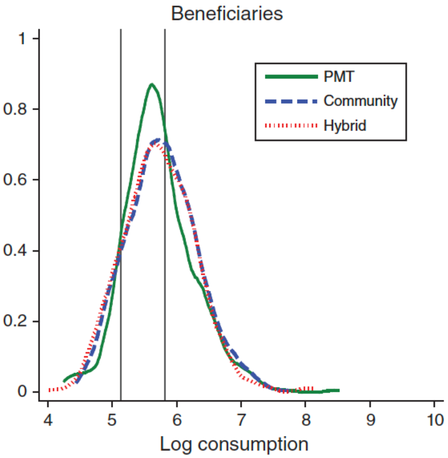
Variable	Obs	Mean	SD
Panel A. Consumption from baseline survey			
Per capita consumption (Rp. 1,000s)	5,753	557.501	602.33
Panel B. Mistargeting variables			
On beneficiary list	5,756	0.30	0.46
Error rate based on consumption	5,753	0.32	0.47
Inclusion error (nonpoor = rich + middle)	3,725	0.20	0.40
Exclusion error (poor = near + very poor)	2,028	0.53	0.50
Error rate based on consumption − rich	1,843	0.14	0.35
Error rate based on consumption − middle income	1,882	0.27	0.44
Error rate based on consumption − near poor	1,074	0.59	0.49
Error rate based on consumption − very poor	954	0.46	0.50
Panel C. Rank correlations between treatment results and …			
Per capita consumption	640	0.41	0.34
Community (excluding subvillage head)	640	0.64	0.33
Subvillage head	640	0.58	0.41
Self-assessment	637	0.40	0.34

Alatas et al (2012): Targeting Results

TABLE 3—RESULTS OF DIFFERENT TARGETING METHODS ON ERROR RATE BASED ON CONSUMPTION

		By income status		By detailed income status				
Sample:	Full population (1)	Inclusion error (2)	Exclusion error (3)	Rich (4)	Middle income (5)	Near poor (6)	Very poor (7)	Per capita consumption of beneficiaries (8)
Community treatment	0.031*	0.046**	0.022	0.028	0.067**	0.49	−0.013	9.933
	(0.017)	(0.018)	(0.028)	(0.021)	(0.027)	(0.038)	(0.039)	(18.742)
Hybrid treatment	0.029*	0.037**	0.009	0.020	0.052**	0.031	−0.008	−1.155
	(0.016)	(0.017)	(0.027)	(0.020)	(0.025)	(0.037)	(0.037)	(19.302)
Observations	5,753	3,725	2,028	1,843	1,882	1,074	954	1,719
Mean in PMT treatment	0.30	0.18	0.52	0.13	0.23	0.55	0.48	366

Alatas et al (2012): Targeting Results



Alatas et al (2012): Elite Capture

- Is the reason that community targeting doesn't dominate PMT elite capture?
- Include elite subtreatment.
- Also estimate whether households connected to the elite are more likely to receive the transfer.

$$
\begin{aligned}
ERROR_{ivk} = \alpha &+ \beta_1 COMMUNITY_{ivk} + \beta_2 HYBRID_{ivk} \\
&+ \beta_3 ELITE_{ivk} + \beta_4 CONN_{ivk} \\
&+ \beta_5 \left(COMMUNITY_{ivk} \times CONN_{ivk}\right) \\
&+ \beta_6 \left(HYBRID_{ivk} \times CONN_{ivk}\right) \\
&+ \beta_7 \left(ELITE_{ivk} \times CONN_{ivk}\right) + \gamma_k + \varepsilon_{ivk}
\end{aligned}
$$

Table 7—Elite Treatments

	Attendance (survey data)	Full sample error rate	Full sample error rate		On beneficiary list	
	(1)	(2)	(3)	(4)	(5)	(6)
Community treatment	0.367***	0.029	0.033	0.048*	0.042*	0.054*
	(0.038)	(0.018)	(0.023)	(0.025)	(0.025)	(0.028)
Hybrid treatment	0.370***	0.027	0.024	0.008	0.025	0.012
	(0.037)	(0.018)	(0.022)	(0.024)	(0.022)	(0.023)
Elite subtreatment	−0.301***	0.004	0.016	−0.013	−0.015	−0.039
	(0.034)	(0.016)	(0.020)	(0.029)	(0.021)	(0.032)
Elite × hybrid				0.062		0.051
				(0.041)		(0.043)
Elite connectedness			−0.025	−0.025	−0.063***	−0.063***
			(0.021)	(0.021)	(0.021)	(0.021)
Elite connectedness × community treatment			−0.015	−0.013	−0.067**	−0.078**
			(0.035)	(0.038)	(0.033)	(0.036)
Elite connectedness × hybrid treatment			0.010	0.010	−0.013	−0.001
			(0.033)	(0.035)	(0.033)	(0.035)
Elite connectedness × elite treatment			−0.029	−0.034	0.041	0.064
			(0.031)	(0.047)	(0.030)	(0.042)
Elite connectedness × elite treatment × hybrid				0.003		−0.047
				(0.063)		(0.060)
Observations	287	5,753	5,753	5,753	5,756	5,756
Mean in PMT treatment	0.11	0.30	0.30	0.30	0.28	0.28

Alatas et al (2012): Alternative Welfare Metrics

$RANKCORR_{vkR} = \alpha + \beta_1 COMMUNITY_{vk} + \beta_2 HYBRID_{vk} + \gamma_k + \varepsilon_{vkR}$,
$RANKCORR_{vkR}$ is rank correlation between targeting rank list and well-being measure R

TABLE 9—ASSESSING TARGETING TREATMENTS USING ALTERNATIVE WELFARE METRICS

	Consumption (r_g) (1)	Community survey ranks (r_c) (2)	Subvillage head survey ranks(r_e) (3)	Self-assessment (r_s) (4)
Community treatment	−0.065**	0.246***	0.248***	0.102***
	(0.033)	(0.029)	(0.038)	(0.033)
Hybrid treatment	−0.067**	0.143***	0.128***	0.075**
	(0.033)	(0.029)	(0.038)	(0.033)
Observations	640	640	640	637
Mean in PMT treatment	0.451	0.506	0.456	0.343

TABLE 12—WHAT IS THE COMMUNITY MAXIMIZING?

	Rank according to welfare metric			Targeting rank list in		
	Community survey ranks (r_c)	Subvillage head survey ranks(r_e)	Self-assessment (r_s)	PMT villages	Community villages	Hybrid villages
	(1)	(2)	(3)	(4)	(5)	(6)
Log per capita consumption	0.176***	0.145***	0.087***	0.132***	0.197***	0.162***
	(0.008)	(0.008)	(0.004)	(0.013)	(0.014)	(0.014)
Panel A. Household demographics						
Log HH size	0.164***	0.134***	0.073***	−0.028	0.154***	0.078***
	(0.011)	(0.010)	(0.006)	(0.019)	(0.019)	(0.021)
Share kids	−0.125***	−0.094***	−0.037***	−0.296***	−0.068*	−0.141***
	(0.021)	(0.021)	(0.012)	(0.035)	(0.041)	(0.039)

Panel B. Ability to smooth shocks						
Elite connected	0.092***	0.044***	0.025***	0.062***	0.051***	0.043***
	(0.008)	(0.009)	(0.005)	(0.016)	(0.015)	(0.015)
Total connectedness	−0.039***	−0.021**	−0.015***	−0.016	−0.019	−0.054***
	(0.010)	(0.009)	(0.005)	(0.017)	(0.017)	(0.019)
Number of family members outside subvillage	0.012***	0.010***	0.006***	0.020***	0.001	0.001
	(0.004)	(0.003)	(0.002)	(0.006)	(0.006)	(0.006)
Participation through work to community projects	0.002	0.021**	0.005	0.000	0.010	0.003
	(0.011)	(0.010)	(0.006)	(0.018)	(0.019)	(0.019)
Participation through money to community projects	0.061***	0.041***	0.024***	0.056***	0.058***	0.034*
	(0.009)	(0.009)	(0.005)	(0.016)	(0.016)	(0.018)
Participation in religious groups	0.027***	0.033***	0.014**	0.033**	0.012	0.029
	(0.010)	(0.010)	(0.006)	(0.016)	(0.017)	(0.017)
Total savings	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Share of savings in a bank	0.096***	0.069***	0.052***	0.121***	0.103***	0.075***
	(0.011)	(0.010)	(0.006)	(0.018)	(0.021)	(0.020)
Debt as share of consumption	0.005***	0.004***	0.002***	0.002	0.007***	0.008***
	(0.001)	(0.001)	(0.000)	(0.002)	(0.001)	(0.001)
Panel C. Discrimination against minorities?						
Ethnic minority	−0.024*	−0.019	−0.003	0.012	−0.051**	−0.011
	(0.014)	(0.014)	(0.008)	(0.026)	(0.025)	(0.024)
Religious minority	0.012	−0.007	−0.014*	−0.018	0.025	0.012
	(0.018)	(0.017)	(0.008)	(0.030)	(0.032)	(0.033)

Panel D. Correcting for earnings ability

HH head with primary education or less	−0.028***	−0.025***	−0.037***	−0.108***	−0.011	−0.066***
	(0.009)	(0.009)	(0.005)	(0.017)	(0.018)	(0.017)
Widow	−0.104***	−0.083***	−0.012	0.009	−0.108***	−0.026
	(0.014)	(0.014)	(0.008)	(0.027)	(0.024)	(0.028)
Disability	−0.045***	−0.037***	−0.026***	−0.079***	0.009	0.012
	(0.016)	(0.014)	(0.008)	(0.027)	(0.026)	(0.027)
Death	−0.041*	−0.031	−0.010	−0.111***	−0.013	−0.059
	(0.025)	(0.025)	(0.015)	(0.042)	(0.048)	(0.043)
Sick	−0.038***	−0.041***	−0.028***	0.007	−0.018	−0.044**
	(0.011)	(0.011)	(0.006)	(0.018)	(0.019)	(0.019)
Recent shock to income	−0.001	−0.005	−0.013**	−0.019	0.009	−0.012
	(0.009)	(0.009)	(0.005)	(0.016)	(0.016)	(0.017)
Tobacco and alcohol consumption	−0.0002***	−0.0002***	−0.0001***	−0.0002***	−0.0002***	−0.0001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	5,337	4,680	5,724	1,814	1,876	1,889

Outline

Targeting in Developing Countries: Who gets the Benefit?

- Barnwal (IGC 2017) *Curbing Leakage in Public Programs* [ET]
- Alatas, Banerjee, Hanna, Olken & Tobias (AER 2012) *Targeting the Poor: Evidence from a Field Experiment in Indonesia*
- Alatas, Banerjee, Hanna, Olken, Purnamasari & Wai-Poi (JPE 2016) *Self-Targeting: Evidence from a Field Experiment in Indonesia*
- Cohen Dupas & Schaner (AER 2015) *Price Subsidies, Diagnostic Tests, and Targeting of Malaria Treatment: Evidence from a Randomized Controlled Trial*
- Banerjee Hanna Olken & Sumarto (WP 2018) *The (lack of) Distortionary Effects of Proxy-Means Tests: Results from a Nationwide Experiment in Indonesia*

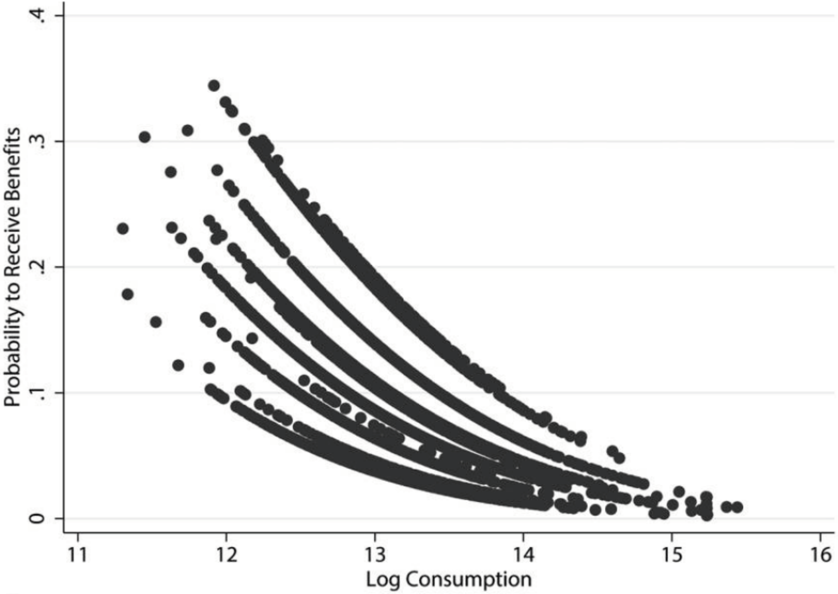
Alatas et al (2016): Overview

- Do application costs (ordeals a la Nichols & Zeckhauser) improve targeting in a developing country context?
- If so, how big should application costs be?
- Implement an experiment in Indonesia to learn about this.
- Use experimental results together with a model to learn about selection margins and effects of alternative policies.

