Lois et Normes:

Incitations, Moralité et Sanctions Sociales

Roland Bénabou

Princeton University

Cours au Collège de France - Nov.-Déc. 2017

(Basé en partie sur des travaux joints avec Jean Tirole)

INTRODUCTION

- People's behavior is shaped by their preferences, by formal incentives (the law, contracts) and by social norms, informal enforcement based on reputation, honor / stigma
- These different channels aspects usually studied separately
  - Economists emphasize incentives. Study norms, but $\neq$ literature
  - Psychologists, sociologists, often skeptical of incentives:
    Said to "crowd out intrinsic motivation," "undermine social norms and values". What does it mean? (When) does it happen?
  - They emphasize instead persuasion, "norms-based interventions"
- Law scholars somewhere in-between: law is a set of incentives, but also reflects, conveys and adapts to the values of society
- Laws, norms interact, shape each other: need to model together
  - When do incentives undermine or strengthen social norms?
  - Optimal setting of taxes, subsidies, laws, in the presence of norms?

Outline

1. General model combining formal & social incentives
   - Understanding their interactions, crowding out or in
   - Empirical evidence: lab and field
2. Honor, stigma and social norms
   - Theoretical predictions on the effectiveness of incentives
   - Empirical evidence: field
3. Optimal incentives with social norms
   - Implications
4. The expressive content of law
   - Empirical evidence: lab

▶ Go to model

I - GENERAL MODEL

1. Actions

- People choose their participation level, *a* (for "altruism'') in some prosocial activity. May be discrete (0/1: giving blood, voting) or continuous (volunteering, recycling)
- If contribute *a* ⇒ incur cost, *C*(*a*): effort, time, resources.
- Incentive rate: you get $ *y* per unit of *a*. Reward, subsidy, tax, etc (← policy, law). Variant: monetary donation *a* ⇒ receive "perks" *y* per $

2. Motivations / preferences

- First part: direct costs and benefits from engaging in pro (or anti) social activity

(intrinsic value + value of money × reward rate) × participation level – cost incurred

$$(v_a + v_y\, y)\; a \;-\; C(a)$$

v_a: valuation of extra public good which you provide + "joy of giving" ⇒ *prosocial orientation*

v_y: valuation of money or private consumption ⇒ *"greed" (or "need")*

- Individual's true "values" are not directly observable by others (sometimes not even accessible to himself). Private information, must be inferred from actions: attribution

The Red Cross
on contributing, volunteering:

“You will be surprised at how good it makes you feel and what a terrific response you will get from loved ones”.

“*Helping others feels good and makes you feel good about yourself”.*

• Second part: social esteem / self-image (reputational concerns)

- Desirable (pleasant, useful) to be perceived as generous, reciprocal, public minded,…
  … and undesirable to be perceived as greedy, interested in money, or as poor.

- What goes for social perceptions goes for self-perception. Judging oneself by actions.

⇒ to people's basic motivations, we add:

+ concern for appearing prosocial × Perceived prosociality, in light of behavior

– concern about appearing greedy × Perceived greed, in light of behavior

or: $r = m_a E(v_a \mid a,y) - m_y E(v_y \mid a,y)$

• *E* is for "expectation": what one can expect your true degree of intrinsic social orientation (or greed) to be, given that you did action *a* for reward *y.*

• *m* is for "image": how much you care about image / self image concerning altruism (or greed, wealth). Depends in particular on how public or salient behavior is.

• People generally differ in their social orientation, i.e., preferences over (v_a , v_y), as well as in their image concerns *(m_a, m_y).*

Summarizing

Three motives for prosocial behavior:

intrinsic + extrinsic + (self) reputational

- Individual will choose his contribution *a* to maximize:

(Intrinsic value + Value of money × Reward rate) × *contribution* – Cost incurred

+ concern for appearing prosocial × (Perceived prosociality | *contribution,* reward)

– concern about appearing greedy × (Perceived greed | *contribution,* reward)

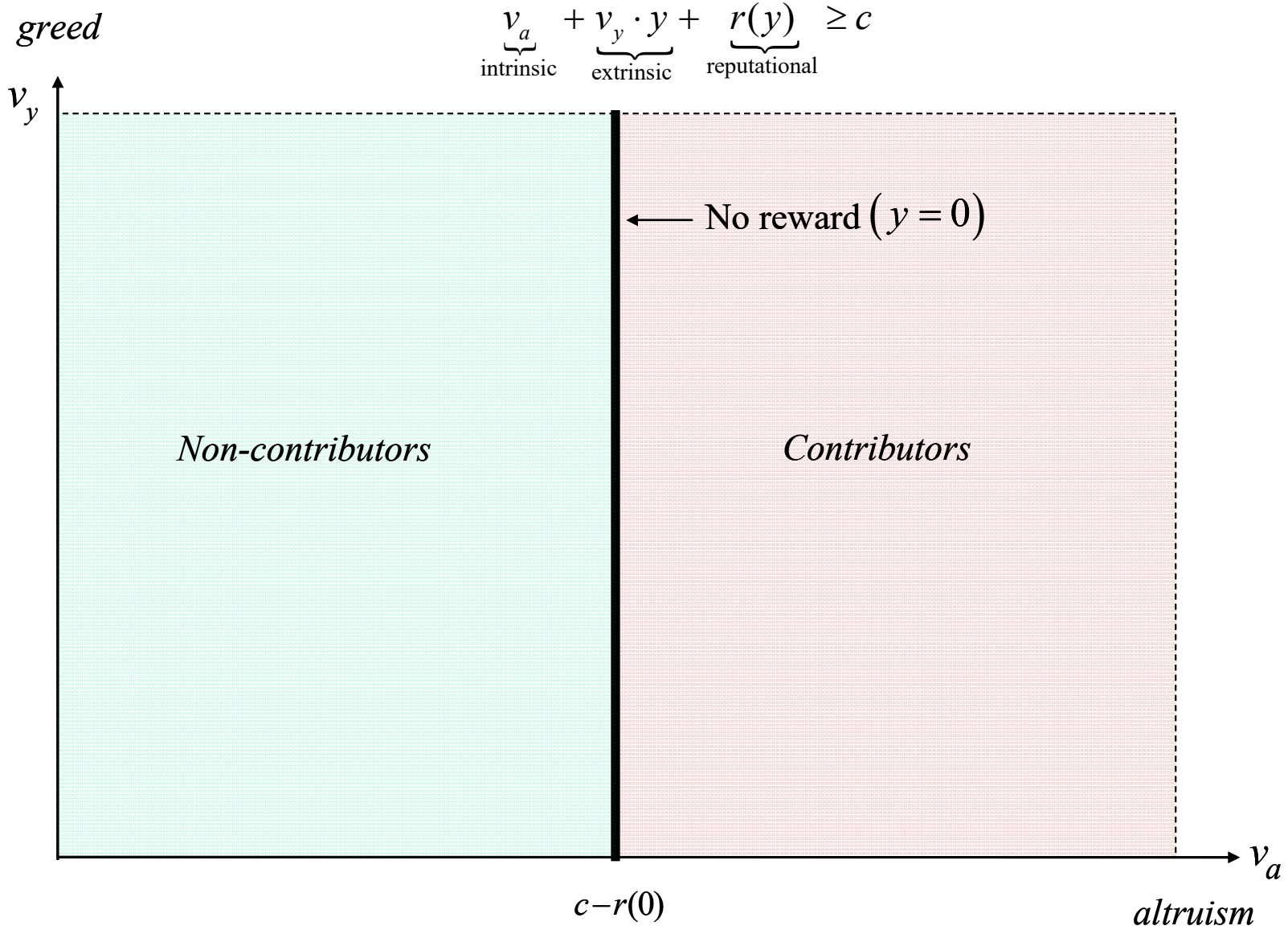
taking into account how his behavior will be *interpreted,* given the context.

$$U = (v_a + v_y y)a - C(a) + m_a E(v_a|a,y) - m_y E(v_y|a,y) + e\bar{a},$$

- $\bar{a}$*:* total supply of public good or externality, resulting from everyone's actions
- Policy parameters: *material rewards* = *y; publicity* = amplifying *m; communication* about what others are doing, or think one should do

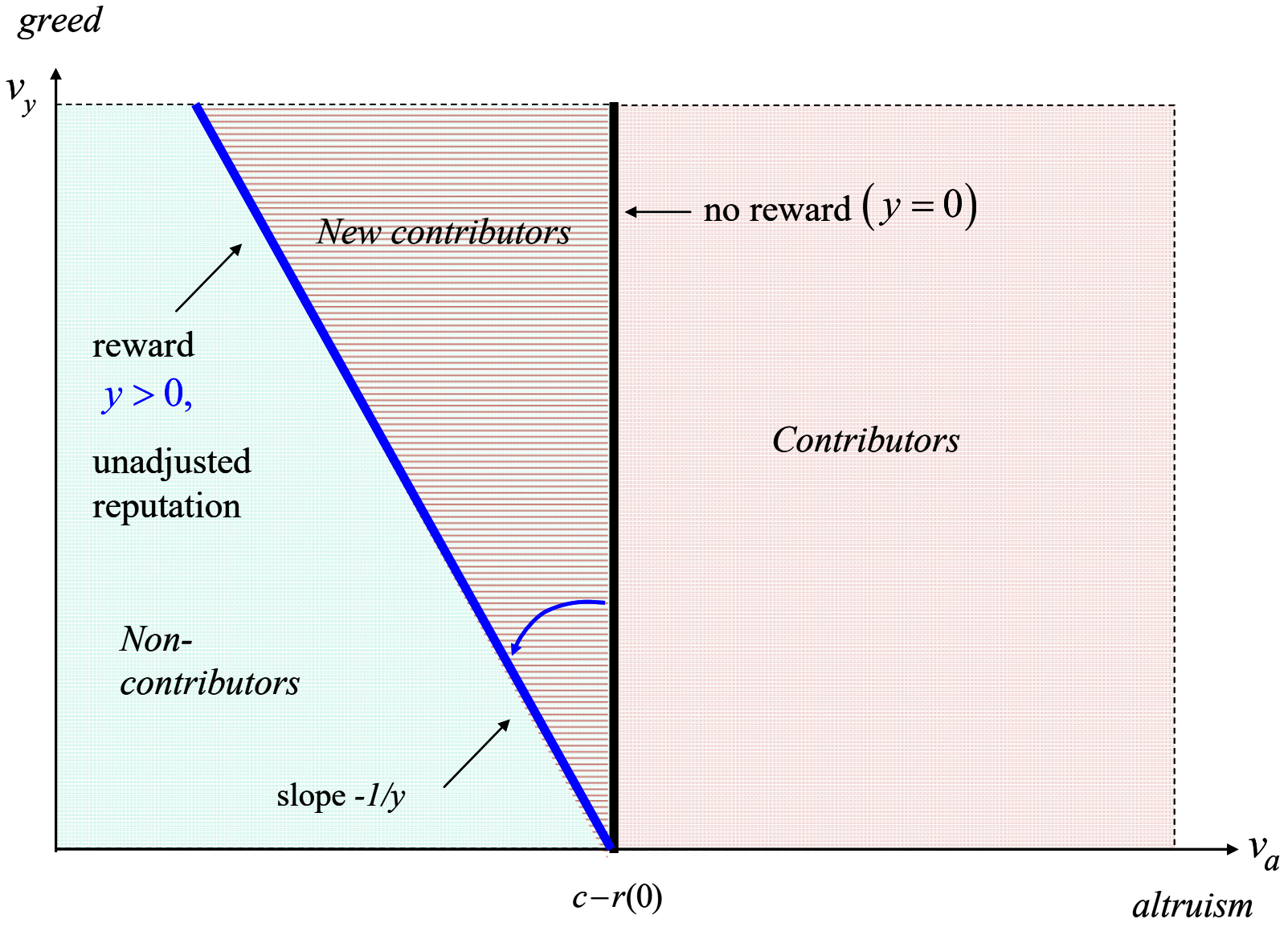
THE IMAGE-SPOILING EFFECT OF REWARDS: BASIC INTUITIONS

- Simple decision: do it (cost = c) or not (cost 0) => contribute if:



❑ Introduce reward $y > 0$. First step: if reputation remains *r(0)*

$$\underbrace{v_a}_{\text{intrinsic}} + \underbrace{v_y y}_{\text{extrinsic}} + \underbrace{r(0)}_{\text{reputational}} \geq c$$

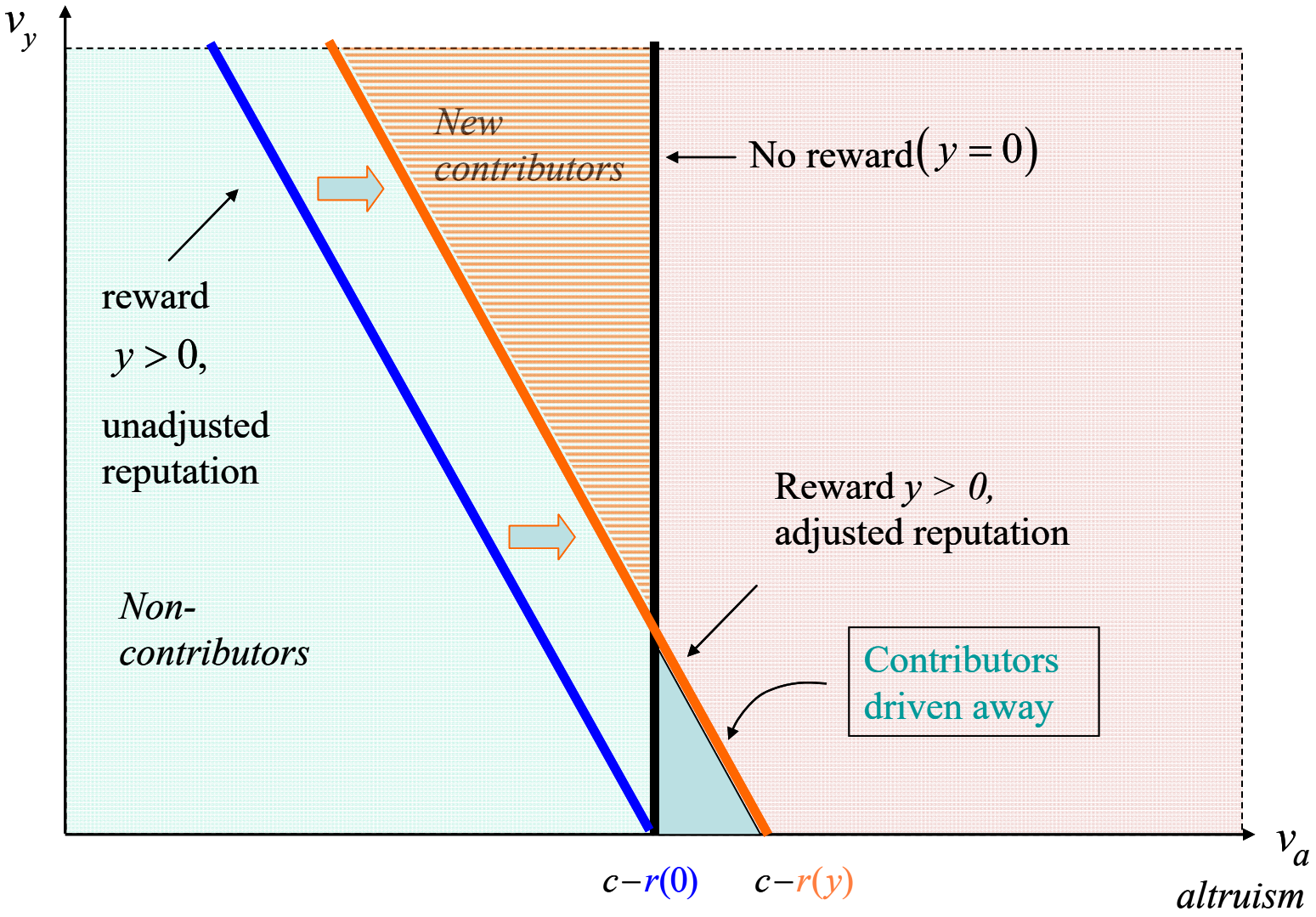


❑ But: compared to original contributors, new ones are:

(i) *less prosocial*; non-contributors also worse, however

(ii) *more greedy*

⇒ reputational value ("meaning of contribution") changes:



Formally: optimally contributes up to the point where:

Marginal Cost = Marginal Benefit

= intrinsic value + extrinsic reward + (self) reputational gain

$$C'(a) = v_a + v_y y + r(a;y,m)$$

"Social meaning" of the act: attribution or signal-extraction problem

- In trying to infer intrinsic motivation v_a, from observed behavior, extrinsic part $v_y y$ acts as a source of noise (uncertainty), which gets "louder", the higher is the incentive rate y.

  ⇒ trying to foster prosocial behavior by increasing y will tend to spoil /crowd out reputational motivation $r(a;y,m)$.

➢ *Classical "overjustification effect" of rewards*

- When people also differ in their image concerns m, the reputational incentive r($a,y;m$) is a further source of noise in inferring v_a or v_y. Wonder if a is done for appearances.

  ⇒ trying to foster prosocial behavior by making glory and shame more observable / public (scaling up the m's) is self-limiting: also has a negative feedback on $r(a;y,m)$.

