Social Preferences or Sacred Values? Theory and Evidence of Deontological Motivations

Daniel Chen & Martin Schonger

ETH

12 March 2014

www.nber.org/~dlchen
www.n.ethz.ch/~scmartin

Introduction

Generations of Social/Moral Preferences

Payoff = material consequences
Domain of preferences:

1. Agent's payoff (homo oeconomicus)
2. Agent's payoff and that of others (e.g. Fehr-Schmidt, pure altruism)
3. Agent's payoff, others' payoffs and social audience (e.g. impure altruism)
4. Agent's payoff, others' payoffs, social audience, and purely internal consequences (this paper)

Research Problem

Research question

"Do people have deontological (duty-based) motivations?"

Note: Revealed preference approach so deontological motivations is an interpretation.

Problem:

Consequentialist and deontological motivations are hard to distinguish in normal circumstances. We identify non-consequentialism by varying the probability of a decision being consequential.

Kant

Moral problem:

Your friend is hiding in your house from a murderer. The murderer arrives and asks you whether your friend is hiding in your house. Assuming you cannot stay silent, should you lie or tell the truth?

Kant (1793) says you must not lie. (Note: Kant considers uncertainty).

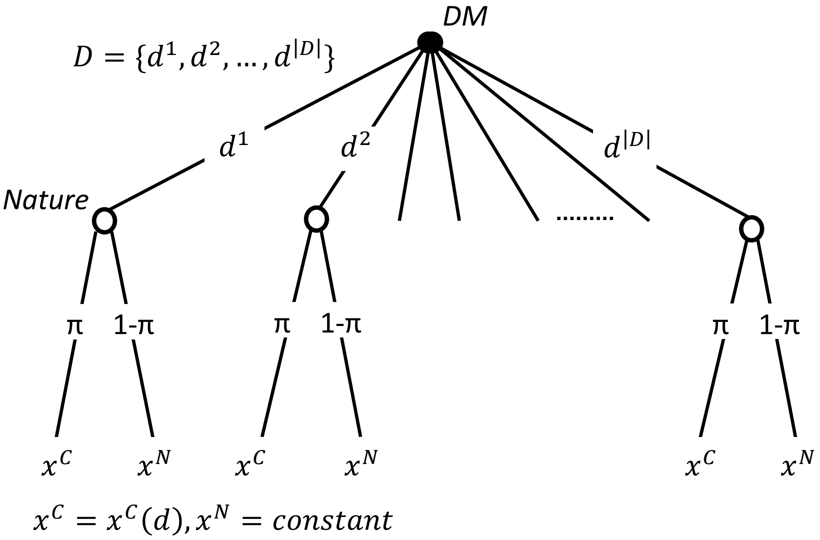
Note:

We are talking here about Kant as a non-consequentialist. The categorical imperative is about what the duties are.

Distinguishing decisions from broad consequences

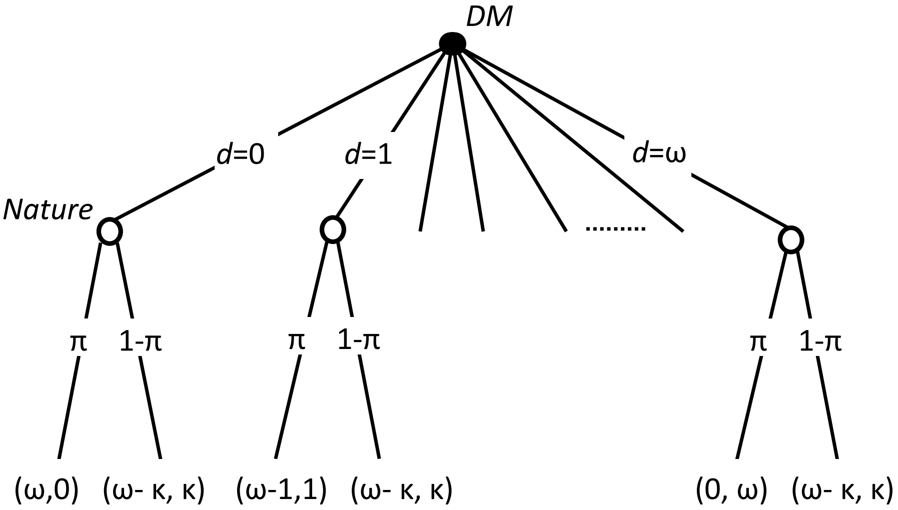
- Thought experiment to separate the decision/action d from the payoffs/prizes/consequences x
- X is the set of payoffs, a generic payoff is $x = (x_1, x_2)$
- consequentialism in a broad sense: anything that is a function of x could be an outcome (reputation, inferences by other player about $DM's$ intention, etc.); we basically implicitly fold these things into the preference over payoffs
- x is a function of the state of nature and decision d
- In state C d becomes common knowledge
- In state N d remains unknown to anyone except DM
- Consequentialism: Preferences are over lotteries defined on X
- Deontological motivations: d matters per se, that is even in state N

General Idea



Lab Implementation

Application: Duty to donate/share (DG)



Application: Duty to tell the truth (die rolling)

Instructions

You will later be asked to roll the die once, write the result on a piece of paper, place it inside an envelope, seal the envelope, and put the envelope on the table in the middle of the room. The experimenter will then spin the wheel of fortune once. Depending on the outcome of the wheel of fortune, your envelope will either be shredded, or privately opened by the experimenter. Here are the details:

If the wheel of fortune stops on 9:
Your envelope will be **privately opened** by the experimenter, who will later pay you privately according to your die outcome:

Die outcome	1	2	3	4	5	6
Resulting payoff	CHF 1.00	CHF 2.00	CHF 3.00	CHF 4.00	CHF 5.00	CHF 0.00

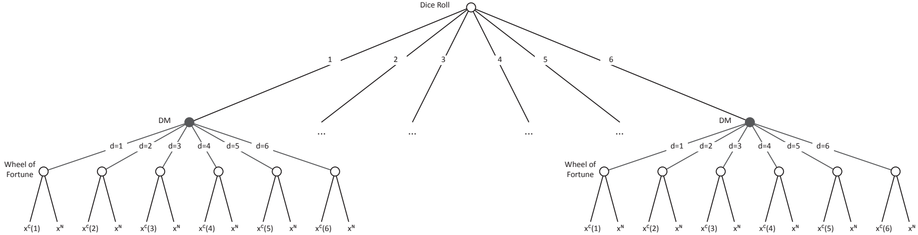
If the wheel of fortune stops on 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, or 16:
Your envelope will not be opened. Instead, it will be **shredded** in front of you. The experimenter will later pay you privately earnings of CHF 5.00.

Once you have read the instructions above, click "Next" to proceed to the understanding questions.

The instructions will remain visible on your screen.

Next

Application: Duty to tell the truth (die rolling)



Invariance

Theory

Goal:

Invariance Theorem.

Specifics:

	State	decision d is	Probability	Payoff vector
d	C	consequential	π	$x^C = x^C(d)$
	N	not consequential	$1-\pi$	$x^N \perp d$

- Decision-maker *DM* takes decision d which determines x^C
- Consequences of the choice: x, and any function thereof (e.g. recipient's opinion about *DM* based on x_2)
- lottery p gives objective probabilities to each outcome x
- let P be the set of all lotteries, generic elements p, q
- set of outcomes X, generic elements x,x',x''

Assumption: Preference relation

DM has a preference relation $\succsim$ over the lotteries on X.