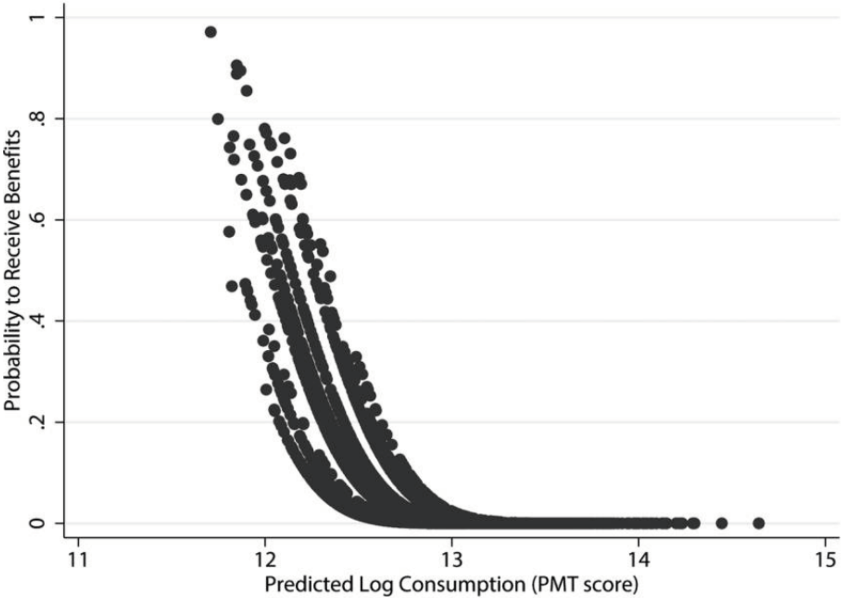
Alatas et al (2016): Setting

- Study a large transfer program in Indonesia: Program Keluarga Harapan (PKH)
- Targeted at households with consumption/capita < 80% of poverty line, and with pregnant woman/children 0-5/children with <9 years of schooling.
- Recipients get \$67–\$250/year (3.5–13% of avg consumption)
- 2.4 million househlods enrolled in 2013.
- Targeting is hard: per-capita consumption is hard to observe.
- The government does a PMT.
  - Every 3 years, stats bureau does a nationwide survey.
  - Potentially eligible households + local leaders' recommendations surveyed. Asked 30 questions.
  - Supplemented with location-level data
  - Generate district-level formula for predicting consumption.

Alatas et al (2016): Targeting



Alatas et al (2016): Targeting



Alatas et al (2016): Experiment Design

- PKH expanded to new areas in 2011
- Work in six districts (2 each in Lampung, South Sumatra, and Central Java)
- randomly select 400 villages stratifying to have 30% urban, 70% rural and by geography.

T1 *Automatic Screening*: Government enumerators make list of potentially poor households. Administer survey. Computer generated scores below cutoff for that district $\rightarrow$ receive transfer

T2 *Self-Targeting:* The test was the same, but households had to go to a central registration station to apply. Meetings held in each village to publicize the program, and emphasize that survey would be verified. On predetermined day, housholds could visit office from 8AM-5PM, get a number and line up for their interview.

- Within Self-targeting treatment also cary distance between village and office.

Alatas et al (2016): Experimental Design

TABLE 1
EXPERIMENTAL DESIGN

	Number of Villages (Households)
Automatic screening	200 (1,998)
Self-targeting:	
Close subtreatment	100 (1,000)
Far subtreatment	100 (1,000)
Total	200 (2,000)

NOTES.—This table provides the number of villages in each treatment cell. The number of households in each cell is shown in parentheses.

Alatas et al (2016): Timing

TIME LINE OF THE EXPERIMENT

	Self-Targeting Villages	Automatic Screening Villages
December 2010 to March 2011	Baseline survey	
January to April 2011	Application process publicized. Registration days: Households that showed up to apply received the PMT interview at the registration site. Verification process: A subset of households received home visits and received another PMT interview.	Prescreen list: Households suggested by village leaders or BPS enumerators were added to the prescreen list. PMT interviews: BPS enumerators conducted home visits and PMT interviews with all prescreened households.
Early August 2011	Midline survey	
Late August 2011	Beneficiary lists were announced to the villages. First round of PKH benefits distributed.	
January to March 2012	Endline survey	

Alatas et al (2016): Model

- Households live for two periods. risk neutral, care about consumption.
- Per-period income of y. Only y^o is observable to the government, so $y = y^o + y^u$.
- Applying for the transfer costs $c(l, y)$ where l is the distance to the office.
- If apply, receive the transfer with probability $\mu(y^o)$, $\mu' \leq 0$.
- Even observable income measured with error by PMT so $\mu(y^o) = \mathbb{P}(y^o + \pi < y^*)$, π is iid noise.
- 2 types of households:
  - sophisticated: Understand $\mu(y^o)$
  - unsophisticated: Know $\lambda(y)$=probability someone with income y receives transfer
- Transfer recipients receive income b in second period. δ is discount factor
- Households also get a utility shock $\varepsilon \sim F(\varepsilon)$ that encourages/discourages them from applying.

Alatas et al (2016): Application decision

- Sophisticated households who apply have expected utility

$$y - c(l, y) + \mu(y^o)\,\delta(y + b) + [1 - \mu(y^o)]\,\delta y + \varepsilon$$

- Unsophisticated households who apply have EU

$$y - c(l, y) + \lambda(y)\,\delta(y + b) + [1 - \lambda(y)]\,\delta y + \varepsilon$$

- If households don't apply, they have EU $y + \delta y$
- Expected gain from applying for the sophisticated hhs is

$$-c(l, y) + \mu(y^o)\,\delta b + \varepsilon \equiv g(y^o, y, l) + \varepsilon$$

- For unsophisticated, it's

$$-c(l, y) + \lambda(y)\,\delta b + \varepsilon \equiv h(y, l) + \varepsilon$$

Alatas et al (2016): Application decision

- Application probabilities:

$$A_s(y^o, y, l) = \mathbb{P}(g(y^o, y, l) > \varepsilon) = 1 - F(-g(y^o, y, l))$$
$$A_u(y, l) = \mathbb{P}(h(y, l) > \varepsilon) = 1 - F(-h(y, l))$$

- Unsophisticated households have consistent beliefs:

$$\lambda(y) = \lambda_{ind}(y) = \frac{\begin{pmatrix} \alpha \int \int \mu(y^o) A_s(y^o, y, l) \vartheta(y^o, l|y)\, dl dy \\ + (1-\alpha) \int \int \mu(y^o) A_u(y, l) \vartheta(y^o, l|y)\, dl dy \end{pmatrix}}{\begin{pmatrix} \alpha \int \int A_s(y^o, y, l) \vartheta(y^o, l|y)\, dl dy \\ + (1-\alpha) \int \int A_u(y, l) \vartheta(y^o, l|y)\, dl dy \end{pmatrix}}$$

where α is proportion of sophisticated hhs

Alatas et al (2016): Application cost

- Simple benchmark:
  - all hhs unsophisticated.
  - Time cost of applying is τl. Wages are $w = \phi y$ so monetary cost is $\tau l \phi y$.
  - No shocks $\varepsilon = 0$

$$\text{apply iff } h(y) = -\tau l \phi y + \delta \lambda(y) b \geq 0$$

- lhs is decreasing in $y \rightarrow$ threshold y^* below which people apply (Nichols & Zeckhauser 1982).
- Now add back in shocks:

$$\text{apply iff } \tau l \phi y - \delta \lambda(y) b \leq \varepsilon$$

- Consider y_1 and $y_2 > y_1$: They show up at relative rates

$$\frac{1 - F(\tau l \phi y_1 - \delta \lambda(y_1) b)}{1 - F(\tau l \phi y_2 - \delta \lambda(y_2) b)}$$

  the higher this is, the better targeted the transfer

Alatas et al (2016): Application cost

- Now consider increasing the ordeal l: Differentiating, targeting improves iff

$$\frac{f\left(\tau l\phi y_2-\delta\lambda\left(y_2\right)b\right)}{1-F\left(\tau l\phi y_2-\delta\lambda\left(y_2\right)b\right)}\tau\phi y_2-\frac{f\left(\tau l\phi y_1-\delta\lambda\left(y_1\right)b\right)}{1-F\left(\tau l\phi y_1-\delta\lambda\left(y_1\right)b\right)}\tau\phi y_1>0$$

- More generally, a sufficient condition for targeting efficiency to be improving as l increases is that the hazard rate is increasing in y

$$\frac{\partial}{\partial y}\frac{f\left(\tau l\phi y-\delta\lambda\left(y\right)b\right)}{1-F\left(\tau l\phi y-\delta\lambda\left(y\right)b\right)}>0$$

- This rules out thick tailed income distributions.
- More generally, shows that single crossing $(c_{ly}>0)$ isn't sufficient to make ordeal optimal

Alatas et al (2016): Nonlinear application cost

- Application costs may affect poor and rich households differently.
- To capture this, put a kink in the application cost: There are two ways to get to the office, by bus or walking.
- Walking: calorie cost γl; time cost $\tau l w$: slow but cheap
- Bus: fixed cost ν; time cost λw: fast but costly: $\lambda < \tau$

$$D = \begin{cases} \text{bus} & \text{if } v + \lambda l \phi y < \gamma l + \tau l \phi y \\ \text{walk} & \text{if } \nu + \lambda l \phi y \geq \gamma l + \tau l \phi y \end{cases}$$

- Apply iff

$$-\min \left\{ \gamma l + \tau l \phi y, \nu + \lambda l \phi y \right\} + \delta \lambda \left(y \right) b \geq \varepsilon$$

- lhs decreasing in y so richer hhs apply less.
- Effect of change in l also stronger for people who walk, who are on average poorer.

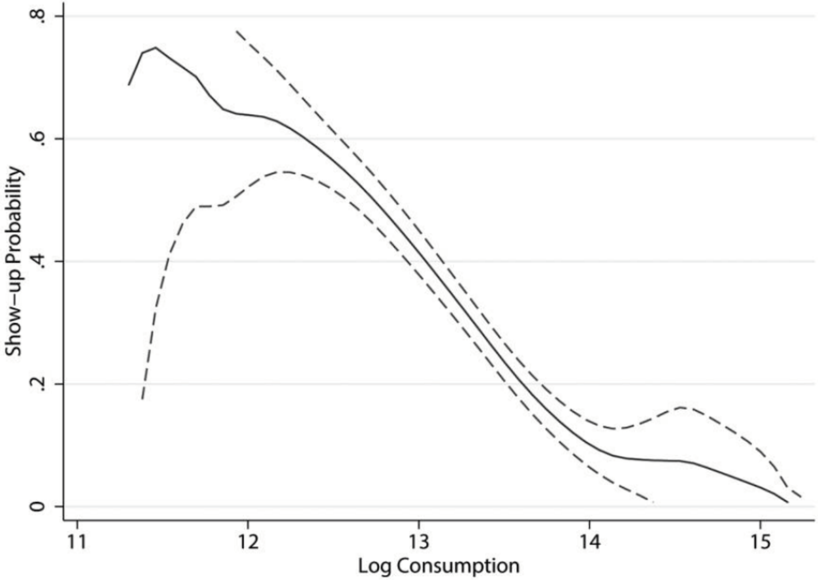
Alatas et al (2016): Sophistication

- All this so far only for unsophisticated households. Perceive pr of transfer as $\lambda\left(y\right)$
- Sophisticated households understand the probability is $\mu\left(y^{o}\right)$
- How does unobserved income y^{u} affect behavior?
  - For sophisticated hhs, only effect is through $c\left(l, y^{o}+y^{u}\right)$
  - For unsophisticated, both through $\lambda\left(y^{o}+y^{u}\right)$ and $c\left(l, y^{o}+y^{u}\right)$

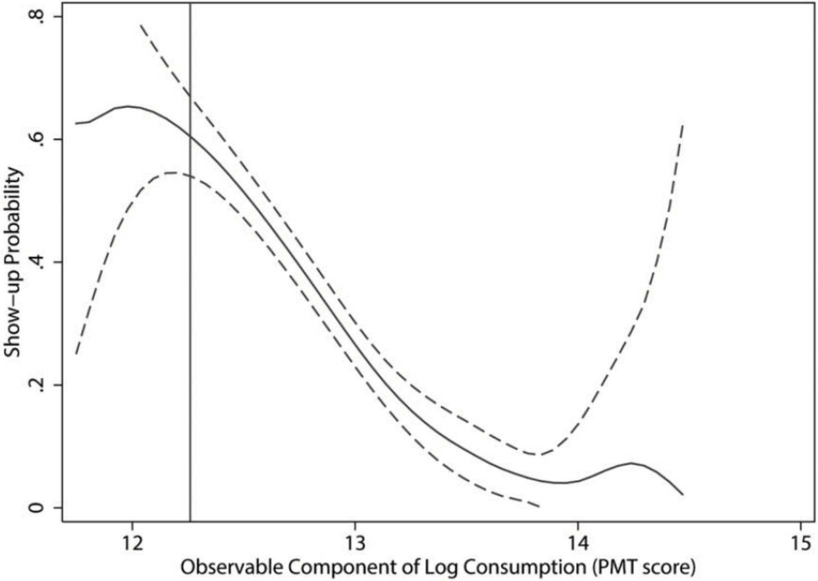
  $\Rightarrow$ expect more selection on unobservables for unsophisticated HHs
- Selection on unobservables is good though when income is poorly observed by PMT:
  - Sophisticated rich households who know they can pass PMT apply
  - unsophisticated rich households unsure they can pass PMT so don't apply

  $\Rightarrow$ benefit of unsophisticated households? Secret PMT formula?