➢ *"Overjustification effect" of publicity / praise and shame*

Material Incentives, Overjustification and Crowding Out

Optimal choice of contribution level $a \Rightarrow$

$$\underbrace{C'(a)}_{\text{marginal cost}} = \underbrace{v_a}_{\text{intrinsic}} + \underbrace{v_y y}_{\text{extrinsic}} + \underbrace{r(a, y; m_a, m_y)}_{\text{reputational}}$$

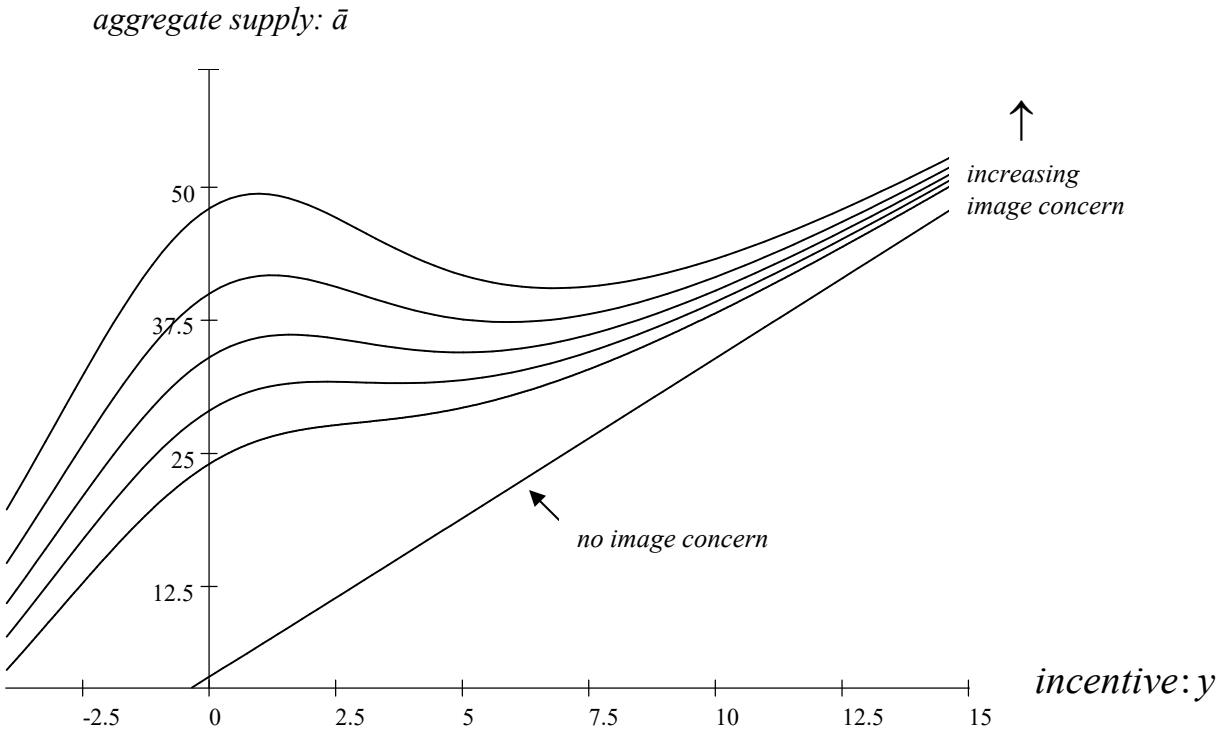


Material Incentives, Overjustification and Crowding Out

Optimal choice of contribution level $a \Rightarrow$

$$\underbrace{C'(a)}_{\text{marginal cost}} = \underbrace{v_a}_{\text{intrinsic}} + \underbrace{v_y y}_{\text{extrinsic}} + \underbrace{r(a, y; m_a, m_y)}_{\text{reputational}}$$

[Proposition](#) *(1) Greater concerns for prosocial reputation (m_a), such as greater visibility, increase contributions. (2) When reputational concern is above some critical value, there will be a range over which incentives are counterproductive: a higher reward reduces the total amount of prosocial behavior.*



Material rewards

Take identical reputational concerns (same (m_a, m_y)) for all agents

Proposition 1 *Equilibrium contributions are:*

$$a = \underbrace{\frac{v_a + v_y y}{k}}_{\text{direct incentive (intrinsic + extrinsic)}} + \underbrace{m_a \left(\frac{1}{1 + y^2 \sigma_y^2 / \sigma_a^2} \right) - m_y \left(\frac{y \sigma_y^2 / \sigma_a^2}{1 + y^2 \sigma_y^2 / \sigma_a^2} \right)}_{\text{reputational incentive: perceived altruism – perceived greed}}$$

- Reward y has usual direct effect, but also acts like an increase in signal-to-noise ratio, making contribution somewhat more likely to be driven by money than by altruism

- Multidimensional heterogeneity / signaling is key for the result.

- Aggregate supply: summing a across individuals $\Rightarrow \bar{a}(y)$.

Image Incentives: Effectiveness and Limits

- Policies based on publicity, prominence, memorability:

  Medals, titles, named buildings, public praise and shame, televised arrests, e-registry, pillory…

- Increases one of the three motives for prosocial behavior…

  … but also amplifies another source of "noise" in inferring people's true preferences (altruism, greed), from their actions: *couldn't they be doing it for the image?*

- When / how this can happen? Model shows:

- Such "doubt" is relevant when (self) image more important to some people than others: must be variability in image concerns (m_a, m_y).
- Making behavior more public / memorable = "scaling up" these image motives:

$$\underbrace{C'(a)}_{\text{marginal cost}} = \underbrace{v_a}_{\text{intrinsic}} + \underbrace{v_y y}_{\text{extrinsic}} + \underbrace{x \cdot r(a, x, y; m_a, m_y)}_{\text{reputational}}$$

- Direct impact: visibility / salience increases incentive to behave well
- Dampening effect: *observers increasingly ascribe behavior to image concerns.*
- New form of overjustification effect: weakens, but never reverses direct impact

Proposition (overjustification and crowding out)

When image concern m_a is above some threshold, there is a range where incentives are counterproductive: compliance $\bar{a}(y)$ is decreasing on $[y_1, y_2]$, and increasing elsewhere.

- Focussed here on the full-crowding-out case, where total supply $\searrow$ as incentive $\nearrow$: more paradoxical, has received more attention
- But, should not be overemphasized:
  - Main message is signaling effect of (receiving) rewards, which can offset much of direct effect $\Rightarrow$ weaker response
  - Can also get crowding-in, in appropriate cases; see later
- Testable implications:
  - People contribute more when observed by others $(d\bar{a}/dy > 0)$, but
  - This should attenuate, or even reverse, when they are (known to be) rewarded for doing it $(d^2\bar{a}/dydm < 0)$
  - Equivalently, effectiveness of incentives is smaller, or even reversed, when both action and reward are observed

Test: charitable donations

- Ariely, Bracha, Meier (2007) "Click for Charity"
- Task: sequentially pressing keys X and Z on the keyboard for up to 5 minutes.
  - Intentionally boring: will do only to earn money
- For every $X - Z$ pair, pay money in participant's name to an assigned charity: 1 cent for each of first 200 pairs, 0.5 cents for each of next 200 pairs, 0.25 cents for each of next 200 pairs,... 0.01 cents for each above 1,200.
- Design: $2 \times 2 \times 2$,with 161 subjects
  - "Good" or "Bad" Charity: American Red Cross, National Rifle Association
  - Incentives: either no payment to self, or same schedule as for charity. Implemented with random draw
  - Private vs. public condition: anonymous, vs. at the end, must tell other participants which charity was assigned to, $ earned for it and for oneself

Figure 1: Effect of Private Incentive for "Good" Charity

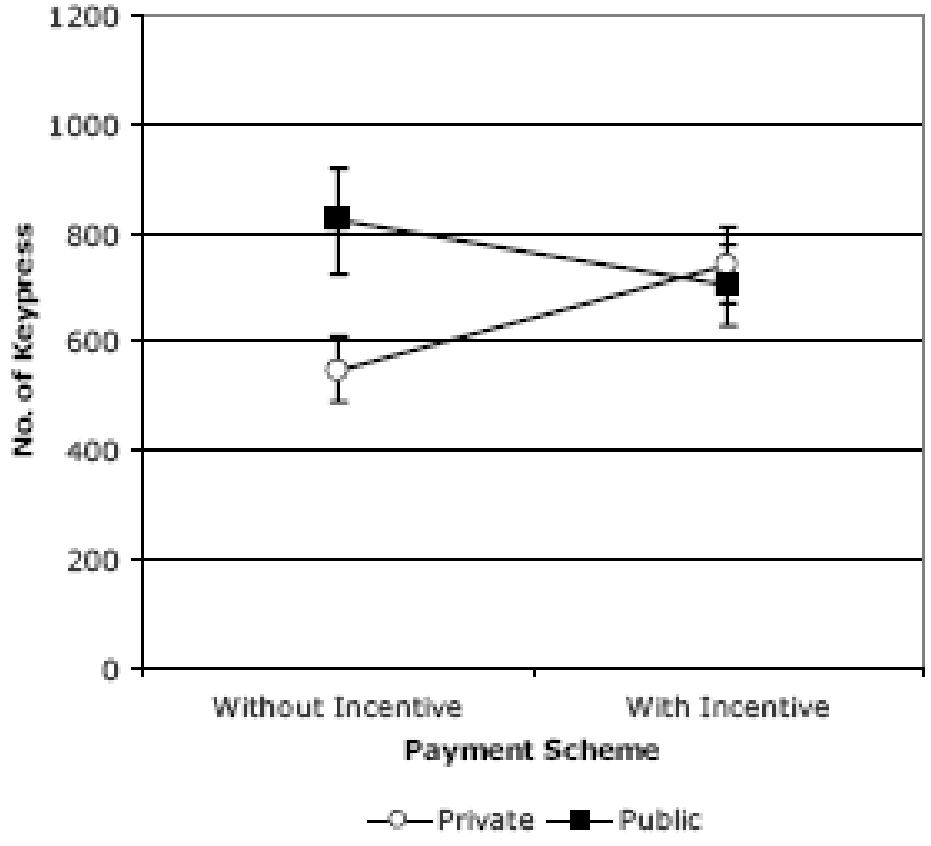
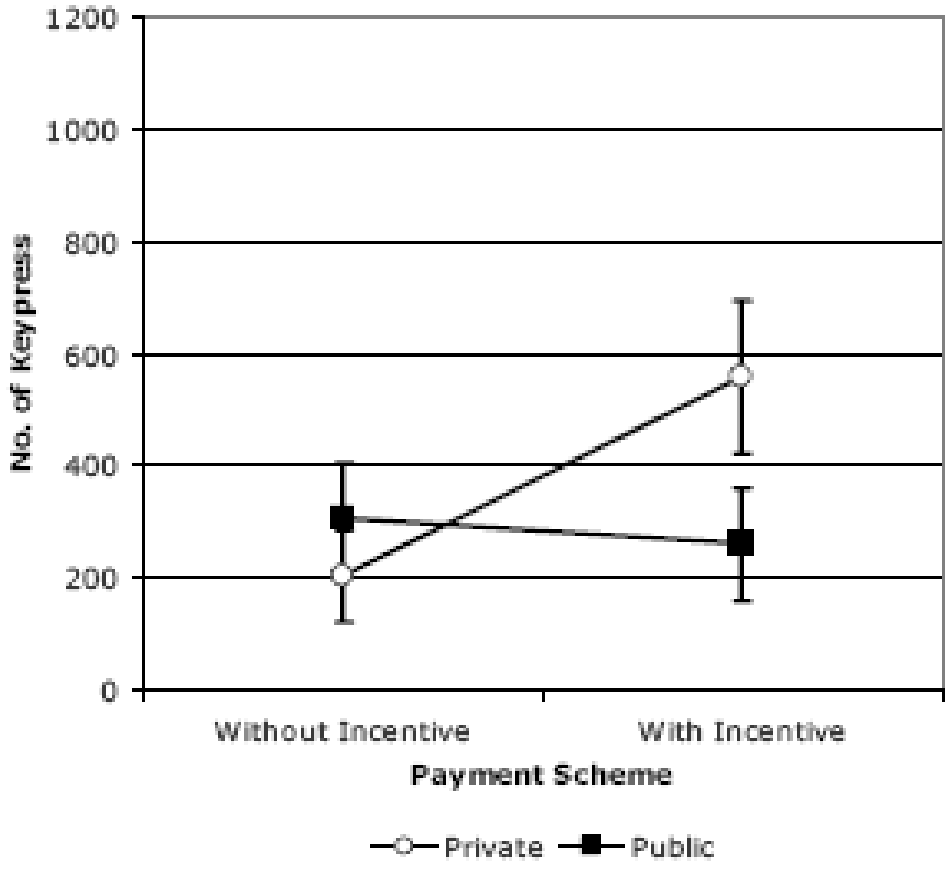


Figure 2: Effect of Private Incentive for "Bad" Charity



Application 1: AIDS prevention

- Ashraf & Bandiera (2014) "No Margin, No Mission?"
- Lusaka, Zambia: collaborate with public health organization that recruits and trains hairdressers and barbers to provide information about HIV prevention and sell condoms in their shops
  - HIV adult- prevalence rates: 14.3%, one of world's highest
  - Aids prevention and condom promotion recognized as public good (gvt. campaigns)
- Experiment: randomly assigns 205 distinct geographical clusters, containing 1,222 agents to:
  - Control group: no rewards ("volunteers")
  - Small or large financial margin: 10% or 90% of each condom sale (restocking)
  - Non-financial scheme: "star" treatment. Gives salon a "thermometer" display, showing condom sales and star stamps on it, one for each sale
- Everyone also receives the same initial training

Recruitment

Become a CARE Promoter!

A great opportunity to help the fight against HIV/AIDS and promote your business!

_____ 2009

Dear Sir/Madam _____ of _____

Society for Family Health (SFH) wishes to invite you to enroll your salon in a CARE female condom promotion program. Your salon would become an official distribution point of the CARE female condom. This represents a great opportunity to improve your business performance through increased visibility and to contribute to the fight against HIV/AIDS in Zambia. **What's SFH?**

SFH is a non-governmental organization whose mission is to improve the health status of Zambians using social marketing techniques, increasing demands and supply of essential health products. Our programs include the promotion of CARE female condoms by hairdressers and barbers.

What's the advantage of joining the program?

As of now, numerous hair salons and barber shops in Lusaka, Chipata, Livingstone, and Kitwe have successfully joined the program. Hairdressers and barbers from these salons and shops tell us that *participating in the program has provided them with the immense satisfaction of helping their community and has attracted additional clients to come to the salon for other services.*

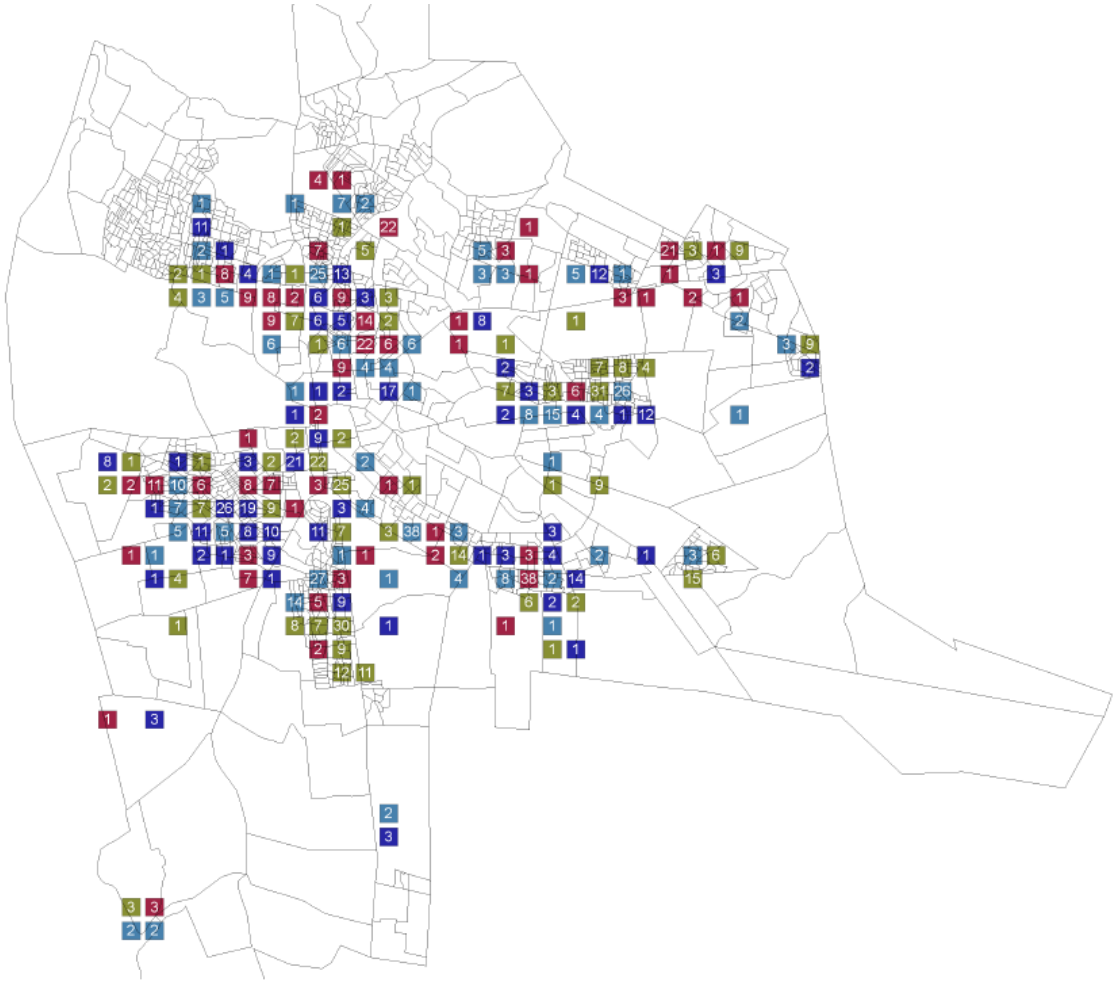
How do I join?

If you are interested in getting involved, we ask you to attend training on HIV/AIDS prevention, adequate use of the female condom and promoting and selling strategies. The training will be held

on _____

at **LUSAKA HOTEL (ON CAIRO ROAD, NEAR KATONDO STREET)** in Lusaka.

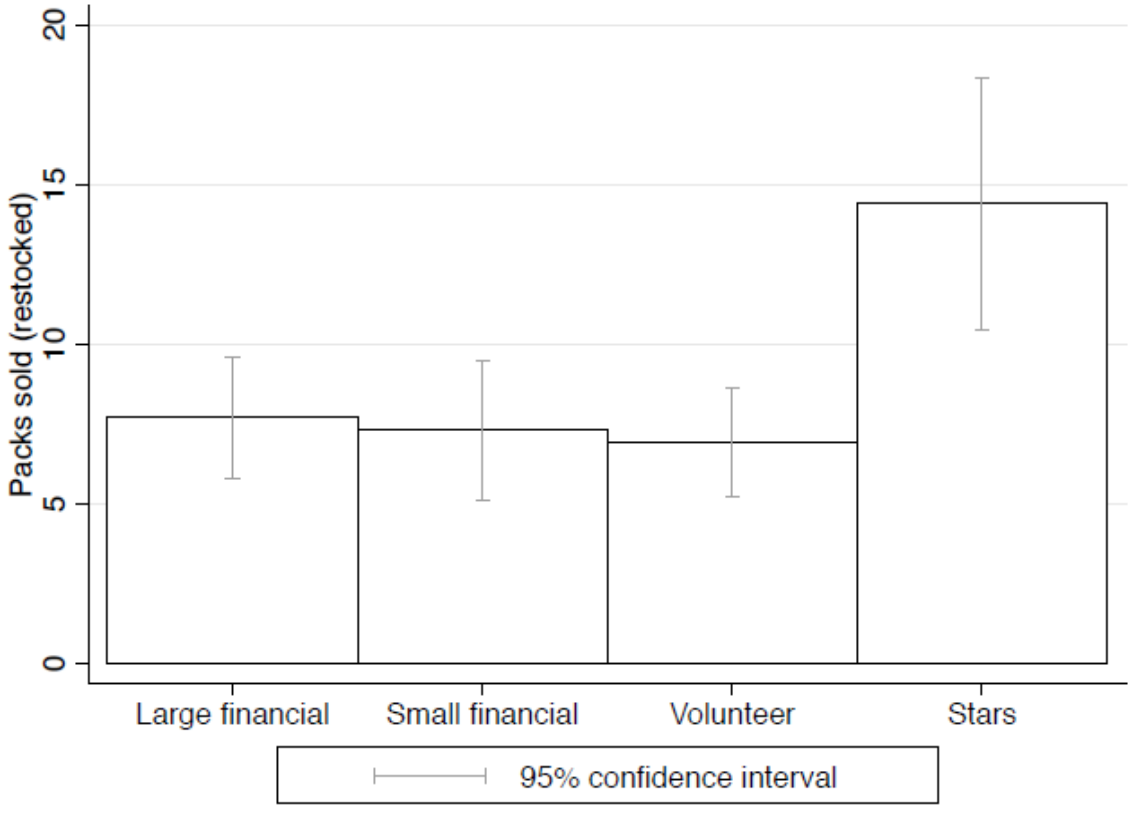
Randomization at neighborhood level



Notes: Treatment groups and volunteer control group are shown by the cell colors. The number of salons attending the training are written in each cell.

