UC Berkeley
Department of Economics
Foundations of Psychology and Economics (219A)

Module II: Social Preferences

Reading list

1. Camerer, C. (2003) "Behavioral Game Theory: Experiments in Strategic Interaction." Princeton University Press (chapter 2).

2. Rabin, M. (1993) "Incorporating Fairness into Game Theory and Economics." *American Economic Review*, 83, pp. 1281-1302.

3. Fehr, E. and K. Schmidt (1999) "A Theory of Fairness, Competition and Co-operation." *Quarterly Journal of Economics*, 114, pp. 817-868.

4. Bolton, G. and A. Ockenfels (2000) "ERC: A Theory of Equity, Reciprocity, and Competition." *American Economic Review*, 90, pp. 166-193.

5. Charness, G. and M. Rabin (2002) "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics*, 117, pp. 817-869.

6. Karni, E. and Z. Safra (2002) "Individual Sense of Justice: A Utility Representation." *Econometrica*, 70, pp. 263–284.

7. Andreoni, J. and J. Miller (2002) "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica*, 70, pp. 737-753.

8. Fisman, R., S. Kariv and D. Markovits (2007) "Individual Preferences for Giving." *American Economic Review*, 97, pp. 1858-1876.

Background

- People often sacrifice their own payoffs in order to increase the payoffs of *anonymous* others.

- They do so even in circumstances that do not engage reciprocity motivations or strategic behavior.

- This has led economists to begin the systematic study of the *distributional preferences* that govern such behavior.

Social preferences theories

- *Social welfare*
  - persons pursue an aggregate of their own payoffs and those of others.
- *Inequality aversion*
  - persons care about differences between their own and others' payoffs.

Template for analysis

- The *dictator game* eliminates strategic behavior and reciprocity motivations and implicates only distributive preferences.

- Choices made by a person $self$ that have consequences for her own payoff and the payoffs of an anonymous $other$.

- Throughout, we denote persons $self$ and $other$ by S and O, respectively, and the associated monetary payoffs by π_S and a π_O.

Given a *nondegenerate* utility function

$$U_S = u_S(\pi_S, \pi_O)$$

that captures the possibility of giving, person $self$ is *selfish* when for any π and π'

$$u_S(\pi) \geq u_S(\pi') \text{ if and only if } \pi_S \geq \pi'_S$$

and otherwise displays some form of *altruism*.

Prototypical social preferences

Charness and Rabin (*QJE*, 2002) propose the following simple formulation

$$U_S(\pi_S, \pi_O) \equiv (\rho r + \sigma q)\pi_O + (1 - \rho r - \sigma q)\pi_S$$

where

$$r = 1 \ (s = 1) \text{ if } \pi_s > \pi_o \ (\pi_s < \pi_o) \text{ and zero otherwise.}$$

Increasing the ratio ρ/σ indicates an increase in concerns for increasing aggregate payoffs rather than reducing differences in payoffs.

(i) *competitive* preferences ($\sigma \leq \rho < 0$) – utility increases in the difference $\pi_S - \pi_O$

(ii) *narrow self-interest* or *selfish* preferences ($\sigma = \rho = 0$) – utility depends only on π_S

(iii) *difference aversion* preferences ($\sigma < 0 < \rho < 1$) – utility is increasing in π_S and decreasing in the difference $\pi_S - \pi_O$

(iv) *social welfare* preferences ($0 < \sigma \leq \rho \leq 1$) – utility is increasing in both π_S and π_O.

Objections and replies

An unpublished working paper concludes

> *This puts the basis of our modeling on unobservable preferences, and raises the specter of extensive ad hoc modeling with a basis primarily in psycho babble.*

Camerer (2003) replies

> *The goal is not to explain every different finding by adjusting the utility function just so; the goal is to find parsimonious utility functions, supported by psychological intuition...*

Experimental design

In a typical dictator game, the problem faced by $self$ is simply allocating a fixed total income between $self$ and $other$.

Person $self$ divides some *endowment* m between $self$ and $other$ in any way he wishes such that

$$\pi_S + \pi_O = m.$$

The dictator game, developed by Andreoni and Miller (*Econometrica*, 2002), allows for m to be spent on π_S and π_O at *price* levels p_S and p_O such that

$$p_S\pi_S + p_O\pi_O = m.$$

This configuration creates *budget sets* over π_S and π_O that allow for the thorough testing for consistency with utility maximization.

Experimental procedures

- A graphical computer interface that allows for the efficient collection of many observations per subject.

- The graphical representation does not force subjects into discrete choices that suggest extreme prototypical preference types.

- It generates a very rich data set well-suited to studying behavior at the level of the individual subject.

Econometric specification

- Our subjects' CCEI scores are sufficiently near one to justify treating the data as utility-generated.

- If choice data satisfy GARP we would ideally like to extract a rationalizing utility function.

- Afriat's theorem tells us that if a rationalizable utility function exists, it can be chosen to be increasing, continuous, and concave.

- The constant elasticity of substitution (CES) utility function is commonly employed in demand analysis.

- The patterns observed in the nonparametric approach suggest that it is appropriate to estimate a CES demand function.

- The CES is useful because attitudes towards giving can be adjusted by means of a single parameter.

The CES utility function is given by

$$U_S = [\alpha(\pi_S)^\rho + (1-\alpha)(\pi_O)^\rho]^{1/\rho}$$

α - the relative weight on $self$ versus $other$.

ρ - the curvature of the altruistic indifference curves.

$\rho > 0$ ($\rho < 0$) indicate preference weighted towards increasing total (reducing differences in) payoffs.

The CES demand function is given by

$$\pi_s(p, m') = \frac{A}{p^r + A} m'$$

where

$$r = -\rho / (\rho - 1)$$

and

$$A = [\alpha / (1 - \alpha)]^{1/(1-\rho)} .$$

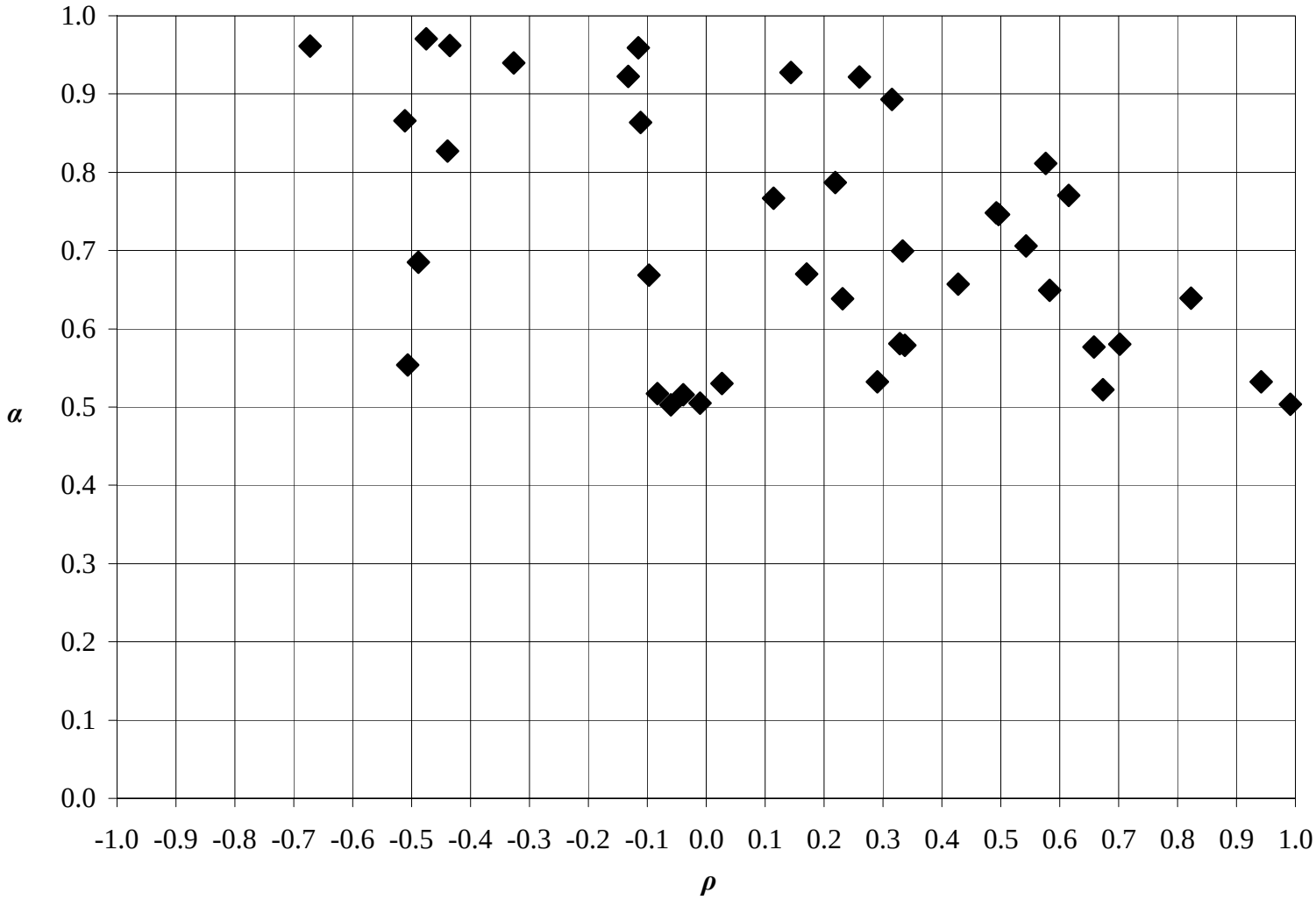
This generates the following individual-level econometric specification for each subject n:

$$\frac{\pi^i_{sn}}{m'^i_n} = \frac{A_n}{(p^i_n)^{r_n} + A_n} + \epsilon^i_n$$

where ϵ^i_n is assumed to be distributed normally with mean zero and variance σ^2_n.