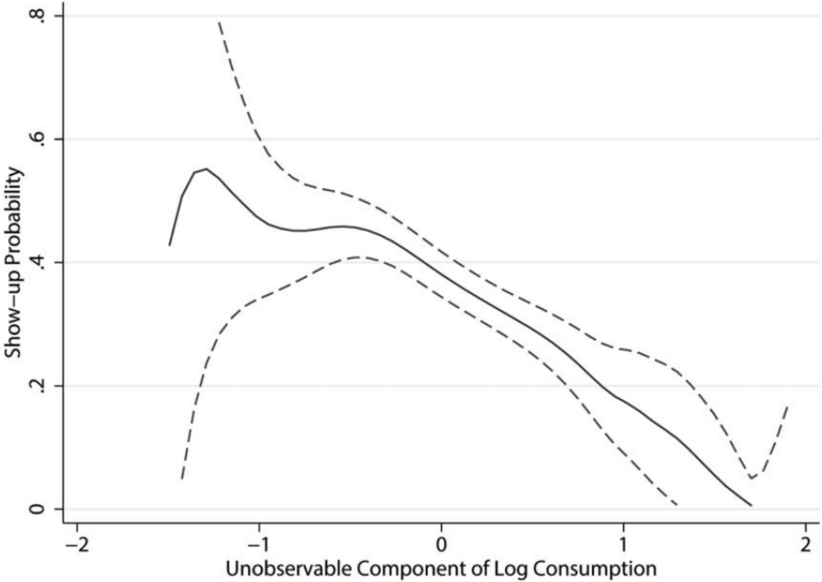
Alatas et al (2016): Self-Selection



Alatas et al (2016): Selection on Observables



Alatas et al (2016): Selection on Unobservables



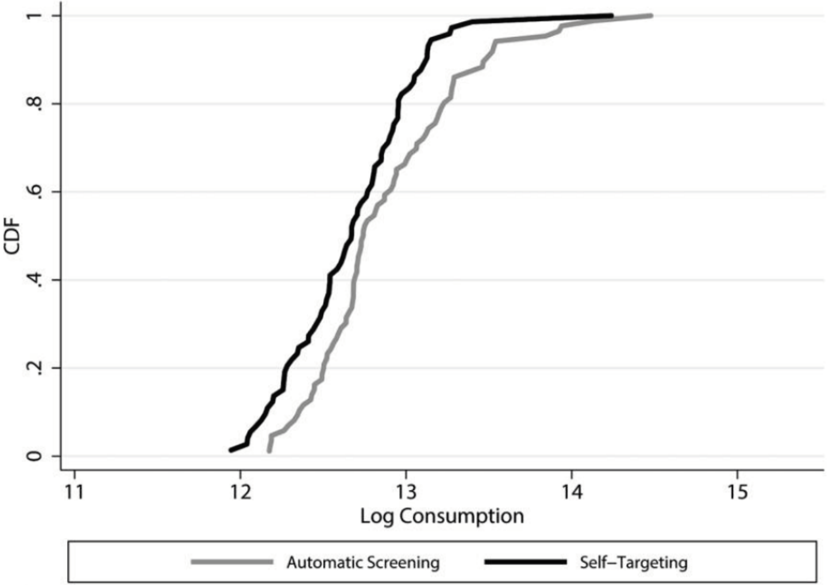
Alatas et al (2016): Self-Selection

- Logit: $\mathbb{P}\left(\text{show-up}_i = 1\right) = \exp\left(\alpha + \gamma y_i^o + \psi y_i^u\right)/1 + \exp\left(\alpha + \gamma y_i^o + \psi y_i^u\right)$

TABLE 4
PROBABILITY OF SHOWING UP AS A FUNCTION OF THE OBSERVED AND UNOBSERVED COMPONENTS OF BASELINE LOG PER CAPITA CONSUMPTION

	SHOWED UP		
	All (1)	Very Poor (2)	Not Very Poor (3)
Observable consumption (y_i^o)	−2.217***	−.325	−2.310***
	(.201)	(1.785)	(.208)
Unobservable consumption (y_i^u)	−.907***	−.775	−.908***
	(.136)	(.581)	(.138)
Stratum fixed effects	No	No	No
Observations	2,000	114	1,886
Mean of dependent variable	.377	.658	.360

Alatas et al (2016): Self-Targeting vs Automatic



Alatas et al (2016): Self-Targeting vs Automatic

	Log Consumption (Beneficiaries; Baseline; OLS) (1)	Log Consumption (Beneficiaries; Baseline + Midline; OLS) (2)	Receives Benefits (Logit) (3)	Error (Logit) (4)	Exclusion Error (Logit) (5)	Inclusion Error (Logit) (6)
	A. No Stratum Fixed Effects					
Self-targeting	−.208***	−.193***	12.142**	−.190	−.506	−.311
	(.076)	(.060)	(4.894)	(.126)	(.402)	(.210)
Log consumption			−1.016***			
			(.280)			
Log consumption × self-targeting			−.964**			
			(.383)			
Observations	159	904	3,996	3,998	249	3,749
Mean of dependent variable	12.78	13.61	.0398	.0870	.880	.0344
	B. With Stratum Fixed Effects					
Self-targeting	−.114	−.175***	15.180***	−.209	−.649	−.331*
	(.077)	(.058)	(5.295)	(.140)	(.441)	(.192)
Log consumption			−1.042***			
			(.283)			
Log consumption × self-targeting			−1.202***			
			(.416)			
Observations	159	904	3,489	3,938	113	3,130
Mean of dependent variable	12.78	13.61	.0456	.0884	.761	.0412

Alatas et al (2016): Changing the Ordeal

- Compare far (average 1.83 km) to close (0.27 km) subtreatments

$$\mathbb{P}\left(\text{show-up}_i = 1\right) = \frac{\exp(\alpha+\beta \text{Close}_v + \gamma y_{vi} + \eta \text{Close}_v \times y_{vi})}{1+\exp(\alpha+\beta \text{Close}_v + \gamma y_{vi} + \eta \text{Close}_v \times y_{vi})}$$

	No Stratum Fixed Effects			With Stratum Fixed Effects		
	(1)	(2)	(3)	(4)	(5)	(6)
Close subtreatment	.205	1.345	.185	.275	.485	.179
	(.146)	(2.841)	(.237)	(.168)	(2.920)	(.314)
Log consumption		−1.434***			−1.446***	
		(.143)			(.144)	
Close subtreatment × log consumption		−.093			−.023	
		(.217)			(.218)	

Alatas et al (2016): Model and Mechanisms

- Use GMM to estimate model parameters.
- Assumptions:
  - $\varepsilon \sim logistic$ with mean v_ε and sd σ_ε
  - Unsophisticated beliefs: probit $\lambda(y) = \Phi(\gamma + \pi y)$
  - Transfer's NPV. Use government credit scheme's interest rate 22%

$\Rightarrow$ 5 parameters to fit: $v_\varepsilon, \sigma_\varepsilon, \alpha, \lambda(y), \gamma, \pi$

- Speciffy
  $c(y_i, l_i) = \text{wage}_i \times (\text{traveltime}_i + \text{waitingtime}) + \text{travelmoney}_i$

Alatas et al (2016): Model and Mechanisms

- Moments:
- Show-up rates in 5 quintiles of consumption and far/close treatments $\Rightarrow$ 10 moments
- Show-up rates in {top/bottom tercile y^o} $\times$ {top/bottom tercile y^u} $\Rightarrow$ 4 moments
- Show-up rates in top/bottom quartiles of distance $\Rightarrow$ 2 moments

$$\mathrm{E}\left[\Phi\left(\gamma+\pi y_i\right)-\text{benefit}_i|\text{show-up}_i=1\right]=0$$

$$\mathrm{E}\left[\left[\Phi\left(\gamma+\pi y_i\right)-\text{benefit}_i\right]\left(y_i-\overline{y}\right)|\text{show-up}_i=1\right]=0$$

$$\mathrm{E}\left[\lambda_{ind}\left(y_i\right)-\text{benefit}_i|\text{show-up}_i=1\right]=0$$

$$\mathrm{E}\left[\left[\lambda_{ind}\left(y_i\right)-\text{benefit}_i\right]\left(y_i-\overline{y}\right)|\text{show-up}_i=1\right]=0$$

Alatas et al (2016): Model Parameters

ESTIMATED PARAMETER VALUES FOR THE MODEL

v_ε	σ_ε	α	γ	π
−79,681	59,715	.50	8.04	−.72
(6,798)	(11,734)	(.07)	(.63)	(.05)

NOTE.—This table reports the estimated mean v_ε and standard deviation σ_ε of the utility shock (ε), the fraction of sophisticated households (α), and the constant γ and log consumption coefficient π in the λ function. The parameters are estimated using two-step feasible GMM. For each step, we choose 100 random initial conditions and minimize the objective function using a trust-region-reflective algorithm. Bootstrapped standard errors, calculated using 100 bootstrap iterations, are in parentheses.

Alatas et al (2016): Mechanisms

	Show-Up Rate (Experimental) (1)	Predicted Show-Up Probability (Model): Baseline Model (2)	$\sigma_\varepsilon = \hat{\sigma}_\varepsilon/2$ (3)	$\sigma_\varepsilon = 0$ (4)	Assuming Same Travel Technology (5)	Constant $\mu(\cdot)$ and $\lambda(\cdot)$ (6)
A. Logistic Regressions						
Close	1.509	−1.365	−1.825	−1.791	−1.367	−1.742
	(2.972)	(3.098)	(3.472)	(3.765)	(2.967)	(2.18)
Log consumption	−1.423***	−1.630***	−2.181***	−2.456***	−1.631***	−.103
	(.148)	(.163)	(.193)	(.204)	(.166)	(.118)
Close × log consumption	−.105	.105	.141	.138	.106	.136
	(.227)	(.238)	(.268)	(.29)	(.228)	(.166)
Observations	1,971	5,913,000	5,913,000	5,913,000	5,913,000	5,913,000
p-value		.522	.483	.509	.513	.391
B. Show-Up Rates						
Above poverty line, far	34.09	34.55	30.04	28.12	34.54	45.89
Above poverty line, close	38.99	37.37	33.11	31.17	37.37	47.15
Below poverty line, far	53.23	71.94	72.94	73.83	71.92	46.53
Below poverty line, close	59.32	65.52	65.81	66.25	65.52	43.84
C. Show-Up Rate Ratios						
Poor to rich ratio, far	1.561	2.082	2.428	2.626	2.082	1.014
	(.213)	(.203)	(.244)	(.262)	(.199)	(.14)
Poor to rich ratio, close	1.522	1.753	1.987	2.126	1.753	.93
	(.169)	(.183)	(.214)	(.221)	(.19)	(.141)
Difference of ratios	.040	.329	.441	.5	.329	.084
	(.268)	(.271)	(.322)	(.34)	(.281)	(.197)
p-value		.448	.338	.288	.456	.893

Alatas et al (2016): Alternative Policies

	Show-Up Rate (Experimental)			Predicted Show-Up Probabilities (Model)					
	Automatic Screening (Scaled) (1)	Automatic Screening (in Sample) (2)	Self-Targeting (3)	Baseline Model (4)	Far Distance + 3 km (5)	Far Distance + 6 km (6)	Far Wait Time × 3 (7)	Far Wait Time × 6 (8)	Perfect Targeting (9)
	A. Program Statistics								
Mean show-up rate (%)	34.62	34.62	37.84	37.93	37.67	37.53	36.79	35.39	5.83
Mean benefit receipt (%)	4.38	4.38	3.64	4.11	4.10	4.09	4.06	4.00	5.83
Mean eligible benefit receipt (%)	.63	.63	.73	.86	.86	.86	.86	.86	5.83
Mean ineligible benefit receipt (%)	3.75	3.75	2.91	3.24	3.23	3.23	3.20	3.14	.00
	B. Average Household Costs for Households That Show Up (Rupees)								
Average cost to households	1,021	1,021	13,674	13,831	15,947	17,218	24,187	37,460	7,621
Average cost to beneficiary households	938	938	12,464	12,797	14,774	16,130	21,987	34,968	7,621
Average cost to nonbeneficiary households	1,033	1,033	13,803	13,957	16,091	17,351	24,459	37,777	. . .
	C. Government Costs and Benefits Paid (Rupees)								
Administrative costs, per household	4,768	31,054	6,764	6,781	6,734	6,710	6,576	6,326	1,042
Expected benefits, per household	332,028	332,028	306,108	353,230	352,305	351,742	349,750	344,826	472,990
	D. Poverty Gap								
Poverty gap under fixed budget (%)	2.736	2.741	2.720	2.724	2.724	2.724	2.724	2.725	2.610
Reduction in poverty gap relative to perfect targeting (%)	29.91	27.42	38.58	36.38	36.71	36.37	36.47	36.26	100.00

Outline

Targeting in Developing Countries: Who gets the Benefit?