- All salons in same neighborhood receive the same treatment, (or left untreated: outside program)

Financial rewards ineffective, image rewards effective



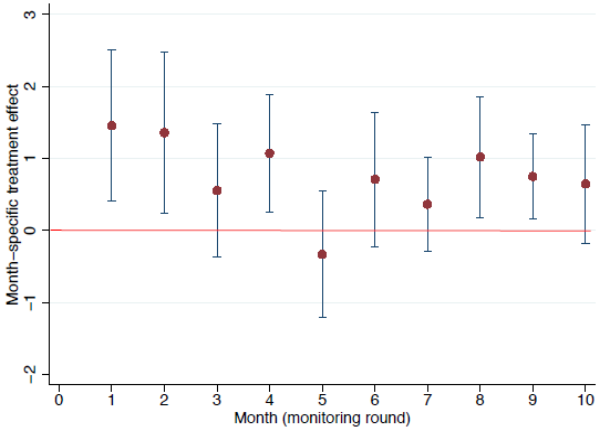
Notes: Each bar measures the average number of packs sold over the year by agents in each of the four groups with 95 percent confidence intervals.

- Agents in the star treatment sell over twice as many condoms as agents in any other group, on average.

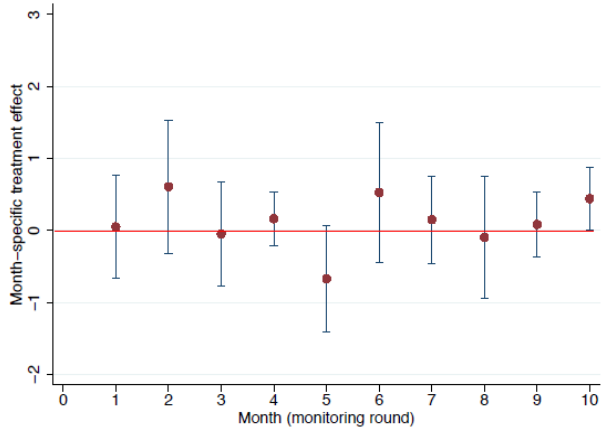
Effects are stable over a full year

Figure 4: Month-specific treatment effects

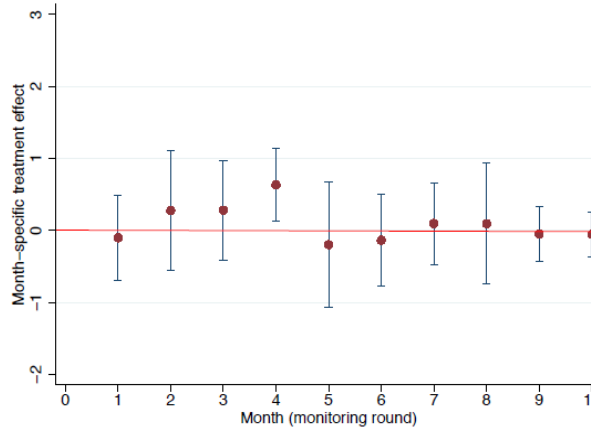
(a) Star reward



(b) Large financial margin



(c) Small financial margin



Different levels of sales reflect different levels of effort

Dependent variable	Total displays	Logbook filled	Promoter attention	Promoter interest	Average standardized effect
Mean in control group	2.285	0.479	2.498	2.111	
Standard deviation in control group	*1.19*	*0.28*	*0.41*	*0.42*	
	(1)	(2)	(3)	(4)	(5)
Large financial reward	0.072	0.028	-0.004	0.024	0.03
	[0.102]	[0.029]	[0.034]	[0.035]	[0.036]
Small financial reward	-0.099	0.008	0.022	0.049	-0.005
	[0.127]	[0.028]	[0.044]	[0.049]	[0.050]
Star reward	0.245**	0.065**	-0.044	0.096**	0.090**
	[0.120]	[0.031]	[0.034]	[0.044]	[0.041]
Controls	yes	yes	yes	yes	yes
R-squared	0.101	0.0234	0.035	0.0605	
Observations	722	722	721	694	726
Large financial = Small financial (p-value)	0.152	0.502	0.516	0.605	0.049
Large financial = Stars (p-value)	0.123	0.219	0.237	0.116	0.133
Small financial = Stars (p-value)	0.0137	0.074	0.12	0.417	0.087

Notes: OLS estimates weighted by the number of observations for each salon. All outcomes are averages are at the salon level across all restocking visits. Standard errors are clustered at the cell level. * p<0.10 ** p<0.05

Different "types" respond differently

Dependent variable is Packs sold (restocked)

Interaction variable	Stylist's dictator game donation is above the median	Stylist's reported work motivation is intrinsic	Stylist's socio-economic status is low	Number of trained salons in the same area is above median
Mean in control group = 6.96				
	(1)	(2)	(3)	(4)
Motivation variable	0.771	-3.631*	-4.126**	-0.983
	[1.531]	[1.958]	[1.610]	[2.302]
Effect of large financial when interaction variable =0	-2.364	-1.66	0.775	2.584
	[1.642]	[2.447]	[2.091]	[2.939]
Effect of small financial when interaction variable =0	1.068	-0.321	-0.077	-0.201
	[1.936]	[2.841]	[1.719]	[2.803]
Effect of stars when interaction variable =0	4.341	3.858	7.016**	2.427
	[2.897]	[3.816]	[2.906]	[3.660]
Effect of large financial when interaction variable =1	3.546	2.63	3.682**	0.223
	[2.490]	[2.228]	[1.839]	[1.741]
Effect of small financial when interaction variable =1	0.383	0.999	4.869*	1.326
	[1.933]	[1.768]	[2.910]	[1.705]
Effect of stars when interaction variable =1	10.010***	10.480***	11.080***	9.144***
	[3.238]	[2.986]	[3.108]	[2.966]
Controls	yes	yes	yes	yes
R-squared	0.073	0.071	0.067	0.073
Observations	765	765	765	765
Large financial: P-value on the interaction term	0.026	0.144	0.301	0.484
Small financial: P-value on the interaction term	0.769	0.686	0.139	0.511
Stars: P-value on the interaction term	0.096	0.134	0.281	0.127

- High v_a types respond only to image rewards. High v_y types (poor) respond to both large financial rewards and to image rewards
- Stars most effective when relevant comparison group is larger (visibility (m_a, m_y))

Application 2: the price of blood

- World Heath Organization prohibits payments for blood donations
  - All countries but Iran prohibit payment for organ donations (liver, kidney, etc.), whether live or cadaveric, in spite of severe and chronic shortages (thousands of deaths each year).
  - Becker-Elías (2007) estimate that donor payments of $15-30,000 would eliminate waiting list for transplants within a few years
  - Same prohibitions and "taboo markets:" human eggs and sperm, surrogate motherhood, etc.

- In practice, a fair amount of disguised payments, and "looking the other way" (out of sight, out of mind):
  - Sale of plasma legal in the US, illegal in Canada. But most plasma used in Canada is bought from the US...
  - Some "indemnities" allowed for egg donors, etc.
  - Any prohibition generates a black market

Arguments, evidence?

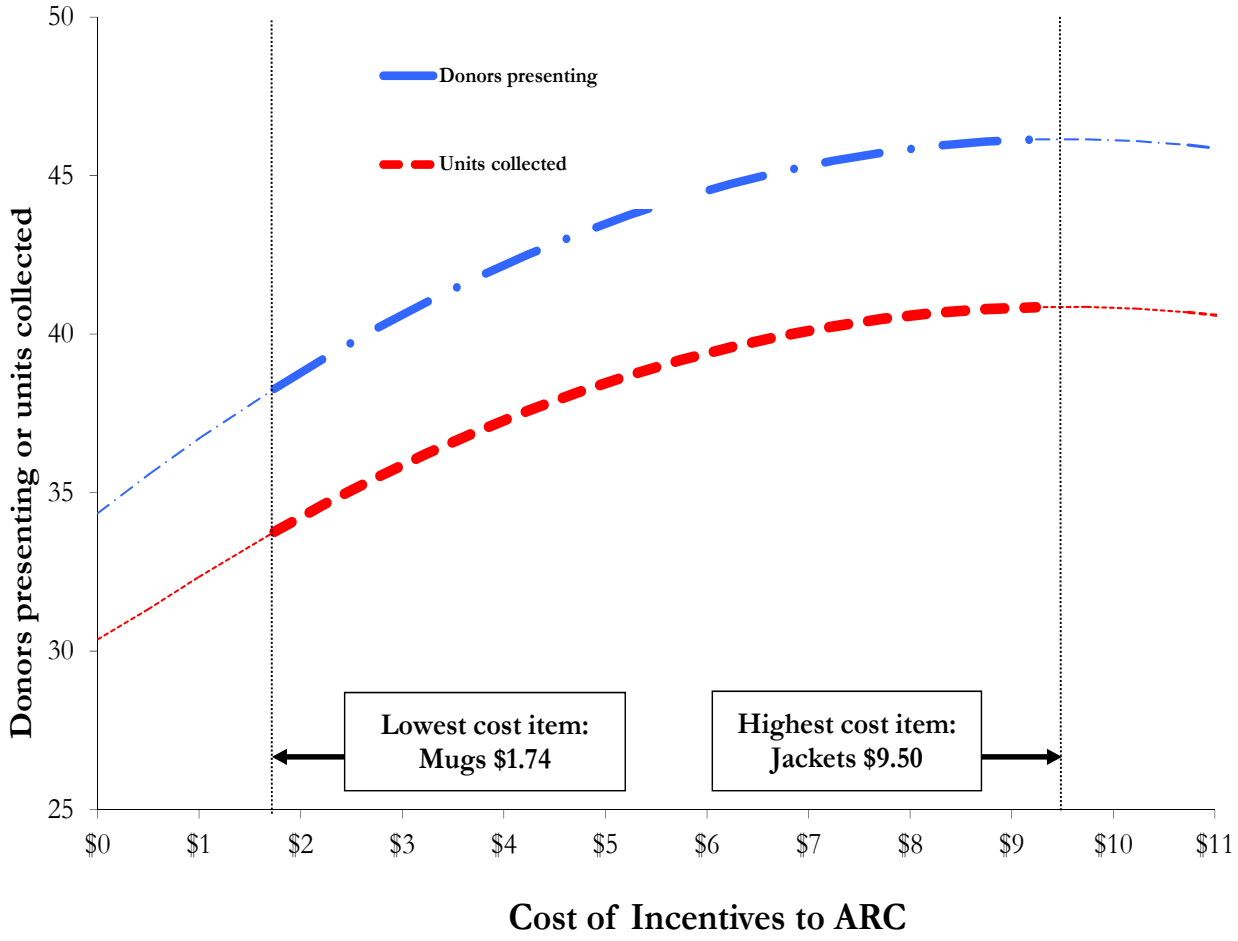
- Influential book by R. Titmuss (1972) "The Gift Relationship: From Human Blood to Social Policies". Strong claims that:
  - Bad for efficiency. Reduces quality of donations (contaminations), and even quantity, creating shortages: people who would do it because it is a virtuous, noble act, no longer will when it becomes a transaction
  - Bad for equity. The poor will sell, the rich will buy
  - Bad for social norms, social order. Will discourage altruism throughout society, "corrupt" moral values, "degrade" human dignity (externalities)
- Evidence?
  - Offered very scant, non-scientific data on blood donations and shortages in UK vs. US, and none at all on social spillovers
- But common type of righteous claim, argument "by fiat"
  - M. Sandel (2012) "What Money Can't Buy: The Moral Limits of Markets"
- Recent theoretical work $\Rightarrow$
  - Lab experiments: suggesting some crowding out of willingness to donate blood. But few, small-scale, non-representative populations, examine only registration to donate...

"Will There be Blood?"

- Lacetera, Macis, Slonim (2012), (2013), (2014)...: series of large-scale studies on effects of actual incentives on actual blood donations...
- Can't offer cash, but:
  1. Small presents: T-shirts, mugs, store coupons (US)
  2. Gifts cards (quasi-cash): \$5, 10, 15 (US) and supermarket vouchers of \$10, 20 (Argentina)
  3. Paid day off work for employees (Italy)
  4. Mention name in local newspaper (Italy, Argentina)
- Some studies are observational (naturally occurring data from regular blood drives), others are large-scale randomized interventions, conducted together with American Red Cross % similar orgs.

▶ Go to data

Lacetera-Macis-Slonim (2012)



Observational analysis of ~14,000 ARC blood drives in N. Ohio.

Incentive items include t-shirts, mugs, coupons…

Exploit "haphazard" within drive variation in availability of incentive items, over time; fixed effect specification

Lacetera-Macis-Slonim (2014)

- LMS 2014: Natural field experiment with ~100,000 subjects in N. Ohio.
- Incentive items are $5-10-15 gift cards

27 drives; 92,722 individuals-wave:
Advertised Reward

9 drives; 35,968 individuals-wave:
Unadvertised Reward

A: ~50% informed of drive in flyer, **and** informed of rewards.

B: ~50% informed of drive in flyer, **not** informed of rewards.

C: all informed of drive in flyer, **not** informed of rewards.

(1) **Direct local effects**: Difference in probability of donation at intervention drive between A and B.

(2) **Spillover effects**: Difference in probability of donation at intervention drive between B and C.

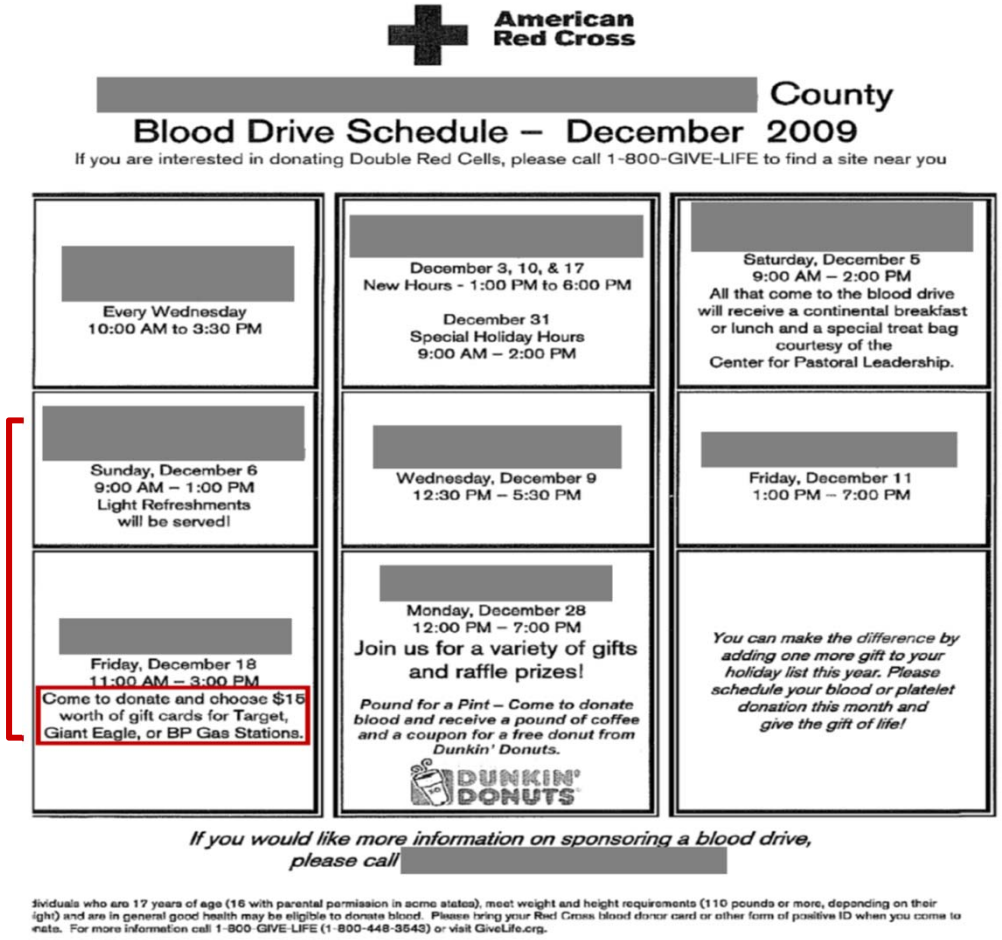
(3) Total local effects: Difference in probability of donation at intervention drive between A and C.
(4) Spatial displacement effects: Difference in probability of donation at non-intervention drive between A and C.
(5) Local heterogeneous effects: Difference in probability of donation at intervention drive between A and C, per donor characteristics.
(6) Heterogeneous displacement effects: Difference in probability of donation at non-intervention drive between A and C, per donor characteristics.

(7)-(8) Post-intervention effects: difference in donation probability and number of donations between A and D in post-intervention period; for all contacted subjects (7) and limited to subjects who donated during the intervention period at the intervention drives (8). Excluded from D are subjects who were also contacted for an Advertised reward drive.

36 drives: **No Reward**

D: select one drive with no reward for each Reward drive's county.

- LMS 2014: Natural field experiment with ~100,000 subjects in N. Ohio.
- Incentive items are $5-10-15 gift cards



American Red Cross

[redacted] County

Blood Drive Schedule – December 2009

If you are interested in donating Double Red Cells, please call 1-800-GIVE-LIFE to find a site near you

[redacted] Every Wednesday 10:00 AM to 3:30 PM	[redacted] December 3, 10, & 17 New Hours - 1:00 PM to 6:00 PM December 31 Special Holiday Hours 9:00 AM – 2:00 PM	[redacted] Saturday, December 5 9:00 AM – 2:00 PM All that come to the blood drive will receive a continental breakfast or lunch and a special treat bag courtesy of the Center for Pastoral Leadership.
[redacted] Sunday, December 6 9:00 AM – 1:00 PM Light Refreshments will be served!	[redacted] Wednesday, December 9 12:30 PM – 5:30 PM	[redacted] Friday, December 11 1:00 PM – 7:00 PM
[redacted] Friday, December 18 11:00 AM – 3:00 PM Come to donate and choose $15 worth of gift cards for Target, Giant Eagle, or BP Gas Stations.	[redacted] Monday, December 28 12:00 PM – 7:00 PM Join us for a variety of gifts and raffle prizes! *Pound for a Pint – Come to donate blood and receive a pound of coffee and a coupon for a free donut from Dunkin' Donuts.* DUNKIN' DONUTS	*You can make the difference by adding one more gift to your holiday list this year. Please schedule your blood or platelet donation this month and give the gift of life!*

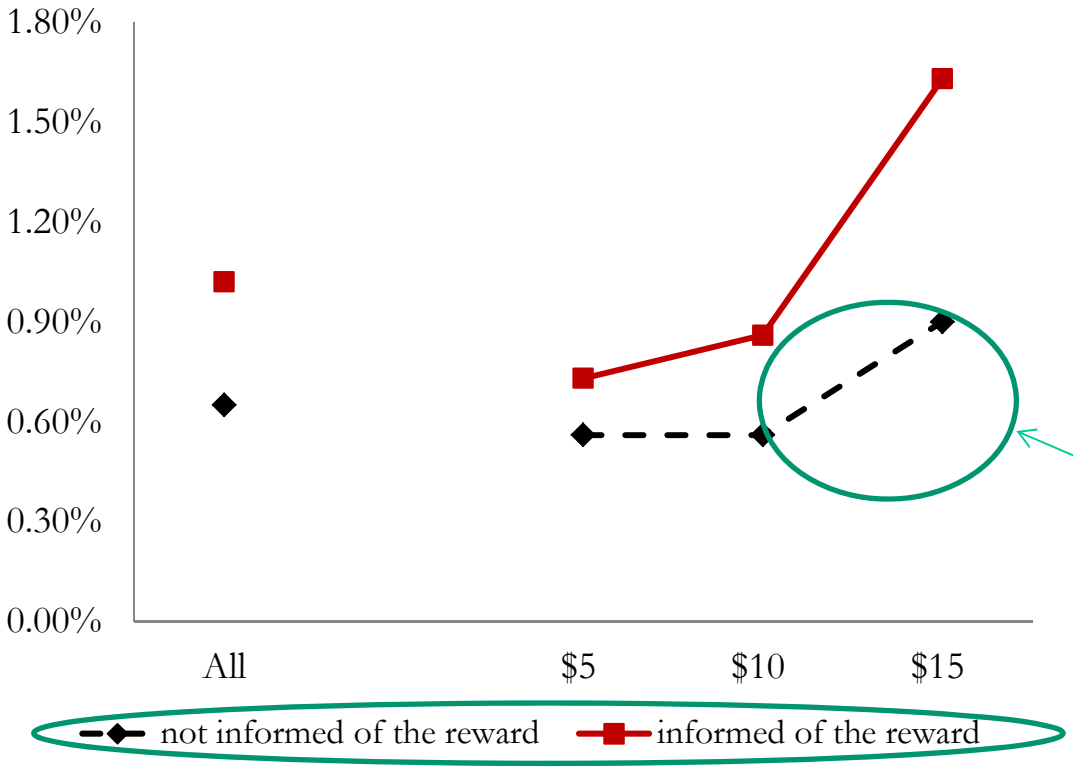
If you would like more information on sponsoring a blood drive, please call [redacted]

[illegible]viduals who are 17 years of age (16 with parental permission in some states), meet weight and height requirements (110 pounds or more, depending on their [illegible]ight) and are in general good health may be eligible to donate blood. Please bring your Red Cross blood donor card or other form of positive ID when you come to [illegible]nate. For more information call 1-800-GIVE-LIFE (1-800-448-3543) or visit GiveLife.org.

