Social Preferences: Theory

Stefanie Stantcheva

Fall 2022

GOALS OF THIS LECTURE

(1) Theory of social preferences: a new, tractable way to capture fairness and justice principles. Applicable to way more than taxation (e.g.: IO problems, trade problems, macro problems).

(2) Empirical evidence on social preferences.

(3) Methodological tool: Online experiments.

Theory

This paper: "Generalized Social Welfare Weights for Optimal Tax Theory" Saez and Stantcheva (2016).

Standard Welfarist Approach: Critiques and Puzzles

- Maximize concave function or weighted sum of individiual utilities.

$$\max_{T(.)} SWF = \max_{T(.)} \int_i \omega_i \cdot u_i$$

- Special case: utilitarianism, $\omega_i = 1$.
- Cannot capture elements important in tax practice:
  - Source of income: earned versus luck.
  - Counterfactuals: what individuals *would* have done absent tax system.
  - Horizontal Equity concerns that go against "tagging."
- Utilitarianism critique: 100% redistribution optimal with concave $u(.)$ and no behavioral responses
- Methodological and conceptual critique: Policy makers use reform-approach rather than posit and maximize objective.

A Novel Approach to Model Social Preferences

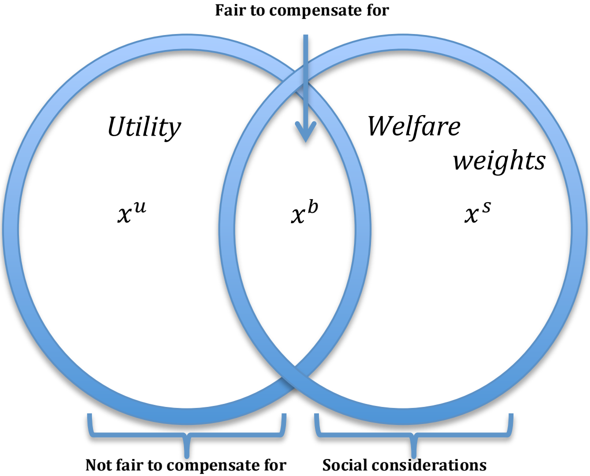
- Tax reform approach: weighs gains and losses from tax changes.

  Change in welfare: $-\int_i g_i \cdot \delta T(z_i)$ with $g_i \equiv g(c_i, z_i; x_i^s, x_i^b)$.

- Replace restrictive social welfare weight by generalized social marginal welfare weights.
  - g_i measures social value of $\$1$ transfer for person i.
  - Specified to directly capture fairness criteria.
  - Not necessarily derived from SWF

Generalized social welfare weights approach

$$u_i = u(c_i - v(z_i; x_i^u, x_i^b)) \qquad g_i = g(c_i, z_i; x_i^s, x_i^b)$$



General Model

- Mass 1 of individuals indexed by i.
- Utility from consumption c_i and income z_i (no income effects):

$$u_i = u(c_i - v(z_i; x_i^u, x_i^b))$$

where x_i^u and x_i^b are vectors of characteristics

- $u(.)$ increasing, v decreasing in z_i.
- Typical income tax: $T(z)$, hence $c_i = z_i - T(z_i)$.
  - More general tax systems, with conditioning variables possible, depending on what is observable and politically feasible.

Generalized social welfare weights approach

Definition

The generalized social marginal welfare weight on individual i is:

$$g_i = g(c_i, z_i; x_i^s, x_i^b)$$

g is a function, x_i^s is a vector of characteristics which only affect the social welfare weight, while x_i^b is a vector of characteristics which also affect utility.