Estimate $\hat{A}_n$ and $\hat{r}_n$ using non-linear tobit maximum likelihood, and use this to infer the values of the CES parameters $\hat{\alpha}_n$ and $\hat{\rho}_n$.

Scatterplot of the CES estimates



Distinguishing social preferences from preferences for altruism

- Distributional preferences may be divided into two qualitatively different types which we call *preferences for altruism* and *social preferences*.

- Social preferences and distributional preferences are used interchangeably in the literature and our usage is not quite standard.

- Nevertheless, the distinctions that we draw are straightforward and capture important differences.

- *Preferences for altruism*
  - tradeoffs between the payoffs to $self$ and the payoffs to $others$.

- *Social preferences*
  - tradeoffs between the payoffs to $others$ (i.e. all persons except $self$).

A common assumption used in demand analysis allows for a clear demarcation between social preferences and preferences for altruism:

Independence *For any* π_S, π'_S, *and profiles* $\pi_O = (\pi_A, \pi_B)$ *and* π'_O

$u_S(\pi_S, \pi_O) > u_S(\pi_S, \pi'_O)$ *if and only if* $u_S(\pi'_S, \pi_O) > u_S(\pi'_S, \pi'_O)$.

If the independence property is satisfied, then the utility function $u_S(\pi_S, \pi_O)$ is (weakly) *separable*.

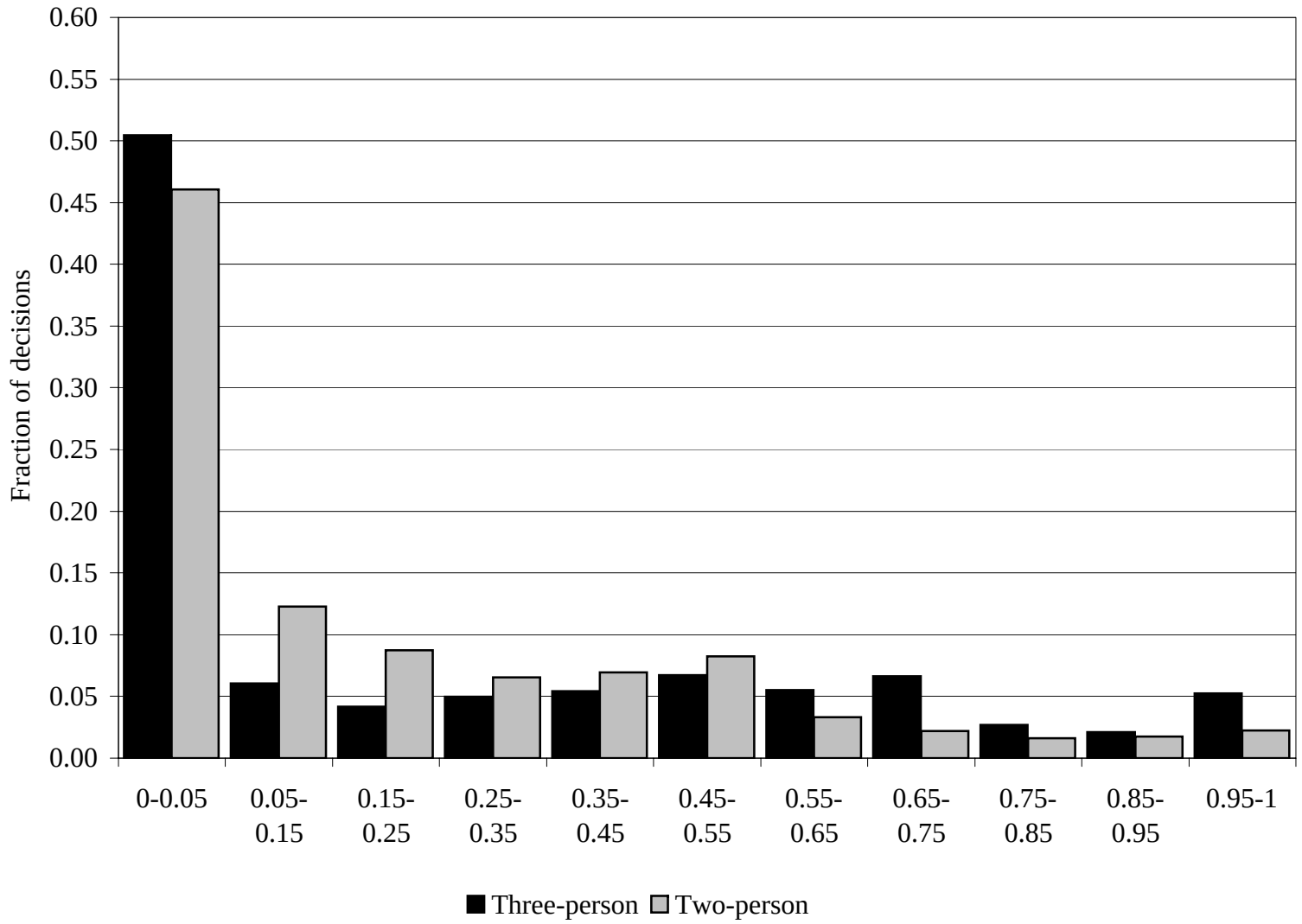
There exists a *subutility* function $w_S(\pi_O)$ and a *macro* function $v_S(\pi_S, w_S)$ with v_S strictly increasing in w_S such that

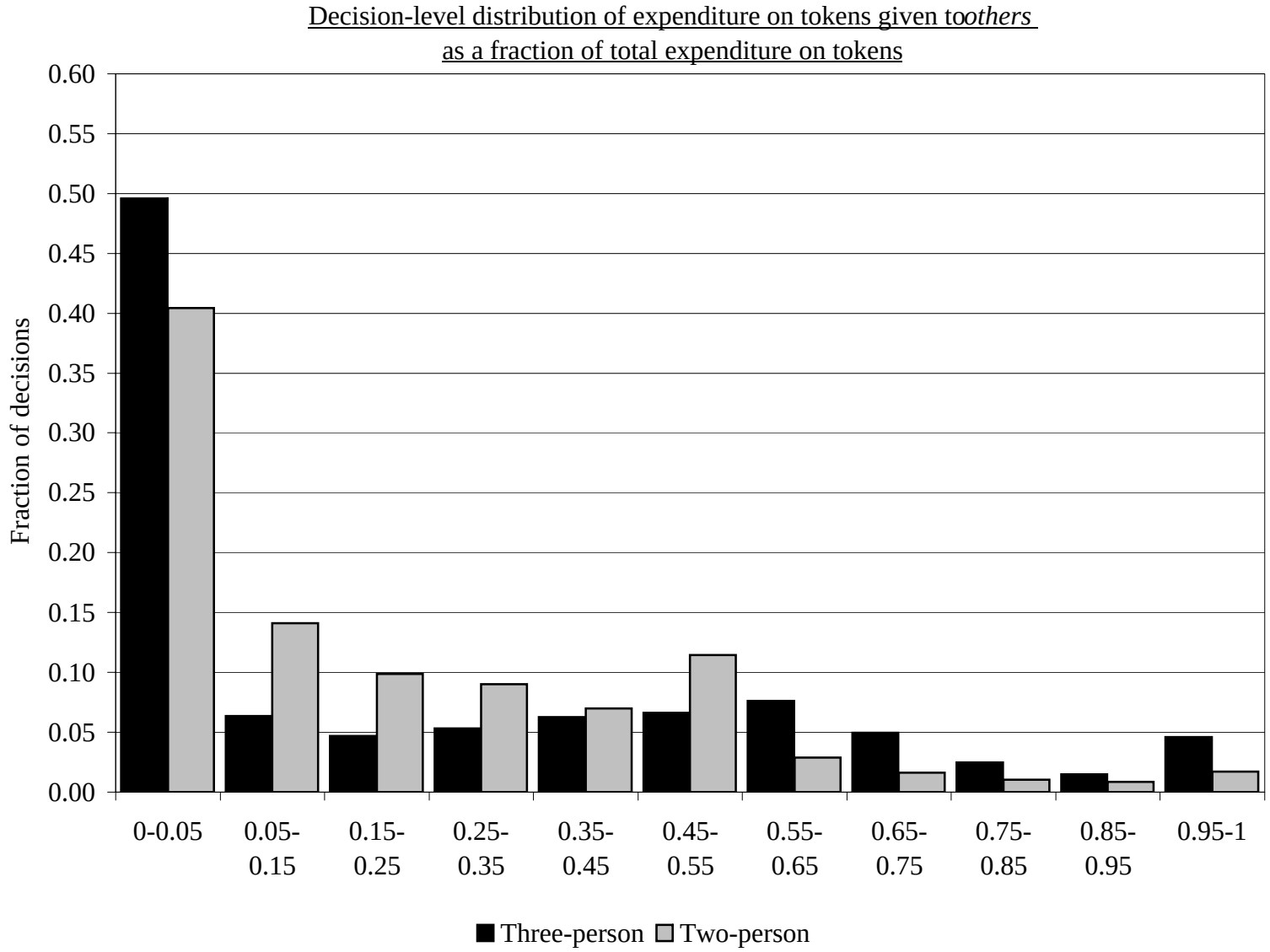
$$u_S(\pi_S, \pi_O) \equiv v_S(\pi_S, w_S(\pi_O)).$$