Note: this implicitly assumes consequentialism.

First-Order Stochastic Dominance

Definition

(FOSD) p first-order stochastically dominates q with respect to $\succsim$ if for all x': $\sum_{x:x'\succsim x} p(x) \leq \sum_{x:x'\succsim x} q(x)$.

Assumption: FOSD

If p FOSD q with respect to $\succsim$, then $p \succsim q$.

Note: Social preferences not monotone, so we define FOSD with respect to ordering induced by preferences.

Strict FOSD

Definition

(Strict FOSD) p strictly FOSD q with respect to $\succsim$ if p FOSD q with respect to $\succsim$, and
there exists x': $\sum_{x:x' \succsim x} p(x) < \sum_{x:x' \succsim x} q(x)$.

Assumption: Strict FOSD

If p strictly FOSD q with respect to $\succsim$, then $p \succ q$.

Notes:

- p strictly FOSD q, then p FOSD q.
- If p FOSD q but not strictly so, and q FOSD p but not strictly so, this does not imply $p = q$.
- Assumption Strict FOSD satisfied does NOT imply the assumption FOSD satisfied.

Invariance Theorem

Theorem

(Invariance) If the DM satisfies the assumptions Preference Relation, FOSD, and Strict FOSD, and there exist $x, x', x'' \in X$ *and* $\pi \in (0;1]$ *such that* $\pi x + (1-\pi)x'' \succcurlyeq \pi x' + (1-\pi)x''$, *then for all* $\pi' \in (0;1]$ *:* $\pi' x + (1-\pi')x'' \succcurlyeq \pi' x' + (1-\pi')x''$.

Paraphrased: "Optimal decision d^* is invariant in the probability."

Proof.

(i) $x \succsim x'$: Suppose not, then $x' \succ x$, and therefore $\pi x' + (1-\pi)x''$ strictly FOSD $\pi x + (1-\pi)x''$. Then by Strict FOSD,
$\pi x' + (1-\pi)x'' \succ \pi x + (1-\pi)x''$, a contradiction.
(ii) Since $x \succsim x'$, $\pi' x + (1-\pi')x''$ first-order stochastically dominates $\pi' x' + (1-\pi')x''$. Thus by FOSD $\pi' x + (1-\pi')x'' \succcurlyeq \pi' x' + (1-\pi')x''$. ∎

Corollary

Expected Utility (EU)

$$E[u(x_1,x_2)] = \pi u(x_1^C, x_2^C) + (1-\pi) u(x_1^N, x_2^N)$$

Since EU implies FOSD (appropriate definition) the theorem has the following corollary:

Corollary

If the *DM* is an expected utility maximizer on P, then the optimal decision is invariant in the probability.

Direct proof of the corollary

Prove the corollary without relying on the theorem:

- $E[u(x_1, x_2, d)] = \pi u(x_1^C, x_2^C) + (1-\pi) u(x_1^N, x_2^N)$
- DM maximizes the objective function given π.
- One choice variable d.
- Denote the indirect objective function by $V(d) = \pi u(\omega - d, d) + (1-\pi) u(\omega - \kappa, \kappa)$
- $V(d)$ is proportional to $u(\omega - d, d)$
- $\Rightarrow \frac{\partial d^*}{\partial \pi} = 0$

Extending the domain of preferences

Extending the domain of preferences

- FOSD & Preferences over lotteries defined on X give us invariance
- If DM varies decision in the probability we consider extended domain, where also d per se matters
- In terms of utility functions we can think of this as having $u(x, d)$ instead of only $u(x)$
- Our thought experiment is a revealed preference approach, we interpret variance in the probability as deontological motivation

Consequentialism

Definition: Consequentialist Preference

A preference is called *consequentialist* if there exists a utility representation u such that $u = u(x)$.

Examples

Homo oeconomicus: $u = u(x_1)$.
Fehr-Schmidt preferences:
$u = u(x_1, x_2) = x_1 - \alpha max\{x_2 - x_1, 0\} - \beta max\{x_1 - x_2, 0\}$,
$\beta < \alpha, 0 \leq \beta < 1$.
Andreoni/Impure altruism: Utility from the total amount of the public good G and her contribution g since social sanctions depend on g.

Deontologicalism

"deontological moralities, unlike most views of consequentialism, leave space for the supererogatory. A deontologist can do more that is morally praiseworthy than morality demands. A consequentialist cannot. For the consequentialist, if one's act is not morally demanded, it is morally wrong and forbidden. For the deontologist, there are acts that are neither morally wrong nor demanded." (Stanford Encyclopedia of Philosophy)

This can be formalized as a lexicographic preference, with deontological before consequentialist motivations.

Definition: Deontological Preference

A preference is called *deontological* if there exist u, f such that $u = u(d)$, and $f = f(x)$, and f.a. $(x, d), (x', d')$: $(x, d) \succsim (x', d')$ if and only if $u(d) > u(d')$ or $[u(d) = u(d')$ and $f(x) \geqq f(x')]$.

Fact

Fact

For purely deontological preferences the optimal decision is constant in the probability.

- If d^* unique, then the above is trivial.
- If d^* not unique, then we are in the second component of the lexicographic preference, then the consequentialism theorem applies.

Can d^* vary in π?

- Under both consequentialist and deontological preferences, d^* invariant to probability.

Definition: Consequentialist-Deontological Preference

A preference is called *consequentialist-deontological* if there exists a utility representation u such that $u = u(x, d)$.

Consequentialist-deontological preferences

Consider additive preferences over x_1, d:

Bernoulli utility function: $u(x_1, d) = f(x_1) + b(d)$; f, b strictly concave

Objective function: $E[u(x_1, d)] = \pi(f(x_1^C) + b(d)) + (1-\pi)(f(x_1^N) + b(d))$

Indirect objective fct.: $V(d) = \pi f(\omega - d) + (1-\pi) f(\omega - \kappa) + b(d)$

FOC: $\frac{\partial V(d)}{\partial d} = -\pi f_1(\omega - d) + b_1(d) = 0$
Second derivative: $\frac{\partial^2 V(d)}{\partial d^2} = \pi f_{11}(\omega - d) + b_{11}(d) < 0$

Using the FOC, by the implicit function theorem:
$\frac{\partial d^*}{\partial \pi} = \frac{f_1(\omega - d^*)}{\pi f_{11}(\omega - d^*) + b_{11}(d^*)} < 0$

Non-Additive Utility

- It is possible to generate examples where $\frac{\partial d^*}{\partial \pi} > 0$
- Let $u = u(x_1, d)$, $u_1, u_2 > 0$ and $u_{11}, u_{22} < 0$ (risk-aversion).
- For sufficiently negative $u_{12}(\omega - d, d)$ we can get $\frac{\partial d^*}{\partial \pi} > 0$.
- SOC: need u_2 sufficiently positive and sufficiently negative u_{22}.
- But these are not interpretable under uncertainty. u_{12} can easily change sign if, for example, you take the log or square or other strictly monotone transformation of the utility function.

Confounds

Ex-ante Fairness (Machina's mom)

Ex-ante fairness

What if people value some kind of ex-ante fairness?