- Recall utility is: $u_i = u(c_i - v(z_i; x_i^u, x_i^b))$
- Characteristics x^s, x^u, x^b may be unobservable to the government.
  - x^b: fair to redistribute, enters utility – e.g. ability to earn
  - x^s: fair to redistribute, not in utility – e.g. family background
  - x^u: unfair to redistribute, enters utility – e.g. taste for work

Aggregating Standard Weights at Each Income Level

Taxes depend on z only: express everything in terms of observable z.
$H(z)$: CDF of earnings, $h(z)$: PDF of earnings [both depend on $T(.)$]

Definition

$\bar{G}(z)$ is the (relative) average social marginal welfare weight for individuals earning at least z:

$$\bar{G}(z) \equiv \frac{\int_{\{i:z_i \geq z\}} g_i}{Prob(z_i \geq z) \cdot \int_i g_i}$$

$\bar{g}(z)$ is the average social marginal welfare weight at z defined so that

$$\int_z^\infty \bar{g}(z') dH(z') = \bar{G}(z)[1 - H(z)]$$

Nonlinear Tax Formula Expressed with Welfare Weights

Proposition

The optimal marginal tax at z:

$$T'(z) = \frac{1 - \bar{G}(z)}{1 - \bar{G}(z) + \alpha(z) \cdot e(z)}$$

$e(z)$: average elasticity of z_i w.r.t $1 - T'$ at $z_i = z$
$\alpha(z)$: local Pareto parameter $zh(z)/[1 - H(z)]$.

Proof follows the same "small reform" approach of Saez (2001): increase T' in a small band $[z, z + dz]$ and work out effect on budget and weighted welfare

Proof

- Reform $\delta T(z)$ increases marginal tax by $\delta\tau$ in small band $[z, z + dz]$.
- Mechanical revenue effect: extra taxes $dz\delta\tau$ from each taxpayer above z: $dz\delta\tau[1 - H(z)]$ is collected.
- Behavioral response: those in $[z, dz]$, reduce income by $\delta z = -ez\delta\tau/(1 - T'(z))$ where e is the elasticity of earnings z w.r.t $1 - T'$. Total tax loss $-dz\delta\tau \cdot h(z)e(z)zT'(z)/(1 - T'(z))$ with $e(z)$ the average elasticity in the small band.
- Net revenue collected by the reform and rebated lump sum is: $dR = dz\delta\tau \cdot \left[1 - H(z) - h(z) \cdot e(z) \cdot z \cdot \frac{T'(z)}{1-T'(z)}\right]$.
- Welfare effect of reform: $-\int_i g_i \delta T(z_i)$ with $\delta T(z_i) = -dR$ for $z_i \leq z$ and $\delta T(z_i) = \delta\tau dz - dR$ for $z_i > z$. Net effect on welfare is $dR \cdot \int_i g_i - \delta\tau dz \int_{\{i:z_i \geq z\}} g_i$.
- Setting net welfare effect to zero, using $(1 - H(z))\bar{G}(z) = \int_{\{i:z_i \geq z\}} g_i / \int_i g_i$ and $\alpha(z) = zh(z)/(1 - H(z))$, we obtain the tax formula.

Linear Tax Formula Expressed with Welfare Weights

The optimal linear tax rate, such that $c_i = z_i \cdot (1 - \tau) + \tau \cdot \int_i z_i$ can also be expressed as a function of an income weighted average marginal welfare weight (Piketty and Saez, 2013).

Proposition

The optimal linear income tax is:

$$\tau = \frac{1 - \bar{g}}{1 - \bar{g} + e} \quad \text{with} \quad \bar{g} \equiv \frac{\int_i g_i \cdot z_i}{\int_i g_i \cdot \int_i z_i}$$

e: elasticity of $\int_i z_i$ w.r.t $(1 - \tau)$.