Fliers inform ‘000s individuals of upcoming drives in a county: date, location, whether rewards offered

Standard ARC procedures; Individuals unaware of experiment

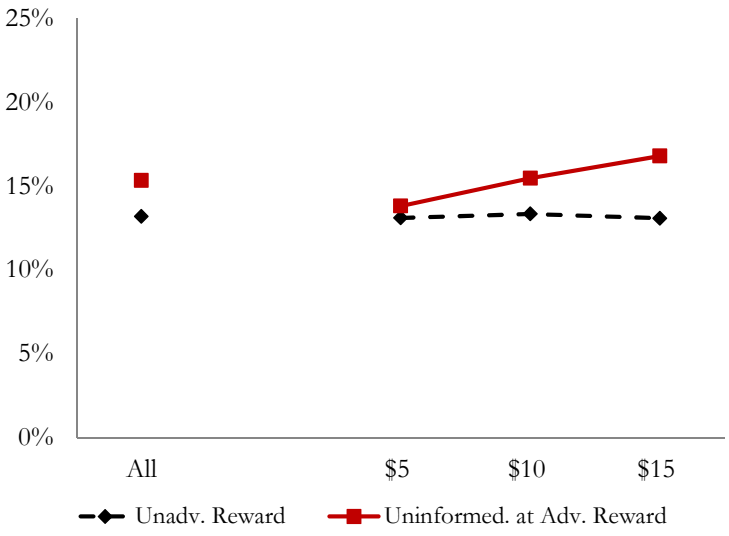
Yield per flyer sent



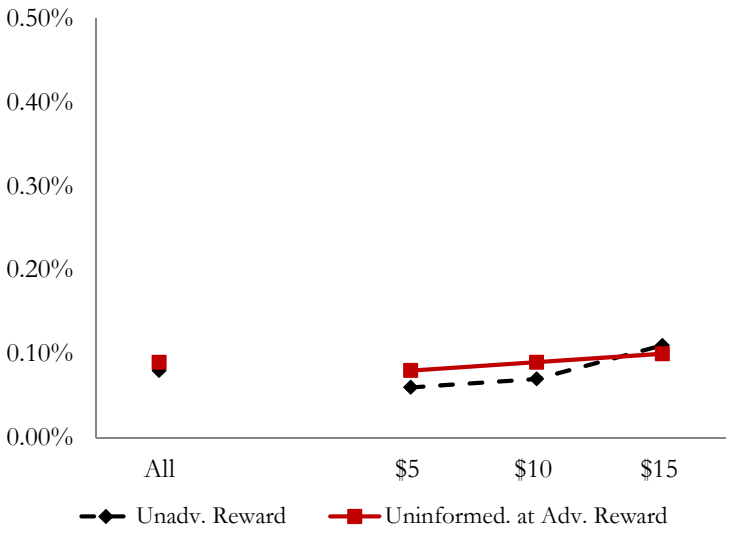
Spillover effect (driven by subjects with previous donations at sites: “neighbours”)

For a given intervention drive, half subjects informed of rewards, half not informed. All would receive gift cards

Donation rates at site



Donors with previous history at intervention sites



Donors without previous history at intervention sites

+: evidence of spatial displacements (toward drives with rewards) and temporal substitution (to enjoy rewards)

Cost-benefit analysis

		Past history at sites		
		$5	$10	$15
	All values are per 100 individuals contacted			
1	Units collected - baseline when no incentives offered [(1)]	13.18	13.18	13.18
2	Donors presenting - baseline when no incentives offered[(2)]	15.14	15.14	15.14
3	Extra units collected when incentives offered[(3)]	--	6.79	6.07
4	Extra donors presenting when incentives offered[(1)(2)]	--	8.20	11.66
5	Total N. of donors presenting when incentives offered	15.14	23.35	26.81
6	$ cost of providing incentives[(4)]	$75.70	$233.5	$402.1
8	$ cost per extra unit collected[(5)]	--	$34.4	$66.2

[(1)] From Table 9, columns 1 and 4.
[(2)] Donors presenting = units collected * 1.149 (donors deferred are 13% of donors presenting, irrespective of the presence of incentives).
[(3)] From Table 9, columns 3 and 6. Note that we used zeros when the coefficients were statistically insignificant.
[(4)] $ value of the incentives * total N. of donors presenting at drives with incentives.
[(5)] Total cost of providing incentives/N. of extra units collected when incentives provided.

- Indicative: hospitals in US charge up to $ 1000 for transfusion of one unit (0.5 l)

Iajya-Lacetera-Macis-Slonim (2013)

- Natural field experiment with ~18,000 subjects in Argentina
- Promote voluntary, undirected donations as opposed to emergency/replacement

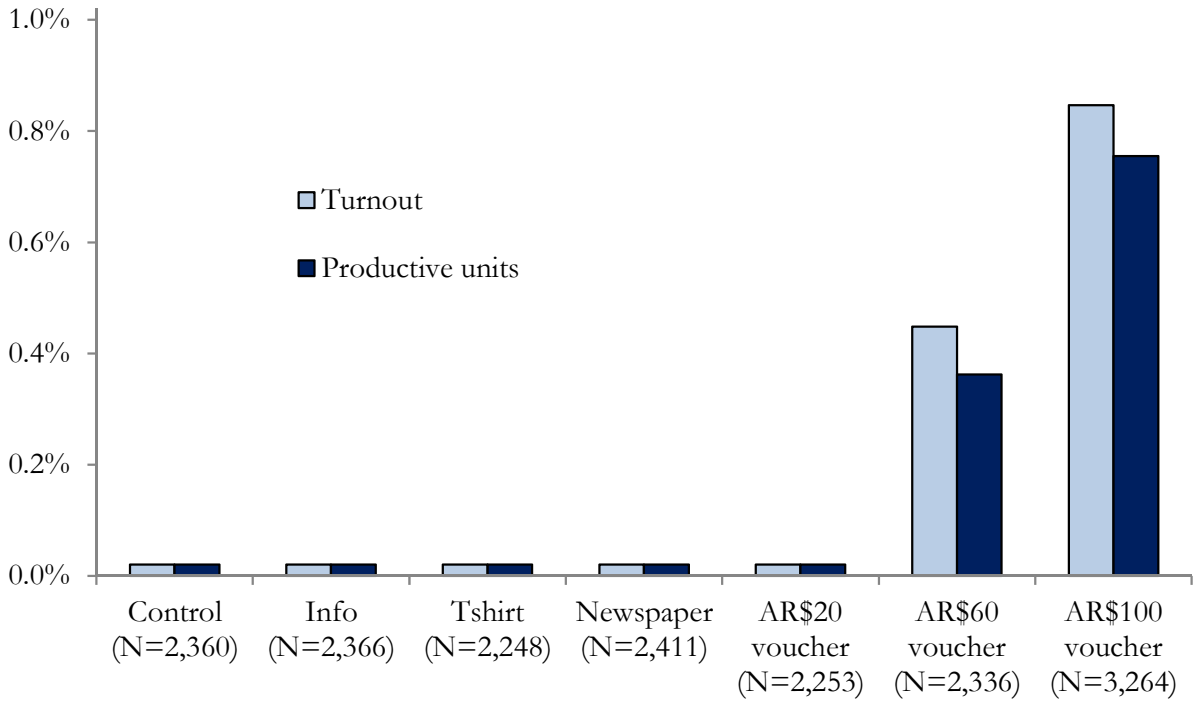
T0: Invitation Allocated: 2,500 Received intervention: 2,360	**T1: Invitation + Information** Allocated: 2,500 Received intervention: 2,366	**T2: T1 + Social recognition: t-shirt** Allocated: 2,500 Received intervention: 2,248	**T3: T1 + Social recognition: newspaper mention** Allocated: 2,500 Received intervention: 2,411	**T4: T1 + AR$20 Voucher** Allocated: 2,500 Received intervention: 2,253	**T5: T1 + AR$60 Voucher** Allocated: 2,500 Received intervention: 2,336	**T6: T1 + AR$100 Voucher** Allocated: 3,500 Received intervention: 3,264

- People paid for showing up, not for actual donation (promotes honest reporting of risks)

Outcomes:

- Blood donor turnout at CMTH – number and share of analyzed
- Actual donations performed– number and share of analyzed
- Usable donations – number and share of analyzed
- Reasons for unsuccessful blood collection: ineligibility, walking away before donating, blood testing positive for infectious diseases – number and share of analyzed

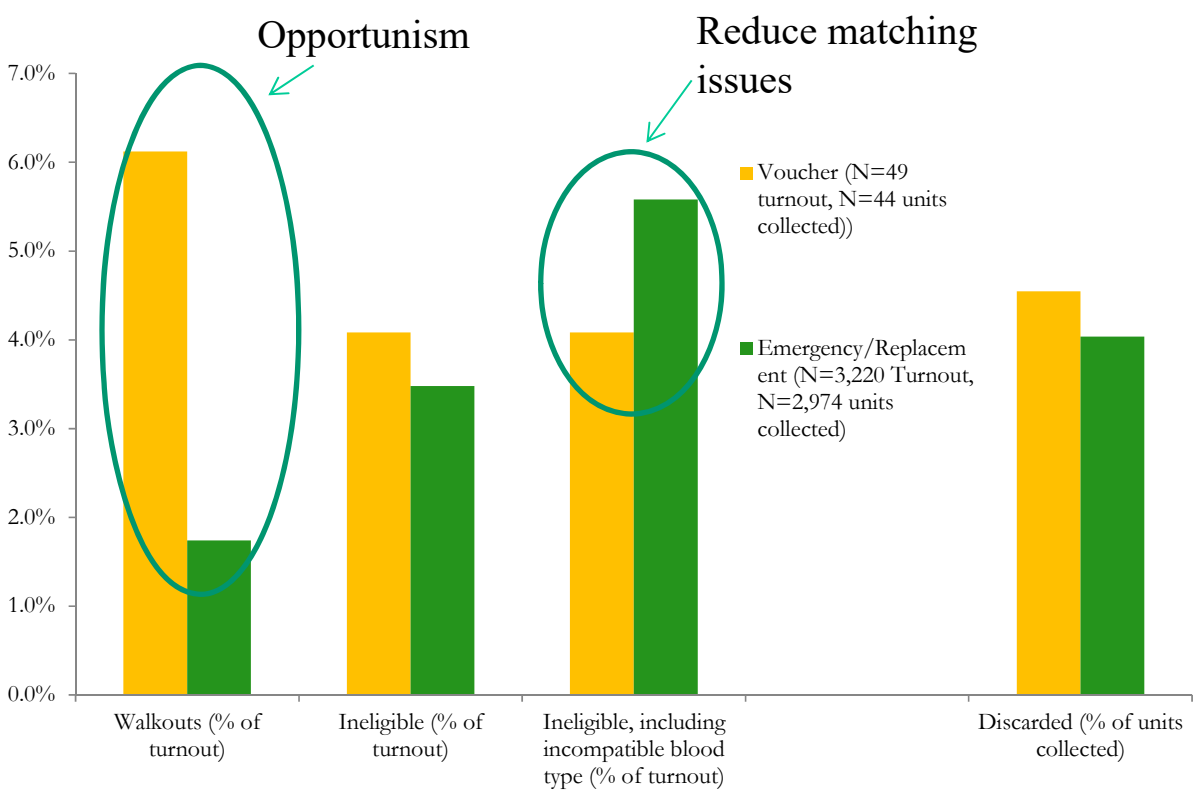
Treatment Effects



ILMS 2013: Natural field experiment with ~18,000 subjects in Argentina

Promote voluntary, undirected donations as opposed to emergency/replacement

Matching and Quality



ILMS 2013: Natural field experiment with ~18,000 subjects in Argentina

Summary: Incentives for Blood Donations

- 18 of the 19 distinct incentive items increased blood donations
- No crowding out of quantity or quality
- Effects increase with the $ amount of the reward
- There are spillovers effects
- Spatial displacement and short-term shifts in timing of donations
- Induce undirected donations in emergency/replacement context (Argentina)
- No long-term effects
- Financial costs generally low, esp. as compared to benefits: ~$30-$50 for one extra unit for vouchers/ gift cards (~$300 for day off: Italy)
- Niessen-Ruenzi et al. (2016): similar findings in Germany with (withdrawal of) cash incentives

III. OPTIMAL INCENTIVES WHEN NORMS ARE PRESENT

- Back to model

WELFARE AND POLICY

What is the overall effect for society from individual contribution (e.g.: I buy an electric car)

I get:

- Cost to individual: $-c$
- Intrinsic value: v_a = how much I value the improvement in public good (e.g., air quality) that this brings about, or pure "joy or giving", satisfaction from doing "the right thing"
- Extrinsic reward: y = I get a subsidy, or avoid a tax
- Improved (self) image: $m_a \times (Honor - Stigma)$

Others get:

- Incentive payments: $-y$ (from taxes or private sources)
- Benefit created by increment to the public good: e (a bit less pollution, cleaner air)
- Loss of self image: Contributors get less honor, and non-contributors more stigma!

  (all loose reputation: SUV owners, but also other electric car owners)

- Key point: pursuit of esteem is a zero-sum game: the average reputation in society remains fixed.

- In sociologists' words, esteem, even self-esteem, is by very nature a *positional good.*

The Graffiti of the Philanthropic Class

❑ Key point: pursuit of esteem is a zero-sum game: what I gain, others lose

- The net social return (what I do not internalize in my decision) to contributing is thus

S = Benefit created for others from increased public good – others' loss in self or social image

= Benefit created for others from increased public good – gain in own self or social image

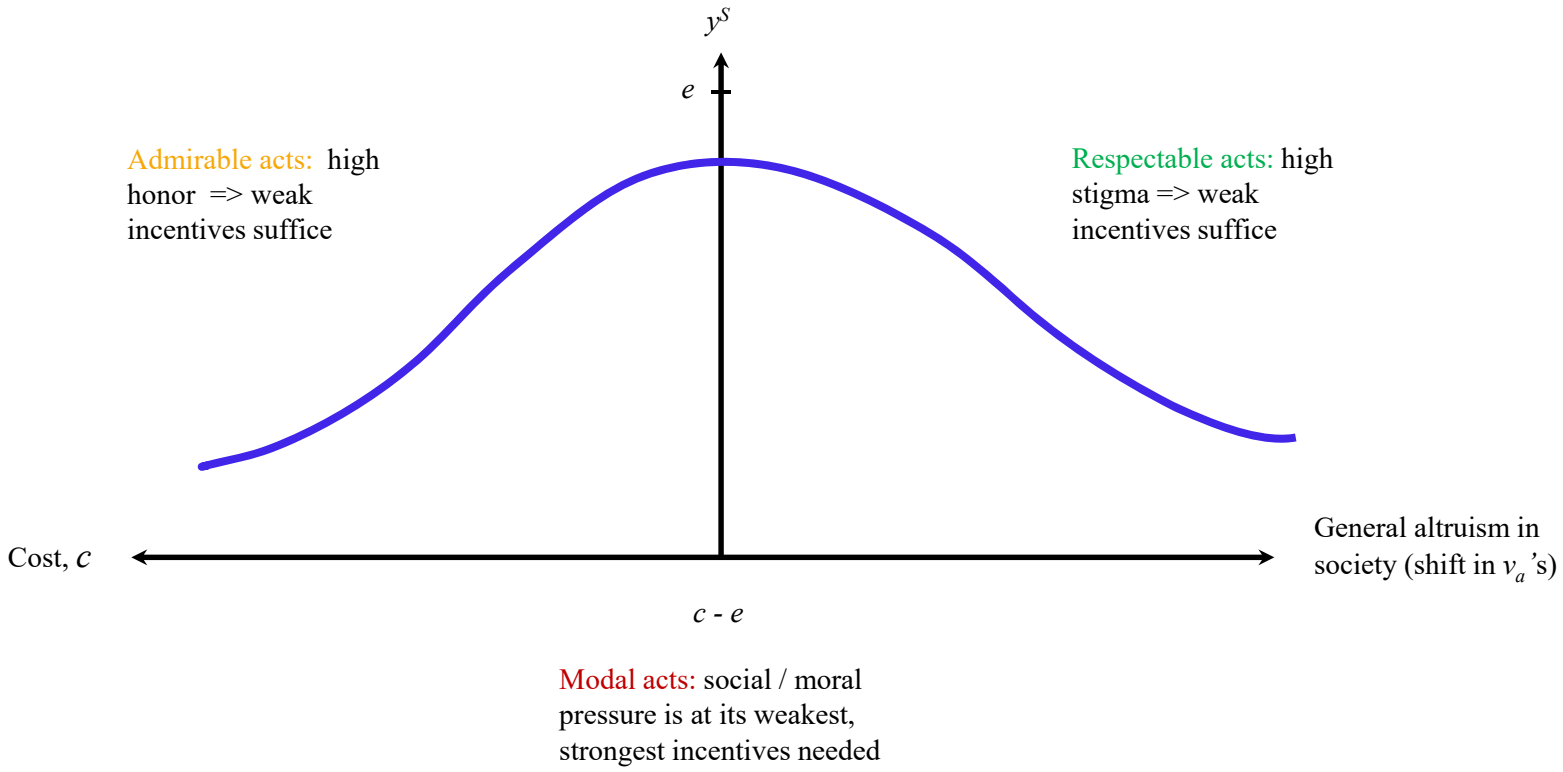
Difference between free-riding effect and reputation-stealing effects ⇒

Proposition *The socially optimal incentive rate is strictly less than the standard subsidy that leads agents to internalize the full public-good value of their contribution. It subtracts the value of image "bought":*

$$y^s = \underbrace{e}_{\textit{Externality}} - \underbrace{m(\mathcal{M}^+ - \mathcal{M}^-)}_{\textit{Reputation Tax}}(c - y^s)$$

❑ Implication: for a typical (bell-shaped) distribution of "values" in society, the optimal incentive rate is bell-shaped as a function of the contribution cost and of the average degree of altruism

Optimal Incentives (subsidy, fine, law…)



$$Optimal\ incentive = social\ value\ of\ contribution - net\ reputational\ gain$$

$$y^s = e - m(\underbrace{\mathcal{M}^+ - \mathcal{M}^-}_{Honor - Stigma})(c - y^s)$$

Policy implications: tax treatment of charitable contributions and externalities

- Deduction rate may be lower than thought (or even be a tax). More concretely:

- Pattern of contributions will be distorted toward those that are most visible (high *m):*

  - Alumni giving to wealthy universities rather than (their or other) high schools, primary schools, preschool programs. Get your name on building, a professorship, etc., at prominent institution, rather than on public school in small town.