Utility function for ex-ante fairness:

$U = f(E[u(x_1)], E[\widetilde{u}(x_2)])$

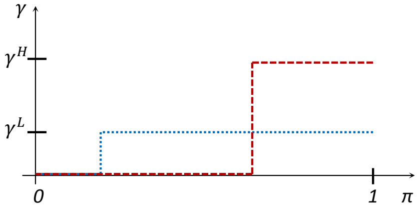
- $u, \tilde{u}$ strictly increasing, concave
- $E[u(x_1)], E[\widetilde{u}(x_2)]$ are normal goods

Fact:

If the *DM* maximizes ex-ante fairness then the sign of $\frac{\partial d^*}{\partial \pi}$ is the same as that of $\kappa - d^*$.

Cognition Costs

The *DM* can compute the optimal decision, but to do so, she incurs a cognition cost $\gamma \geq 0$, or otherwise she can make a heuristic (fixed) choice $\bar{d}$ for which (normalized) costs are 0.



Note:

Cognition cost model predicts that time spent on the survey also changes as d changes with π.

“Internal” consequences

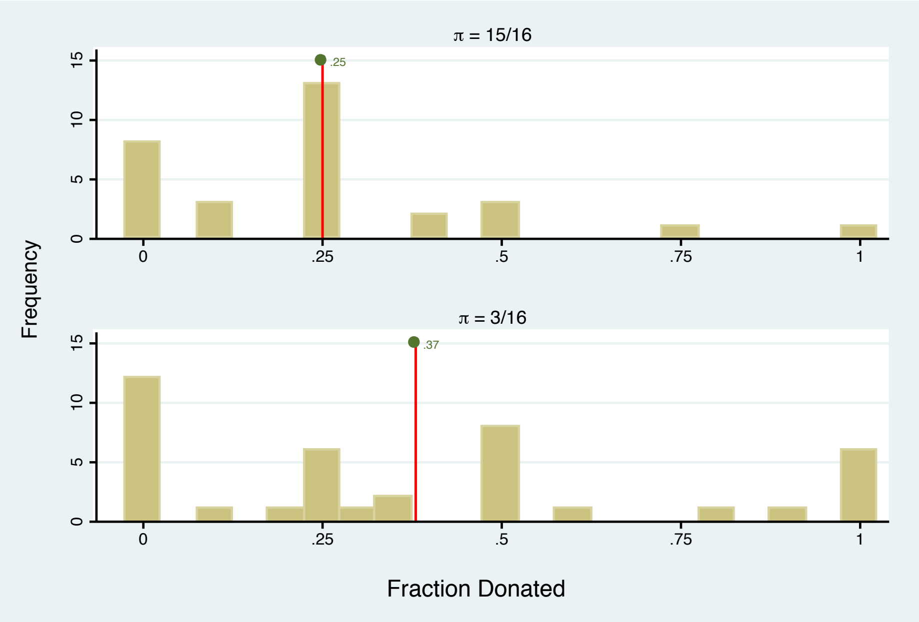
- Revealed preference method cannot distinguish between different internal “consequences”
- Even purely deontological preferences likely have some neurobiological “consequence”
- Self-signaling (Benabou and Tirole 2006)
- Or, what drives deontological motivations, is it conscience or guilt or ?

Experiments

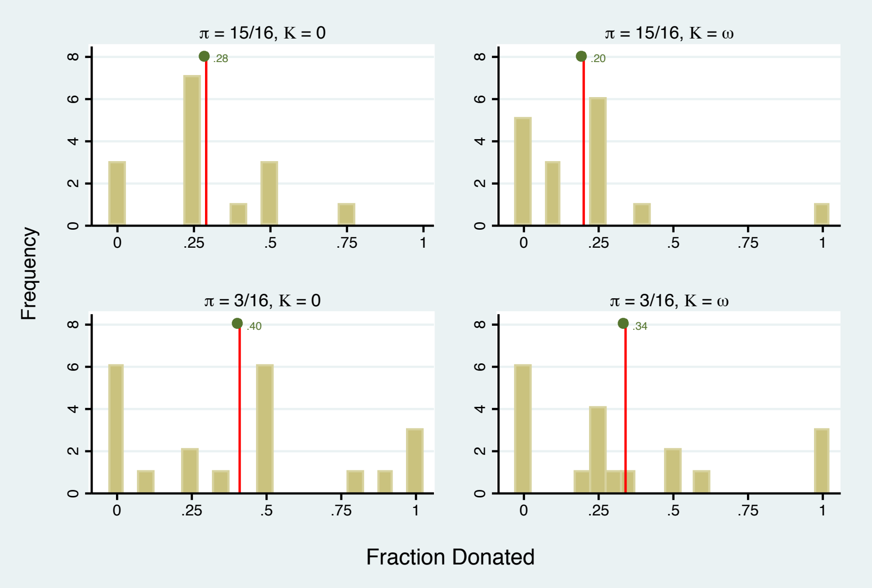
(Potential) Experimental Settings

1. Laboratory: Shredding
    1. Zurich - Donation
    2. Germany - Donation
    3. Germany -Truth-telling (planned)
2. Online: Authority
    1. LISS: population representative panel, authority shreds, heterogeneity (planned)
3. Online:
    1. AMT: online, Red Cross, encryption, large N, structural estimates of trade-offs
    2. AMT: online, Red Cross, $\omega = 50c$, experimenter observes, structural estimates of trade-offs

Donation (κ pooled)

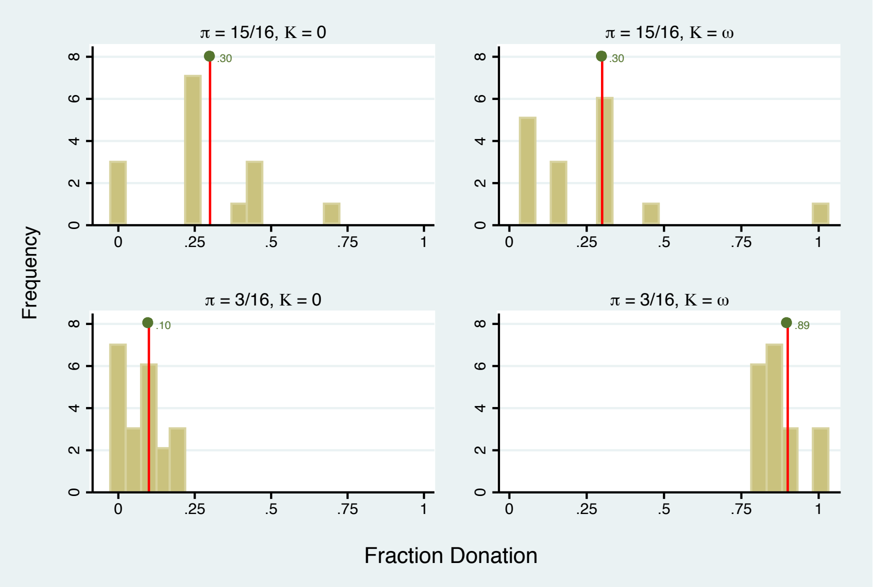


Donation (by κ)