Applying Standard Formulas with Generalized Weights

- Individual weights need to be "aggregated" up to characteristics that tax system can conditioned on.
  - E.g.: If $T(z, x^b)$ possible, aggregate weights at each $(z, x^b) \to \bar{g}(z, x^b)$.
  - If standard $T(z)$, aggregate at each z: $\bar{G}(z)$ and $\bar{g}(z)$.
- Then apply standard formulas. Nests standard approach.
- If $g_i \geq 0$ for all i, (local) Pareto efficiency guaranteed.
- Can we back out weights? Optimum $\Leftrightarrow$ max $SWF = \int_i \omega_i \cdot u_i$ with *Pareto weights* $\omega_i = g_i / u_{ci} \geq 0$ where g_i and u_{ci} are evaluated at the optimum allocation
  - Impossible to posit correct weights ω_i without *first* solving for optimum

1. Optimal Tax Theory with Fixed Incomes

Modelling fixed incomes in our general model.

- Focus on redistributive issues.
- $z = z_i$ is fixed for each individual (fully inelastic labor supply).
- Concave uniform utility $u_i = u(c_i)$

Standard utilitarian approach.

- Optimum: $c = z - T(z)$ is constant across z, full redistribution.
- Is it acceptable to confiscate incomes fully?
- Very sensitive to utility specification
- Heterogeneity in consumption utility? $u_i = u(x_i^c \cdot c_i)$

1. Tax Theory with Fixed Incomes: Generalized Weights

Definition

Let $g_i = g(c_i, z_i) = \tilde{g}(c_i, z_i - c_i)$ with $\tilde{g}_c \leq 0$, $\tilde{g}_{z-c} \geq 0$.

i) Utilitarian weights: $g_i = g(c_i, z_i) = \tilde{g}(c_i)$ for all z_i, with $\tilde{g}(\cdot)$ decreasing.

ii) Libertarian weights: $g_i = g(c_i, z_i) = \tilde{g}(z_i - c_i)$ with $\tilde{g}(\cdot)$ increasing.

- Weights depend negatively on c – "ability to pay" notion.
- Depend positively on tax paid – taxpayers contribute socially more.
- Optimal tax system: weights need to be equalized across all incomes z:

$$\tilde{g}(z - T(z), T(z)) \text{ constant with } z$$

1. Tax Theory with Fixed Incomes: Optimum

Proposition

The optimal tax schedule with no behavioral responses is:

$$T'(z) = \frac{1}{1 - \tilde{g}_{z-c} / \tilde{g}_c} \quad \text{and} \quad 0 \leq T'(z) \leq 1. \tag{1}$$

Corollary

Standard utilitarian case, $T'(z) \equiv 1$. Libertarian case, $T'(z) \equiv 0$.

- Empirical survey shows respondents indeed put weight on both disposable income and taxes paid.
- Between the two polar cases, $g(c,z) = \tilde{g}(c - \alpha(z-c)) = \tilde{g}(z - (1+\alpha)T(z))$ with $\tilde{g}$ decreasing.
- Can be empirically calibrated and implied optimal tax derived.

1. Libertarianism and Rawlsianism

Libertarianism:

- Principle: "Individual fully entitled to his pre-tax income."
- Morally defensible if no difference in productivity, but different preferences for work.
- $g_i = g(c_i, z_i) = \tilde{g}(c_i - z_i)$, increasing ($x_i^s$ and x_i^b empty).
- Optimal formula yields: $T'(z_i) \equiv 0$.

Rawlsianism:

- Principle: "Care only about the most disadvantaged."
- $g_i = g(u_i - \min_j u_j) = 1(u_i - \min_j u_j = 0)$, with $x_i^s = u_i - \min_j u_j$ and x^b is empty.
- If least advantaged people have zero earnings independently of taxes, $\bar{G}(z) = 0$ for all $z > 0$.
- Optimal formula yields: $T'(z) = 1/[1 + \alpha(z) \cdot e(z)]$ (maximize demogrant $-T(0)$).