- Barnwal (IGC 2017) *Curbing Leakage in Public Programs* [ET]
- Alatas, Banerjee, Hanna, Olken & Tobias (AER 2012) *Targeting the Poor: Evidence from a Field Experiment in Indonesia*
- Alatas, Banerjee, Hanna, Olken, Purnamasari & Wai-Poi (JPE 2016) *Self-Targeting: Evidence from a Field Experiment in Indonesia*
- Cohen Dupas & Schaner (AER 2015) *Price Subsidies, Diagnostic Tests, and Targeting of Malaria Treatment: Evidence from a Randomized Controlled Trial*
- Banerjee Hanna Olken & Sumarto (WP 2018) *The (lack of) Distortionary Effects of Proxy-Means Tests: Results from a Nationwide Experiment in Indonesia*

Cohen et al (2015): Overview

- Usually, the targeting tradeoff is that people who know they are ineligible may try to mimic the deserving types in order to gain access to the transfer.
- What if incentives are aligned (government and households agree on who should receive the transfer) but households don't know whether they're eligible?
- Here: Malaria treatments: artemisinin combination therapies (ACT)
  - Huge benefits if have malaria.
  - no direct benefits if don't have malaria, people don't learn real reason they're sick, speeds up development of parasite's resistence.
  - But people who are sick don't know for sure whether they have malaria (or something else) so many people take malaria treatments just in case.
- Experiment in Kenya to test impact of
  - better diagnosis technology
  - subsidies for ACTs

Cohen et al (2015): Setting

- Malaria causes 200 million illnesses, kills 600K people a year
- Many countries (ncluding Kenya) provide ACTs for free at public health facilities if diagnosed with malaria. But...
  - diagnosis often incorrect
  - stockouts common
  - Have to pay fees, travel far, line up, etc...
- Many households go to private drugstores to get ACTs or other over-the-counter medications (40–97% of the market!)
- Large subsidies to ACTs to improve access. Subsidy $\sim 95\%$ of cost

Cohen et al (2015): Model

- When households receive an illness shock they pick an action

$$a \in \begin{cases} h & \text{seek diagnosis at a formal health facility} \\ s & \text{buy ACTs at a shop} \\ n & \text{buy non-ACT drugs or do nothing} \end{cases}$$

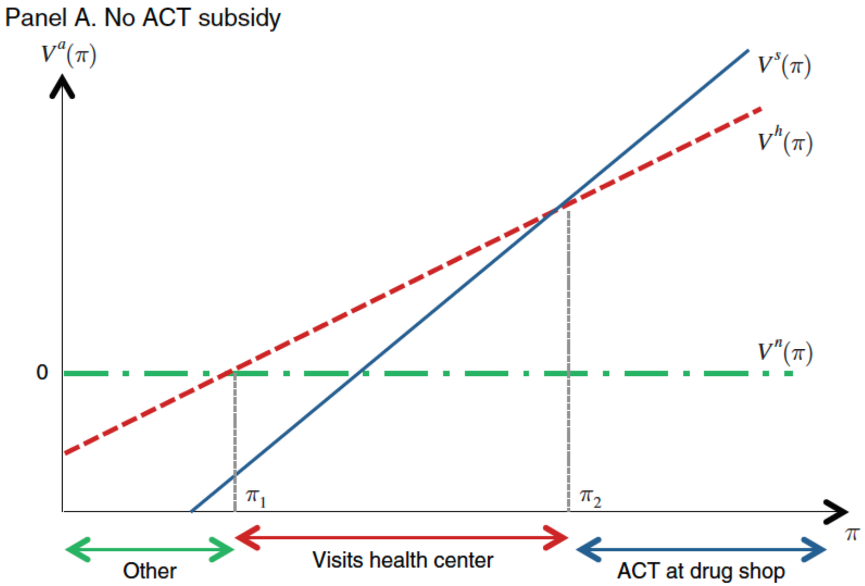
- Households who fall ill form a subjective probability that the illnes is malaria with probability π

$$\begin{aligned} V^a(\pi) &= \pi\left[U_P^a(\pi) - p_P^a(\pi)\right] + (1-\pi)\left[U_N^a(\pi) - p_N^a(\pi)\right] \\ &= \pi V_P^a(\pi) + (1-\pi) V_N^a(\pi) \end{aligned}$$

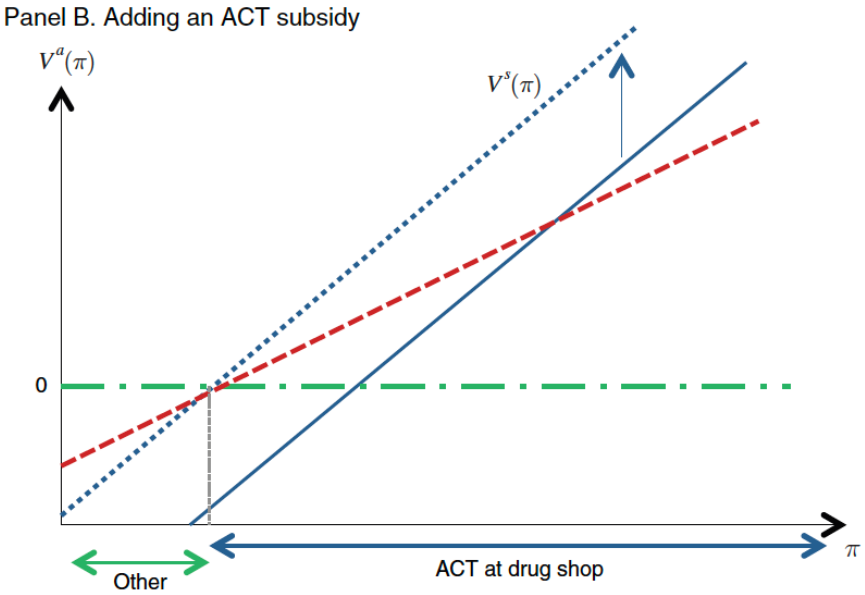
  where P denotes malaria-positive, N malaria negative.
- Assume value of acting increasing with π: $\partial\left(V^a(\pi) - V^n(\pi)\right)/\partial\pi > 0$ for $a \in \{h, s\}$
- Go to the drug shop iff

$$V^s(\pi) > \max\left\{V^h(\pi), V^n(\pi)\right\}$$

Cohen et al (2015): Model



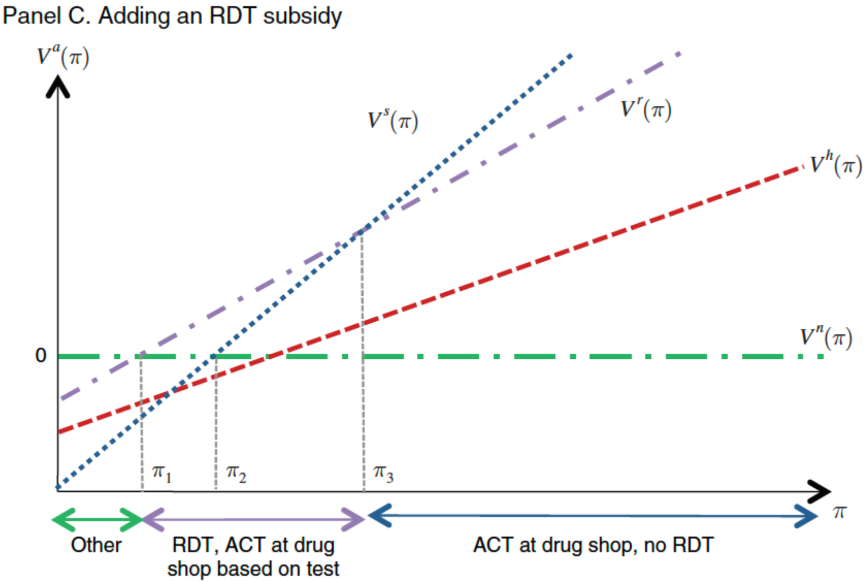
Cohen et al (2015): Adding an ACT Subsidy



Cohen et al (2015): Model

- Effects of subsidy:
  - More access: More people get ACTs
  - Worse targeting: People induced to use ACTs have lower π
  - Better targeting possible if lots of poor people with high π can't afford ACTs.
- What about Retail Diagnosis Test (RDT) to improve accuracy of π?
  - introduce $V^r(\pi)$: Value of taking the RDT and then getting ACT if positive.
  - $V^r(\pi) > V^s(\pi)$ at low π since V^s relatively more attractive as π increases.

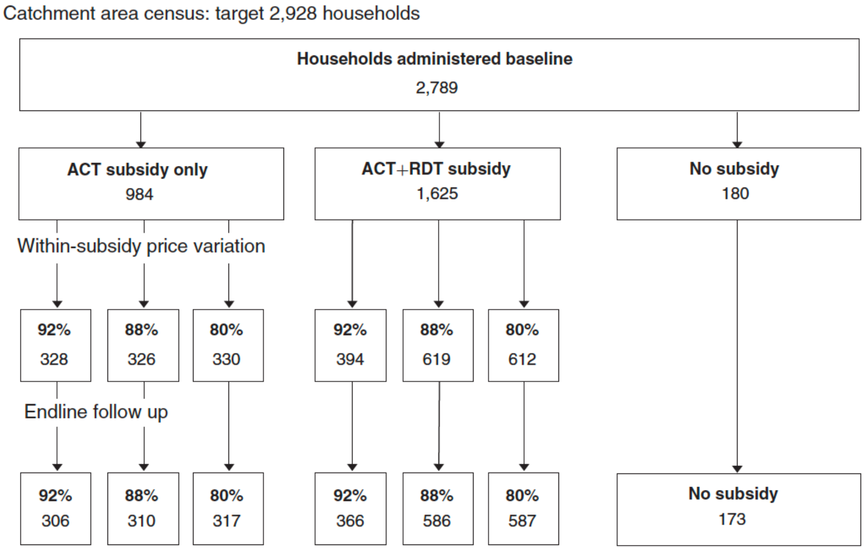
Cohen et al (2015): Effect of RDT



Cohen et al (2015): Experimental Design

- Experiment in Western Kenya in May–December 2009
- Sample 4 rural drug shops. Sample all households in 4 km catchment radius
- Ecery household interviewed for baseline survey
- At the end of the interview, households get 2 ACT vouchers and 2 RDT vouchers if applicable.
- Vouchers redeemable at the local drug shop
- Enumerators explain what RDT is and how it works

Cohen et al (2015): Experimental Design



Cohen et al (2015): Balance

	Control group mean (1)	Regression coefficients and standard errors: 92 percent ACT subsidy (T1) (2)	88 percent ACT subsidy (T2) (3)	80 percent ACT subsidy (T3) (4)	RDT subsidy (T4) (5)	Joint test: all subsidies = 0 (6)	Observations (7)
Characteristics of interviewed household head							
Female	0.867	0.017	0.029	0.040	0.010	1.25	2,789
	[0.341]	(0.029)	(0.028)	(0.028)	(0.012)	{0.287}	
Age (years)	41.7	−1.98	−3.22**	−2.44*	0.185	1.61	2,646
	[17.3]	(1.46)	(1.44)	(1.45)	(0.626)	{0.170}	
Education (years)	5.10	0.141	0.381	0.151	0.169	1.17	2,774
	[4.00]	(0.343)	(0.341)	(0.342)	(0.161)	{0.323}	
Literate	0.575	0.047	0.050	0.027	0.000	0.621	2,782
	[0.496]	(0.042)	(0.042)	(0.042)	(0.020)	{0.647}	
Married	0.783	−0.015	0.004	0.006	−0.015	0.514	2,784
	[0.413]	(0.035)	(0.035)	(0.034)	(0.016)	{0.725}	
Subsistence farmer	0.589	0.052	0.039	0.059	−0.005	0.612	2,787
	[0.493]	(0.042)	(0.042)	(0.042)	(0.019)	{0.654}	
Number dependents	4.12	−0.263	−0.096	−0.077	0.021	0.809	2,663
	[2.78]	(0.223)	(0.221)	(0.222)	(0.098)	{0.519}	

Cohen et al (2015): Balance

	Control group mean (1)	92 percent ACT subsidy (T1) (2)	88 percent ACT subsidy (T2) (3)	80 percent ACT subsidy (T3) (4)	RDT subsidy (T4) (5)	Joint test: all subsidies = 0 (6)	Observations (7)
		Regression coefficients and standard errors					
Household characteristics							
Number members	5.48	−0.354	−0.233	−0.197	0.024	0.885	2,789
	[2.77]	(0.217)	(0.214)	(0.215)	(0.092)	{0.472}	
Fraction adults (ages 14+)	0.623	−0.035*	−0.048***	−0.024	0.002	2.23*	2,337
	[0.235]	(0.020)	(0.019)	(0.020)	(0.009)	{0.063}	
Acres land	2.72	−0.660**	−0.601*	−0.571*	0.197*	1.63	2,250
	[3.69]	(0.330)	(0.327)	(0.324)	(0.117)	{0.164}	
Distance from drug shop (km)	1.68	0.012	0.012	0.002	0.010	0.523	2,788
	[0.917]	(0.023)	(0.022)	(0.022)	(0.011)	{0.719}	
Distance from closest clinic (km)	6.57	−0.018	−0.036	−0.043	0.044*	0.796	2,785
	[2.47]	(0.060)	(0.059)	(0.059)	(0.027)	{0.528}	
Baseline malaria knowledge and health practices							
Number bednets	1.77	−0.031	−0.060	0.028	0.005	0.476	2,784
	[1.43]	(0.120)	(0.121)	(0.120)	(0.057)	{0.753}	
Share HH members slept under net	0.561	0.023	0.006	0.030	−0.012	0.612	2,661
	[0.397]	(0.034)	(0.034)	(0.034)	(0.017)	{0.654}	
Only mosquitoes transmit malaria	0.517	0.045	0.011	0.024	−0.020	0.842	2,789
	[0.501]	(0.042)	(0.042)	(0.042)	(0.020)	{0.499}	

Cohen et al (2015): Balance

Heard of ACTs	0.399	0.016	0.017	0.030	0.001	0.197	2,771
	[0.491]	(0.042)	(0.041)	(0.042)	(0.020)	{0.940}	
ACT is preferred antimalarial	0.207	−0.023	−0.029	−0.049	−0.002	0.978	2,771
	[0.406]	(0.034)	(0.034)	(0.033)	(0.015)	{0.418}	
Heard of RDTs	0.128	0.039	0.020	0.021	−0.011	0.682	2,786
	[0.335]	(0.030)	(0.029)	(0.029)	(0.014)	{0.604}	
Treats water regularly	0.408	−0.036	−0.018	0.004	0.023	1.13	2,779
	[0.493]	(0.041)	(0.041)	(0.041)	(0.019)	{0.339}	
Number of presumed malaria episodes last month	1.20	0.015	−0.008	−0.029	0.033	0.200	2,789
	[1.22]	(0.102)	(0.103)	(0.103)	(0.050)	{0.939}	
Cost per episode (among those seeking care)							
Total cost (US $)	1.63	0.140	−0.040	−0.217	0.131	0.725	1,319
	[1.86]	(0.293)	(0.250)	(0.238)	(0.174)	{0.575}	
Sample size in treatment	180	328	326	330	1,625		

Notes: The first column shows average values of characteristics for the control group. Columns 2–5 show regression coefficients and standard errors on indicated treatment groups (the omitted category is the control group). All regressions include a full set of strata dummies. Column 6 shows F-statistics and p-values from a test of whether the three ACT subsidy coefficients are jointly equal to zero. Standard deviations are in brackets, standard errors are in parentheses, and p-values are in braces. All tests are based on heteroskedasticity robust standard errors. The exchange rate at the time of the study was around 78 Ksh to US$1.