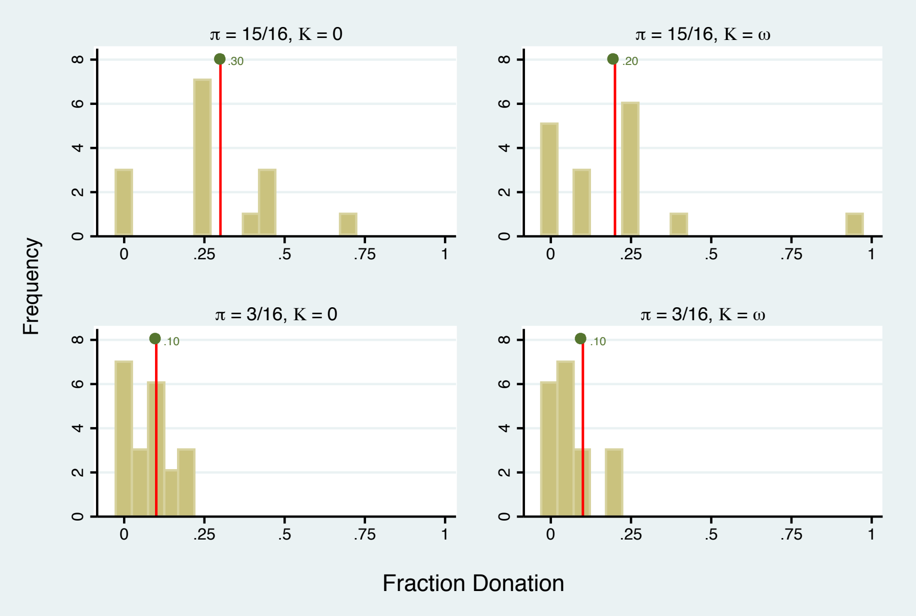
Recall: ex-ante => sign $\frac{\partial d^*}{\partial \pi}$ =sign $(\kappa - d^*)$

Expected Income $E(x_2)$ (by κ)



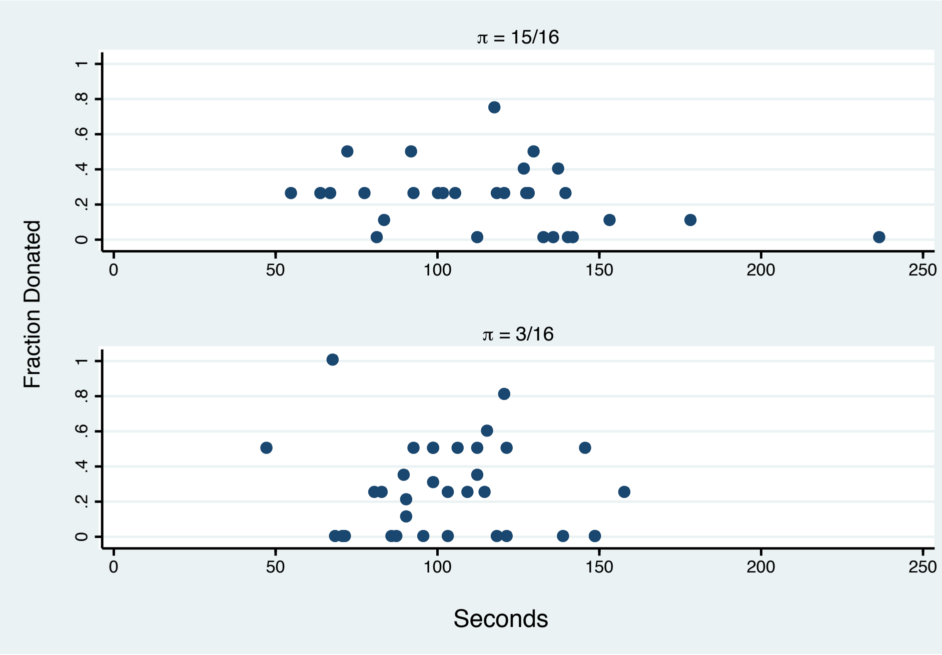
Recall: ex-ante => sign $\frac{\partial d^*}{\partial \pi}$ =sign $(\kappa - d^*)$

Expected Giving (πd^*) (by κ)



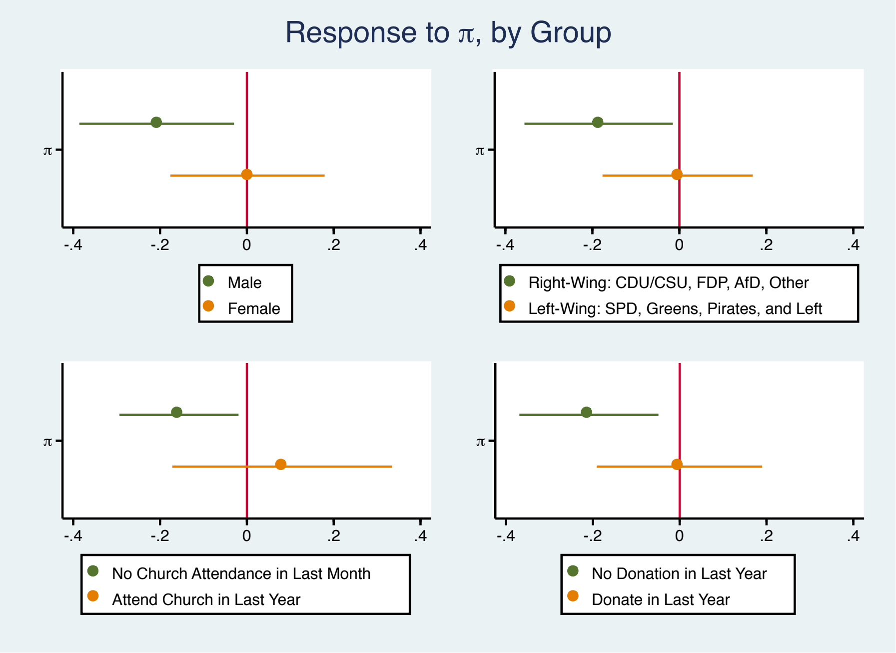
Recall: ex-ante => sign $\frac{\partial d^*}{\partial \pi}$ =sign $(\kappa - d^*)$