3. Transfers and Free Loaders: Setting

- Behavioral responses closely tied to social weights: biggest complaint against redistribution is "free loaders."
- Generalized welfare weights can capture "counterfactuals."
- Consider linear tax model where τ funds demogrant transfer.
- $u_i = u(c_i - v(z_i; \theta_i)) = u(c_{z_i} - \theta_i \cdot z_i)$ with $z_i \in \{0, 1\}$.
- Individuals can choose to not work, $z = 0$, $c_i = c_0$.
- If they work, earn $z = \$1$, consume $c_1 = (1 - \tau) + c_0$.
- Cost of work θ, with cdf $P(\theta)$, is private information.
- Individual: work iff $\theta \leq c_1 - c_0 = (1 - \tau)$.
- Fraction working: $P(1 - \tau)$.
- e: elasticity of aggregate earnings $P(1 - \tau)$ w.r.t $(1 - \tau)$.

3. Transfers and Free Loaders: Optimal Taxation

Apply linear tax formula:

- $\tau = (1 - \bar{g})/(1 - \bar{g} + e)$
- In this model, $\bar{g} = \int_i g_i z_i / (\int_i g_i \cdot \int_i z_i) = \bar{g}_1 / [P \cdot \bar{g}_1 + (1 - P) \cdot \bar{g}_0]$ with: $\bar{g}_1$ the average g_i on workers, and $\bar{g}_0$ the average g_i on non-workers.

Standard Approach:

- $g_i = u'(c_0)$ **for all non-workers** so that $\bar{g}_0 = u'(c_0)$.
- Hence, approach does not allow to distinguish between the deserving poor and free loaders.
- We can only look at actual situation: work or not, not "why" one does not work.
- Contrasts with public debate and historical evolution of welfare

3. Transfers and Free Loaders: Generalized Welfare Weights

- Distinguish people according to what would have done absent transfer.
- **Workers**: Fraction $P(1-\tau)$. Set $g_i = u'(c_1 - \theta_i)$.
- **Deserving poor**: would not work even absent any transfer: $\theta > 1$. Fraction $1 - P(1)$. Set $g_i = u'(c_0)$.
- **Free Loaders**: do not work because of transfer: $1 \geq \theta > (1-\tau)$. Fraction $P(1) - P(1-\tau)$. Set $g_i = 0$.
- Cost of work enters weights – fair to compensate for (i.e., not laziness).
- Average weight on non-workers $\bar{g}_0 = u'(c_0) \cdot (1 - P(1))/(1 - P(1-\tau)) < u'(c_0)$ lower than in utilitarian case.
- Reduces optimal tax rate not just through e but also through $\bar{g}_0$.

3. Transfers and Free Loaders: Remarks and Applications

- Ex post, possible to find suitable Pareto weights $\omega(\theta)$ that rationalize same tax.
  - $\omega(\theta) = 1$ for $\theta \leq (1 - \tau^*)$ (workers)
  - $\omega(\theta) = 1$ for $\theta \geq 1$ (deserving poor)
  - $\omega(\theta) = 0$ for $(1 - \tau^*) < \theta < 1$ (free loaders).
- But: these weights depend on optimum tax rate τ^*.
- Other applications:
  - **Desirability of in-work benefits** if weight on non-workers becomes low enough relative to workers.
  - **Transfers over the business cycle**: composition of those out of work depends on ease of finding job.

2. Equality of Opportunity: Setting

- Standard utility $u(c - v(z/w_i))$ with w_i ability to earn
- w_i is result of i) family background $B_i \in \{0, 1\}$ (which individuals not responsible for) and ii) merit (which individuals are responsible for) = rank r_i conditional on background.
- Advantaged background gives earning ability w advantage: $w(r_i|B_i = 1) > w(r_i|B_i = 0)$
- Society is willing to redistribute across backgrounds, but not across incomes conditional on background.
- $\Rightarrow$ Conditional on earnings, those coming from $B_i = 0$ are more meritorious [because they rank higher in merit]
- $\bar{c}(r) \equiv (\int_{(i:r_i=r)} c_i)/Prob(i : r_i = r)$: average consumption at rank r.
- $g_i = g(c_i; \bar{c}(r_i)) = 1(c_i \leq \bar{c}(r_i))$

2. Equality of Opportunity: Results

- Suppose government cannot condition taxes on background.
- $\bar{G}(z)$: **Representation index**: % from disadvantaged background earning $\geq z$ relative to % from disadvantaged background in population.
- Implied Social Welfare function as in Roemer et al. (2003).
- $\bar{G}(z)$ decreasing since harder for those from disadvantaged background to reach upper incomes.
- If at top incomes, representation is zero, revenue maximizing top tax rate.
- Justification for social welfare weights decreasing with income not due to decreasing marginal utility (utilitarianism).