Cohen et al (2015): Data

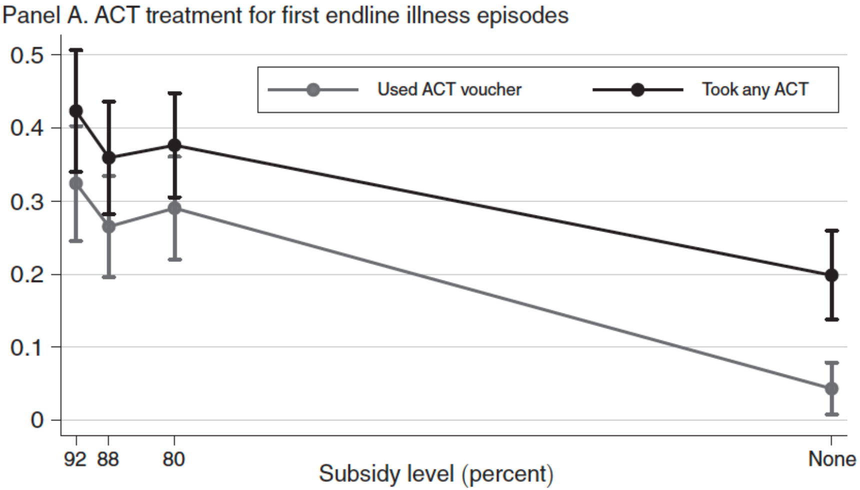
- 3 data sources

1. Administrative data from drug shop. Captured by surveyors posted at the 4 shops every single day. Contains 1,700 drug shop visits over 4 months.
   1.1 Also administer “surprise ADTs” to random subset of people who redeem ACT voucher (to measure true malarial status)
2. Endline survey data from 4 months after vouchers distributed. Includes recall data on all illnesses, where/what treatment sought.
3. Symptoms database: 1-year after vouchers, surveyors did unannounced household survey. Ask if anyone is ill and collect all symptoms and administer RDT. Use these to construct “predicted” malaria scores (proxy for π)

Cohen et al (2015): ACT Acces

	Took ACT (1)	Took ACT from drug shop (2)	Took ACT from health center (3)	Visited drug shop (4)	Visited health center (5)	Sought no care (6)	Took malaria test (7)	Took antibiotic (8)
Panel A. Pooled impact								
Any ACT subsidy	0.187***	0.222***	−0.038	0.167***	−0.079*	−0.096***	−0.014	−0.072**
	(0.038)	(0.031)	(0.030)	(0.046)	(0.042)	(0.036)	(0.038)	(0.034)
Panel B. Impact by subsidy level								
B1. ACT subsidy = 92 percent	0.225***	0.249***	−0.024	0.159***	−0.055	−0.110***	−0.031	−0.046
	(0.053)	(0.046)	(0.037)	(0.058)	(0.053)	(0.042)	(0.048)	(0.043)
B2. ACT subsidy = 88 percent	0.161***	0.217***	−0.056	0.167***	−0.070	−0.097**	−0.042	−0.062
	(0.050)	(0.043)	(0.037)	(0.058)	(0.052)	(0.042)	(0.047)	(0.040)
B3. ACT subsidy = 80 percent	0.178***	0.206***	−0.035	0.173***	−0.106**	−0.085*	0.023	−0.100***
	(0.048)	(0.042)	(0.035)	(0.054)	(0.047)	(0.045)	(0.046)	(0.038)
p-value: B1 = B2 = B3 = 0	0.000***	0.000***	0.498	0.004***	0.164	0.048**	0.533	0.066
p-value: B1 = B2 = B3	0.531	0.723	0.660	0.968	0.535	0.846	0.362	0.304
DV mean (control group)	0.190	0.071	0.119	0.488	0.286	0.226	0.214	0.185
Observations	631	631	631	631	631	631	631	631

Cohen et al (2015): Subsidy Level



Cohen et al (2015): Targeting

$$pos_h = \beta_0 + \beta_1 ACT88_h + \beta_2 ACT80_h + \varepsilon_h$$

TABLE 3—IMPACT OF RETAIL SECTOR ACT SUBSIDY ON ACT TARGETING

	Actual malaria status (1)	Predicted positivity (2)	Predicted positivity (3)
A. ACT subsidy = 88 percent	0.187**	0.112***	0.111**
	(0.081)	(0.042)	(0.053)
B. ACT Subsidy = 80 percent	0.182**	0.107**	0.040
	(0.084)	(0.043)	(0.052)
p-value: A = B = 0	0.038**	0.012**	0.104
p-value: A = B	0.955	0.906	0.179
DV mean (ACT 92 percent, no RDT)	0.563	0.424	0.422
Observations	190	189	178
Data source	Admin.	Admin.	Endline

Cohen et al (2015): Mechanism

	Used first voucher for patient under 14 (1)	Used first voucher for patient 14 or older (2)
Panel A. Does the ACT subsidy level reallocate ACTs across dosage groups?		
A. ACT subsidy = 88 percent	0.035	−0.057**
	(0.035)	(0.027)
B. ACT subsidy = 80 percent	0.031	−0.080***
	(0.034)	(0.026)
p-value: A = B = 0	0.540	0.007***
DV mean (ACT 92 percent, no RDT)	0.268	0.171
Observations	984	984
Subsample	All households	All households

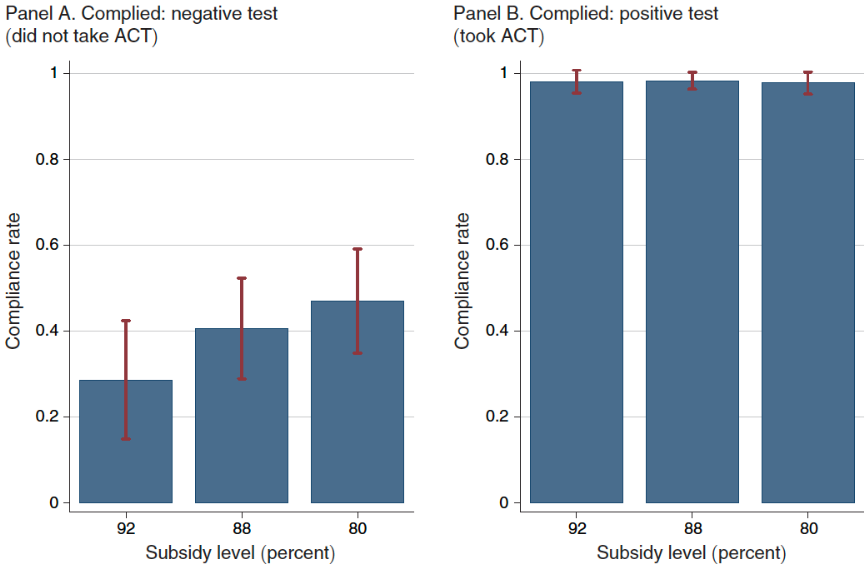
Cohen et al (2015): RDT

	Visited drug shop	Visited health center	Sought no care	Took malaria test	Took RDT test	Took microscopy test	Took ACT	Took antibiotic
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Across all ACT subsidy levels								
RDT subsidy	0.004	−0.013	0.010	0.216***	0.215***	−0.014	0.018	0.020
	(0.026)	(0.022)	(0.018)	(0.023)	(0.017)	(0.018)	(0.026)	(0.017)
DV mean (no RDT)	0.657	0.212	0.123	0.207	0.076	0.125	0.389	0.110
Panel B. By ACT subsidy level								
RDT subsidy × 92% ACT subsidy	−0.005	−0.018	0.029	0.258***	0.263***	−0.019	0.002	0.004
	(0.048)	(0.042)	(0.032)	(0.044)	(0.034)	(0.034)	(0.050)	(0.033)
RDT subsidy × 88% ACT subsidy	0.026	−0.045	0.007	0.252***	0.229***	0.000	0.042	−0.016
	(0.046)	(0.041)	(0.030)	(0.039)	(0.030)	(0.032)	(0.044)	(0.030)
RDT subsidy × 80% ACT subsidy	−0.012	0.023	−0.003	0.152***	0.166***	−0.021	0.016	0.070**
	(0.043)	(0.035)	(0.033)	(0.040)	(0.029)	(0.030)	(0.041)	(0.028)
88% ACT subsidy	−0.006	−0.002	0.014	−0.013	0.004	−0.016	−0.067	−0.011
	(0.058)	(0.052)	(0.038)	(0.048)	(0.032)	(0.041)	(0.058)	(0.038)
80% ACT subsidy	0.009	−0.041	0.020	0.050	0.028	0.007	−0.058	−0.047
	(0.055)	(0.047)	(0.040)	(0.049)	(0.032)	(0.040)	(0.056)	(0.035)
p-value: RDT terms jointly = 0	0.938	0.612	0.832	0.000***	0.000***	0.851	0.787	0.079*
DV mean (ACT 92%, No RDT)	0.667	0.222	0.104	0.194	0.069	0.125	0.444	0.125

Cohen et al (2015): RDT and targeting

	Household sought treatment at drug shop (1)	Surprise RDT reveals that patient is malaria-positive: Sample: patients who visited drug shop (2)	Sample: patients who bought subsidized ACT at drug shop (3)	Proportion that redeemed RDT voucher, conditional on seeking treatment at drug shop (4)
Panel A. Across all ACT subsidy levels				
RDT subsidy	0.025	0.009	0.081**	0.818
	(0.026)	(0.039)	(0.039)	
Panel B. By ACT subsidy level				
RDT subsidy × 92% ACT subsidy	0.028	0.127*	0.163**	0.792
	(0.045)	(0.070)	(0.070)	
RDT subsidy × 88% ACT subsidy	0.052	−0.058	0.018	0.837
	(0.044)	(0.063)	(0.062)	
RDT subsidy × 80% ACT subsidy	−0.010	−0.047	0.061	0.818
	(0.047)	(0.068)	(0.067)	
DV mean (ACT 92%, no RDT)	0.429	0.556	0.563	—
Observations	1,776	755	687	573

Cohen et al (2015): RDT compliance



Cohen et al (2015): Alternative Subsidy Schemes

	No subsidy (1)	ACT 92 percent subsidy (2)	ACT 88 percent subsidy (3)	ACT 80 percent subsidy (4)	ACT 80 percent + RDT subsidy (5)
Experimental estimates of access and drug shop targeting					
Total share taking ACT	0.190	0.415	0.351	0.369	0.385
Share taking ACT at drug shop	0.071	0.320	0.288	0.278	0.303
Share taking ACT at health center	0.119	0.095	0.063	0.084	0.078
Targeting at drug shop	1.000	0.563	0.750	0.745	0.806
Assumptions for estimates of under- and over-treatment					
Share of illness episodes that are malaria[a]	0.386	0.386	0.386	0.386	0.386
Targeting at health center (medium)[b]	0.750	0.750	0.750	0.750	0.750
Targeting at health center (high)	1.000	1.000	1.000	1.000	1.000
Targeting at health center (low)	0.650	0.650	0.650	0.650	0.650
Under- and over-treatment: Preferred estimates (assuming medium targeting at health center)					
Overall targeting	0.844	0.606	0.750	0.747	0.795
Over-treatment	0.048	0.266	0.143	0.152	0.129
Under-treatment	0.583	0.347	0.317	0.287	0.207

Outline

Targeting in Developing Countries: Who gets the Benefit?

- Barnwal (IGC 2017) *Curbing Leakage in Public Programs* [ET]
- Alatas, Banerjee, Hanna, Olken & Tobias (AER 2012) *Targeting the Poor: Evidence from a Field Experiment in Indonesia*
- Alatas, Banerjee, Hanna, Olken, Purnamasari & Wai-Poi (JPE 2016) *Self-Targeting: Evidence from a Field Experiment in Indonesia*
- Cohen Dupas & Schaner (AER 2015) *Price Subsidies, Diagnostic Tests, and Targeting of Malaria Treatment: Evidence from a Randomized Controlled Trial*
- Banerjee Hanna Olken & Sumarto (WP 2018) *The (lack of) Distortionary Effects of Proxy-Means Tests: Results from a Nationwide Experiment in Indonesia*

Banerjee et al (2018): Overview

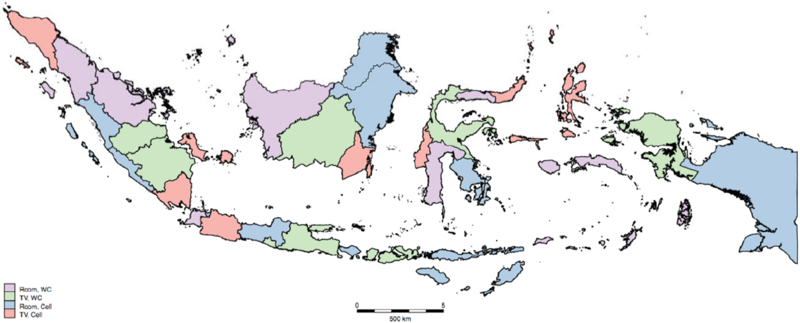
- Proxy-means tests are a very common way to target transfers.
- We might expect that these tests induce behavioral responses, distorting consumption baskets.
- This paper:
  - add questions to Indonesia's poverty census.
  - Survey households 6 months later to ask about consumption $\Rightarrow$ reported response
  - Look at administrative data on sales to assess $\Rightarrow$ actual response

Banerjee et al (2018): Setting

- Indonesia conducts nationwide censuses of the poor approximately every 3 years.
- Proxy-means testing determines households' eligibility for transfer programs from cash transfers to health insurance.
- Most recent census: June-August 2015, covered 25 mn households, 92 mn individuals.
- Experiment: Add 2 questions to the survey:
  1. either: Q on flat-screen TV ownership or # rooms in the house
  2. either: Q on # active SIM cards or whether have "swan neck" toilet
- Households probably believe that these additional questions will be used to determine benefit eligibility
- The questions were *not* actually used in the PMT formula.
- Randomized across the 34 provinces (stratifying by 5 regions).