Time Spent



π = 15/16
π = 3/16
Fraction Donated
Seconds
0
.2
.4
.6
.8
1
0
50
100
150
200
250

Heterogenous Treatment



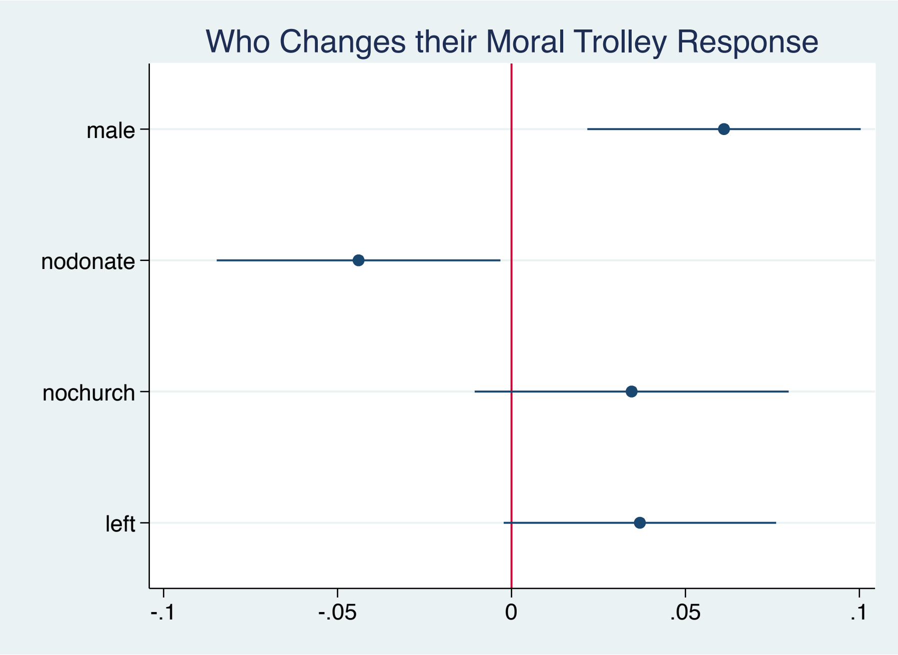
Moral Trolley

You are a solider and you and a few other soldiers are taken prisoner by the enemy. After a year in captivity, your group has tried to escape, but was caught. The enemy overseer decided to hang you and your group. At the gallows he loosens the noose around your neck and proclaims that if you pull the chair under one of the other soldiers in your group, you and the other **19 soldiers** in your group will be released. If you do not, everyone will be hung. The warden says this seriously and will keep his promise. Would you remove the chair in this situation?

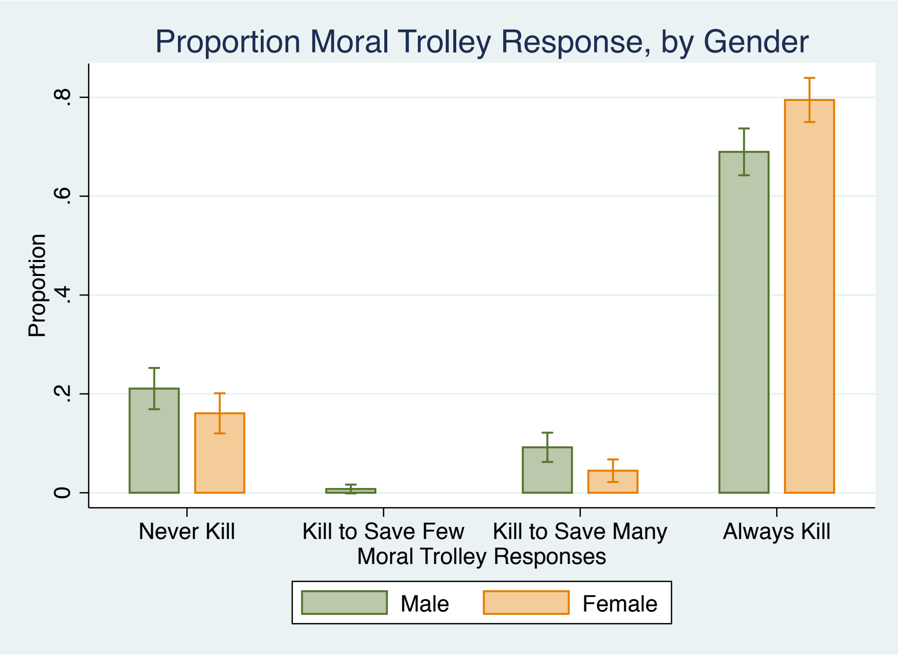
Moral Trolley

You are a solider and you and a few other soldiers are taken prisoner by the enemy. After a year in captivity, your group has tried to escape, but was caught. The enemy overseer decided to hang you and your group. At the gallows he loosens the noose around your neck and proclaims that if you pull the chair under one of the other soldiers in your group, you and the other **5 soldiers** in your group will be released. If you do not, everyone will be hung. The warden says this seriously and will keep his promise. Would you remove the chair in this situation?

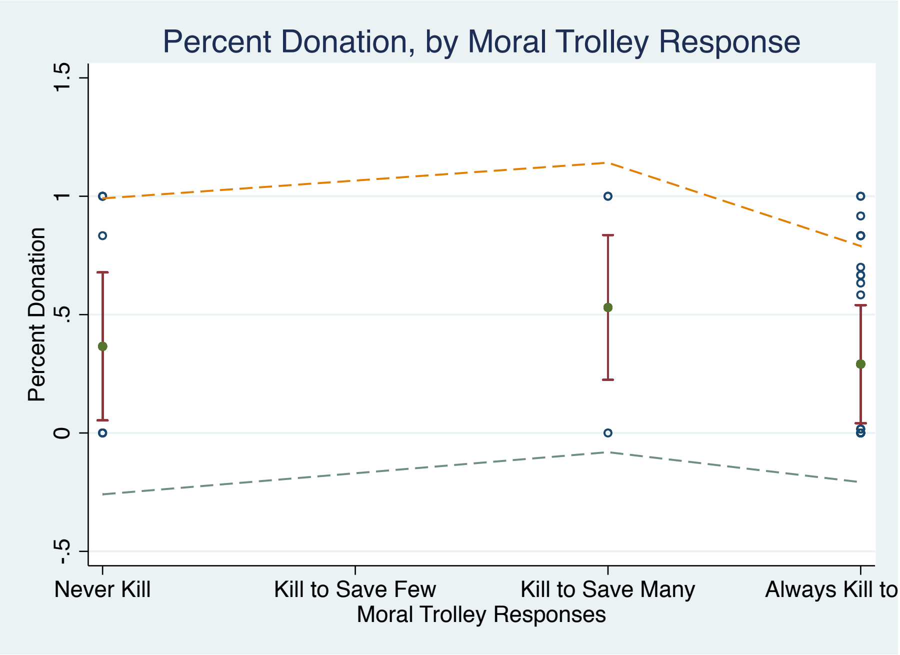
Moral Trolley



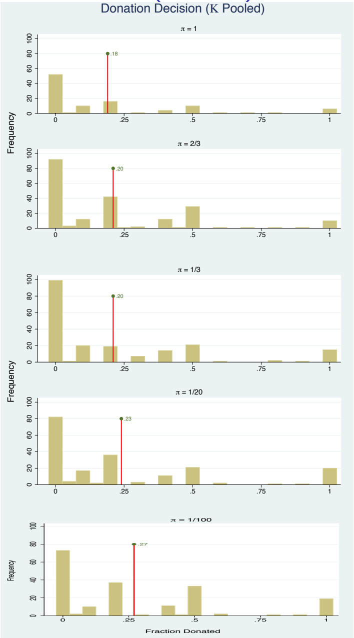
Moral Trolley



Moral Trolley



Donation (AMT)



Red: Mean

Econometric Specifications

- Basic regression: $d_i = \beta_0 + \beta_1 Treatment_i + \beta_2 X_i + \varepsilon_i$
- Wilcoxon test
- Sensitivity d_i^* to π as predicted from demographics
  - $d^*{}_i = \beta_0 \pi_i + \beta_1 \mathbf{X}_i \pi_i + \alpha \mathbf{X}_i + \varepsilon_i$
- Structural estimation (assume specific mixed preference and get estimates on relative importance)