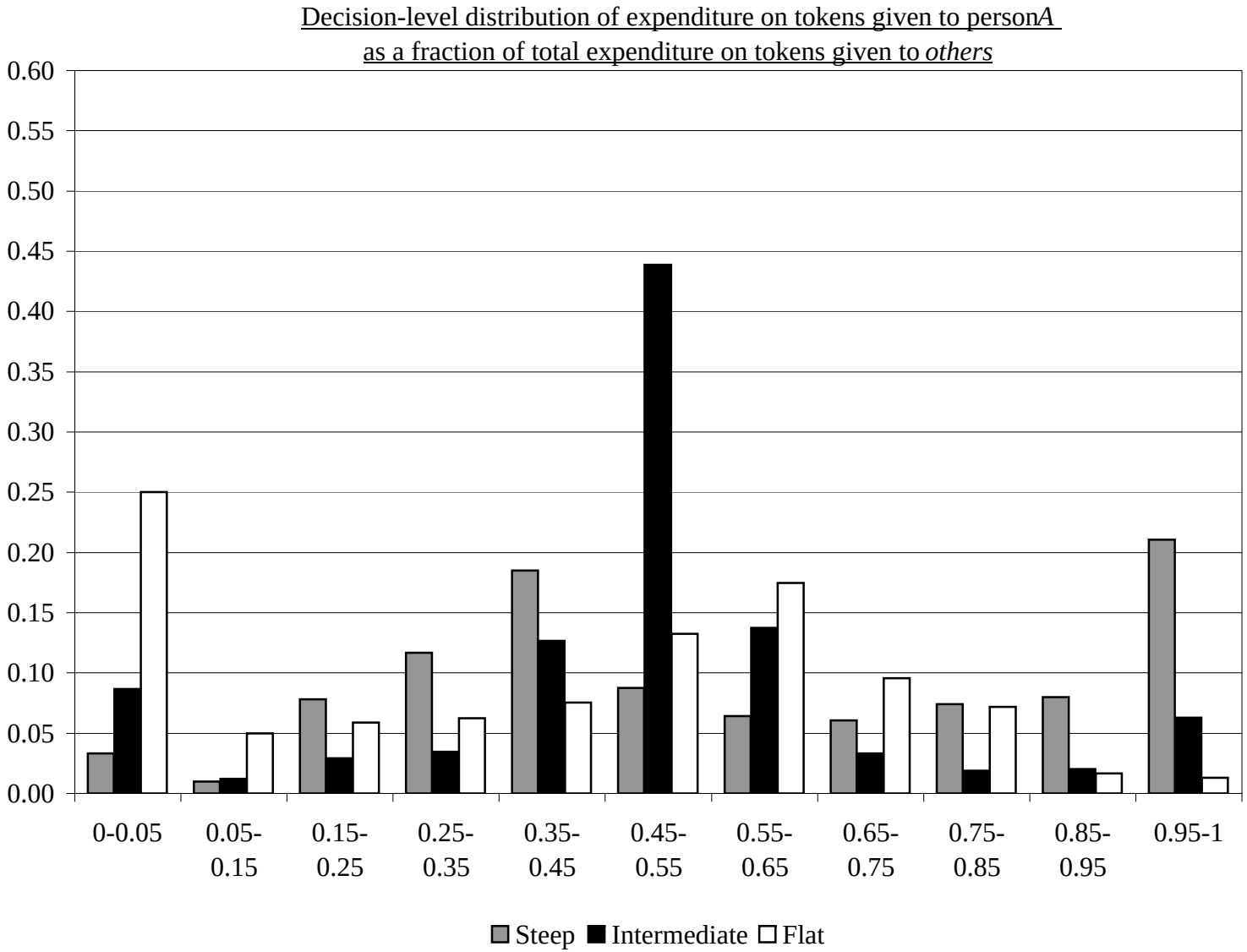
- This formulation makes it possible to represent distributional preferences in a particularly convenient manner.

- The macro function v_S represents preferences for altruism, whereas the subutility function w_S represents social preferences.

- Separability imposes convenient (but specific and quit restrictive) patterns on demand behavior (Karni and Safra 2002).


Decision-level distribution of tokens given to*others* as a fraction of total tokens kept and given
Fraction of decisions
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0-0.05
0.05-0.15
0.15-0.25
0.25-0.35
0.35-0.45
0.45-0.55
0.55-0.65
0.65-0.75
0.75-0.85
0.85-0.95
0.95-1
Three-person
Two-person


Decision-level distribution of expenditure on tokens given to *others* as a fraction of total expenditure on tokens
Fraction of decisions
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0-0.05
0.05-0.15
0.15-0.25
0.25-0.35
0.35-0.45
0.45-0.55
0.55-0.65
0.65-0.75
0.75-0.85
0.85-0.95
0.95-1
Three-person
Two-person


Decision-level distribution of expenditure on tokens given to person A as a fraction of total expenditure on tokens given to others
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0-0.05
0.05-0.15
0.15-0.25
0.25-0.35
0.35-0.45
0.45-0.55
0.55-0.65
0.65-0.75
0.75-0.85
0.85-0.95
0.95-1
Steep
Intermediate
Flat

Econometric specification

Suppose that w_S and v_S are members of the CES family:

$$w_S(\pi_O) = [\alpha' (\pi_A)^{\rho'} + (1 - \alpha')(\pi_B)^{\rho'}]^{1/\rho'}$$

and

$$v_S(\pi_S, w_S) = [\alpha (\pi_S)^{\rho} + (1 - \alpha) [w_s (\pi_O)]^{\rho}]^{1/\rho}$$

A family of CES functions that embed preferences for altruism and social preferences in a particularly convenient manner

$$U_S = [\alpha(\pi_S)^{\rho} + (1 - \alpha)[\alpha'(\pi_A)^{\rho'} + (1 - \alpha')(\pi_B)^{\rho'}]^{\rho/\rho'})]^{1/\rho}$$

The solution to the subutility maximization problem is given by

$$\pi_A(p_O, m_O) = \left[\frac{g'}{(p_B/p_A)^{r'} + g'}\right] \frac{m_O}{p_A}$$

where

$$r' = -\rho' / \left(1 - \rho'\right),$$

$$g' = \left[\alpha' / \left(1 - \alpha'\right)\right]^{1/\left(1-\rho'\right)}$$

and $m_O = p_O \pi_O$ is the total expenditure on tokens given to $others$.