Banerjee et al (2018): Randomization

Figure 1: Map of Randomization



Notes: This map shows the treatment assignment of each of Indonesia's 34 provinces.

Banerjee et al (2018): Data

- Use 3 main datasets

1. household survey data (SUSENAS) for March 2016 and March 2017 (300k households each). Worked with Statistics agency to add relevant questions.
2. Monthly TV sales of flat-screen TVs from January 2013–December 2016. Data aggregated to randomization arm$\times$month.
3. Yearly active SIM cards by province for 2015–2017 from Ministry of Information and Communications.

Banerjee et al (2018): Effect on Self-Reported Assets

- Estimate

$$Asset_{hdp} = \beta_0 + \beta_1 TVTreat_p + \beta_2 CellTreat_p + \mathbf{X}'_{dp}\boldsymbol{\gamma} + \alpha_r + \varepsilon_{hdp}$$

- $Asset_{hdp}$ is self-reported asset for HH H in district d in province p.

- Randomization ensures balance in expectation. To improve precision

  1. Include strata fixed effects α_r

  2. Code 1,388 asset variables from 2007-2015 SUSENAS surveys and average by district $\times$ urbanization. Use double-LASSO to select controls (Belloni, Chernozhukov & Hansen, 2014)

Banerjee et al (2018): Effect on Self-Reported Assets

Table 1: Treatment Effect on Self-Reported Asset Acquisition

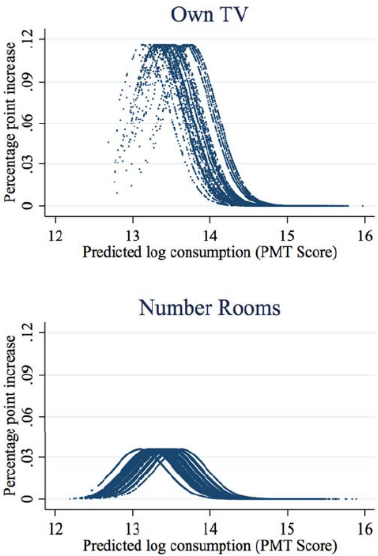
	(1) Own TV	(2) Nb. Sim Cards	(3) Nb. Rooms	(4) Own WC
	Panel A: 2016 Outcome Data			
TV Treatment	-0.0171	-0.00319		
	(0.00447)	(0.0318)		
Cell Treatment	-0.00265	0.0125		
	(0.00463)	(0.0277)		
Room Treatment			-0.140	-0.000209
			(0.179)	(0.00468)
WC Treatment			0.128	0.00587
			(0.160)	(0.00468)
Observations	291,414	291,414	291,414	291,414
Controls	Lasso	Lasso	Lasso	Lasso
Strata FE	YES	YES	YES	YES
Dep. Variable Mean	0.110	2.183	6.150	0.672
FWER adjusted p-value	0.005	0.735	0.735	0.584

Banerjee et al (2018): Effect on Self-Reported Assets

	Own TV	Nb. People with Phones	Nb. Rooms	Own WC
	Panel B: 2017 Outcome Data			
TV Treatment	-0.00463	-0.0169		
	(0.00524)	(0.0330)		
Cell Treatment	0.00505	-0.0217		
	(0.00473)	(0.0319)		
Room Treatment			-0.196	-0.00118
			(0.167)	(0.00602)
WC Treatment			0.0466	0.00433
			(0.152)	(0.00600)
Observations	297,276	297,276	297,276	297,276
Controls	Lasso	Lasso	Lasso	Lasso
Strata FE	YES	YES	YES	YES
Dep. Variable Mean	0.116	1.957	6.229	0.696
FWER adjusted p-value	0.823	0.823	0.722	0.823

Banerjee et al (2018): Effect on Self-Reported Assets

- Why only lie for TV? Redo PMT adding in these 4 attributes.



Banerjee et al (2018): Effect on Asset Acquisition

- To estimate impact on TV sales estimate

$$LogSales_{mg} = \beta_0 + \beta_1 TVTreat \times Post_{mg} + \beta_2 CellTreat \times Post_{mg} \\ + \beta_3 Post_{mg} + \alpha_g \times m + \varepsilon_{gm}$$

  where $LogSales_{mg}$ is sales in month m in randomization group g (TV-cell, TV-toilet, room-cell, room-toilet)

- Need to think about the time-series structure of the data

1. Use Newey-West (1987) standard errors with 3 lags
2. panel-corrected model with AR(1) errors.

Banerjee et al (2018): Effect on Asset Acquisition

Table 2: Treatment Effect on Actual Television Sales

	(1) Log Sales	(2) Log Sales	(3) Log Sales
TV Treatment x Post	0.0540	0.0540	0.0517
	(0.0563)	(0.0806)	(0.0475)
Cell Treatment x Post	0.190	0.190	0.0771
	(0.0563)	(0.0806)	(0.0505)
Observations	192	192	192
Model/Standard Errors	Robust	Newey	Panel-Corrected AR(1)
Dep. Variable Mean	10.77	10.77	10.77
Treat coeff = -0.97 F-Statistic	331	161.3	463.3
Treat coeff = -0.97 P-Value	0	0	0

Banerjee et al (2018): Effect on Asset Acquisition

- To estimate impact on SIM card ownership estimate

$$LogSubscribers_p = \beta_0 + \beta_1 CellTreat_p + \beta_2 TVTreat_p + \varepsilon_p$$

in province-level data separately in each year.

Banerjee et al (2018): Effect on Asset Acquisition

Table 3: Treatment Effect on SIM Card Ownership

	(1) Log Subscribers	(2) Log Subscribers	(3) Log Subscribers
	Panel A: 2015 Data		
Cell Treatment	-0.225	-0.102	-0.106
	(0.406)	(0.146)	(0.153)
TV Treatment	-0.258	-0.192	-0.189
	(0.418)	(0.156)	(0.148)
Observations	34	34	34
Log population control	N	Y	Y
Strata FE	N	N	Y
Dep. Variable Mean	14.95	14.95	14.95
	Panel B: 2016 Data		
Cell Treatment	-0.251	-0.129	-0.135
	(0.401)	(0.148)	(0.159)
TV Treatment	-0.249	-0.184	-0.175
	(0.414)	(0.159)	(0.152)
Observations	34	34	34
Log population control	N	Y	Y
Strata FE	N	N	Y
Dep. Variable Mean	15.17	15.17	15.17

Banerjee et al (2018): Effect on Asset Acquisition

	Panel C: 2017 Data		
Cell Treatment	-0.173	-0.0478	-0.0529
	(0.403)	(0.129)	(0.109)
TV Treatment	-0.0445	0.0227	0.0121
	(0.408)	(0.127)	(0.103)
Observations	34	34	34
Log population control	N	Y	Y
Strata FE	N	N	Y
Dep. Variable Mean	15.55	15.55	15.55

Notes: This table provides estimates of the treatment effects of the different targeting questions in the PBDT on actual active SIM card subscribers. We have yearly, province level data from 2015 to 2017. All regressions are estimated using OLS, with robust standard errors.

Outline

Outline

Transfer Design: What is the Benefit?

- [ET] Baird McIntosh & Özler (QJE 2011) *Cash or Condition? Evidence from a Cash Transfer Experiment*
- Baird McIntosh & Özler (QJE 2011) *Cash or Condition? Evidence from a Cash Transfer Experiment*
- Cunha, De Giorgi & Jayachandran (2017) *The Price Effects of Cash Versus In-Kind Transfers*

[ET] Baird et al (2011): Overview

- RCT in Malawi with CCT and UCT treatment + control.
- Are conditions good or bad?
  - +CCT - market failures (like underinvestment in education or health) can be addressed with conditions
  - +CCT - political palatability to middle and upper class programs
  - +UCT - What is the marginal contribution of CCTs? If unclear, assume that rational poor will would use money in a way that benefits them most.
  - Do CCTs end up distorting other behaviors
- In addition to just observing education, also observe: human capital formation (english, math, cognitive skills), marriage, childbearing.
- Marriage and schooling practically mutually exclusive in study

[ET] Baird et al (2011): Overview 2

- Trade off inherent to CCT programs
  - Dropout rates decline in both UCT and CCT, effect of UCT is 43% as large.
  - CCT had significantly better comprehension and math but not UCT
  - But UCT reduced marriage (27%) and pregnancy (44%) rates more
- Past studies might have overlooked effects of denying benefits to those who fail to satisfy conditions

[ET] Baird et al (2011): Study Design

- Malawi (2008 GNI = \$760\$, secondary school enrollment =24%, and 81% living in rural areas)
- Zomba district, stratification on distance from Zomba city up to 16km.
- 176 EAs selected and each dwelling visited to make a full-listing of never-married females.
- Treatment status at the EA level
  - CCT: 46 EAs,
  - UCT: 27 EAs,
  - 15 with no cash transfers (control?)
  - 88 EAS for control
- UCTs paid the same amount as with the conditional transfer.
- Lottery held to determine transfer amount (\$1,2,3,4,5 to girls), (\$4,6,8,10) to parents

[ET] Baird et al (2011): Estimation Strategy

$$Y_i = T_i^C \gamma^C + T_i^U \gamma^U + X_i \beta + \epsilon_i$$

- Y_i- outcome variable for individual i
- $T_i^c$and T_i^Ubinary indicators ofr offers to be in the CCT and the UCT arms
- $X_i$vector of baseline characteristics: household asset index, highest grade attended, dummy variable for 'having started sexual activity' and dummy variables for age - chosen because they are strongly predictive of schooling outocmes and so improve precision.
- cluster errors at the EA level

[ET] Baird et al (2011): Enrolment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Program impacts on *self-reported* school enrollment								
	Dependent variable: =1 if enrolled in school during the relevant term							
	Year 1: 2008			Year 2: 2009			Year 3: 2010	
	Term 1	Term 2	Term 3	Term 1	Term 2	Term 3	Total terms (6 terms)	Term 1, post-program
Conditional treatment	0.007	0.019*	0.041**	0.049***	0.056***	0.061***	0.233***	0.005
	(0.011)	(0.011)	(0.017)	(0.017)	(0.018)	(0.019)	(0.070)	(0.025)
Unconditional treatment	0.034***	0.051***	0.054***	0.072***	0.095***	0.101***	0.406***	0.074***
	(0.010)	(0.011)	(0.018)	(0.021)	(0.022)	(0.021)	(0.079)	(0.026)
Mean in the control group	0.958	0.934	0.900	0.831	0.800	0.769	5.191	0.641
Number of observations	2,087	2,087	2,086	2,087	2,087	2,087	2,086	2,086
Prob > F(Conditional = Unconditional)	0.006	0.012	0.460	0.299	0.102	0.098	0.038	0.028
Panel B: Program impacts on *teacher-reported* school enrollment								
Conditional treatment	0.043***	0.044***	0.061***	0.094**	0.132***	0.113***	0.535***	0.058*
	(0.015)	(0.016)	(0.018)	(0.041)	(0.035)	(0.039)	(0.129)	(0.033)
Unconditional treatment	0.020	0.038**	0.018	0.027	0.059	0.033	0.231*	0.001
	(0.015)	(0.017)	(0.023)	(0.038)	(0.037)	(0.039)	(0.136)	(0.036)
Mean in the control group	0.906	0.881	0.852	0.764	0.733	0.704	4.793	0.596
Number of observations	2,023	2,023	2,023	852	852	852	852	847
Prob > F(Conditional = Unconditional)	0.173	0.732	0.067	0.076	0.014	0.020	0.011	0.108

[ET] Baird et al (2011): Misreporting

	Dependent variable	
	(1)	(2)
	Core respondents over-reporting	Teachers over-reporting
Conditional treatment	−0.093*	−0.021
	(0.052)	(0.035)
Unconditional treatment	−0.001	−0.014
	(0.058)	(0.038)
Mean in the control group	0.170	0.052
Number of observations	325	325
Prob > F(Conditional = Unconditional)	0.02	0.79

[ET] Baird et al (2011): Attendance

	Dependent variable: Fraction of days respondent attended school				
	(1)	(2)	(3)	(4)	(5)
	Term 1, 2009	Term 2, 2009	Term 3, 2009	Overall 2009	Term 1, 2010
Conditional treatment	0.139***	0.014	0.169**	0.080**	0.092**
	(0.045)	(0.033)	(0.085)	(0.035)	(0.041)
Unconditional treatment	0.063	0.038	0.118	0.058	−0.038
	(0.056)	(0.033)	(0.102)	(0.037)	(0.053)
Mean in the control group	0.778	0.849	0.688	0.810	0.801
Number of observations	284	285	192	319	211
Prob > F(Conditional = Unconditional)	0.129	0.334	0.358	0.436	0.010

[ET] Baird et al (2011): Attainment

	Dependent variable			
	(1)	(2)	(3)	(4)
	English test score (standardized)	TIMMS math score (standardized)	Non-TIMMS math score (standardized)	Cognitive test score (standardized)
Conditional treatment	0.140***	0.120*	0.086	0.174***
	(0.054)	(0.067)	(0.057)	(0.048)
Unconditional treatment	−0.030	0.006	0.063	0.136
	(0.084)	(0.098)	(0.087)	(0.119)
Number of observations	2,057	2,057	2,057	2,057
Prob > F(Conditional= Unconditional)	0.069	0.276	0.797	0.756

[ET] Baird et al (2011): Marriage & Pregnancy

	Dependent variable			
	(1)	(2)	(3)	(4)
	=1 if ever married		=1 if ever pregnant	
	Round 2	Round 3	Round 2	Round 3
Conditional treatment	0.007	−0.012	0.013	0.029
	(0.012)	(0.024)	(0.014)	(0.027)
Unconditional treatment	−0.026**	−0.079***	−0.009	−0.067***
	(0.012)	(0.022)	(0.017)	(0.024)
Mean in the control group	0.043	0.180	0.089	0.247
Number of observations	2,087	2,084	2,086	2,087
Prob > F(Conditional = Unconditional)	0.024	0.025	0.265	0.003

[ET] Baird et al (2011): Making sense of results

Consider 3 groups:

1. UCT compliers: - UCT would be sufficient to keep girls enrolled in school - receive payments in both UCT and CCT case - differe only on behavior changes on the intensive margin arising from conditionality
2. CCT compliers - enrolled in CCT but not in CCT: conditionality generates a differential impact on enrollment by lowering the opportunity cost of schooling
3. noncompliers - those who dropout of school in either treatment arm.