Pooled Results

	Ordinary Least Squares					
	(1)	(2)	(3)	(4)	(5)	(6)
	d^*		Expected Income $E(x_2)$		Expected Giving (πd^*)	
Mean dep. var.	0.30		0.39		0.12	
% Consequential (π)	-0.176*	-0.159*	-0.259**	-0.278***	0.212***	0.219***
	(0.0978)	(0.0855)	(0.108)	(0.0802)	(0.0484)	(0.0452)
K Fixed Effects	N	Y	N	Y	N	Y
Observations	71	71	71	71	71	71
R-squared	0.045	0.292	0.077	0.506	0.218	0.339

Notes: Standard errors in parentheses. Raw data shown in Figures 1 and 2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Pooled Results

	Ordinary Least Squares					
	(1)	(2)	(3)	(4)	(5)	(6)
	d^*		Expected Income $\mathrm{E}(x_2)$		Expected Giving ($\boldsymbol{\pi} d^*$)	
Mean dep. var.	0.23		0.34		0.07	
% Consequential ($\boldsymbol{\pi}$)	-0.0725**	-0.0684*	-0.224***	-0.219***	0.194***	0.213***
	(0.0288)	(0.0390)	(0.0334)	(0.0299)	(0.0132)	(0.0181)
K Fixed Effects	N	Y	N	Y	N	Y
Controls	N	Y	N	Y	N	Y
Observations	902	900	902	900	902	900
R-squared	0.007	0.059	0.048	0.604	0.194	0.214

Notes: Standard errors in parentheses. Raw data shown in Figure 3. Controls include indicator variables for gender, American, Indian, Christian, Atheist, aged 25 or younger, and aged 26-35 as well as continuous measures for religious attendance and accuracy in the lock-in data entry task. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Disaggregated Results

	Ordinary Least Squares							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Decision (d)		Decision (d)		Decision (d)		Decision (d)	
	K = Unknown		$K = 10$¢		$K = 0$¢		$K = 50$¢	
Mean dep. var.	0.26		0.22		0.20		0.22	
% Consequential ($\boldsymbol{\pi}$)	-0.0778	-0.0654	-0.0525	-0.0321	-0.0711	-0.0708	-0.0644	-0.0675
	(0.0523)	(0.0523)	(0.0526)	(0.0536)	(0.0464)	(0.0466)	(0.0462)	(0.0456)
Male		-0.0909**		-0.0474		0.0108		0.0178
		(0.0399)		(0.0430)		(0.0395)		(0.0362)
American		0.0241		-0.0539		0.0838		0.117*
		(0.0524)		(0.0539)		(0.0664)		(0.0598)
Indian		-0.0672		-0.0785		-0.0673		-0.0626
		(0.0566)		(0.0560)		(0.0630)		(0.0590)
Christian		-0.0295		0.0584		-0.0215		-0.000293
		(0.0483)		(0.0503)		(0.0494)		(0.0479)
Atheist		-0.0188		0.00480		0.0113		-0.0927
		(0.0644)		(0.0649)		(0.0802)		(0.0725)
Religious Services Attendance		-0.00614		0.000508		0.00367		-0.00546
		(0.0145)		(0.0156)		(0.0137)		(0.0137)
Ages 25 or Under		-0.0207		-0.122**		-0.0109		-0.113**
		(0.0518)		(0.0570)		(0.0493)		(0.0474)
Ages 26-35		0.00271		-0.110*		-0.00105		-0.111**
		(0.0548)		(0.0593)		(0.0493)		(0.0480)
Own Errors		-0.000192		-0.000186		0.000220		-0.000148
		(0.000193)		(0.000163)		(0.000194)		(0.000143)
Observations	260	260	218	218	256	255	271	270
R-squared	0.009	0.069	0.005	0.081	0.009	0.052	0.007	0.097

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Non-Parametric Tests

	Wilcoxon-Mann-Whitney 2-sided test (p-values)		
	(1)	(2)	(3)
Thresholds	K Unknown or 10¢	K = 0¢ or 50¢	K Pooled
$\pi = 1$ vs. $\pi \leq 0.67$	0.91	0.05	0.11
$\pi \geq 0.67$ vs. $\pi \leq 0.33$	0.07	1.00	0.20
$\pi \geq 0.33$ vs. $\pi \leq 0.05$	0.05	0.10	0.01
$\pi \geq 0.05$ vs. $\pi = 0.01$	0.15	0.02	0.01

	π Pooled
$K \geq$ 10¢ vs. K = 0¢	0.40
K = 50¢ vs. $K \leq$ 10¢	0.11

Non-Parametric Tests

	Wilcoxon-Mann-Whitney 2-sided test (p-values)
Thresholds	Pooled
$\boldsymbol{\pi} = 3/16$ vs. $\boldsymbol{\pi} = 15/16$	0.16
$K = 0$ vs. $K = \text{Max}$	0.16

Who responds to π?

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Ordinary Least Squares									
	Decision (d)									
Mean dep. var.	0.23									
% Consequential (π)	-0.100**	-0.0493	-0.124**	-0.0500	-0.0522	-0.0774	-0.0618	-0.0548	-0.0839**	-0.0190
	(0.0494)	(0.0429)	(0.0506)	(0.0436)	(0.0403)	(0.0616)	(0.0467)	(0.0443)	(0.0407)	(0.126)
π * Male	0.0612									0.0490
	(0.0577)									(0.0611)
π * American		-0.0675								0.0370
		(0.0627)								(0.0911)
π * Indian			0.0990*							0.0426
			(0.0574)							(0.0963)
π * Christian				-0.0599						-0.0658
				(0.0632)						(0.0783)
π * Atheist					-0.133					-0.145
					(0.0837)					(0.108)
π * Religious Services Attendance						0.00394				-0.00739
						(0.0210)				(0.0224)
π * Ages 25 or Under							-0.0149			-0.0815
							(0.0576)			(0.0787)
π * Ages 26-35								-0.0386		-0.0878
								(0.0597)		(0.0808)
π * Own Errors									0.000402	0.000319
									(0.000299)	(0.000307)
K Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	900	900	900	900	900	900	900	900	900	900
R-squared	0.061	0.061	0.063	0.060	0.062	0.059	0.059	0.060	0.061	0.068

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Structural Estimation

Homo oeconomicus

Example

Homo oeconomicus and Duty Bliss point

$u(x_{1,},d) = \lambda(x_{DM}) + (1-\lambda)(-(\delta - d)^2) = \lambda(\omega - d) + (1-\lambda)(-(\delta - d)^2)$

FOC: $0 = \pi\lambda(-1) + 2(1-\lambda)(\delta - d)$
Linear: $d^* = \frac{-\lambda}{2(1-\lambda)}\pi + \delta$
Estimate of -0.073 implies that $\lambda = 0.13$.
Many people donate more than the bliss point of 25%

Fehr-Schmidt

Example

Fehr-Schmidt preferences and deontological bliss point

$u(x_{1,}, x_2, d) =$
$\lambda(x_1 - \alpha max\{x_2 - x_1, 0\} - \beta max\{x_1 - x_2, 0\}) + (1 - \lambda)(-(\delta - d)^2).$
GMM:
$$E\left[\pi\left(1[\tfrac{\omega}{2} > d]\left[d - \pi\frac{\lambda(2\beta-1)}{2(1-\lambda)} - \delta\right] + 1[\tfrac{\omega}{2} \leq d]\left[d - \pi\frac{\lambda(-2\alpha-1)}{2(1-\lambda)} - \delta\right]\right)\right] = 0$$