2. Equality of Opportunity vs. Utilitarian Tax Rates

	Equality of Opportunity			Utilitarian (log-utility)	
	Fraction from low background (=parents below median) above each percentile	Implied social welfare weight G(z) above each percentile	Implied optimal marginal tax rate at each percentile	Utilitarian social welfare weight G(z) above each percentile	Utilitarian optimal marginal tax rate at each percentile
	(1)	(2)	(3)	(4)	(5)
Income percentile					
z= 25th percentile	44.3%	0.886	53%	0.793	67%
z= 50th percentile	37.3%	0.746	45%	0.574	58%
z= 75th percentile	30.3%	0.606	40%	0.385	51%
z= 90th percentile	23.6%	0.472	34%	0.255	42%
z= 99th percentile	17.0%	0.340	46%	0.077	54%
z= 99.9th percentile	16.5%	0.330	47%	0.016	56%

Chetty *et al.* (2013) intergenerational mobility data for the U.S.

Above 99^{th} percentile, stable representation, hence stable tax rates.

Optimal tax rate lower than in utilitarian case.

3. Poverty Alleviation: Setting

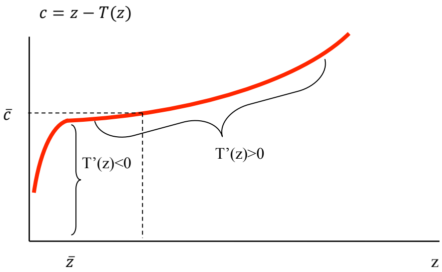
- Poverty gets substantial attention in public debate.
- Poverty alleviation objectives can lead to Pareto dominated outcomes:
  - Besley and Coate (1992) and Kanbur, Keen, and Tuomala (1994).
  - Intuition: disregard people's disutility from work.
- Generalized welfare weights can avoid pitfall of Pareto inefficiency.
- $\bar{c}$: poverty threshold. "Poor": $c < \bar{c}$.
- $u_i = u(c_i - v(z_i / w_i))$.
- $\bar{z}$: (endogenous) pre-tax poverty threshold: $\bar{c} = \bar{z} - T(\bar{z})$.
- Poverty gap alleviation: care about shortfall in consumption.
- $g_i = g(c_i, z_i; \bar{c}) = 1 > 0$ if $c_i < \bar{c}$ and $g_i = g(c_i, z_i; \bar{c}) = 0$ if $c_i \geq \bar{c}$.
- $\Rightarrow \bar{g}(z) = 0$ for $z \geq \bar{z}$ and $\bar{g}(z) = 1/H(\bar{z})$ for $z < \bar{z}$.
- $\Rightarrow \bar{G}(z) = 0$ for $z \geq \bar{z}$ and $1 - \bar{G}(z) = \frac{1/H(\bar{z})-1}{1/H(z)-1}$ for $z < \bar{z}$.