Depending on the numbers of 2 vs 1, a CCT or a UCT may be more beneficial.

[ET] Baird et al (2011): Discussion

- Only way CCTs can reduce marriage rates at follow-up is by averting dropouts
- UCTs avoid marriage and pregnancy rate also among drop-outs, (in this experiment, this group is much larger).
- Heterogeneity - increasing transfer amounts or varying the recipient in the household doesn't have effects on CCTs, but they do on on UCTs (performance on test scores).
- Actual amount of transfers made per person is 19% lower in the CCT arm
- Poverty a cause of school drop out – income effects from UCT might be more effective than school CCTS in reducing teen marriage and pregnancy.
- Income support might improve outcomes more cost-effectively than CCTs.

Outline

Transfer Design: What is the Benefit?

- [ET] Baird McIntosh & Özler (QJE 2011) *Cash or Condition? Evidence from a Cash Transfer Experiment*
- Baird McIntosh & Özler (QJE 2011) *Cash or Condition? Evidence from a Cash Transfer Experiment*
- Cunha, De Giorgi & Jayachandran (2017) *The Price Effects of Cash Versus In-Kind Transfers*

Baird et al (2011): Overview

- Should cash transfers come with conditions?
  - CCT: Market failures lead to underinvestment in education/health, conditions make transfers easier to "sell" politically
  - UCT: Conditions uneffective, and very costly to enforce
- Conditions are common around the world (attend school, attend clinics for checkups, government work) but
  - are they effective at increasing targeted behavior?
  - What other behaviors do they end up distorting?
- Explore these questions in an experiment in Malawi

Baird et al (2011): Setting

- Work in Zomba District in southern Malawi
- Sample 176 of the 550 Enumeration Areas (EAs) in 3 strata. 29 in Zomba city 119 within 16 km, 28 "far rural".
- Survey to get census of never-married females aged 13-22. Those in school at baseline (87%) are the target population for the study.
- Randomly sample, stratifying by age and stratum, to get 2,907 schoolgirls.

Baird et al (2011): Experiment

T1: CCT arm (46 EAs). 12/2007 & 1/2008. offered parents monthly transfer on condition regularly attend school. Transfer amount to the parent randomly varied, \$4, \$6, \$8, \$10/month, and to the schoolgirl \$1, \$2, \$3, \$4, \$5. Paid school fees.

T2: UCT arm (27 EAs). Identical offers, but no requirement to attend school

- Controls: (88 EAs).
- Track attendance, other outcomes for 2008, 2009

Baird et al (2011): Attrition

	Dependent variable					
	(1)	(2)	(3)	(4)	(5)	(6)
	=1 if surveyed in Round 3	=1 if surveyed in all three Rounds	=1 if took educational tests	=1 if information found in Round 2 survey	=1 if information found in Round 3 school survey	=1 if legible ledger found
Conditional treatment	0.020	0.021	0.029*	0.033	−0.000	0.116*
	(0.015)	(0.030)	(0.016)	(0.024)	(0.027)	(0.064)
Unconditional treatment	0.021	0.030	0.035*	−0.029	0.014	0.061
	(0.019)	(0.024)	(0.020)	(0.053)	(0.028)	(0.077)
Mean in the control group	0.946	0.893	0.929	0.890	0.935	0.378
Number of observations	2,284	2,284	2,284	2,284	983	821
Prob > F(Conditional = Unconditional)	0.965	0.797	0.801	0.246	0.627	0.513

Baird et al (2011): Enrolment

Panel A: Program impacts on *self-reported* school enrollment								
	Dependent variable: =1 if enrolled in school during the relevant term							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Year 1: 2008			Year 2: 2009			Year 3: 2010	
	Term 1	Term 2	Term 3	Term 1	Term 2	Term 3	Total terms (6 terms)	Term 1, post-program
Conditional treatment	0.007	0.019*	0.041**	0.049***	0.056***	0.061***	0.233***	0.005
	(0.011)	(0.011)	(0.017)	(0.017)	(0.018)	(0.019)	(0.070)	(0.025)
Unconditional treatment	0.034***	0.051***	0.054***	0.072***	0.095***	0.101***	0.406***	0.074***
	(0.010)	(0.011)	(0.018)	(0.021)	(0.022)	(0.021)	(0.079)	(0.026)
Mean in the control group	0.958	0.934	0.900	0.831	0.800	0.769	5.191	0.641
Number of observations	2,087	2,087	2,086	2,087	2,087	2,087	2,086	2,086
Prob > F(Conditional = Unconditional)	0.006	0.012	0.460	0.299	0.102	0.098	0.038	0.028
Panel B: Program impacts on *teacher-reported* school enrollment								
Conditional treatment	0.043***	0.044***	0.061***	0.094**	0.132***	0.113***	0.535***	0.058*
	(0.015)	(0.016)	(0.018)	(0.041)	(0.035)	(0.039)	(0.129)	(0.033)
Unconditional treatment	0.020	0.038**	0.018	0.027	0.059	0.033	0.231*	0.001
	(0.015)	(0.017)	(0.023)	(0.038)	(0.037)	(0.039)	(0.136)	(0.036)
Mean in the control group	0.906	0.881	0.852	0.764	0.733	0.704	4.793	0.596
Number of observations	2,023	2,023	2,023	852	852	852	852	847
Prob > F(Conditional = Unconditional)	0.173	0.732	0.067	0.076	0.014	0.020	0.011	0.108

Baird et al (2011): Misreporting

	Dependent variable	
	(1)	(2)
	Core respondents over-reporting	Teachers over-reporting
Conditional treatment	−0.093*	−0.021
	(0.052)	(0.035)
Unconditional treatment	−0.001	−0.014
	(0.058)	(0.038)
Mean in the control group	0.170	0.052
Number of observations	325	325
Prob > F(Conditional = Unconditional)	0.02	0.79

Baird et al (2011): Attendance

	Dependent variable: Fraction of days respondent attended school				
	(1)	(2)	(3)	(4)	(5)
	Term 1, 2009	Term 2, 2009	Term 3, 2009	Overall 2009	Term 1, 2010
Conditional treatment	0.139***	0.014	0.169**	0.080**	0.092**
	(0.045)	(0.033)	(0.085)	(0.035)	(0.041)
Unconditional treatment	0.063	0.038	0.118	0.058	−0.038
	(0.056)	(0.033)	(0.102)	(0.037)	(0.053)
Mean in the control group	0.778	0.849	0.688	0.810	0.801
Number of observations	284	285	192	319	211
Prob > F(Conditional = Unconditional)	0.129	0.334	0.358	0.436	0.010

Baird et al (2011): Attainment

	Dependent variable			
	(1)	(2)	(3)	(4)
	English test score (standardized)	TIMMS math score (standardized)	Non-TIMMS math score (standardized)	Cognitive test score (standardized)
Conditional treatment	0.140***	0.120*	0.086	0.174***
	(0.054)	(0.067)	(0.057)	(0.048)
Unconditional treatment	−0.030	0.006	0.063	0.136
	(0.084)	(0.098)	(0.087)	(0.119)
Number of observations	2,057	2,057	2,057	2,057
Prob > F(Conditional= Unconditional)	0.069	0.276	0.797	0.756

Baird et al (2011): Marriage & Pregnancy

	Dependent variable			
	(1)	(2)	(3)	(4)
	=1 if ever married		=1 if ever pregnant	
	Round 2	Round 3	Round 2	Round 3
Conditional treatment	0.007	−0.012	0.013	0.029
	(0.012)	(0.024)	(0.014)	(0.027)
Unconditional treatment	−0.026**	−0.079***	−0.009	−0.067***
	(0.012)	(0.022)	(0.017)	(0.024)
Mean in the control group	0.043	0.180	0.089	0.247
Number of observations	2,087	2,084	2,086	2,087
Prob > F(Conditional = Unconditional)	0.024	0.025	0.265	0.003

Baird et al (2011): Decomposition

- How to rationalize these results? Imagine 3 strata of schoolgirls:

1. UCT Compliers: UCT is sufficient to keep them in school. Differences in program impact must be due to intensive margin responses to conditionality
2. CCT Compliers: Enrolled under CCT but not UCT. Conditionality lowers opportunity cost of schooling.
3. Noncompliers: Never enrol. Only receive transfers under UCT.

- Overall effects depend on sizes of the three strata and effects in each group.

Baird et al (2011): Strata Sizes

	(1) Enrolled	(2) Not enrolled	(3) Total
Control, %	1.7	46.9	19.9
(row %)	(59.8)	(40.2)	(100.0)
Conditional treatment, %	0.5	50.8	16.0
(row %)	(69.2)	(30.8)	(100.0)
Unconditional treatment, %	0.3	25.2	10.1
(row %)	(60.5)	(39.5)	(100.0)
Total, %	1.1	44.2	17.2
(row %)	(62.7)	(37.3)	(100.0)

Baird et al (2011): Enrolment and Marriage

	Dependent variable			
	(1)	(2)	(3)	(4)
	=1 if enrolled term 1 2010	=1 if ever married	=1 if ever married	=1 if ever married
	All	All	Enrolled	Not enrolled
Conditional treatment	0.058*	−0.026	−0.012	0.033
	(0.034)	(0.037)	(0.015)	(0.097)
Unconditional treatment	−0.000	−0.088***	−0.011	−0.159**
	(0.036)	(0.030)	(0.010)	(0.067)
Mean in the control group	0.598	0.199	0.017	0.469
Sample size	844	844	490	354
Prob > F(Conditional = Unconditional)	0.099	0.106	0.857	0.088

Baird et al (2011): Age Heterogeneity

	Dependent variable			
	(1)	(2)	(3)	(4)
	Total number of terms enrolled (school survey)	Standardized English test score	=1 if ever married	=1 if ever pregnant
Conditional treatment	0.467***	0.141*	−0.023	−0.008
	(0.159)	(0.073)	(0.017)	(0.028)
Unconditional treatment	0.257	−0.116	−0.051**	−0.059***
	(0.157)	(0.102)	(0.020)	(0.020)
=1 if Over 15	−0.786***	−0.546***	0.122***	0.176***
	(0.244)	(0.058)	(0.026)	(0.027)
Conditional treatment * Over 15	0.290	0.017	0.037	0.104*
	(0.291)	(0.089)	(0.056)	(0.054)
Unconditional treatment * Over 15	0.103	0.245**	−0.067	−0.032
	(0.255)	(0.110)	(0.042)	(0.046)
Number of unique observations	852	2,057	2,084	2,087
Prob > F(Conditional = Unconditional)	0.095	0.031	0.188	0.067
Prob > F(Conditional * Older = Unconditional * Older)	0.364	0.059	0.097	0.027

Baird et al (20110: Transfer Amount Elasticities

	Dependent variable			
	(1)	(2)	(3)	(4)
	Total number of terms enrolled (school survey)	Standardized English test score	=1 if ever married	=1 if ever pregnant
Conditional treatment, individual amount	0.024	−0.032	−0.002	0.006
	(0.051)	(0.029)	(0.008)	(0.012)
Unconditional treatment, individual amount	−0.048	−0.019	−0.016	0.013
	(0.064)	(0.038)	(0.011)	(0.013)
Conditional treatment, household amount	−0.027	−0.000	0.001	0.005
	(0.035)	(0.016)	(0.007)	(0.010)
Unconditional treatment, household amount	0.081***	−0.058**	−0.017**	−0.002
	(0.031)	(0.029)	(0.007)	(0.009)
Conditional treatment, minimum transfer amounts	0.572***	0.202*	−0.011	0.001
	(0.213)	(0.118)	(0.044)	(0.052)
Unconditional treatment, minimum transfer amounts	0.094	0.175	0.001	−0.089*
	(0.167)	(0.132)	(0.040)	(0.050)
Number of unique observations	852	2,057	2,084	2,087
Prob > F(Conditional = Unconditional), individual amount	0.390	0.788	0.300	0.702
Prob > F(Conditional = Unconditional), household amount	0.025	0.082	0.069	0.614
Prob > F(Conditional = Unconditional), minimum amount	0.046	0.877	0.834	0.203

Outline

Transfer Design: What is the Benefit?