The solution to the macro utility maximization problem is then given by

$$\pi_S(p, m) = \left[\frac{g}{q^r + g}\right] \frac{m}{p_S}$$

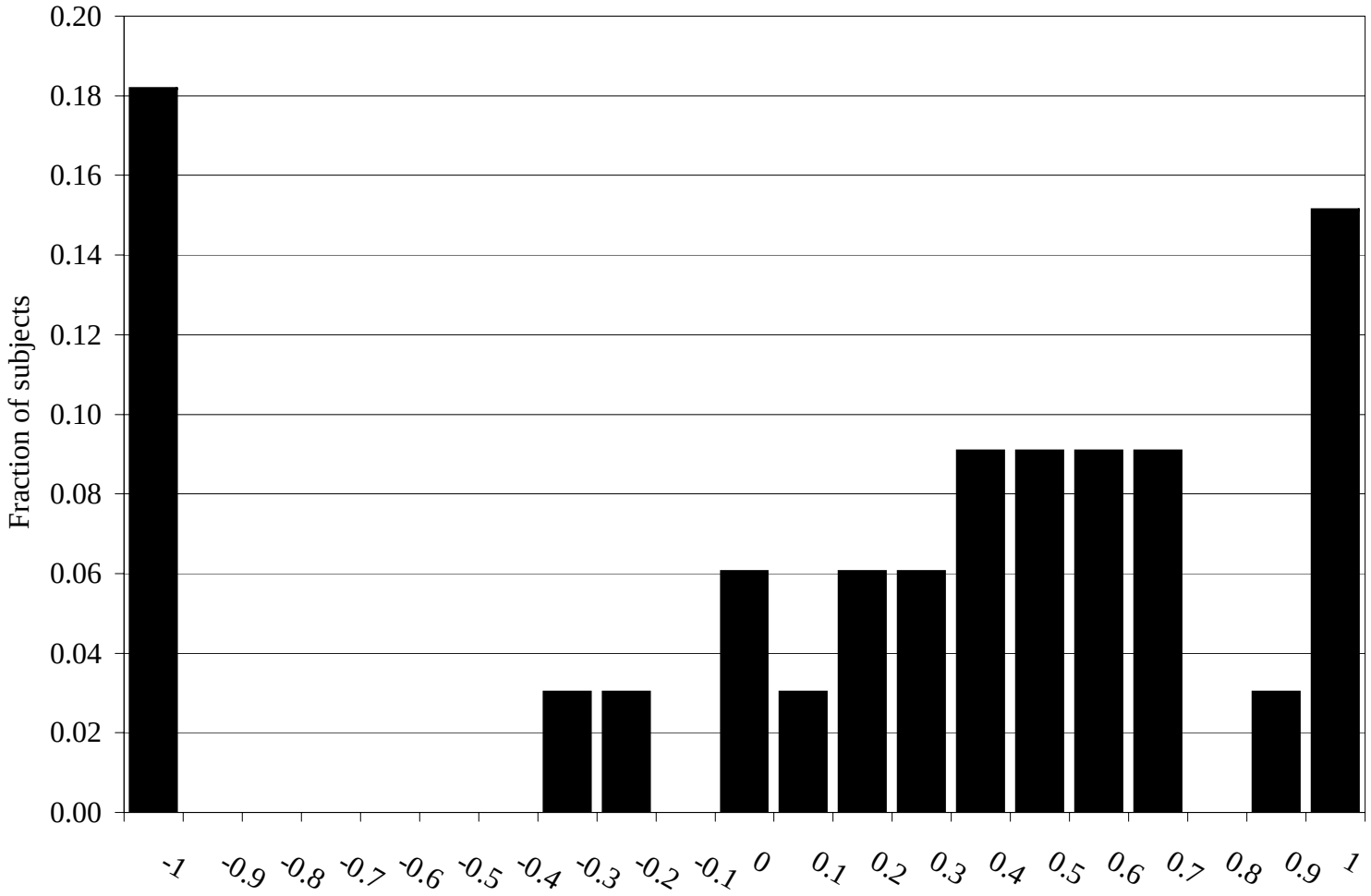
where

$$r = -\rho/(1 - \rho),$$

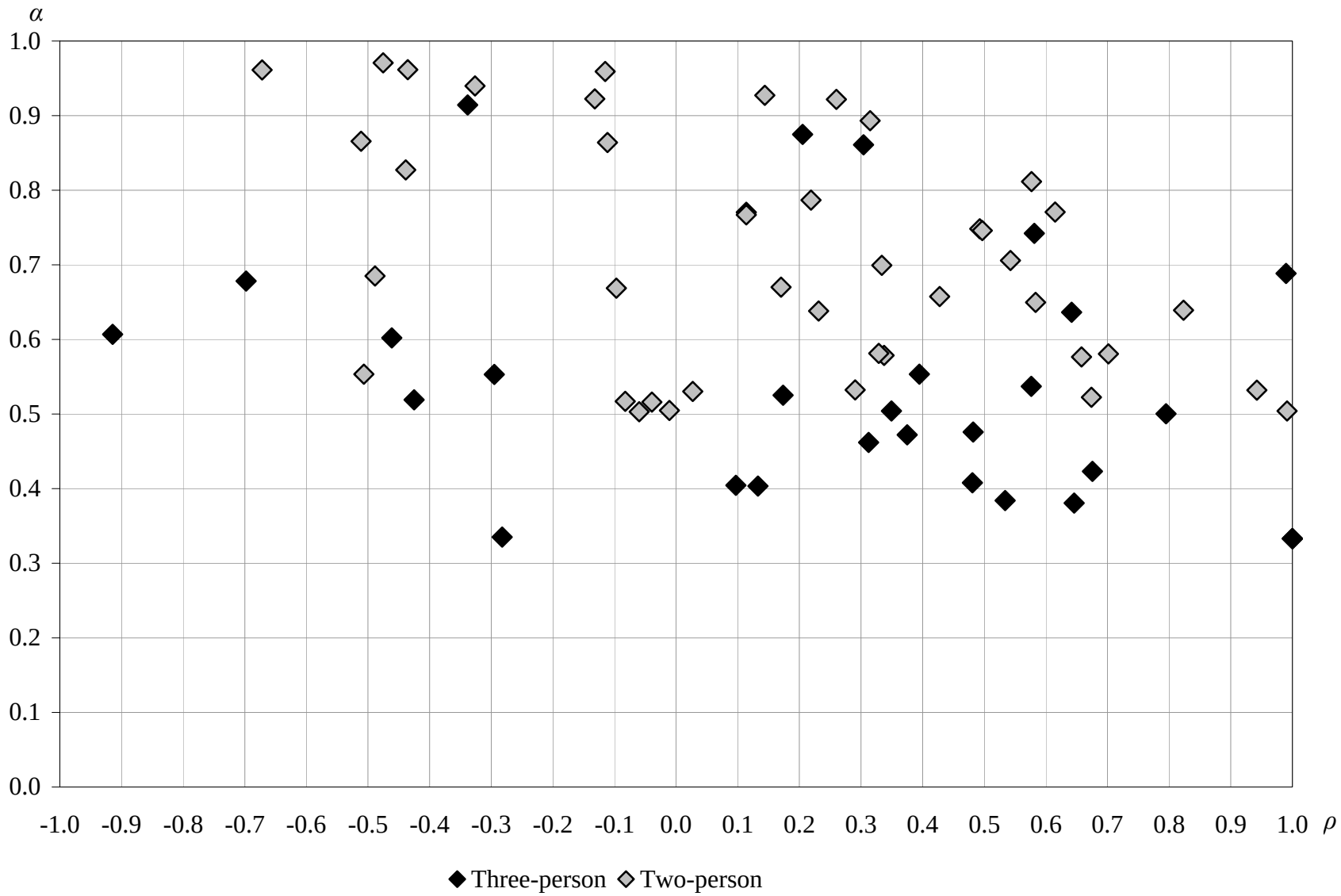
$$g = [\alpha/ (1 - \alpha)]^{1/(1-\rho)}$$

and q is a *weighted relative price of giving*.

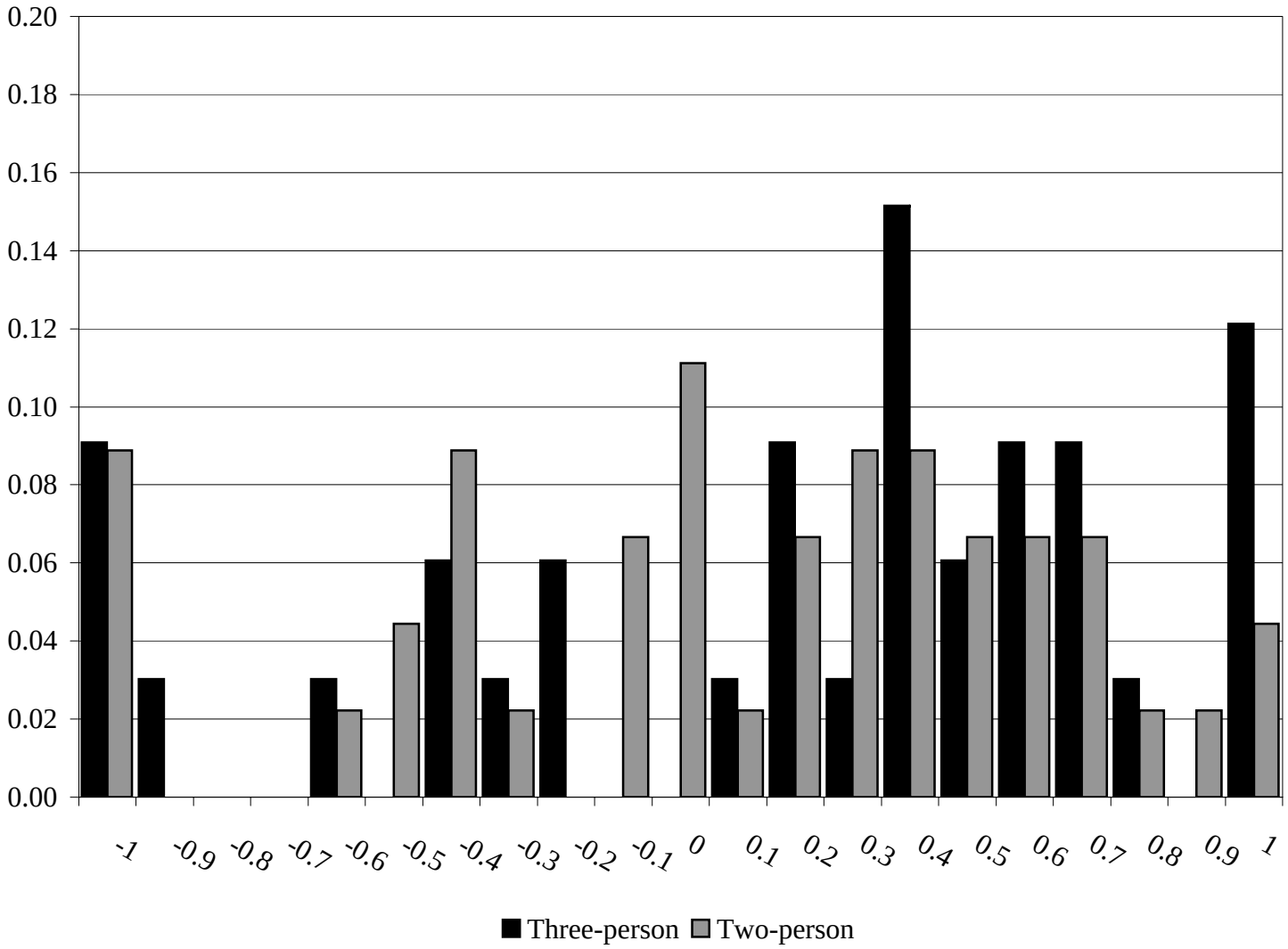
The distribution of the subutility CES parameter ρ'