  - Giving to big hospitals, museums, etc., rather than rural clinics, vaccination programs in 1/3 world, etc. [NYT on new Chinese elite's giving patterns]

  -D. Ariely: putting solar panels on your roof vs. new furnace or insulation

- This gets worse with sponsor competition, e.g., between NGO's universities, etc. ⇒ kind of arm's race in image seeking. Show can be worse for welfare than a monopoly.

⇒ Tax deductibility of different contributions or prosocial actions should, to the extent possible, vary (inversely) with the publicity / image value inherent to them

- Same for ethical funds, fair trade, "green" products: the premium you pay also buys you social and self image… and confers stigma / bad conscience on others.

Other uses of same money may do more social good but don't have those image private benefits and social externalities (e.g., give to orphanage) ⇒ too little of them.

II. HONOR, STIGMA AND SOCIAL NORMS

- Back to model

II -HONOR, STIGMA, AND SOCIAL NORMS

- What makes a behavior socially or morally unacceptable is often the very fact that "it is just not done". But in other times, other places: "everyone does it".

  [choosing surrender over death, not going to church, not voting, divorce, welfare dependency, minor tax evasion, conspicuous modes of consumption,....]

- People contribute more, behave "better" when they know / see that others do

  [public goods, fundraising, voting; helping strangers, Salvation Army…]

- Often explained and modeled by some form of untargeted "reciprocity": interdependent preferences. (Not whole story, some evidence against it)

- In other situations, people will try to distinguish themselves from "what most people do", and reap social/self esteem from it: heroism, organ donation. identity.

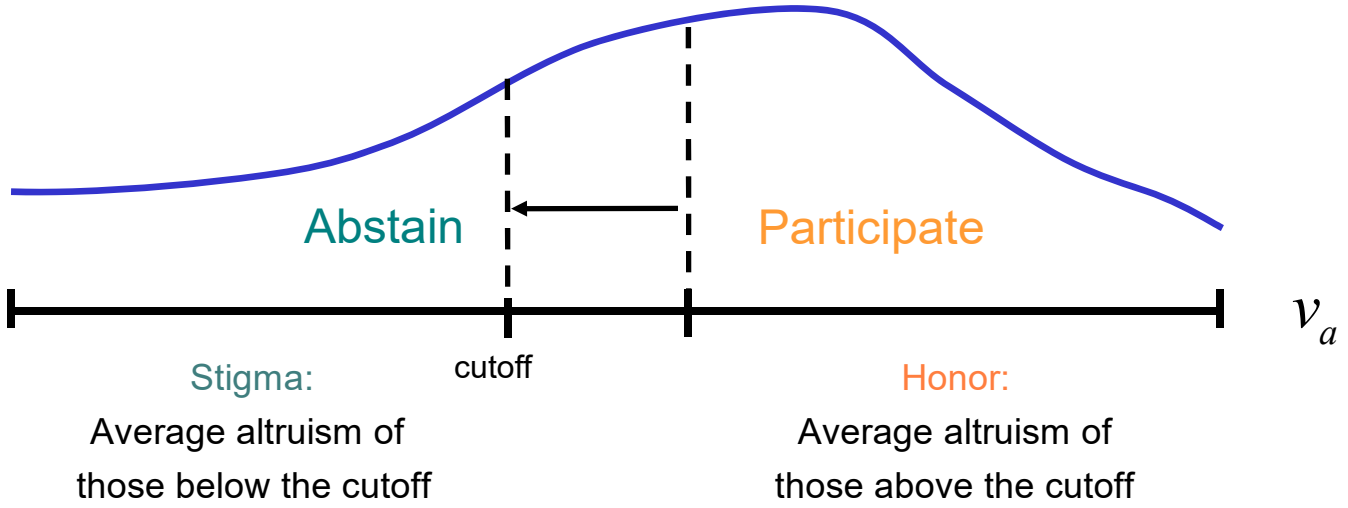
- Social or norms arise endogenously from the interplay of honor and stigma.

- When does the fact that others contribute more increase or decrease the pressure (social, moral) on me to do so? (complements vs. substitutes, conformity vs. distinction).

- Policy implications: in each case, what effects of incentives, laws? How should set?

Honor and Stigma

- Simplify: yes/no decision (blood, vote) & focus on differences in prosocial orientation only (not greed / need) => only image concern measured by $m = m_a$
- Fix incentive. Who participates, and how are they seen?



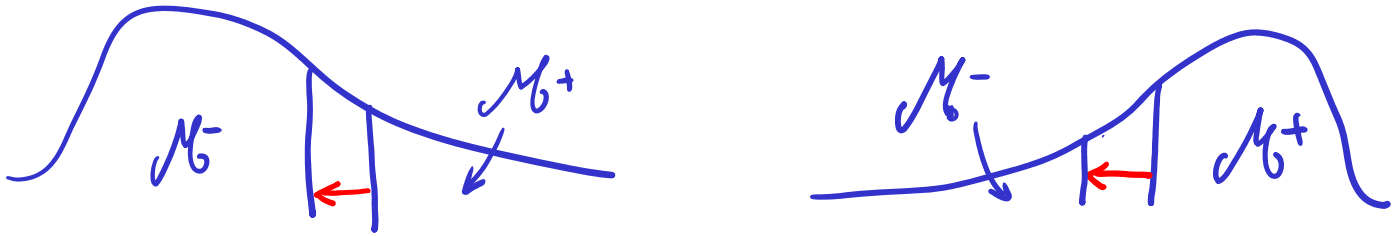
- People who contribute are those with

$$intrinsic\ altruism + extrinsic\ incentive + image\ concern \times [Honor - Stigma] \geq cost$$

- When more people participate, *honor declines* but *stigma worsens!* ⇒ Social / moral pressure to participate may decrease or increase ("multiplier" $\gtrless 1$)
- Key difference between behaviors in which quest for honor versus avoidance of stigma is main driver of behavior; role of (endogenous) initial prevalence

Role of the distribution of values in society

- Expect honor-seeking considerations to dominate when there are only a few heroic or saintly types, whom the mass of more ordinary individuals would like to be identified with.
  - Heroism, organ donation…
  - Other activities for which being prosocial has high cost



- Expect stigma-avoidance considerations to dominate when population includes only a few "bad apples" with very low intrinsic values, which most agents are eager to differentiate themselves from.
  - Serious crime, spousal abuse…
  - Other activities for which being prosocial has high cost

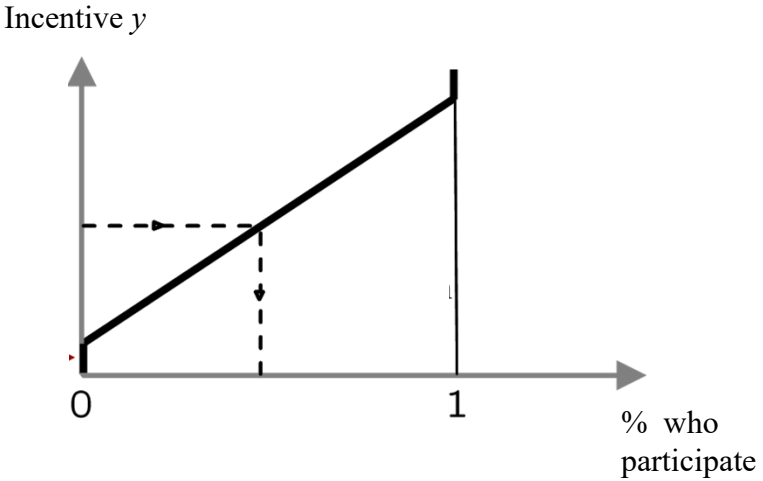
Proposition *When do we get multiple social norms? What do incentives do?*

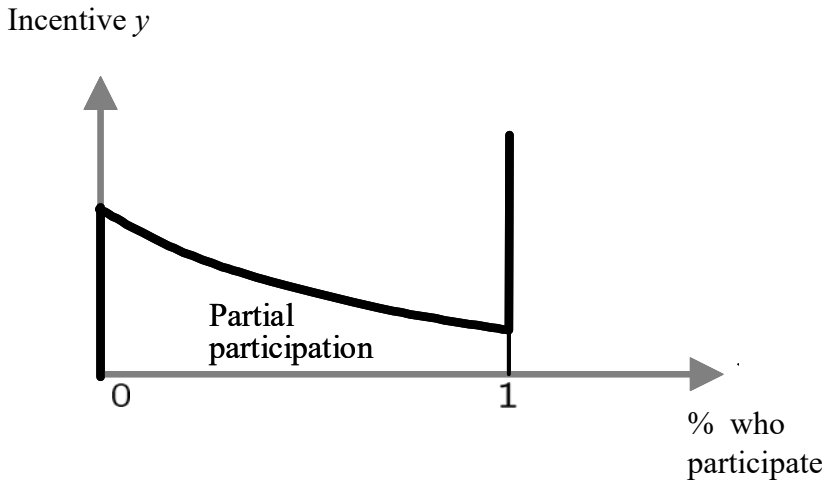
When honor motive is dominant:

- Individuals' decisions are substitutes
- Incentives → partial crowding out
  (still work, but weakened)

When stigma motive is dominant:

- Individual's decisions are complements
- Crowding in => small incentives can have large effects, even shift norms
- Multiple norms may coexist





This occurs when:

- Most people are "mediocre", or cost c very high: only rare "saintly" types with v_a well above most others (heroism, organ donation)
- There are possible "excuses" for not participating, and / or one can do it without being noticed (⇒ weak stigma)

This occurs when

- Most people are "OK", or cost c fairly low: only a few "rotten apples" with v_a well below most others (crime, child neglect)
- There are possible selfish reasons for behaving well (e.g., fear of the law), and/or participation may go unnoticed (⇒ weak honor)

Implications

- Material incentives (prizes, fines, law) not very effective to spur "admirable" prosocial behaviors (honor driven): $y\nearrow$ weakens motivation for the upper tail (partial crowding out). Better to give visibility, image rewards ($m\nearrow$)

- More effective to strengthen "respectable" behaviors (stigma-driven): $y\nearrow$ intensifies social pressure in the lower tail (crowding in). Small changes in incentives can now have large effects, shift social norms, when cost is fairly low and non-compliance easily observable

  - Ireland: 33¢ tax on plastic shopping bags + awareness campaign*: "Within weeks, plastic bag use dropped 94%. Within a year, nearly everyone had bought reusable cloth bags, keeping them in offices and in the backs of cars."*

  - How did it work? *"Plastic bags were not outlawed, but carrying them became socially unacceptable -on par with wearing a fur coat or not cleaning up after one's dog."*

- Positive implications: effect of material incentives should generally be stronger where compliance / socially approved behavior is already more prevalent => Tests

- Normative implications: optimal level of tax or subsidy on activities with positive /or negative externalities. See next.

Testable Implications

- When a normatively approved behavior is sufficiently prevalent, stigma-avoidance rather than honor-seeking will be the dominant attributionnal concern $\Rightarrow$ formal incentives will have more powerful effects on compliance (crowding-in) .

- When a normatively approved behavior is sufficiently rare, honor-seeking rather than stigma-avoidance will be the dominant attributionnal concern $\Rightarrow$ formal incentives will have weaker effects on compliance (partial crowding-out)

- Cross effect prediction: response to increased incentive (da/dy) to $y \nearrow$ is larger, the greater the initial level of compliance $(\bar{a})$

- Prevalence of good or bad behavior is, of course, endogenous
  - But exogenous / experimentally manipulable factors shift it: visibility of action m, its cost c, historical practices and norms

Application 3: the economics of identity

1. Identity: personal, social, gender, ethnic, cultural, political, national...
   - Traditionally a topic for sociologists and social psychologists, but now considerable amount of work by economists as well
   - Theoretical (Akerlof-Kranton 2000, Bénabou-Tirole 2011, Bisin-Verdier 2016), experimental (Chen-Li 2009), empirical (Bisin-Pattachini-Verdier-Zénou 2011)
   - Is identity a Kantian-like, priceless "dignity," or a malleable element of choice, shaped by costs and benefits?
2. Affirmative-action policies toward disadvantaged minorities
   - US of course, but also India ("scheduled castes and tribes"), Canada and Australia ("first people"), China ("ethnic minorities")
   - What are their effects, incidence, how do people respond?
   - Clearly a setting where powerful social norms and economic motives are at play, and interact

- Jia and Persson (2017): "Ethnicity in Children and Mixed Marriages: Theory and Evidence from China". Use and test the BT model

Ethnic minorities in China

- In 2010: Han (1.2 billion) + 55 minorities (105 million)
- Great regional dispersion: minorities = 0.3% (Jiangxi) to 94% (Tibet)
  - Variations within region as well: prefecture level
- Substantial affirmative-action style policies by national and provincial governments
  - Exceptions to the one-child policy (which got tightened in 1980)
  - University entrance exam: extra points for minority children (1977$\mapsto$)
  - Some employment priority
  - These policies rolled out / implemented gradually across places
- Parents in mixed couples choose the official identity of their child
  - It is then fully official and public information: appears on all his/her documents, school, etc.
- Economic incentives are clear. Will check what the social norm is, then study how the two interact

Jia-Persson (2017)

- Data sources
  - 1% samples of 1982 and 1990 censuses, 0.095% sample of 2000 census, 1% sample of 2005 mini-census
  - Information on demographics and socioeconomic status for about 25 million people
  - Outcomes: choose minority identity/ethnicity for child or not (individual level)
  - Incentives: (1) Pre-post 1980: tightening of family-planning policies

    (2) Actual differential fertility of *HM* vs. *HH* couples

    (3) Average score difference in entrance exam for minorities
  - Social-group features: region, prefecture, subgroup, individual levels
- Prevailing social norm: children, especially sons, should have the ethnic identity of their father $\Rightarrow$
  - No tradeoff for *MH* couples: father is Minority, mother is Han
  - Real action will be for *HM* couples: father is Han, mother is Minority

Model predictions and empirical strategy

1. Following the rollouts and gradual spread of the various affirmative action benefits, the fraction of *HM*'s choosing mother's (minority) identity for their child should increase..
2. It should increase more (greater sensitivity to material incentives) where, initially, prevalence of the social norm (choosing Han identity) was high, i.e. fraction choosing mother's identity was low $\Rightarrow$
   - Compare prefectures where that fraction was $< 50\%$ vs. $> 50\%$. Other cutoffs as well, from 10% to 90%
   - Compare prefectures ranking in each of the 4 quartiles according to that fraction: effect of incentives should be smallest in 4th quartile than in any other, and decreasing from 2d to 3d to 4th quartile
3. It should increase less where the cost of "doing the right thing" (choosing Han identity) is higher:
   - Child is a son (see dialogue)
   - Mother is from a religious minority, hence so will be the child (greater "distance" to father's identity; see dialogue)

▶ Go to data

Figure A1 Anecdotal Evidence on Ethnic Choice

(a) Example 1:

Notes: This discussion comes from http://www.babytree.com/ask/detail/3690549, which shows that parents are thinking about both social motives and ethnic policies (especially the option of having more children for their child if they choose minority for their child).

(b) Example 2:

作者: 中二门　时间: 2010-5-18 13:17
标题: 我为孩子选择了汉族，放弃了少数民族
前一阵去给孩子办户口，我是汉族，孩子她妈是少数民族，我对警察说孩子入汉族。
警察大姐好好把我教育了一番，说中考的时候少一分就多一操场人，她家孩子要是有6分，稳上四中了，说要对孩子负责。
最后我还是坚持入了汉族。我是想孩子成长过程中要靠自己的实力，而不是靠特权和特殊的照顾。我可以花钱让她去上好的培训班，也可以给她辅导，而不是直接把分给她。希望孩子长大后能理解她的父母。

Zhongermen: "I went to register the birth of my child a while ago. I am a Han man and my wife is a minority. I told the police that I want my child to be a Han. The police kindly suggested that I should choose minority for the child. She said that one score lower implies an extra playground of competitors in the high-school entrance exam and that I should be responsible for my child's future. But I insisted on choosing Han in the end. I hope that my child's future will reply on his own ability, not ethnic favors."

作者: fh2315　时间: 2010-5-18 13:27
我觉得无所谓
如果没有什么特别的信仰的话。

fh2315: "Not a big deal if the minority is not religious."

作者: claetitia　时间: 2010-5-18 15:49
告诉你吧加分什么的你要是瞧不上的话
至少少数民族还可以生2个孩子

claetia: "Well, if you despise the ethnic favor for extra scores, minorities can at least have more children!"

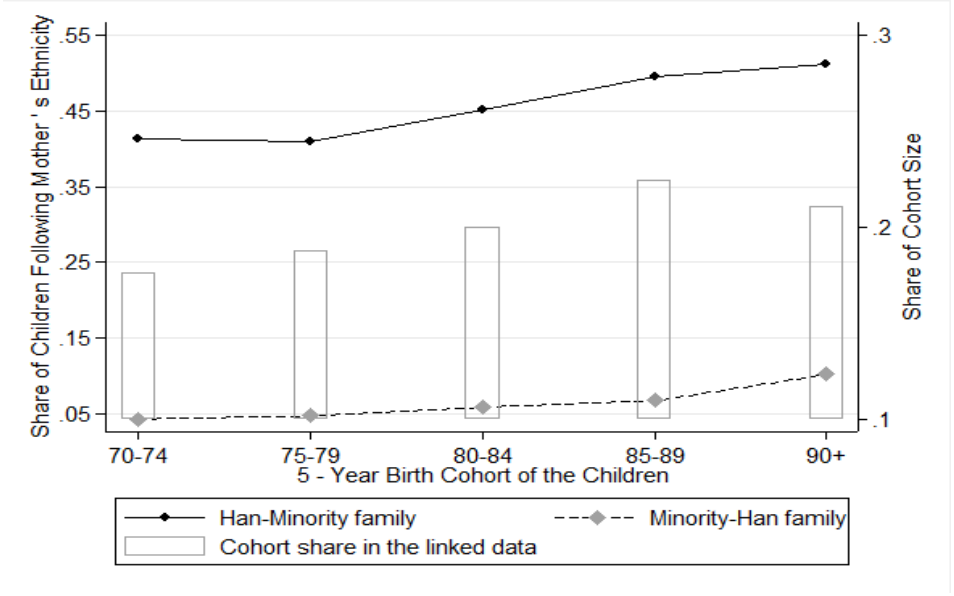
作者: 麻爪　时间: 2010-5-19 00:04
我是少数民族，女儿随我，原因很简单，姓随爸爸，民族就随妈妈吧，虽然是人数多的少数民族但我还是以自己的民族为骄傲的，所以希望女儿也是，与加分无关

Magua: "I am a minority and my child follows my ethnicity. The reason is simple…Even though I belong to a minority group whose population size is large, I am proud of my ethnicity. So I hope that my child is also [proud of my ethnicity]. This has nothing to do with extra scores.