Linear: $d^* = \frac{\lambda(-2\alpha-1)}{2(1-\lambda)}\pi + \frac{\lambda(\alpha+\beta)}{(1-\lambda)}\pi 1[\frac{\omega}{2} > d] + \delta$

Fehr-Schmidt

Table 8: Trading Off Consequentialist-Deontological Motivations (AMT Experiment)

	OLS	IV	IV
	(1)	(2)	(3)
		Decision (d)	
Mean dep. var.		0.23	
% Consequential (π)	-0.239***	-0.363***	-0.368***
	(0.0249)	(0.0548)	(0.139)
$\pi * 1(d \geq w/2)$	0.870***	1.516***	1.542**
	(0.0412)	(0.250)	(0.714)
Constant (Duty Bliss Point)	0.251***	0.249***	0.249***
	(0.0116)	(0.0131)	(0.0134)
IV	N	π, Indian	π, Age $\leq$ 25
Observations	902	902	902
R-squared	0.336	0.155	0.140

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Fehr-Schmidt

Example

Fehr-Schmidt preferences and deontological bliss point

$u(x_{1,}, x_2, d) =$
$\lambda(x_1 - \alpha max\{x_2 - x_1, 0\} - \beta max\{x_1 - x_2, 0\}) + (1-\lambda)(-(\delta - d)^2)$.

$\frac{\lambda(2\beta-1)}{2(1-\lambda)} = -0.36$ (I) and $\frac{\lambda(-2\alpha-1)}{2(1-\lambda)} = 1.16$ (II). 2 equations and 3 unknowns.
Same problem: $\lambda < 0$ if $\alpha > \beta > 0$.
Linear form of Fehr-Schmidt leads to bang-bang solution. Either consequentialist donates nothing, but many people donate more than bliss point, or, consequentialist donates 50-50, which is too much so $\lambda < 0$.

Encryption

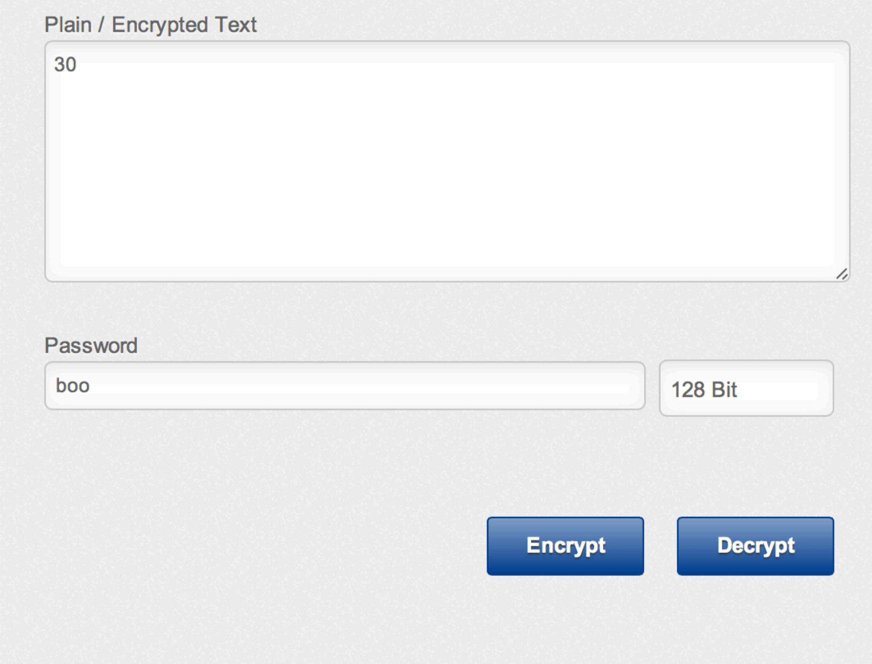
AES ENCRYPTOR

This is a free online AES encryptor/Decryptor. The AES encryptor or AES encoder tool is a simple and user friendly text file encryption program that uses the Advanced Encryption Standard (AES) specification. The AES encryption algorithm secures the sensitive and unclassified information from illegal access.

Plain / Encrypted Text

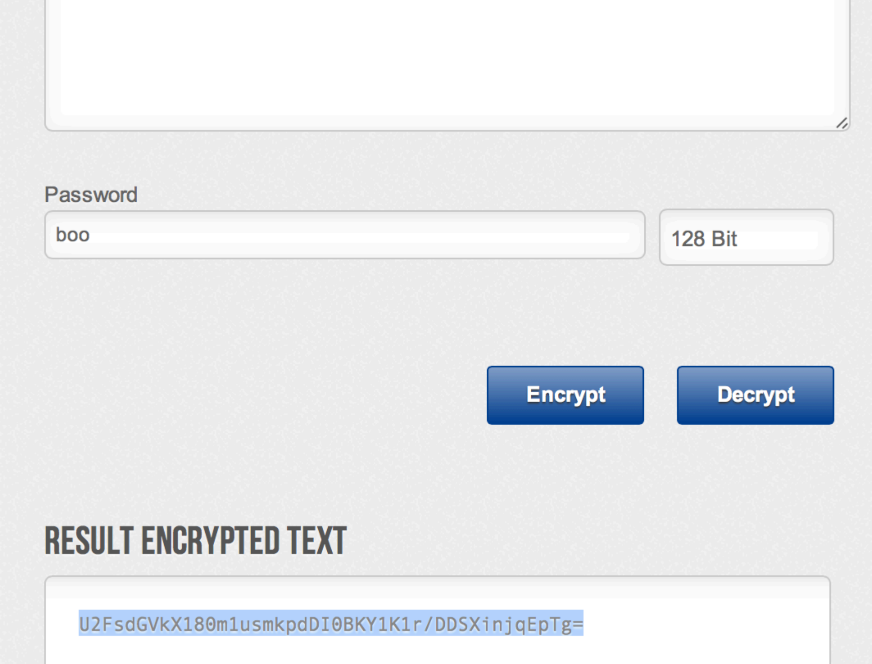
30

Encryption



Plain / Encrypted Text
30
Password
boo
128 Bit
Encrypt
Decrypt

Encryption


Password
boo
128 Bit
Encrypt
Decrypt
RESULT ENCRYPTED TEXT
U2FsdGVkX180m1usmkpdDI0BKY1K1r/DDSXinjqEpTg=

Wheel

Please click on this link. On this site, a new ball is picked several times a minute. Each result has an ID. We will wait for result **72745**, which will be generated at **16:41:20 (time zone: UTC)**. Please pay attention to the color of the ball for that result.

Once that result has been generated, a button will appear below. Please wait...

Decryption

Task

Result 72745 was a green ball. So, we will split the money according to your decision. Please enter below the password you used on the encryption site. This will allow us to decrypt the offer you submitted earlier, and find out how much you wanted to give to Red Cross.