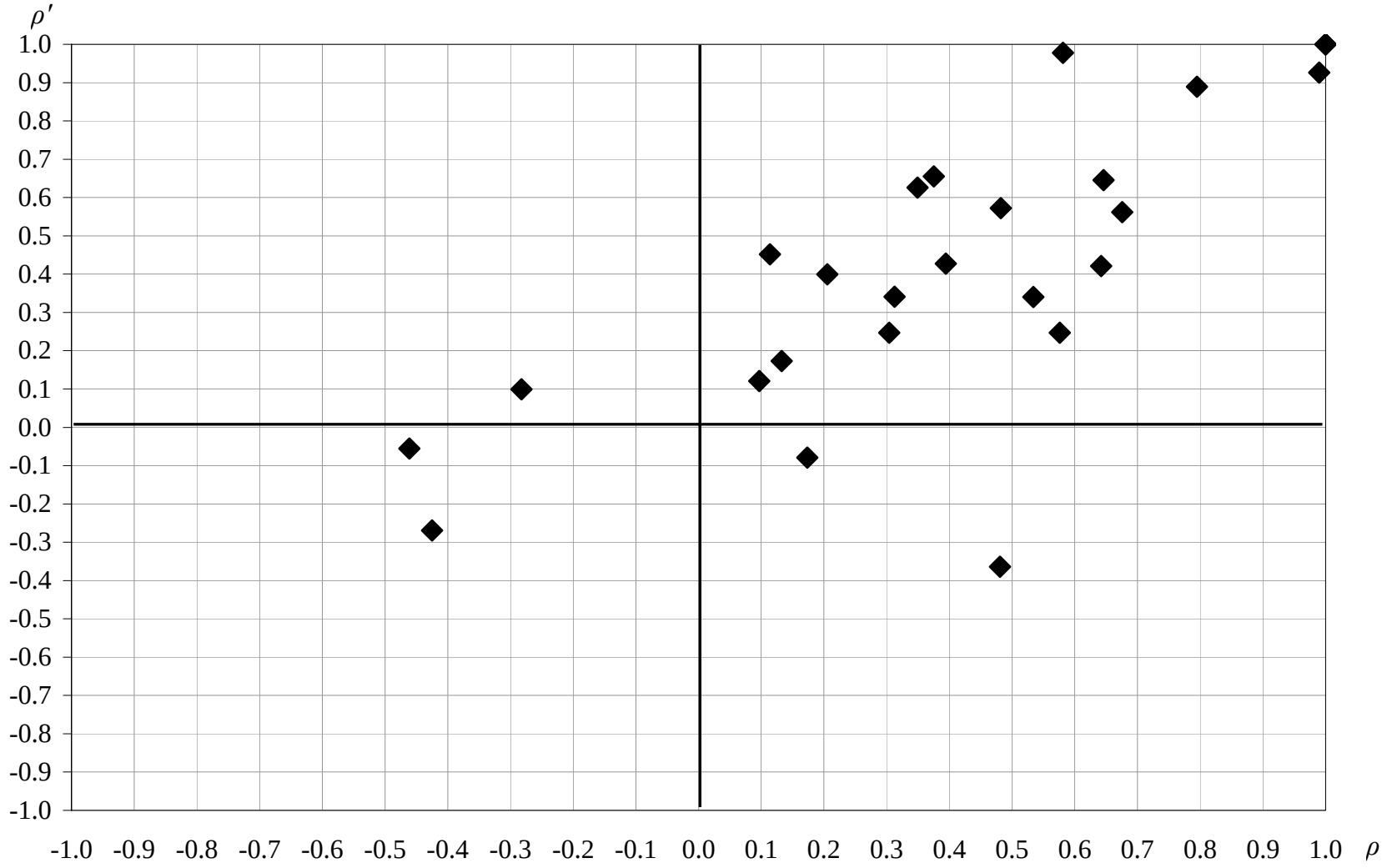
Scatterplot of the CES estimates ρ and α in the three- and two-person experiments



The distribution of the CES parameter ρ in the three- and two-person experiments



Scatterplot of the CES estimates ρ and ρ'



Takeaways

- Both social preferences and preferences for altruism are highly heterogeneous, ranging from utilitarian to Rawlsian.

- In spite of this heterogeneity across subjects, there exists a strong positive within-subject correlation.

- A strong correlation between the equality-efficiency tradeoffs subjects make in their altruistic and social preferences.

Moral preferences

Harsanyi and Rawls argue for theories of *social justice* (equivalently, *fairness*) based on the choices that agents would make for society in the *original position*, behind a *veil of ignorance*.

> *. . . without knowing their own social and economic positions, their own special interests in the society, or even their own personal talents and abilities (or their lack of them).* – Harsanyi (1975) –

Harsanyi and Rawls come to quite different conclusions, not because they view the original position differently, but because they treat uncertainty quite differently (Rawls denies orthodox decision theory).

Harsanyi's (1953, 1955) model for moral value judgments

Suppose an agent wants to make a moral value judgment about the relative merits of two alternative social systems.

> *. . . act in such a way as if he assigned the same probability to his occupying each social position under either system. . .*
>
> *. . . then, he would clearly satisfy the impartiality and impersonality requirements to the fullest possible degree.* – Harsanyi (1978) –

The agent has two different sets of preferences: *personal preferences* and *moral preferences* (preferences in the original position).

Two observations

[1] Both Harsanyi and Rawls insist that moral preferences must conform to certain rationality requirements, and hence must have a special form – as opposed to personal preferences, which merely reflect taste.

[2] Harsanyi and Rawls – and many other writers – view the original position as a purely hypothetical environment, and hence view moral preferences as a purely intellectual construct.

Our point of departure from the work of Harsanyi and Rawls – and the enormous literature they spawned – comes from two observations:

[1] Choice behavior/preferences *behind* the veil of ignorance can be decomposed into choice behavior/preferences *in front of* the veil of ignorance:

> choices that involve only personal consumption under uncertainty and choices that involve social consumption – but no uncertainty.

[2] Choices behind the veil of ignorance *can* be presented – and choices in the other two environments as well – in a controlled *laboratory* setting.

$\Rightarrow$ The linkage between preferences behind and in front of the veil of ignorance provides new ways of interpreting the theory of justice:

> not just as a *normative* theory, but also as a *descriptive* theory and even as a *prescriptive* theory.

$\Leftarrow$ This linkage means that moral preferences *cannot* occupy such a privileged position – modulo certain assumptions, they are completely determined by risk preferences and social preferences.

Template for analysis

- Consider choice behavior by a single agent in each of three environments.
- Each choice has consequences for *self* (the agent) and for an (unknown) *other*.
- We consider only environments that involve binary choices and equiprobable lotteries.
- The results extend to more general choices and lotteries, and to unknown probabilities as well.

Consider lotteries over outcomes $[a, b]$, where a is consumption for *self* and b is consumption for *other*.

For our purposes, it suffices to consider binary lotteries with equal probabilities:

$$(.5)[a, b] + (.5)[c, d]$$

where $a, b, c, d \geq 0$. Write $\mathcal{L}$ for the space of all such lotteries, and identify $\mathcal{L}$ with the convex cone $\mathbb{R}^4_+$.

Define closed convex subcones of $\mathcal{L}$:

$$\begin{aligned}
\mathcal{R} &= \{(.5)[a,0] + (.5)[c,0]\}, \\
\mathcal{S} &= \{(.5)[a,b] + (.5)[a,b]\}, \\
\mathcal{M} &= \{(.5)[a,b] + (.5)[b,a]\}.
\end{aligned}$$

We can interpret choice in each of the environments as choice in one of the corresponding cones by making an obvious identification:

– Risk: identify $\mathbb{R}^2_+$ with $\mathcal{R}$ by

$$(x, y) \mapsto (.5)[x, 0] + (.5)[y, 0].$$

– Social: identify $\mathbb{R}^2_+$ with $\mathcal{S}$ by

$$(x, y) \mapsto (.5)[x, y] + (.5)[x, y].$$

– Moral: identify $\mathbb{R}^2_+$ with $\mathcal{M}$ by

$$(x, y) \mapsto (.5)[x, y] + (.5)[y, x],$$

which coincides *exactly* with Harsanyi's (1953, 1955) formalization of the original position.

Research questions

[1] What is the relationship between moral preferences and personal/social (altruistic) preferences?

[2] How can behavior behind [Harsanyi's] veil of ignorance be characterized experimentally?

[3] Is behavior behind a veil of ignorance consistent with the utility maximization model?

[4] Can the underlying moral preferences be recovered from observed choices?

Assumptions

Given a preference relation $\succeq$ on $\mathcal{L}$, write $\succeq_{\mathcal{R}}$, $\succeq_{\mathcal{S}}$, $\succeq_{\mathcal{M}}$ for its restrictions to $\mathcal{R}$, $\mathcal{S}$, $\mathcal{M}$, respectively.

$[i]$ $\succeq$ satisfies the usual requirements: completeness, transitivity, reflexivity, continuity, and the Sure Thing Principle.

$[ii]$ $\succeq$ satisfies (weak) *independence*:

$$\begin{aligned} [a,b] \;\succeq\; & \mathcal{S}[a',b'] \text{ and } [c,d] \succeq_{\mathcal{S}} [c',d'] \\ \Rightarrow\; & (.5)[a,b] + (.5)[c,d] \succeq (.5)[a',b'] + (.5)[c',d'] \end{aligned}$$

(*not* the usual independence axiom and does not have the usual consequences).

Next, we make two assumptions about *social* preferences:

$[iii]$ **Worst outcome**: $[a, b] \succeq_{\mathcal{S}} [0, 0]$ for every $[a, b] \in \mathcal{S}$.

$[iv]$ **Self-regarding**: for each outcome $[a, b]$ there is an outcome $[s, 0]$ such that $[s, 0] \succeq_{\mathcal{S}} [a, b]$.

$[i]$ and $[ii]$ are rationality requirements (should not necessarily be given any philosophical interpretation).

$[iii]$ and $[iv]$ limit the extent to which the *self* is (respectively) spiteful or altruistic toward *other*; they seem very natural requirements but they are not entirely innocuous.

Result: Every preference relation $\succeq$ on $\mathcal{L}$ that satisfies $[i]$-$[iv]$ is determined by its restrictions $\succeq_{\mathcal{R}}$and $\succeq_{\mathcal{S}}$.

Proof: Fix an outcome $[x, y]$. Because $\succeq_{\mathcal{S}}$ is self-regarding, there is some s such that $[s, 0] \succeq_{\mathcal{S}} [x, y]$.