3. Optimal Tax Schedule that Minimizes Poverty Gap

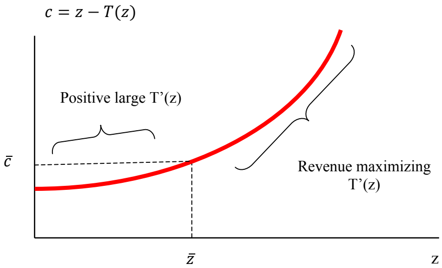
Proposition

$$T'(z) = \frac{1}{1+\alpha(z)\cdot e(z)} \quad \text{if} \quad z > \bar{z}$$

$$T'(z) = \frac{1}{1+\alpha(z)\cdot e(z)\cdot \frac{1/H(z)-1}{1/H(\bar{z})-1}} \quad \text{if} \quad z \leq \bar{z}$$



(a) Direct poverty gap minimization



(b) Generalized weights approach

Evidence against utilitarianism

- Respondents asked to compare families w/ different combinations of z, $z - T(z)$, $T(z)$.
- Who is most deserving of a $1000 tax break?
- **Both disposable income and taxes paid matter** for deservedness
  - Family earning $40K, paying $10K in taxes judged more deserving than family earning $50K, paying $10K in taxes
  - Family earning $50K, paying $15K in taxes judged more deserving than family earning $40K, paying $5K in taxes
- **Frugal vs. Consumption-loving** person with same net income

Consumption-lover more deserving	Frugal more deserving	Taste for consumption irrelevant
4%	22%	74%

Which of the following two individuals do you think is most deserving of a $1,000 tax break?

Individual A earns $50,000 per year, pays $10,000 in taxes and hence nets out $40,000. She greatly enjoys spending money, going out to expensive restaurants, or traveling to fancy destinations. She always feels that she has too little money to spend.

Individual B earns the same amount, $50,000 per year, also pays $10,000 in taxes and hence also nets out $40,000. However, she is a very frugal person who feels that her current income is sufficient to satisfy her needs.

- ○ Individual A is most deserving of the $1,000 tax break
- ○ Individual B is most deserving of the $1,000 tax break
- ○ Both individuals are exactly equally deserving of the tax $1,000 break

>>

Source: survey in Saez and Stantcheva (2013)

Does society care about effort to earn income?

- **Hard-working vs. Easy-going person with same net income**
- "*A earns \$30,000 per year, by working in two different jobs, 60 hours per week at \$10/hour. She pays \$6,000 in taxes and nets out \$24,000. She is very hard-working but she does not have high-paying jobs so that her wage is low.*"
- "*B also earns the same amount, \$30,000 per year, by working part-time for 20 hours per week at \$30/hour. She also pays \$6,000 in taxes and hence nets out \$24,000. She has a good wage rate per hour, but she prefers working less and earning less to enjoy other, non-work activities.*"

Hardworking more deserving	Easy-going more deserving	Hours of work irrelevant conditional on total earnings
43%	3%	54%

Do people care about "Free Loaders" and Behavioral Responses to Taxation?

Starting from same benefit level, which person most deserving of more benefits?

	Disabled unable to work	Unemployed looking for work	Unemployed not looking for work	On welfare not looking for work
Average rank (1-4)	1.4	1.6	3.0	3.5
% assigned 1st rank	57.5%	37.3%	2.7%	2.5%
% assigned last rank	2.3%	2.9%	25%	70.8%

Calibrating Social Welfare Weights

- Calibrate $\tilde{g}(c, T) = \tilde{g}(c - \alpha T)$
- 35 fictitious families, w/ different net incomes and taxes
- Respondents rank them pair-wise (5 random pairs each)

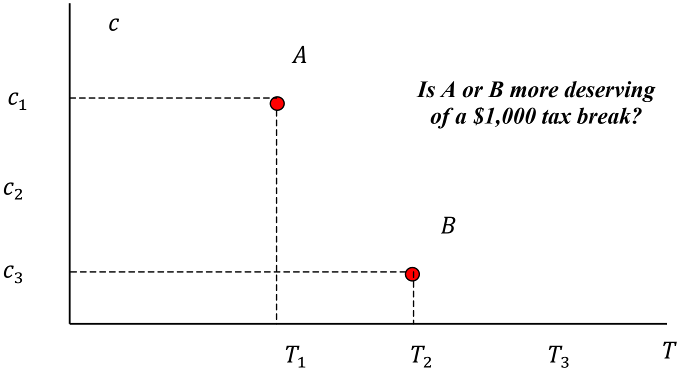
Which of these two families is most deserving of the $1,000 tax break?