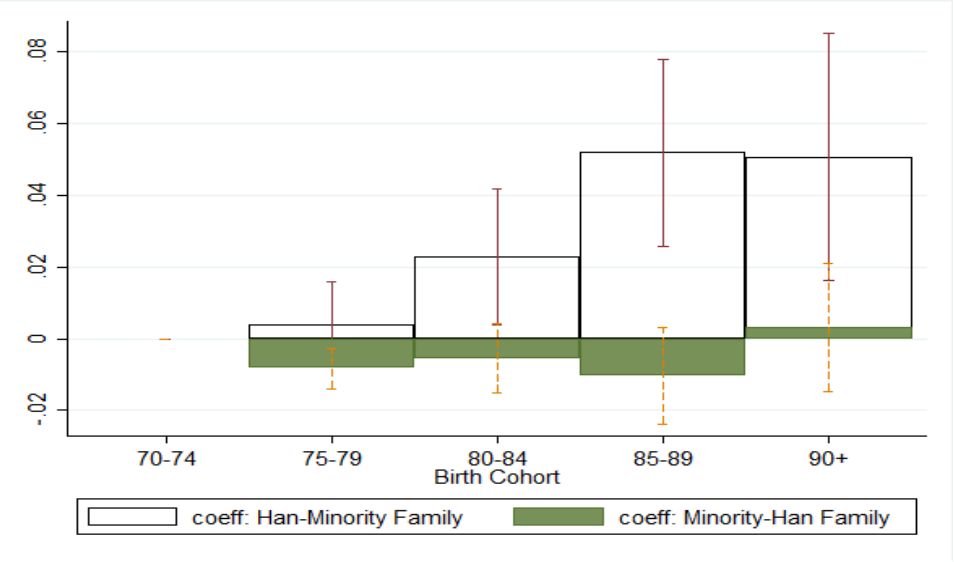
Notes: This discussion comes from http://jzb.com/bbs/thread-335421-1-1.html?action=printable, which shows that both honor and stigma are discussed in making the ethnic choices for the children.

Figure 1 Share of Children Following Mother's Ethnicity by Type of Mixed Marriage and Birth Cohort

(a) Aggregate Data: Share of Children



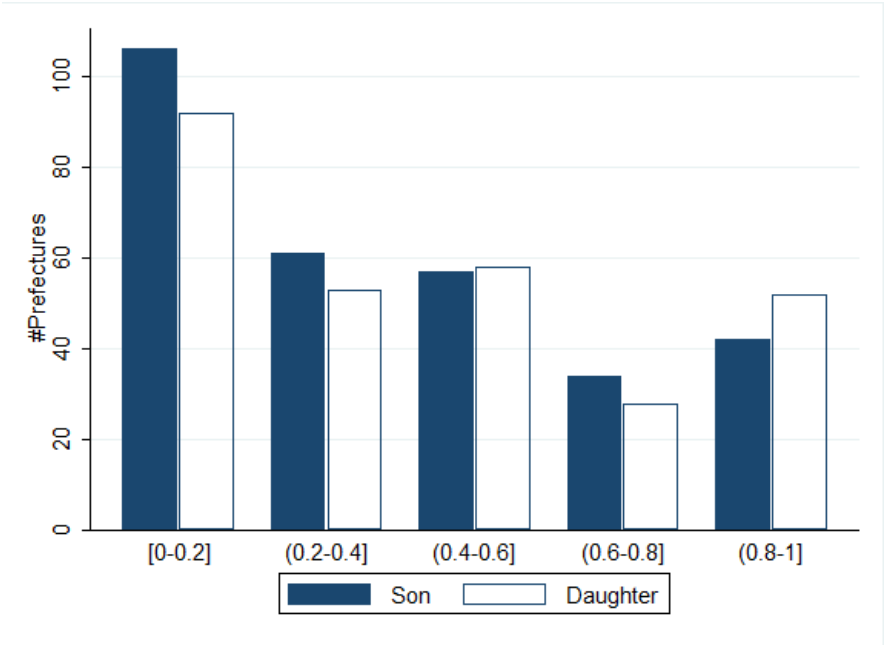
(b) Individual Data: Probability of Following Mother's Ethnicity



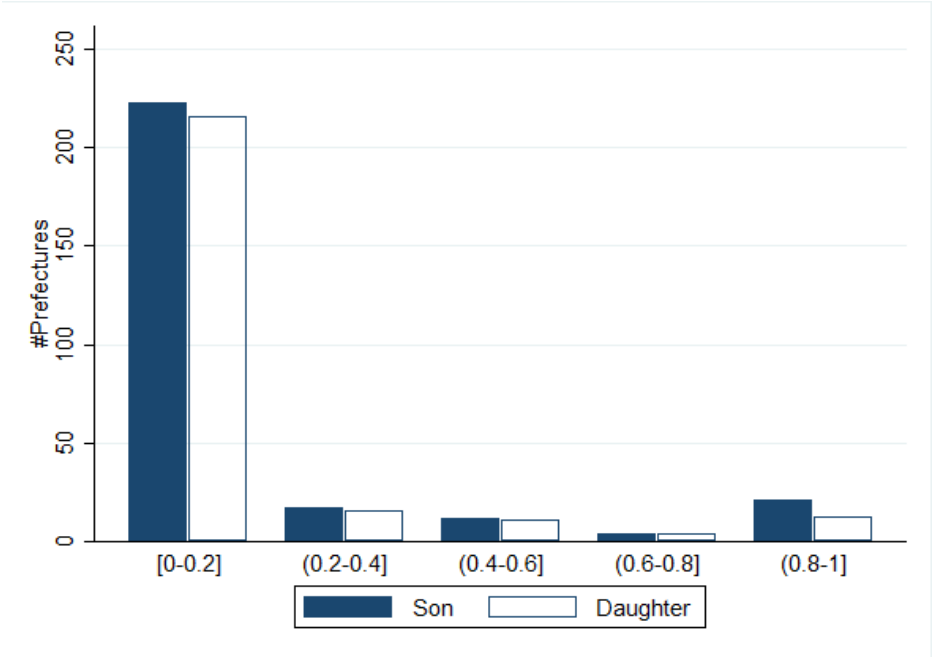
Notes: This figure shows two facts using aggregate and individual data: As in **F1,** children are more likely to follow their mother's ethnicity in Han-Minority families; as in **F2,** an increasing share of children following their mother's ethnicity in Han-Minority families after 1980. Figure (b) visualizes the results in columns (3) and (6) of Appendix Table A.2. It shows the probability of having a minority child in two types of mixed marriages over time, using those born during 1970-74 as the comparison group. The bars indicate 95% confidence intervals.

Figure 3 Distribution of the Share of Children Following Mother's Ethnicity in Mixed Marriages across Marriage Type and Prefectures (for those born in 1970-74)

(a) Share of Children with Mother's (Minority) Ethnicity in Han-Minority Marriages

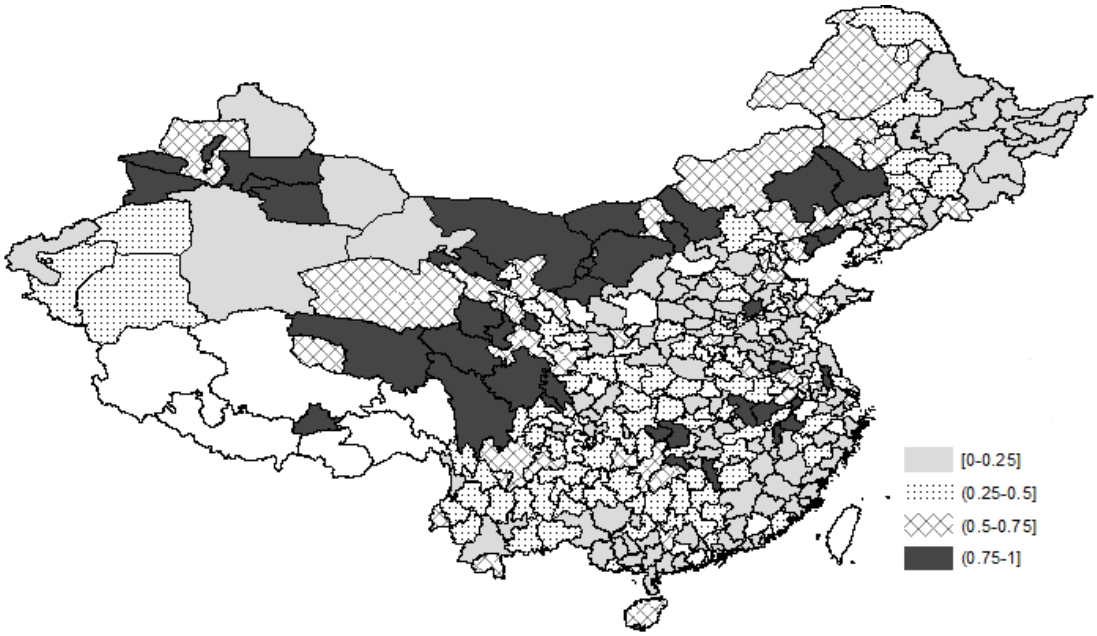


(b) Share of Children with Mother's (Han) Ethnicity in Minority-Han Marriages



Notes: Panel (a) shows that the share of children following mother's ethnicity varies a great deal across regions for Han-minority marriages and that sons are slightly less likely to do so. Panel (b) shows that the children of minority-Han couples in most prefectures almost never follow their mother's ethnicity. Our empirical analysis focuses on Han-minority families.

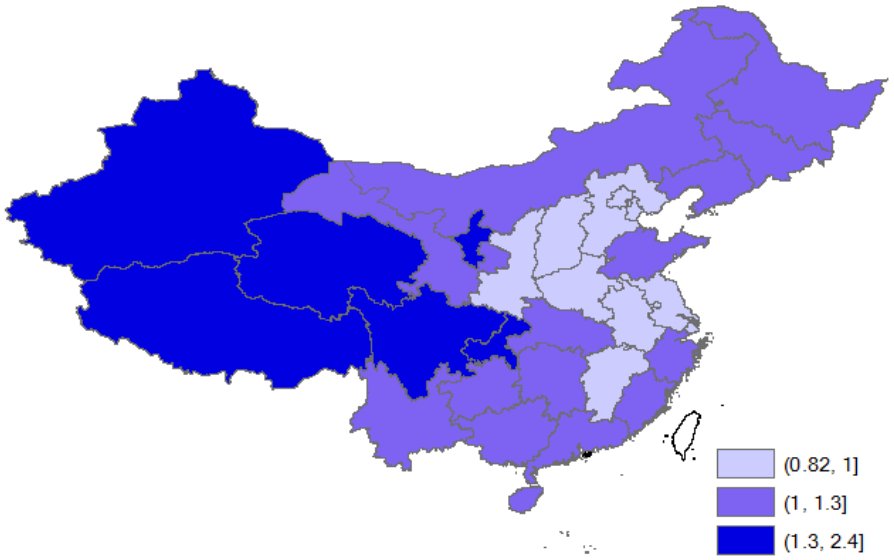
Figure 4 Spatial Variation in the Share of Children with Mother's Ethnicity in Han-minority Marriages (for those born in 1970-74)



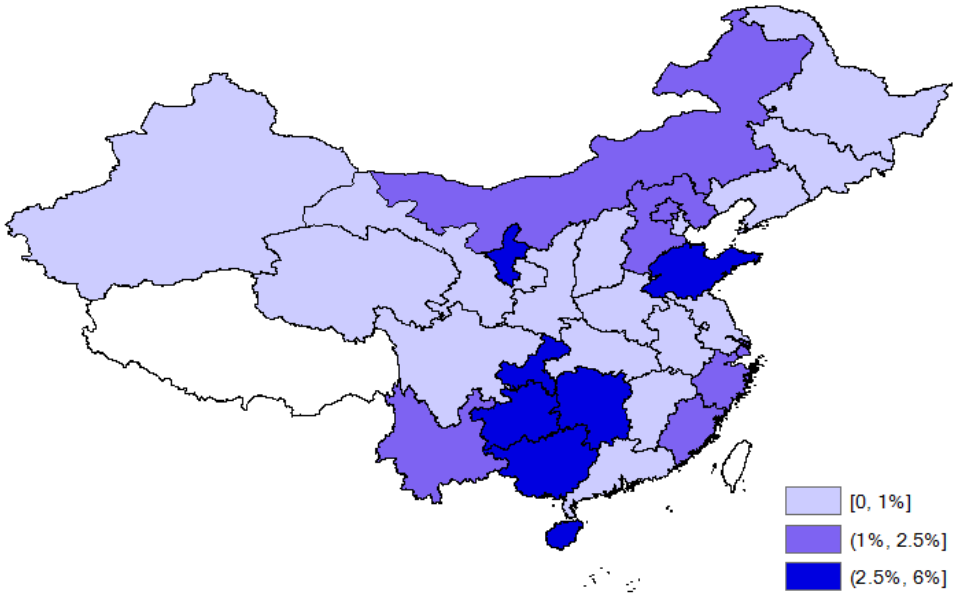
Notes: This figure maps the average probability of children following mother's ethnicity born during 1970-74 in Han-minority families across prefectures. A set of province fixed effects explains only about 36% of the variation across prefectures. Our empirical analysis exploits only within-province variation. We also control for province-by-year fixed effects to allow for flexible (non-parametric) time trends across provinces.

Figure A2 Spatial Variation in Ethnic Policies

(a) Total fertility ratio between Minority and Han women born in 1955-59

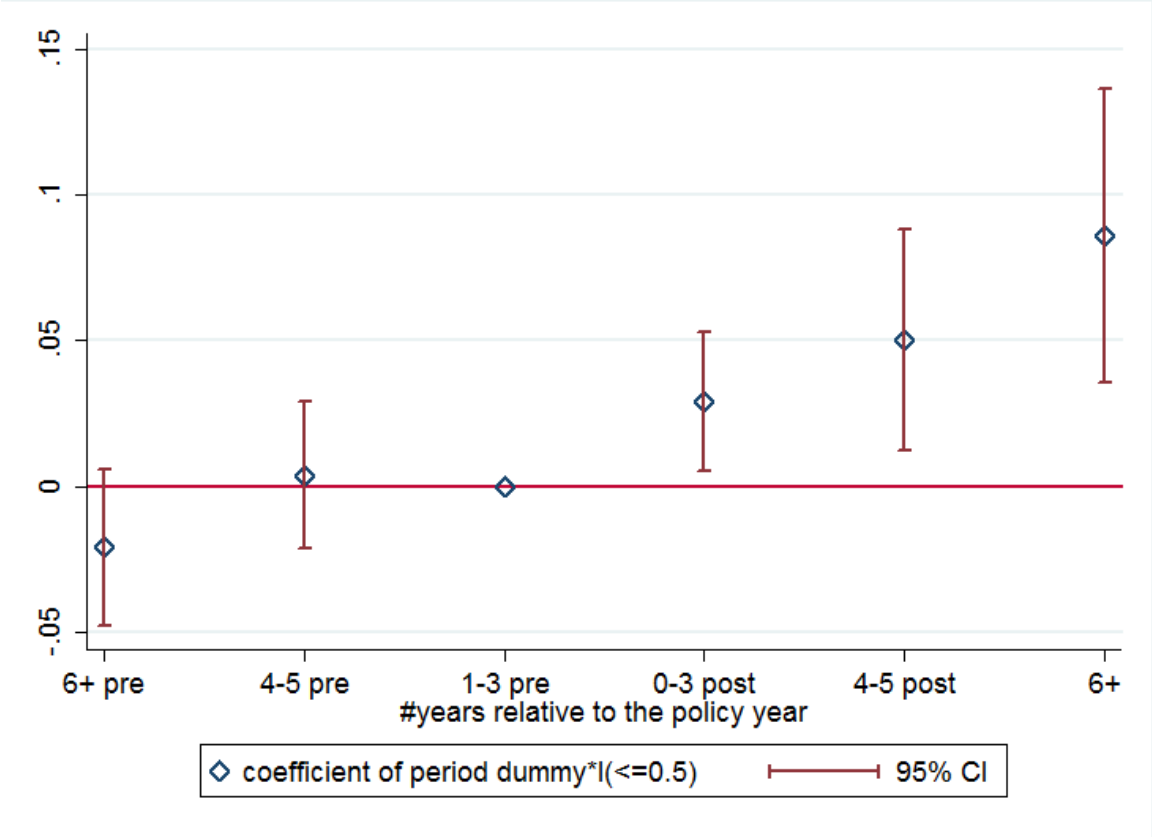


(b) Extra scores (relative to provincial cutoff) for minorities in 2000



Notes: The two figures present the cross-sectional variation in our measures of ethnic policies: extra fertility and extra scores. They show that the two types of benefits are not closely correlated at the cross-sectional level, with an insignificant correlational coefficient of 0.06. The data for extra scores in Tibet are not available.

Figure 5: The Dynamic Effects of Material Benefits (b) * Social Motives on the Probability of Mixed-Marriage Children Following Mother's Ethnicity



Notes: These figures plot the dynamic impacts of material benefits (measured by the rollout of family-planning policy) interacting with social motives (measured by the 1970-74 share of children following mother's ethnicity falling below a 0.5 cutoff). The reference period is 1-3 years before the introduction of the policy . The dimonds indicate the estimates in column (1) of Table A6 and the bars indicate 95% confidence intervals. The figure shows that (1) there are no similar positive effects before the policy, and (2) the effect is increasing over time, which is consistent with the dynamic extension of the model in Appendix A1.