Define the *selfish equivalent* of $[x, y]$ by

$$\sigma[x, y] = \inf\{s : [s, 0] \succeq_{\mathcal{S}} [x, y]\}.$$

Continuity and worse outcome guarantee that $[\sigma[x, y], 0] \sim_{\mathcal{S}} [x, y]$, and by construction,

$$[a, b] \sim_{\mathcal{S}} [\sigma[a, b], 0] \text{ and } [c, d] \sim_{\mathcal{S}} [\sigma[c, d], 0].$$

independence guarantees that

$$(.5)[a,b] + (.5)[c,d] \sim (.5)[\sigma[a,b],0] + (.5)[\sigma[c,d],0].$$

Hence

$$\begin{array}{rcl} (.5)[a,b] + (.5)[c,d] & \succeq & (.5)[a',b'] + (.5)[c',d'] \\ & \Updownarrow & \\ (.5)[\sigma[a,b],0] + (.5)[\sigma[c,d],0] & \succeq & \mathcal{R}(.5)[\sigma[a',b'],0] + (.5)[\sigma[c',d'],0] \end{array}$$

which decomposes preferences over $\mathcal{L}$ into preferences over $\mathcal{S}$ (selfish equivalents) and preferences over $\mathcal{R}$, as desired.

Given a linear budget constraint, we identify choice behavior in the Social Choice environment as

- *selfish* if the choice subject to every budget constraint is of the form $[y, 0]$ – giving nothing to *other*.

- *symmetric* if (a, b) is chosen subject to $px+qy \leq w$ iff (b, a) is chosen subject to the mirror-image budget constraint $qx + py \leq w$.

Corollary I: If the preference relation $\succeq$ satisfies $[i]$ and $[ii]$ and choice behavior in the $\mathcal{S}$ is selfish then choice behavior in $\mathcal{R}$ coincides with choice behavior in $\mathcal{M}$.

Proof: Monotonicity and continuity guarantee that purely selfish behavior implies that $[x, 0] \sim_{\mathcal{S}} [x, y]$ for every x, y. independence implies that

$$(.5)[y, 0] + (.5)[x, 0] \sim (.5)[x, y] + (.5)[y, x].$$

It follows immediately that $\succeq_{\mathcal{R}}$ and $\succeq_{\mathcal{M}}$ coincide from whence choices in the Risk and Veil of Ignorance environments coincide, as asserted.

Corollary II: If the preference relation $\succeq$ satisfies $[i]$ and $[ii]$ and choice behavior in $\mathcal{S}$ is symmetric, then choice behavior in $\mathcal{S}$ coincides with choice behavior in $\mathcal{M}$.

Proof: Suppose that (a, b) is chosen from some budget set B for the Social Choice environment, so that (b, a) is chosen in the mirror image budget set B'.

Say that (c, d) is chosen from the budget set B for the Veil of Ignorance environment, and that $(c, d) \neq (a, b)$.

Because $(c, d) \in B$, it follows that

$$(.5)[c, d] + (.5)[d, c] \succ_{\mathcal{M}} (.5)[a, b] + (.5)[b, a].$$

independence implies that

$$[c, d] \succ_{\mathcal{S}} [a, b] \text{ or } [d, c] \succ_{\mathcal{S}} [b, a],$$

which is inconsistent with the fact that (a, b) (resp. (b, a)) is chosen from the budget set B (resp. B').

It follows that risk attitude is irrelevant in the Veil of Ignorance environment, as asserted.

Experimental analysis

- Subjects in the experiments were recruited from all classes at UCLA and Yale Law School.

- Each decision problem is presented as a choice from a two-dimensional budget line.

- A choice (x, y) from the budget line represents an allocation between accounts x, y (corresponding to the horizontal and vertical axes).

- Choices are made through a simple point-and-click design using a graphical computer interface.

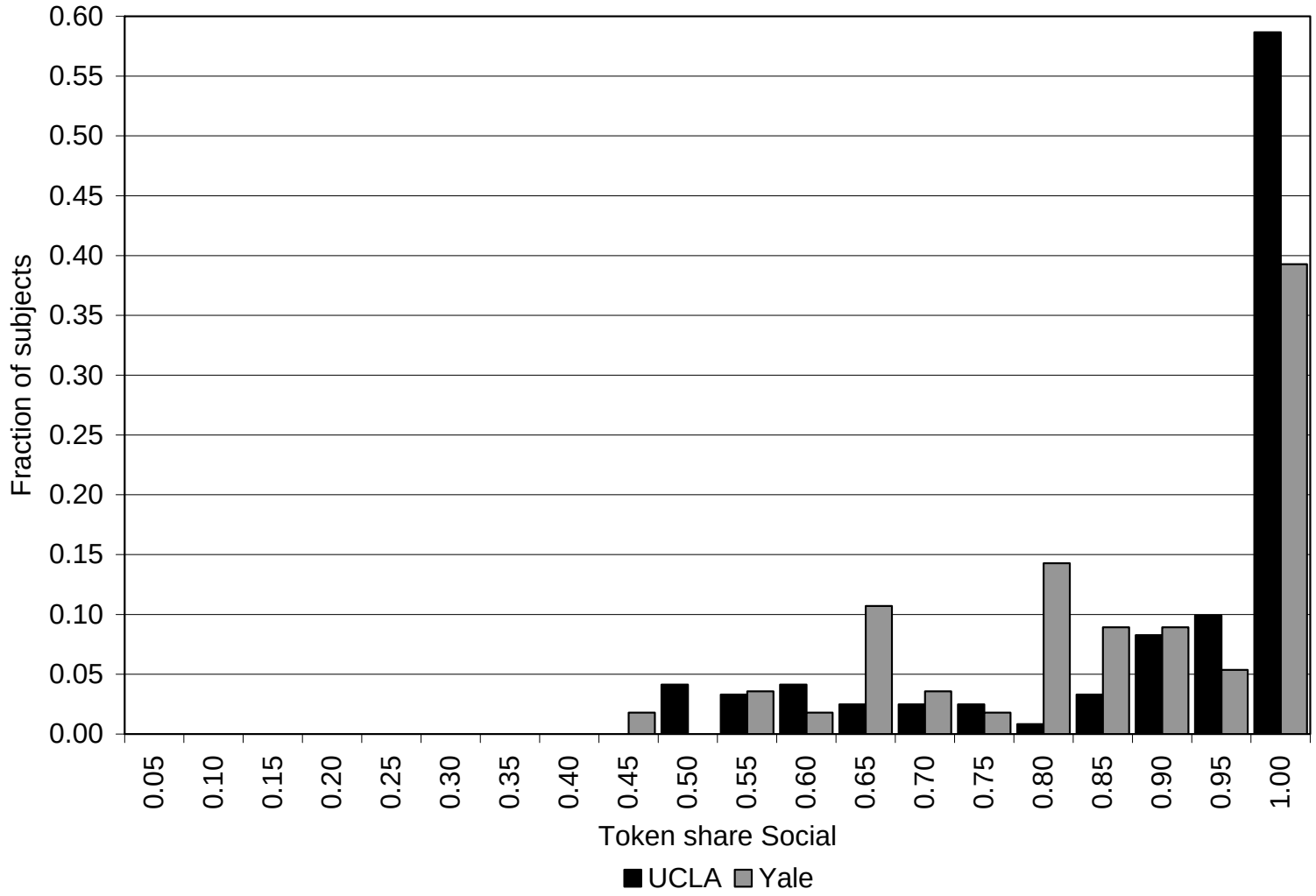
The actual payoffs of a particular choice in a particular environment/treatment are determined by the allocation to the x and y accounts:

- Risk: involves only pure risk; it is identical to the (symmetric) risk experiment of Choi, Fisman, Gale & Kariv (*AER*, 2007).

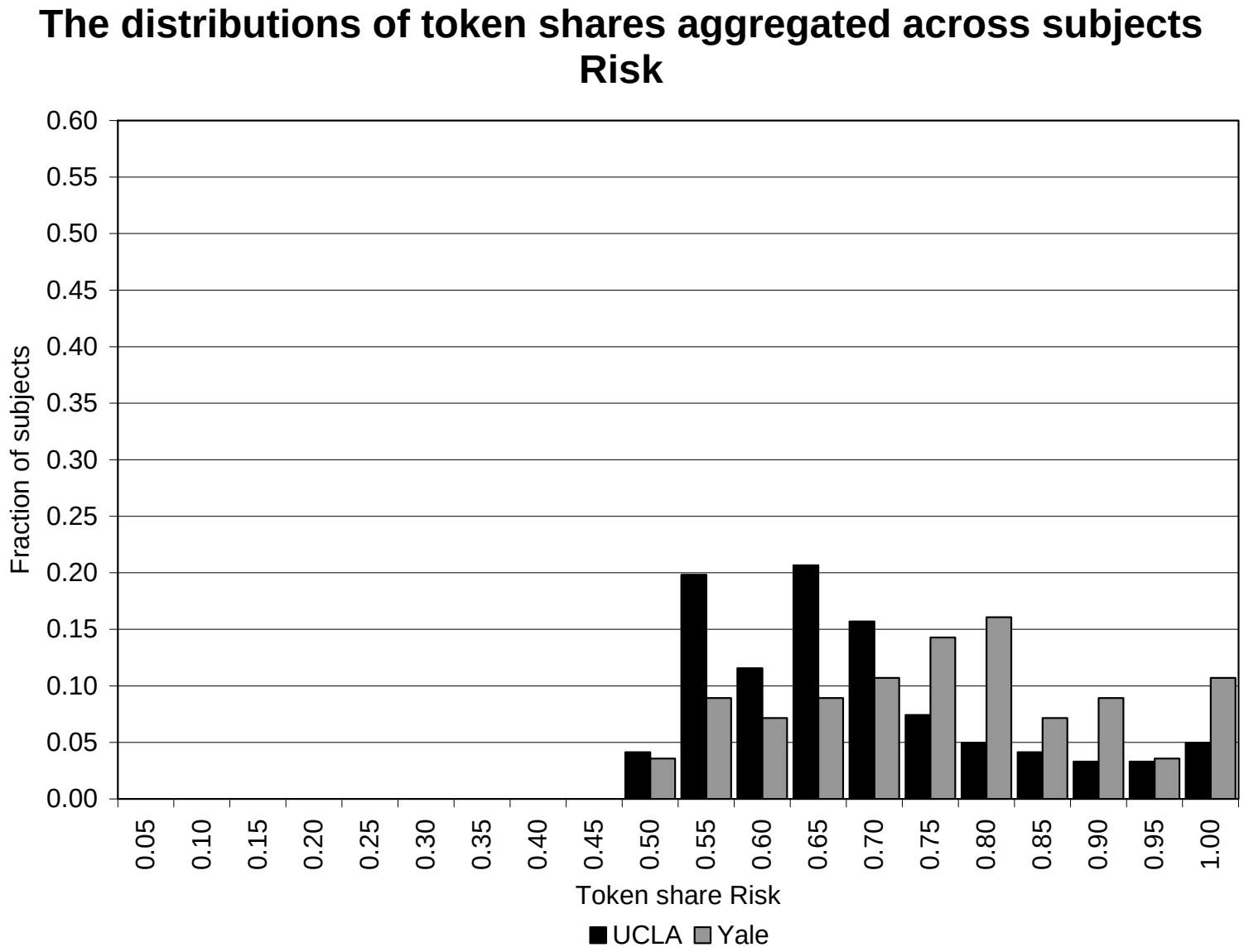
- Social Choice: involves only altruism; it is identical to the (linear) two-person dictator experiment of Fisman, Kariv & Markovits (*AER*, 2007).