Password:

Submit

Conclusion

Conclusion

- Formal interpretation of major moral philosophies
- Revealed preferences method to detect deontological motivations (instead of, e.g. moral trolley vignettes)
- The direction of the decision changes gives insight into the location of the optimand for one's greatest duty.
- Evidence suggests that people have both consequentialist and deontological motivations
- Contingent valuation, psychological vignettes vs. consequential decisions
- Random lottery method for moral decisions may be inappropriate
- Consequentialist policy responses [welfare economics] to 'sacred values'
- Charitable fund-raising strategies
- Measure *mens rea* (intention) vs. *actus reus* (act) when it matters for certain legal settings

Last slide

Thank you!

We appreciate comments now or by email.

Daniel Chen: chendan@ethz.ch
Martin Schonger: mschonger@ethz.ch

Latest version of paper at:
http://nber.org/∼dlchen/papers/deontological.pdf

Mas-Colell

Mas-Colell restricts attention to money where monotonicity is presumably satisfied:

In the remaining sections, we focus on the special case in which the outcome of a risky choice is an amount of money (or any other one-dimensional measure of consumption). This case underlies much of finance and portfolio theory, as well as substantial areas of applied economics.

Mas-Colell definition of FOSD and its implication:

Definition 6.D.1: The distribution $F(\cdot)$ *first-order stochastically dominates* $G(\cdot)$ if, for every nondecreasing function $u: \mathbb{R} \to \mathbb{R}$, we have

$$\int u(x)\, dF(x) \geq \int u(x)\, dG(x).$$

Proposition 6.D.1: The distribution of monetary payoffs $F(\cdot)$ first-order stochastically dominates the distribution $G(\cdot)$ if and only if $F(x) \leq G(x)$ for every x.

Back to FOSD .

Kant on lying

- Kant does discuss uncertainty
- Kant says one must not lie (one may remain silent)

"It is indeed possible that after you have honestly answered Yes to the murderer's question as to whether the intended victim is in the house, the latter went out unobserved and thus eluded the murderer, so that the deed would not have come about. However, if you told a lie and said that the intended victim was not in the house, and he has actually (though unbeknownst to you) gone out, with the result that by so doing he has been met by the murderer and thus the deed has been perpetrated, then in this case you may be justly accused as having caused his death. For if you had told the truth as best you knew it, then the murderer might perhaps have been caught by neighbors who came running while he was searching the house for his intended victim, and thus the deed might have been prevented. [...] To be truthful (honest) in all declarations is, therefore, a sacred and unconditionally commanding law of reason that admits of no expediency whatsoever."

- Kant (1799) : "On a supposed right to lie because of philanthropic concerns".

Strict and Weak FOSD

Example

Preferences that satisfy strict FOSD but violate weak FOSD.

An indivisible treat that mom can allocate to one of her two children x or y. Mom would like to be exactly fair, thus her most preferred lottery is $(x; \frac{1}{2}, y; \frac{1}{2})$, she is indifferent between all other lotteries.
For all $\pi, \pi' \in [0;1] \setminus \frac{1}{2}$: $(x; \pi, y; 1-\pi) \sim (x; \pi', y; 1-\pi')$ and $(x; \frac{1}{2}, y; \frac{1}{2}) \succ (x; \pi, y; 1-\pi)$. These preferences are complete and transitive.
Strict FOSD is trivially satisfied since there is no lottery that strictly first-order stochastically dominates another lottery.
However, axiom WFOSD is violated: $(x; \frac{2}{3}, y; \frac{1}{3})$ weakly first order-stochastically dominates $(x; \frac{1}{2}, y; \frac{1}{2})$, but $(x; \frac{1}{2}, y; \frac{1}{2}) \succ (x; \frac{2}{3}, y; \frac{1}{3})$.

Continuity and FOSD

Definition

$\succsim$ is **continuous** if for all $p, q, r \in P$ the sets $\{\alpha \in [0,1] : \alpha p + (1-\alpha) q \succsim r\}$ and $\{\alpha \in [0,1] : r \succsim \alpha p + (1-\alpha) q\}$ are closed in [0,1].

Note that: $\{\alpha \varepsilon [0,1] : x \succsim \alpha x + (1-\alpha) y\} = [0; \frac{1}{2}) \cup (\frac{1}{2}, 1]$. But continuity is not enough.

Example

Again mom would like to be fair, but now between two unfair lotteries she prefers the one that is more fair. For all $\pi, \pi' \in [0;1]$: $\pi \cdot (1-\pi) \geq \pi' \cdot (1-\pi')$ if and only if $(x; \pi, y; 1-\pi) \succsim (x; \pi', y; 1-\pi')$. Strict FOSD and continuity are satisfied. But WFOSD is violated: $(x; \frac{2}{3}, y; \frac{1}{3})$ weakly first order-stochastically dominates $(x; \frac{1}{2}, y; \frac{1}{2})$, but $(x; \frac{1}{2}, y; \frac{1}{2}) \succ (x; \frac{2}{3}, y; \frac{1}{3})$.

Rich Domain, Continuity and FOSD

However, if there are also two outcomes $x, y \in X$ such that $x \succ y$, then strict FOSD implies weak FOSD.

Proof.

Suppose p weakly first-order stochastically dominates q. We need to show that $p \succsim q$.

Suppose not, that is $q \succ p$. Since X is finite there exists an $\overline{x}$, $\underline{x}$ such that for all x: $\overline{x} \succsim x$, and an $x \succsim \underline{x}$. By RICH $\overline{x} \succ \underline{x}$.

At least one of the following three cases is satisfied: (i) $\overline{x} \succ q$, (ii) $p \succ \underline{x}$ or (iii) $q \succsim \overline{x} \succ \underline{x} \succsim p$.

(i) Since p weakly first-order stochastically dominates q, and $\overline{x} \succ q$, for any $\alpha > 0$ the lottery $\alpha\overline{x} + (1-\alpha)p$ strictly first-order stochastically dominates q. But then $\{\alpha : \alpha\overline{x} + (1-\alpha)p \succsim q\} = (0,1]$, a violation of continuity. (ii) and (iii) similarly.

∎

FOSD is not Sure-Thing Principle

Savage's Sure-Thing Principle is not FOSD. If invariant, then probability does not matter.

7 The sure-thing principle

A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant to the attractiveness of the purchase. So, to clarify the matter for himself, he asks whether he would buy if he knew that the Republican candidate were going to win, and decides that he would do so. Similarly, he considers whether he would buy if he knew that the Democratic candidate were going to win, and again finds that he would do so. Seeing that he would buy in either event, he decides that he should buy, even though he does not know which event obtains, or will obtain, as we would ordinarily say. It is all too seldom that a decision can be arrived at on the basis of the principle used by this businessman, but, except possibly for the assumption of simple ordering, I know of no other extralogical principle governing decisions that finds such ready acceptance.

Machina and Schmeidler (1992) formulation of Sure-Thing Principle is about $\frac{\partial d^*}{\partial \kappa} = 0$.

What Machina and Schmeidler (1992) call Eventwise Montonicity is FOSD and invariance.

FOSD is not Independence

(IND) $\succsim$ satisfies independence if for all lotteries p, q, r in $P : p \succcurlyeq q \Leftrightarrow \alpha p + (1-\alpha) r \succcurlyeq \alpha q + (1-\alpha) r$.

Example

Cumulative Prospect Theory (Rank Dependent Expected Utility) satisfies FOSD and allows for Allais Paradox, but not Independence.

If the cardinality of the outcome space is 2, that then independence is as weak an axiom as first-order stochastic dominance.