- [ET] Baird McIntosh & Özler (QJE 2011) *Cash or Condition? Evidence from a Cash Transfer Experiment*
- Baird McIntosh & Özler (QJE 2011) *Cash or Condition? Evidence from a Cash Transfer Experiment*
- Cunha, De Giorgi & Jayachandran (2017) *The Price Effects of Cash Versus In-Kind Transfers*

Cunha et al (2017): Overview

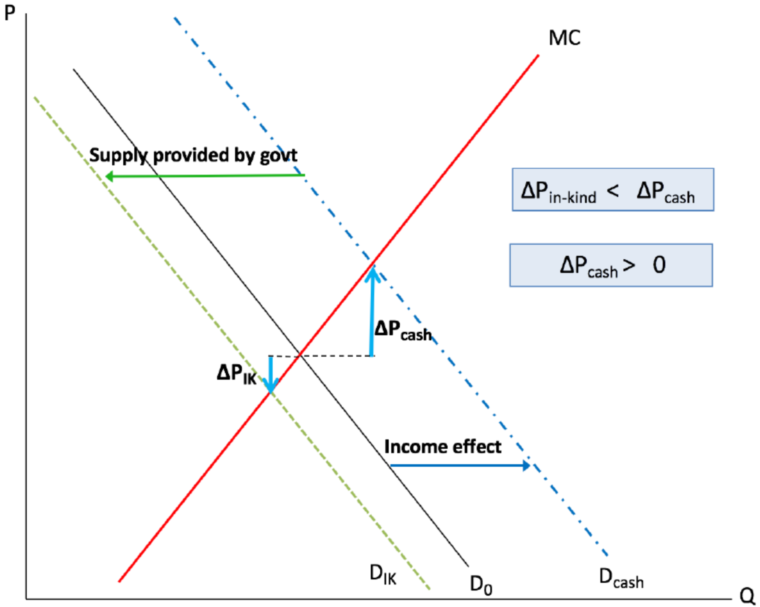
- Should transfers be cash or in-kind?
- In-kind transfers could do better at targeting.
- Cash transfers have lower admin costs, and give recipients freedom to choose.
- Also: transfers may affect prices.
  - cash transfers increase demand for normal goods $\rightarrow$ prices increase
  - in-kind transfers increase demand for normal goods, but also increase supply $\rightarrow$ lower prices under in-kind transfer
- Test this in an experiment in Mexico.

Cunha et al (2017): Conceptual Framework

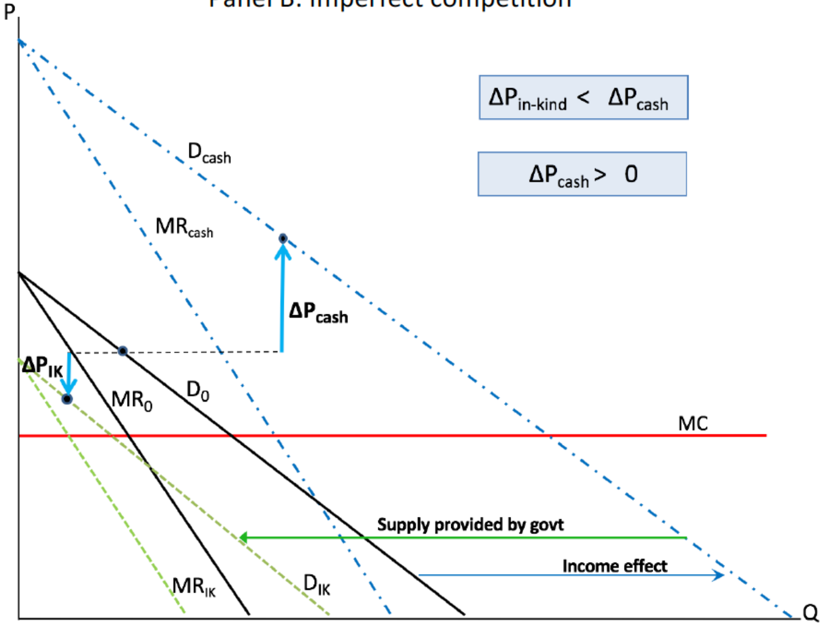
- Small open economy $\rightarrow$ no price effects. Prices pinned down by world prices
- Think of partially-closed village economies. Prices depend on local conditions.
- Local suppliers are shopkeepers, supplying packaged food bought from outside the village.
- Case 1: Perfect competition
  - Cash transfer of X_{Cash} $\rightarrow$ demand shifts out, prices go up: $\partial p / \partial X_{Cash} > 0$
  - In-kind transfer with value X_{InKind} $\rightarrow$ same demand shift. Influx of supply $\rightarrow$ residual demand facing local suppliers shifts left.

$$\frac{\partial p}{\partial X_{InKind}} - \frac{\partial p}{\partial X_{Cash}} < 0$$

Panel A: Perfect competition



Panel B: Imperfect competition



Cunha et al (2017): Conceptual Framework

- Case 3: Cournot-Nash.
- N firms with constant MC c and linear demand $p = d - Q$.
- Equilibrium prices: $p = (d + Nc) / (N + 1)$
- Transfer changes d. Δd larger for cash transfer. Then $\Delta p/p = \Delta d / (d + Nc)$
- In a more general Cournot model, can show dependence on degree of competition

$$\frac{\partial^2 p}{\partial N \partial X_{Cash}} < 0$$

$$\frac{\partial}{\partial N}\left(\frac{\partial p}{\partial X_{InKind}} - \frac{\partial p}{\partial X_{Cash}}\right) > 0$$

Cunha et al (2017): Context

- Study Programa de Apoyo Alimentario (PAL) in Mexico
- Operates in 5,000 very poor, rural villages in Mexico.
- Villages eligible if population under 2.5K, and classified as highly marginalized.
- Typically poorer than Progresa/Oportunidades villages.
- PAL provides monthly in-kind allotment:
  - 7 basics (corn flour, rice, beans, pasta, biscuits, fortified milk powder, vegetable oil)
  - 2-4 supplements (canned tuna/sardines, lentils, corn starch, chocolate powder, cereal)
- Allotment inframarginal for most households (consumption > allotment)

Cunha et al (2017): Experimental Design

- During roll-out of the program, 208 villages randomly selected
- Each village randomly assigned to in-kind, cash, or control.
- In-kind transfer of MX$150 is $\sim 18\%$ of households' food expenditures.
- Cash transfer is $\sim 8\%$ of recipients' income, 7% increase in village income.
- Value of the in-kind transfer: Market value is MX$ 206
  - 116 pesos-worth consumed (infra-marginal)
  - 35 pesos of extra consumption of transferred goods (extra-marginal, value at 2/3 discount)
  - 55 pesos not consumed, presumably resold (extra-marginal, assume transaction costs of selling 2/3 of value)
  - $\rightarrow$ value to recipients is 146 pesos

Cunha et al (2017): PAL transfer

Item	Type (1)	Amount per box (kg) (2)	Value per box (pre-program, in pesos) (3)	Calories, as % of total box (4)	Village change in supply (ΔSupply) (5)
Corn flour	basic	3	15.7	20%	1.00
Rice	basic	2	12.7	12%	0.61
Beans	basic	2	21.0	13%	0.29
Fortified powdered milk	basic	1.92	76.2	17%	8.62
Packaged pasta soup	basic	1.2	16.2	8%	0.93
Vegetable oil	basic	1 (lt)	10.4	16%	0.25
Biscuits	basic	1	18.7	8%	0.81
Lentils	supplementary	1	10.3	2%	3.73
Canned tuna/sardines	supplementary	0.6	14.8	2%	1.55
Breakfast cereal	supplementary	0.2	9.3	1%	0.90

Cunha et al (2017): Data and Empirical Strategy

- Want to estimate regressions like

$$p_{gsv} = \alpha + \beta_1 InKind_v + \beta_2 Cash_v + \phi p_{gv,t-1} + \sigma I_{gv} + \varepsilon_{gsv}$$

  where p_{gsv} is price of good g at store s in village v, I_{gv} indicates imputed price

- Data from surveys of stores and households. Baseline data from 2003Q4 & 2004Q1. Follow-up survey in 2005Q4.

- Survey data on prices of 66 food items from stores, markets.

- Price data missing for 19% of village-goods. Impute from household survey (expenditure/quantity).

- End up with 360 stores in 194 villages and 12,940 good-village-store observations

Cunha et al (2017): Price Effects

	All PAL goods	Basic PAL goods only	All PAL goods	Basic PAL goods only	All PAL goods	Basic PAL goods only
Outcome =	price	price	price	price	Δprice	Δprice
	(1)	(2)	(3)	(4)	(5)	(6)
In-kind	-0.037*	-0.033	-0.036*	-0.033	-0.062**	-0.025
	(0.020)	(0.020)	(0.020)	(0.020)	(0.029)	(0.024)
Cash	0.002	0.014	0.003	0.012	0.000	0.039
	(0.023)	(0.027)	(0.023)	(0.026)	(0.031)	(0.029)
Lagged normalized unit value	0.027	0.127***				
	(0.021)	(0.042)				
Observations	2,335	1,617	2,335	1,617	2,335	1,617
***Effect size:* In-kind - Cash**	**-0.039****	**-0.047****	**-0.038****	**-0.045****	**-0.063****	**-0.064****
H_0: In-kind = Cash (p-value)	**0.02**	**0.04**	**0.03**	**0.04**	**0.02**	**0.02**

Cunha et al (2017): Persistence

	All PAL goods		Basic PAL goods only	
Outcome =	price	price	price	price
	(1)	(2)	(3)	(4)
In-kind	-0.031	-0.056**	-0.038	-0.056
	(0.022)	(0.026)	(0.031)	(0.035)
In-kind x Above median length of treatment	-0.021	-0.011	-0.022	-0.018
	(0.034)	(0.035)	(0.040)	(0.043)
Above median length of treatment	0.004	-0.002	0.018	0.013
	(0.028)	(0.029)	(0.033)	(0.035)
In-kind x Development index		-0.047**		-0.037
		(0.022)		(0.023)
Development index		0.036**		0.039**
		(0.015)		(0.016)
Observations	1,818	1,665	1,258	1,150

Cunha et al (2017): Heterogeneity

	Below-median development	Above-median development	All villages	Villages with market power	Villages without market power	Below-median price correlation	Above-median price correlation
Outcome =	price	price	price	price	price	price	price
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
In-kind	-0.036	-0.033	-0.033	-0.048*	-0.005	-0.060**	-0.019
	(0.031)	(0.027)	(0.021)	(0.025)	(0.021)	(0.028)	(0.028)
Cash	0.015	-0.007	-0.007	0.007	-0.005	0.002	-0.014
	(0.032)	(0.037)	(0.032)	(0.029)	(0.026)	(0.032)	(0.032)
Development index below-median x In-kind			-0.006				
			(0.018)				
Development index below-median x Cash			0.018				
			(0.031)				
Market power village x In-kind							
Market power village x Cash							
Price correlation below-median x In-kind							
Price correlation below-median x Cash							
Observations	1,094	1,210	2,304	1,733	602	1,115	1,220
***Effect size:* In-kind - Cash**	**-0.051****	**-0.027**	**-0.026**	**-0.055*****	**0.000**	**-0.063*****	**-0.006**
H_0: In-kind = Cash (p-value)	**0.02**	**0.37**	**0.37**	**0.01**	**1.00**	**0.01**	**0.81**

Cunha et al (2017): Effects on Producers

Outcome =	Farm profits	Farm costs	ln(Expenditure per capita)	ln(Expenditure per capita)	Asset index	Asset index
	(1)	(2)	(3)	(4)	(5)	(6)
In-kind	143.87	134.01	0.115**		0.084	
	(89.839)	(119.511)	(0.046)		(0.075)	
Cash	186.16*	345.32**	0.064		-0.040	
	(106.082)	(140.378)	(0.052)		(0.106)	
Producer x In-Kind			0.001	-0.018	0.077	0.055
			(0.060)	(0.046)	(0.115)	(0.088)
Producer x Cash			0.087	0.015	0.266*	0.229**
			(0.068)	(0.051)	(0.142)	(0.109)
Producer			-0.161***	-0.003	-0.308***	-0.007
			(0.050)	(0.036)	(0.092)	(0.071)
Control for pre-period outcome?	yes	yes	yes	yes	yes	yes
Village FE				yes		yes
Observations	4,924	5,038	5,534	5,534	5,571	5,571
***Effect size:* In-kind - Cash**	**-42.29**	**-211.31***	**0.050**		**0.124**	
***H_0:* In-kind = Cash (p-value)**	**0.67**	**0.08**	**0.25**		**0.20**	
***Effect size:* Producer x In-Kind - Producer x Cash**			**-0.086**	**-0.033**	**-0.189**	**-0.174***
***H_0:* Producer x In-Kind = Producer x Cash (p-value)**			**0.13**	**0.47**	**0.13**	**0.07**

Outline

Open Questions

- How do transfers interact with labor market formality? Should transfers be made conditional on having (previously held) formal work?
- How can governments reduce the administrative burden of enforcing conditionality/eligibility? What role does technology play in this?
- What types of transfer programs face the fewest political economy challenges? In what circumstances will governments adopt what types of policies?
- Who should administer transfer programs? What level of government? How much to rely on self-reporting?
- How much do eligibility requirements reduce incentives to graduate out of eligibility? To stay in rural areas?