- Veil of Ignorance: involves equiprobable binary lotteries over symmetric pairs of consumption for *self* and for *other*.

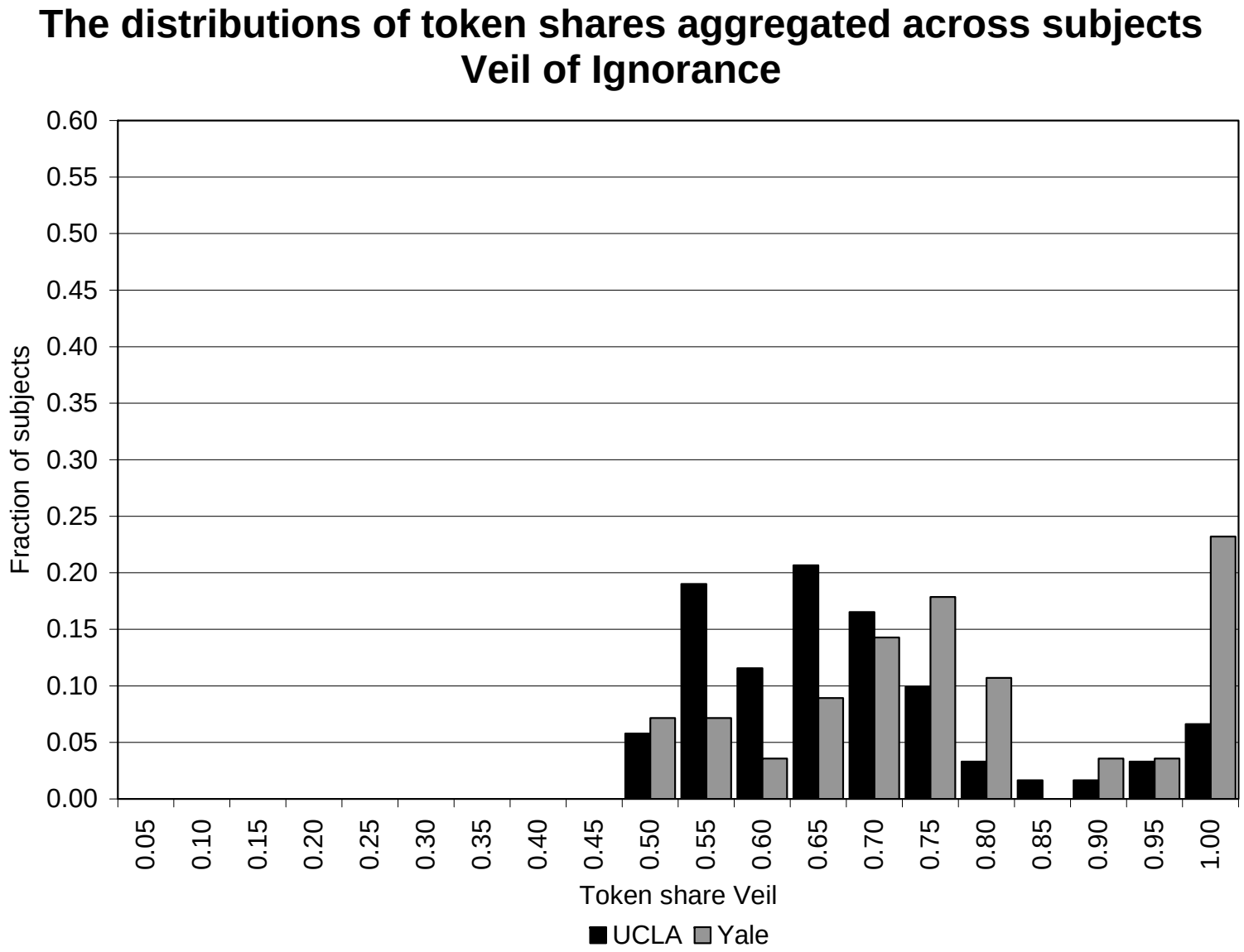
The distributions of token shares aggregated across subjects
Social Choice



The tokens kept as a fraction of the sum of the tokens kept and given to other.

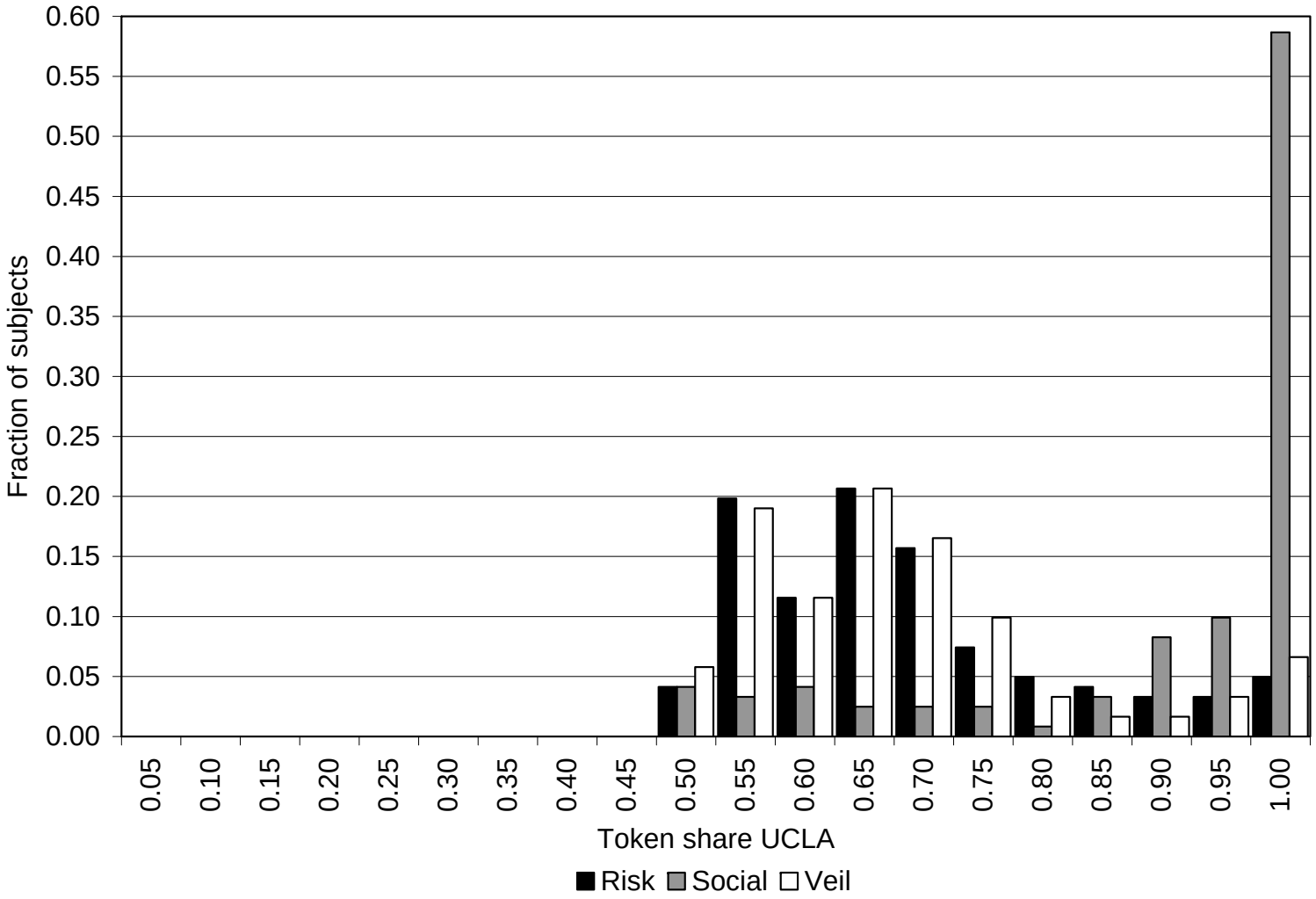


The fraction of tokens allocated to the cheaper account.