- Family earns $100,000 per year, pays $50,000 in taxes, and hence nets out $50,000
- Family earns $25,000 per year, pays $1,250 in taxes, and hence nets out $23,750

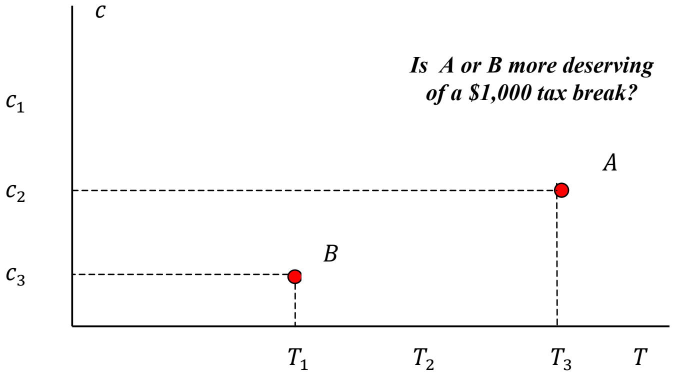
Which of these two families is most deserving of the $1,000 tax break?

- Family earns $50,000 per year, pays $2,500 in taxes, and hence nets out $47,500
- Family earns $500,000 per year, pays $170,000 in taxes, and hence nets out $330,000

Eliciting Social Preferences



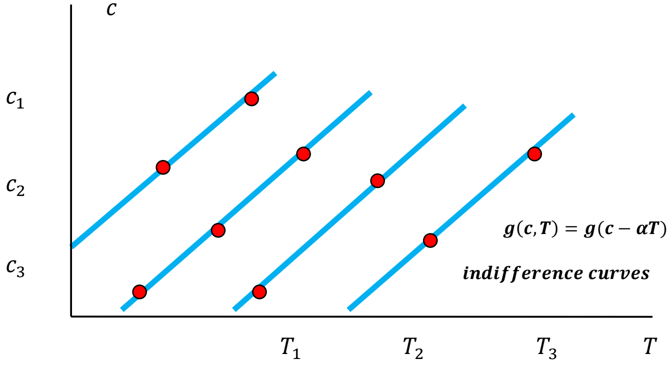
Eliciting Social Preferences



Eliciting Social Preferences

$S_{ijt} = 1$ if i ranked 1st in display t for respondent j, δT_{ijt} is difference in taxes, δc_{ijt} difference in net income for families in pair shown.

$$S_{ijt} = \beta_0 + \beta_T \delta T_{ijt} + \beta_c \delta c_{ijt} \qquad \alpha = \frac{\delta c}{\delta T}|_S = -\frac{\beta_T}{\beta_c} = -slope$$



Eliciting Social Preferences

	Probability of being deemed more deserving in pairwise comparison					
Sample	Full	Excludes cases with income of $1m	Excludes cases with income of $500K+	Excludes cases with income $500K+ and $10K or less	Liberal subjects only	Conservative subjects only
	(1)	(2)	(3)	(4)	(5)	(6)
d(Tax)	0.0017***	0.0052***	0.016***	0.015***	0.00082***	0.0032***
	(0.0003)	(0.0019)	(0.0019)	(0.0022)	(0.00046)	(0.00068)
d(Net Income)	-0.0046***	-0.0091***	-0.024***	-0.024***	-0.0048***	-0.0042***
	(0.00012)	(0.00028)	(0.00078)	(0.00094)	(0.00018)	(0.00027)
Number of observations	11,450	8,368	5,816	3,702	5,250	2,540
Implied α	0.37	0.58	0.65	0.64	0.17	0.77
	(0.06)	(0.06)	(0.07)	(0.09)	(0.12)	(0.16)
Implied marginal tax rate	73%	63%	61%	61%	85%	57%