Table 2A Baseline Results for P1: The Impact of Material Benefits (b) and Social Motives on the Probability of Mixed-marriage Children Following Mother's Ethnicity (Results using pre-post family planning policy to proxy material benefits *b*)

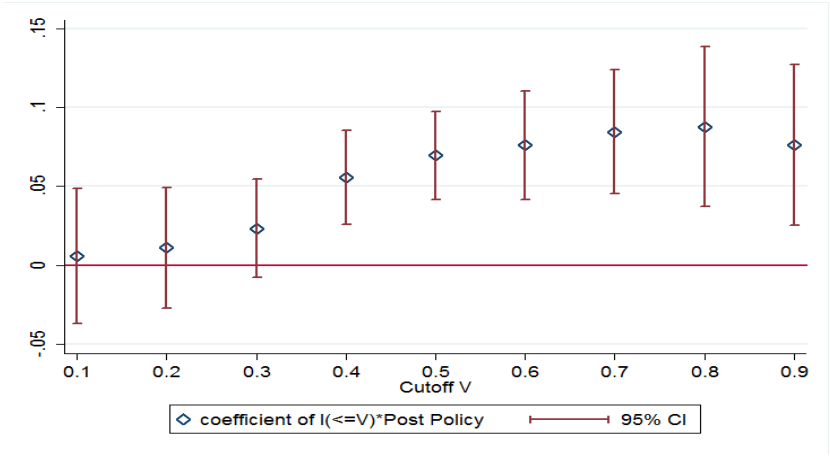
Dependent Variable: Child Following Mother's Ethnicity=0/1 (mean: 0.47)

	(1)	(2)	(3)	(4)	(5)	(6)
I(≤0.5)**b*(Post Policy)		0.072***	0.069***	0.071***	0.074***	0.070***
		(0.019)	(0.019)	(0.020)	(0.022)	(0.014)
		[0.030]	[0.031]	[0.031]	[0.027]	[0.024]
b(Post Policy)	0.078***	0.031**	0.035**			
	(0.011)	(0.014)	(0.013)			
	[0.019]	[0.014]	[0.013]			
Prefecture FE	Y	Y	Y	Y	Y	Y
Wife Ethnicity FE			Y	Y	Y	Y
Birth Year FE				Y	Y	Y
Controls**b*					Y	Y
Province FE*Year FE						Y
Observations	121,908	121,908	121,908	121,908	108,914	108,914
R-squared	0.276	0.277	0.290	0.292	0.299	0.334

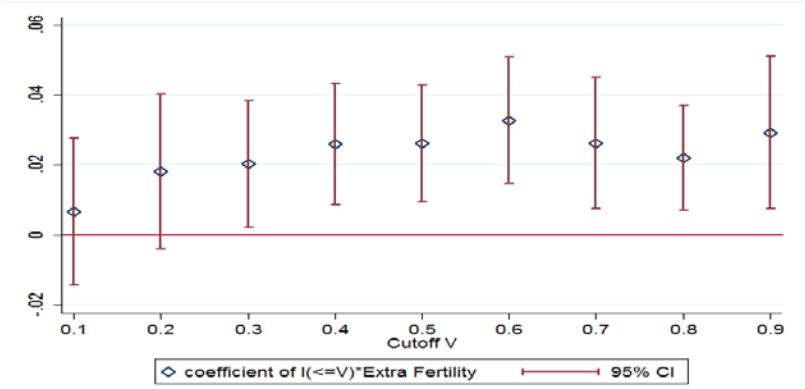
Notes: This table shows the results using the provincial roll-out of family-planning polices to measure material benefits. The cutoff (0.5) is defined by the share of minority children in Han-Minority families in the 1970-74 birth cohort. Controls include couples' charateristics (education-level fixed effects and 5-year birth-cohort fixed effects, for both husband and wife) and prefecture characteristics (listed in panel (d) of Table 1). The data come from three censuses and a mini census from 1982-2005. Standard errors in parentheses are clustered at the prefecture level, while those in brackets are clustered at the province level. Significance: ***, 1%, **, 5%, *, 10%.

Figure A3 Results Using Different Cutoffs

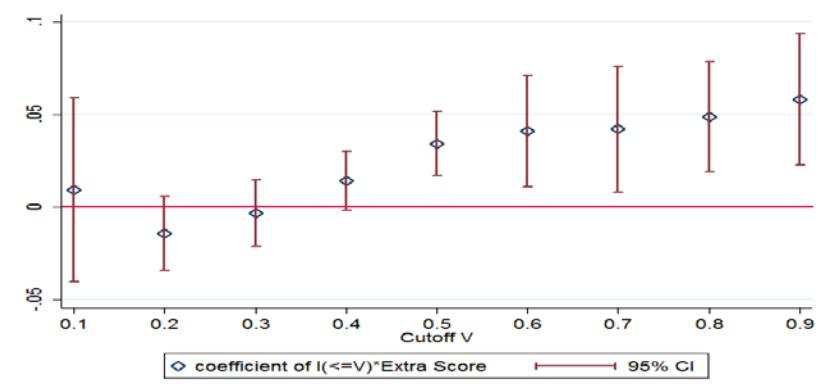
(a) Coefficient of I(<=V)*b(Post Policy)



(b) Coefficient of I(<=V)*b(Extra Fertility)



(c) Coefficient of I(<=V)*b(Extra Score)



Notes: This figure plots the results for testing prediction P1 while using different cutoff values for the share of minority children, ranging from 0.1 to 0.9. The econometric specification is the same as that in column (6) of Table 2A. The diamonds indicate the estimates and the bars through each dot indicates 95% confidence intervals.

Table 2B Baseline Results for P1: The Impact of Material Benefits (*b*) and Social Motives on the Probability of Mixed-marriage Children Following Mother's Ethnicity (with additional fertlity and additional exam scores for minorities proxying material benefits *b*)

Dependent Variable: Child Following Mother's Ethnicity=0/1 (mean: 0.47)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
I(≤0.5)* *b* (Extra Fertility)		0.027***	0.027***				0.022**
		(0.009)	(0.009)				(0.009)
b (Extra Fertility)	0.034***						
	(0.005)						
I(≤0.5)* *b* (Extra Scores)					0.052***	0.034***	0.022***
					(0.009)	(0.009)	(0.008)
b (Extra Scores)				0.044***			
				(0.007)			
Prefecture FE	Y	Y	Y	Y	Y	Y	Y
Wife Ethnicity FE		Y	Y		Y	Y	Y
Birth Year FE		Y	Y		Y	Y	Y
Controls*b*			Y			Y	Y
Province FE*Year FE			Y			Y	Y
Observations	107,903	107,903	96,874	124,938	124,938	111,944	96,873
R-squared	0.277	0.295	0.341	0.269	0.287	0.329	0.342

Notes: This table reports the results using two proxies for individual material incentives: extra fertility for minority in the previous birth cohort and extra scores for minority (relative to the provincial college acceptance cutoffs) in the college entrance exam. Both proxies are standardized. The cutoff is defined by the share of minority children in Han-Minority families in the cohort of 1970-74 births. Controls include couples' charateristics (education level fixed effects and 5-year birth-cohort fixed effects, for both husband and wife) and prefecture characteristics (listed in panel (d) of Table 1). The data come from three censuses and a mini census from 1982-2005. Standard errors are clustered at the prefecture level. Significance: ***, 1%, **, 5%, *, 10%.

Table 3 Results for P1': The Interaction Effects by Quartiles on the Probability of Mixed-marriage Children Following Mother's Ethnicity

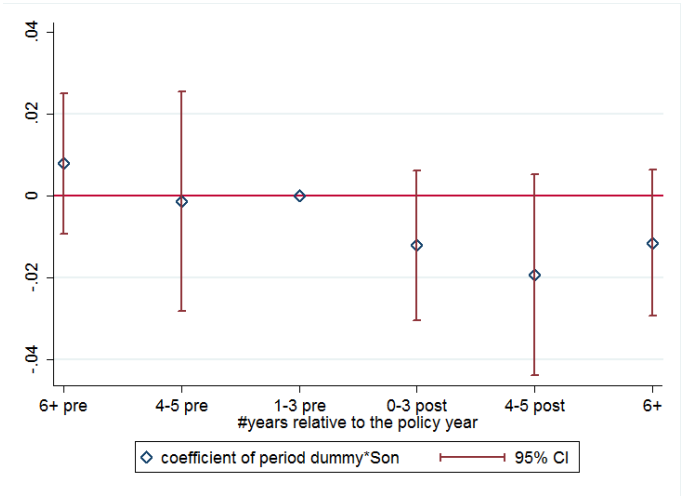
Dependent Variable: Child Following Mother's Ethnicity=0/1 (mean: 0.47)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Rollout of Family Planning			Extra Fertility for Minority			Extra Exam Scores for Minority		
Proxy for material benefits *b*	Post Policy = 0/1			(lagged cohort)			(share of cutoff score)		
I(0-0.25)*b	0.097***	0.099***	0.094***	0.030**	0.036***	0.032***	0.052***	0.058***	0.040**
	(0.024)	(0.023)	(0.029)	(0.014)	(0.013)	(0.011)	(0.016)	(0.017)	(0.016)
I(0.25-0.5)*b	0.145***	0.144***	0.119***	0.044**	0.041***	0.032***	0.077***	0.088***	0.063***
	(0.027)	(0.026)	(0.027)	(0.017)	(0.013)	(0.009)	(0.015)	(0.016)	(0.014)
I(0.5-0.75)*b	0.079***	0.080***	0.052*	0.032***	0.025**	0.009	0.030	0.034**	0.029*
	(0.025)	(0.024)	(0.027)	(0.009)	(0.011)	(0.011)	(0.018)	(0.017)	(0.016)
Prefecture FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Wife Ethnicity FE		Y	Y		Y	Y		Y	Y
Birth Year FE		Y	Y		Y	Y		Y	Y
Controls*b			Y			Y			Y
Province FE * Year FE			Y			Y			Y
p-value	0.014	0.019	0.000	0.514	0.219	0.048	0.003	0.000	0.0003
Observations	121,908	121,908	108,914	107,903	107,903	96,874	124,938	124,938	111,944
R-squared	0.278	0.293	0.334	0.277	0.296	0.341	0.270	0.288	0.329

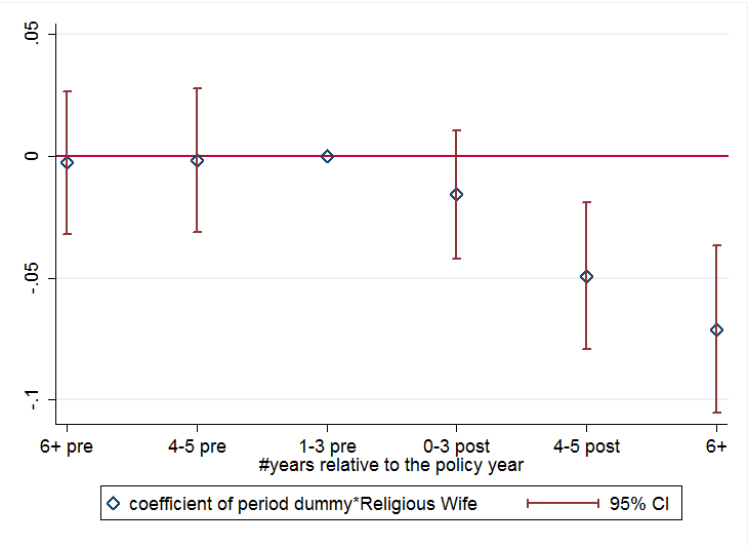
Notes: According to Prediction P1', the interaction effects of material benefits and social motives estimated for the first three quartiles should be larger than that for the fourth quartile. Further, the effect for the second quartile should be larger than that for the third quartile – the *p*-values refer to tests for a difference between the effects in the second and third quartiles. Controls include couples' charateristics (education-level fixed effects and 5-year birth-cohort fixed effects, for both husband and wife) and prefecture characteristics (listed in panel (d) of Table 1). Standard errors are clustered at the prefecture evel. Significance: ***, 1%, **, 5%, *, 10%.

Figure 6: The Dynamic Effects of Material Benefits (*b*)*Intrinsic Costs (*e*) on the Probablity of Mixed-Marriage Children Following Mother's Ethnicity

(a) The Effect of Material Benefits (*b*)*Son (vs. Daughter)



(b) The Effect of Material Benefits (*b*)* Religious (vs. Non-religious) Wife



Notes: These figures plot the dynamic impacts of material benefits (measured by the rollout of family planning policy) interacted with intrinsic costs (proxied by son and religious wife). The reference period is 1-3 years before the policy. The diamonds indicate the estimates in columns (3) and (5) of Table A6 and the bars indicate 95% confidence intervals. The figures show that there are no systematic trends before the introduction of the policy, and that the effects are consistent with model Prediction P2.

Table 6 Results for P2: The Interaction of Material Benefits (*b*) and Intrinsic Costs (*e*) on the Probability of Mixed-marriage Children Following Mother's Ethnicity

Dependent Variable: Child Following Mother's Ethnicity=0/1

	(1)	(2)	(3)	(4)	(5)	(6)
b(Post Policy)*Son	-0.017***	-0.008				
	(0.005)	(0.005)				
b(Extra Fertility)*Son			-0.004***	-0.003**		
			(0.001)	(0.002)		
b(Extra Score)*Son					0.001	0.002
					(0.002)	(0.002)
Son	-0.000	-0.009**	-0.007**	-0.011***	-0.011***	-0.015***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Prefecture FE	Y	Y	Y	Y	Y	Y
Wife Ethn. FE		Y		Y		Y
Birth Year FE		Y		Y		Y
Controls*b		Y		Y		Y
Province FE*Year FE		Y		Y		Y
Observations	122,835	109,250	108,528	97,100	122,803	109,227
R-squared	0.278	0.334	0.278	0.341	0.276	0.334
	(7)	(8)	(9)	(10)	(11)	(12)
b(Post Policy)*Religious Wife	-0.044***	-0.026**				
	(0.015)	(0.013)				
b(Extra Fertility)* Relig. Wife			-0.016***	-0.011***		
			(0.004)	(0.004)		
b(Extra Score)*Relig. Wife					-0.027***	-0.047***
					(0.008)	(0.008)
Religious Wife	0.204	0.127	0.072***	0.129	0.071***	0.130
	(0.206)	(0.283)	(0.013)	(0.276)	(0.013)	(0.283)
Prefecture FE	Y	Y	Y	Y	Y	Y
Wife Ethn. FE		Y		Y		Y
Birth Year FE		Y		Y		Y
Controls*b		Y		Y		Y
Province FE*Year FE		Y		Y		Y
Observations	122,835	109,250	108,528	97,100	122,803	109,227
R-squared	0.293	0.334	0.280	0.341	0.277	0.335

Notes: This table shows that the effect of the policy change tends to be smaller when the child is a son or when the wife belongs to a religious ethnicity, consistent with Prediction P2 of the model. Controls include couples' charateristics (education level fixed effects and 5-year birth-cohort fixed effects, for both husband and wife) and prefecture characteristics (listed in panel (d) of Table 1). Standard errors are clustered at the prefecture level. Significance: ***, 1%, **, 5%, *, 10%

Table 8 Alternative Explanation: Censoring (Shares between 0.3 and 0.7 only)

Dependent Variable: Mixed-marriage Child Following Mother's Ethnicity=0/1

	(1)	(2)	(3)	(4)	(5)	(6)
I(≤0.5)*b(Post Policy)		0.076***				
		(0.017)				
b(Post Policy)	0.108***					
	(0.021)					
I(≤0.5)*b(Extra Fertility)				0.022		
				(0.014)		
b(Extra Fertility)			0.042***			
			(0.010)			
I(≤0.5)*b(Extra Score)						0.031***
						(0.010)
b(Extra Score)					0.051***	
					(0.010)	
Prefecture FE	Y	Y	Y	Y	Y	Y
Wife Ethnicity FE		Y		Y		Y
Birth Year FE		Y		Y		Y
Controls*b		Y		Y		Y
Province FE*Year FE		Y		Y		Y
Observations	54,345	48,480	47,286	42,258	54,345	48,480
R-squared	0.093	0.195	0.084	0.200	0.088	0.195

Notes: This table shows the baseline results on a sample restricting the share of Mixed-marriage children following mother's ethnicity for those born during 1970-74 lies between 0.3 and 0.7. It shows that censoring of the room for change should not be a critical concern. Controls include couples' charateristics (education level fixed effects and 5-year birth-cohort fixed effects, for both husband and wife) and prefecture characteristics (listed in panel (d) of Table 1). Standard errors are clustered at the prefecture level. Significance: ***, 1%, *, 5%, *, 10%.

Table A1 Fact F1: HM-Families versus MH-Families

Dependent Variable: Following Mother's Ethnicity = 0/1

	(1)	(2)	(3)	(4)
HM-Marriage	0.475***	0.447***	0.448***	0.449***
	(0.028)	(0.028)	(0.028)	(0.028)
Prefecture FE		Y	Y	Y
Birth Year FE			Y	Y
Provincial Trends				Y
Observations	235,930	235,930	235,930	235,930
R-squared	0.260	0.370	0.371	0.382

Notes: This table shows that fact F1 in Figure 1 also holds at the individual level. Provincial trends indicate provincial-birth year linear trends. Standard errors are clustered at the prefecture level. Significance: ***, 1%, **, 5%, *, 10%.

Table A2 Fact F2: Ethnicity of Children by Cohorts

Dependent Variable: Following Mother's Ethnicity = 0/1

	(1)	(2)	(3)	(4)	(5)	(6)
Born 1975-79	-0.002	0.017***	0.003	0.004*	0.002	-0.008***
	(0.009)	(0.005)	(0.006)	(0.002)	(0.002)	(0.003)
Born 1980-84	0.040**	0.048***	0.020**	0.016***	0.015***	-0.005
	(0.015)	(0.008)	(0.010)	(0.003)	(0.003)	(0.005)
Born 1985-90	0.086***	0.089***	0.048***	0.024***	0.020***	-0.010
	(0.017)	(0.011)	(0.013)	(0.004)	(0.004)	(0.007)
Born 1990+	0.108***	0.109***	0.047***	0.059***	0.047***	0.003
	(0.024)	(0.015)	(0.018)	(0.006)	(0.005)	(0.009)
Prefecture FE		Y	Y		Y	Y
Provincial Trends			Y			Y
Observations	124,940	124,940	124,940	110,020	110,020	110,020
R-squared	0.008	0.272	0.277	0.007	0.082	0.086

Notes: This table shows that fact F2 in Figure 1 also holds at the individual level. Provincial trends indicate provincial-birth year linear trends. Standard errors are clustered at the prefecture level. Significance: ***, 1%, **, 5%, *, 10%.

Table A3 Differences across Marriages

	HH	MM	HM	MH	
#Couples	6436486	417089	90704	81570	
Share in total marriages	91.60%	5.90%	1.30%	1.20%	
HM Share for a minority woman					1.3/(1.3+5.9)=18%
MH Share or a minority man					1.2/(1.2+5.9)=17%
Husband Edu-Wife Edu	0.27	0.26	0.21	0.23	
Husband Age-Wife Age	2.41	2.72	2.8	2.48	

Notes: This table describes the marriage patterns among all married couples in the four censuses (1982, 1990, 2000 and 2005). This sample includes all the couples in the data, while our analysis on mixed marriages focuses on those with children born between 1970 and 2005.

Table A4. Correlation of Prefecture Characteristics

	Mother Ethnicity Prob.	Minority Pop Share	High school	Borderland	#Children for Minority
Prob. of Children following mother's ethnicity in HM marriages (1970-74 cohort)	1				
Minority Pop Share 1982	0.24	1			
High-school edu.+ Pop Share 1982	0.04	-0.26	1		
Borderland	-0.14	0.15	0.05	1	
# Children for Minority Women (aged 40+)	0.11	0.25	-0.30	-0.01	1

Notes: This table presents bivariate correlations between prefecture characteristics. In particular, the probability of having a minority child in HM families is positively correlated with the minority-population share. This correlation rejects a "scarcity" effect, whereby children are less likely to be minority in regions with a higher share of minority population because a more or less fixed set of material benefits get diluted by population. This is consistent with the fact that ethnic policies are not set via a fixed quota system.

IV. THE EXPRESSIVE FUNCTION OF LAW

- Large (informal) literature arguing that laws have a dual role:
  - Not just a menu with "prices" for good or bad behaviors
  - Also express society's values: what it approves of or chooses to punish, how it chooses to punish; this expressive function is important
  - Injunctive vs. descriptive norm
- Expressive considerations used to argue for both
  - Tougher laws (even inefficiently so), e.g. prison vs. fines, reparations, or community service
  - Gentler hand, e.g. limiting severity of sanctions: length of sentences, corporal punishments, torture, shaming, death penalty
- Other examples
  - Prohibition / legalization of flag burning
  - Gay marriage vs. equivalent civil union. Earlier: Georgia's anti-sodomy law, unenforced but remained on the books; antimiscegenation laws
  - No price / market for organs, adoption, "repugnant transactions"

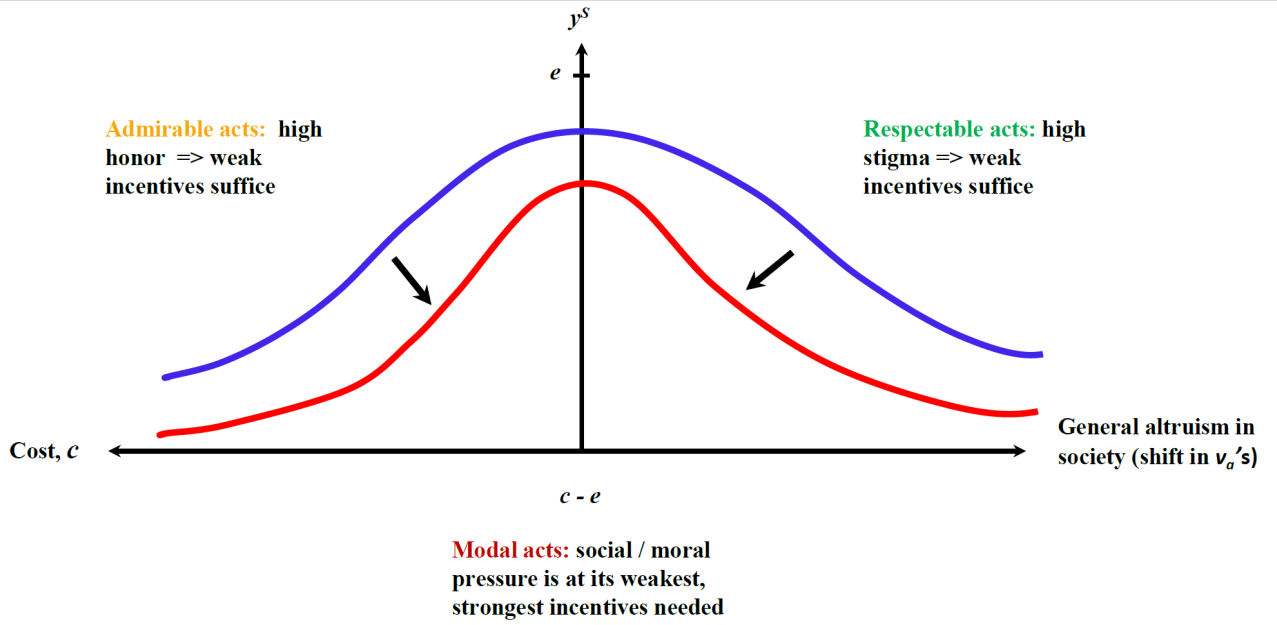
Modeling expressive law: asymmetric information

- Planner, legislator, has information on average preference of society, or "community standards": distribution of underlying values v_a's
  - Has observed behavior of a representative sample: $\bar{a}$ = past tax compliance, polluting activities, drug use, black market for X...
- Choice of law, incentives, will then inevitably convey message about it
  - Indeed, saw earlier that optimal y^s depends on those "societal values"
- Individuals in society only have broad sense of whether behavior in question is rare and admirable (honor-driven), or common and merely respectable (stigma-driven)
- Observing the laws, incentives that are set, make inferences about "what kind of a society" (or peer group, company,etc.) they live in
- This affects the norm: what they expect others do, and how they expect to be judged by them
  - Indeed, saw earlier that strength of honor and stigma depend on the underlying distribution of people's "values" (the v_a's)

Proposition (law expressing societal standards)

Whether the prosocial action is common or rare,

1. *Principal always sets weaker incentives when has private information about population's compliance, or strength of norms:* $y^{AI} < y^{FI}$.
2. *As a result, participation / compliance is lower than under full information. But principal economizes on costly incentives (fines, subsidies, enforcement) through such signaling.*



Blue: symmetric information between incentive designer and individuals

Red: designer has more information than individuals about earlier compliance or general degree of prosocial orientation in the population

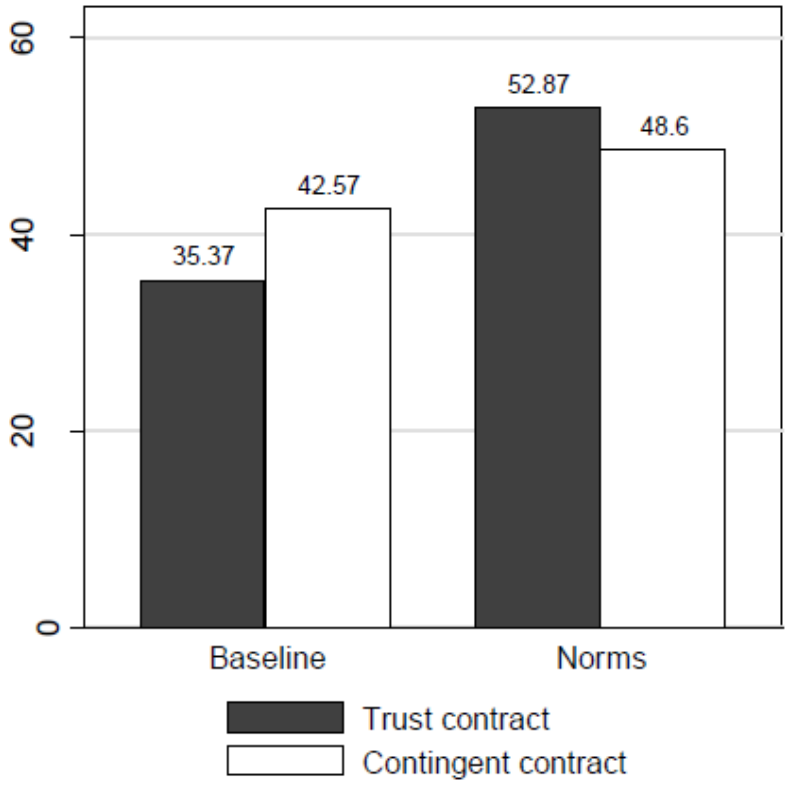
Danilov and Sliwka (2013) "Can Contracts Signal Social Norms?"

- Agent chooses "effort" $a \in [0, 100]$, at cost $C(a) = a^2/2$.
- Principal ("employer") earns 12 Euros with probability a, otherwise 0
- Principal chooses between:
  - "Trust contract": unconditional wage of 5 Euros
  - "Contingent" or incentive contract: agent gets bonus $b = 5$ Euros if, and only if Principal does receive the 12 Euros
- Agent's efforts elicited for both contracts (strategy method)
- Two informational conditions, payoffs unchanged:
  - "Baseline" = uninformed Principal: as described above
  - "Norms" = informed Principal: before choosing contract, Principal sees decisions taken
    by 10 agents from previous baseline condition.

    Agent knows Principal selecting his contract has seen such information.

Agents' effort behavior

- Agents respond to Principal's choice of Trust contrast with more (voluntary) effort only when the latter acts with knowledge of earlier play $\Rightarrow$reveals good news about the norm

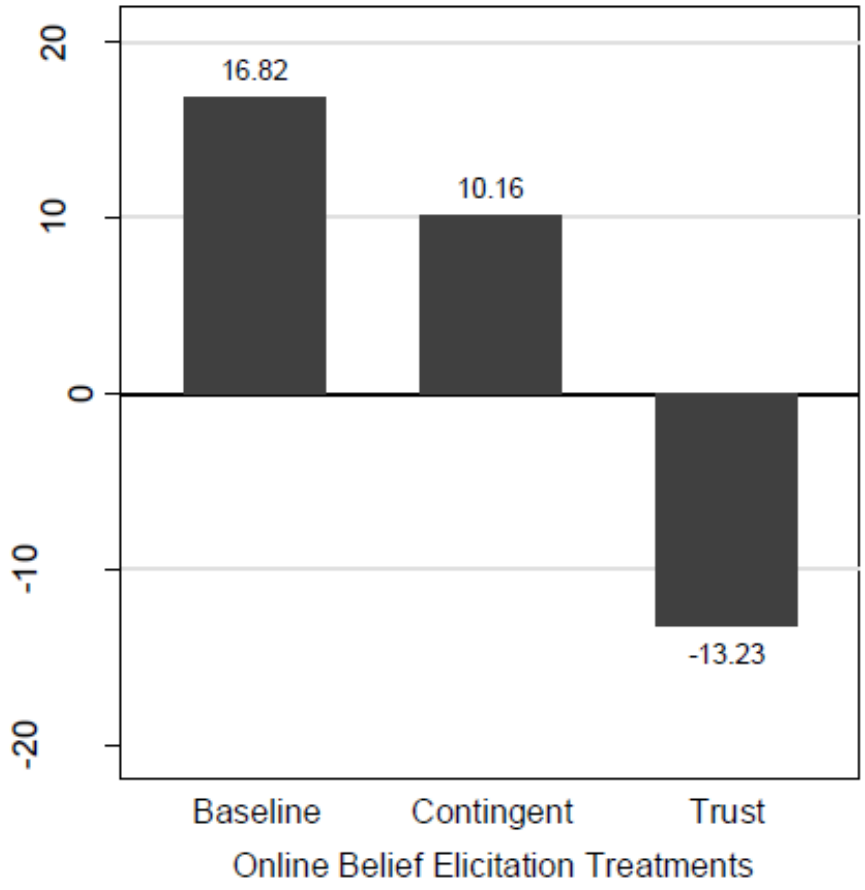


Average Effort for the Trust and Contingent Contracts

- No significant $\neq$for contingent contract

Inferences made from about principal's contract choice

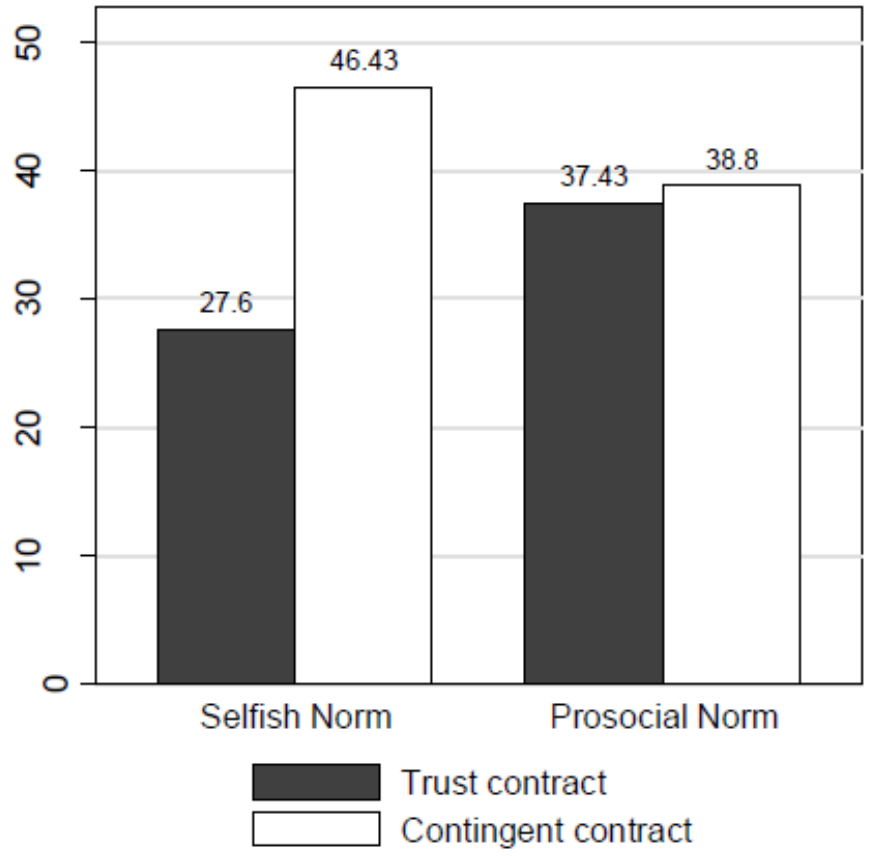
- Outside observers correctly predict that Principals who chose Trust contract must have seen higher effort in a previous group that worked under Trust than in (another) that worked under Contingent. Thus, they understand that choice of Trust by informed principal means that the latter observed good news about the norm



Average Difference in Estimated Efforts

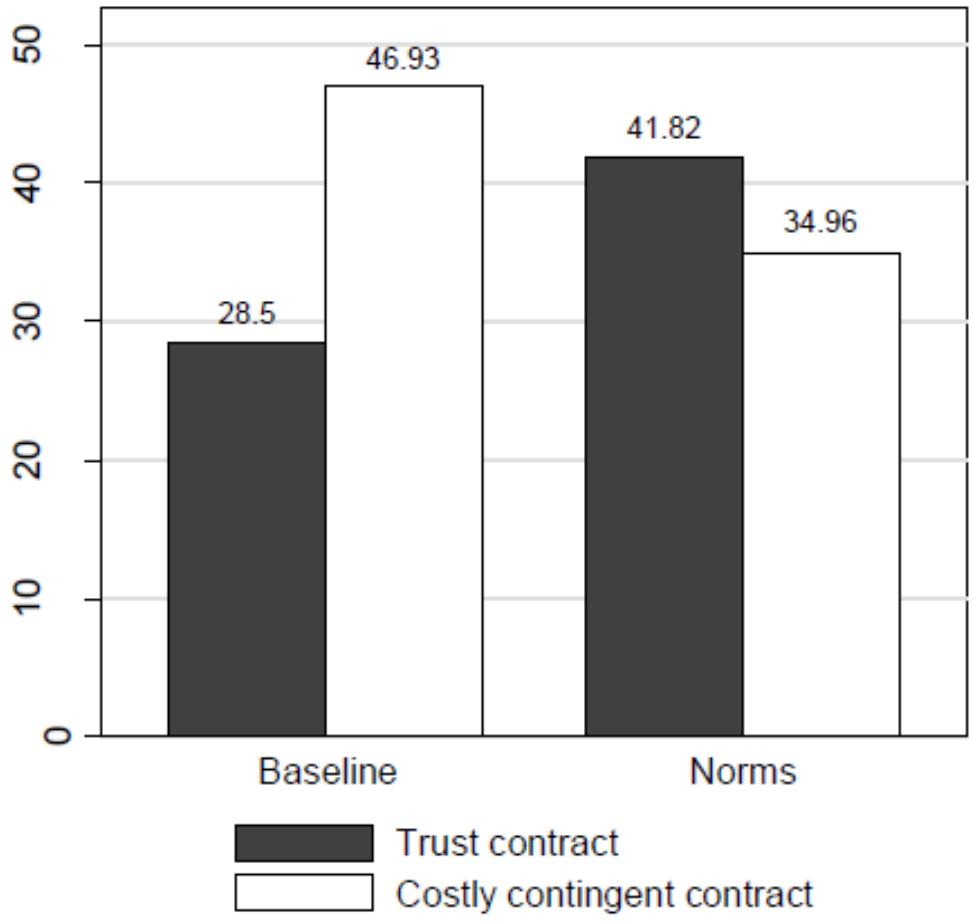
Effects of agents' beliefs about others

- Agents who saw play of an earlier group that had high effort, versus low effort, work more under the Trust contract: they saw good news about the norm directly
  - No significant difference under the contingent contract



Average Effort, "Induced Norms" Treatment

Varying the strength of the Principal's signal



Average Efforts, When Contingent Contract is Costly

- When Principal must pay to choose the contingent contract (making that choice a stronger signal), her doing so when informed of previous play now reduces their effort, whereas when she is uninformed it has the standard, positive effect

Spillovers across spheres of behavior

- Recall question. Now two activities, a and b, both 0 - 1 decisions,
- Informal interactions: individual's $a-$behavior is observed by other private citizens, but not by principal / gvt.
  - Cooperating, helping, public goods contributions, not rent-seeking
  - Activity done privately, observable but not verifiable by court:

$$y_a = 0, \quad m_a > 0$$

- Formal interactions: individual's $b-$behavior is observed by principal / gvt., but not by other private citizens
  - Transactions involving principal: paying / evading taxes, bureaucrats' honesty or corruption; employee productivity. Or, peers less able than principal to sort through excuses:

$$y_b = y > 0, \quad m_b = 0$$

- For simplicity, a person has same degree of prosociality in both activities: $v_a = v_b = v$ (just need correlated distributions)

The expressive spillovers of law

- Raising by \$1 the incentive rate y for realm-b behavior has same social benefit as before (induces some more compliance), but social cost now includes:
  - Standard: must pay that extra \$1 to all who were complying anyway
  - New: people infer that they face "worse" society, hence lower community standards and weaker social enforcement on realm-a behavior $\Rightarrow$ lower compliance

Proposition (expressive spillovers)

Let the norms-enforced behavior (a) be of a relatively common (hence, stigma-driven) nature:

1. *Principal always sets weaker incentives for the incentivized action (b) when has private information about population compliance, or average preferences: $y^{AI} < y^{FI}$*
2. *As a result, compliance on b is always lower than under full information; compliance on a is unchanged*

Why economists are unpopular

- Frequent resistance to economists' positive and normative messages about power of / need for incentives and markets
- "Putting a price on everything": valuing scare resources, versus stating bad news about human nature: low altruism v_a, high greed v_y

1. Society may just not want to hear bad news about itself.
   - Often does not. Wishful thinking about moral identity, just-world beliefs, groupthink, ideology... Saw a lot of evidence
2. Economists may be focussing more on b -type behaviors, where incentives are more easily implementable and social norms weak
   - Perhaps less attention to or data on a -type behaviors, in which incentives are unavailable and social norms are strong
   - Making salient a dim view of human nature, by stating or signaling that strong incentives are effective or needed in b, undermines the social norms in a. Increases need for incentives there, but might be less cost-effective at achieving compliance
3. Now, with theory in place: it becomes an empirical question

When expressiveness strengthens the law

- Planner, legislator, may be informed not (just) about how people tend to behave, i.e. compliance, but about the consequences this has, i.e. the importance of resulting externalities (e) :
  - How damaging are CO_2 emissions, how much good \$1 can do in poor countries, negative externalities from drunk driving, drugs, how important to firm is quality / customer service, etc.
- If people's intrinsic motivation is "consequentialist," principal will then want the law to signal that "this is really important"

Proposition (law expressing magnitude of externalities)

Whether the prosocial action is of common or rare:

1. *The principal always sets stronger incentives when has private information about the importance of externalities:* $y^{AI} > y^{FI}$
2. *As a result, participation / compliance is higher than under full information, and so are enforcement costs.*

Concluding Thoughts I

1. Laws and norms shape each other, and behavior
   - "Admirable" acts: few people do, honor motive important
     Material incentives $\rightsquigarrow$ partial or full crowding out
   - "Respectable" acts: most people do, stigma motive important
     Material incentives $\rightsquigarrow$ crowding in
   - Empirical predictions, wide variety of tests

2. Optimal incentives in the presence of norms
   - Social or self esteem is a positional good. Prosocial actions are inefficiently distorted toward the most visible
   - Optimal incentives (Pigou-Ramsey) adjusted by "reputation tax"

Concluding Thoughts II

- Expressive role weakens optimal incentives when they are informative about society's general "goodness": previous compliance, degree of prosocial orientation, etc.
- It strengthens them when they are informative about magnitude of externalities that people are intrinsically motivated to remedy

4. Expressive spillovers: what is signaled about "societal values" by law or incentives bearing on one activity carries over to people's attitudes and behavior in others $\Rightarrow$ affects norms $(+/-)$
   - Resistance to economic discourse, evidence on incentives: wishful thinking / willful blindness, or real adverse informational spillovers?
   - A lot of work nowadays on incentives/markets and moral attitudes, taboo tradeoffs, "repugnance," and policy implications
   - Also on interactions of social norms and (vanishing) privacy (Ali & Bénabou 2016)

5. On all these topics of Laws $\rightleftarrows$ Norms, progress is again via constant back-and-forth between formal theory, experiments and empirics