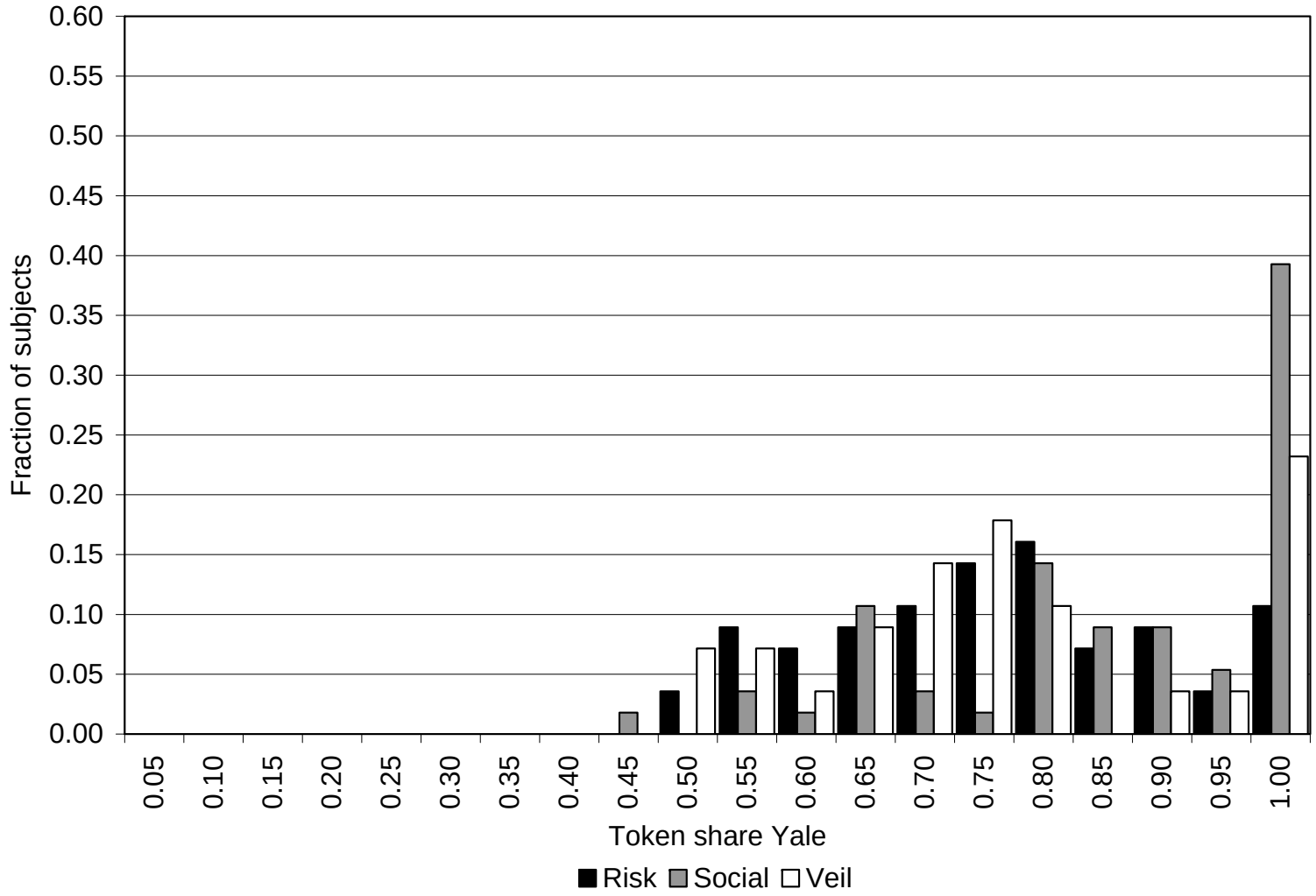
The fraction of tokens allocated to the cheaper account.

The distributions of token shares aggregated across subjects
UCLA



Social: fraction of tokens kept by self. Risk and Veil: fraction of tokens allocated to the cheaper account.

The distributions of token shares aggregated across subjects
Yale



Social: fraction of tokens kept by self. Risk and Veil: fraction of tokens allocated to the cheaper account.

Individual behavior

- The aggregate data tell us little about the choice behavior of individual subjects.

- Scatterplots of all choices of illustrative subjects – each entry plots $y/(x+y)$ as a function of $\log(p_x/p_y)$ in a particular treatment.

- There is no taxonomy that allows us to classify all subjects unambiguously.

- The characteristic of all our data is striking regularity *within* subjects and heterogeneity *across* subjects.

Testing the theory

- Many selfish subjects seem to display the same choice behaviors in the Risk and Veil of Ignorance environments, but a substantial number do not.

- Because of the nature of the data, "flexible" functional forms do not provide a plausible fit for the data.

- No satisfactory formulation to explain the "switching" between stylized behavior patterns exhibited by many subjects.

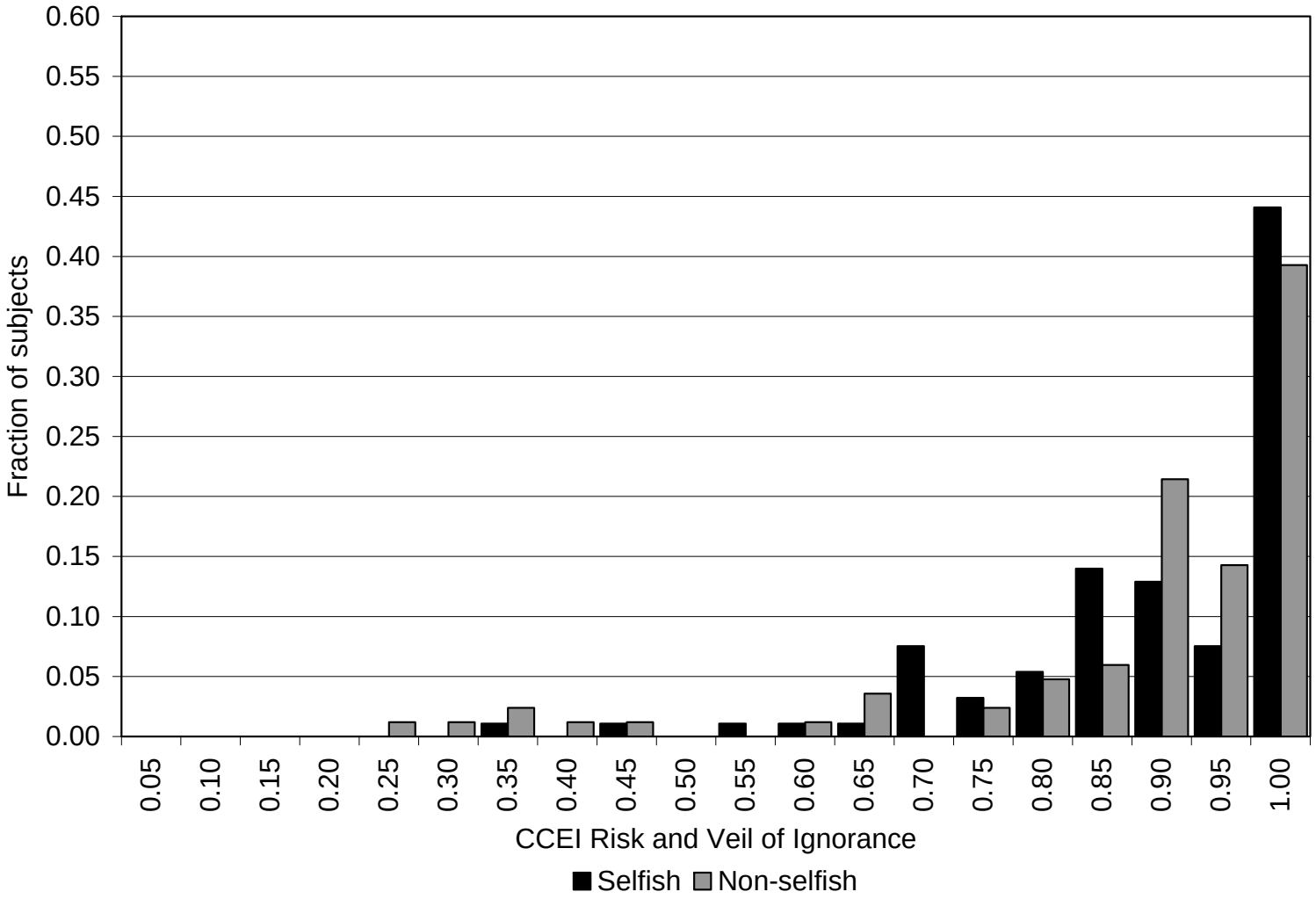
- Parametric approaches may be possible – keeping in mind that individual behaviors are extremely heterogeneous.

Non-parametric econometric approaches

Revealed preference

- The ratio of the CCEI score for the combined data set to the *minimum* of the CCEI scores for the separate data sets.
- A measure of the extent to which choice behaviors in any two environments coincide.
- Unfortunately, this test is weak – cannot discriminate between Risk and Veil of Ignorance behavior of selfish and non-selfish subjects.

The distributions of CCEI scores for the combined data set



Kolmogorov-Smirnov type tests

- A two-sample Kolmogorov-Smirnov tests of the equality of distributions of token and budget shares.

- The test is sensitive to differences in both location and shape of the empirical cumulative distribution functions of the two samples.

- Generalize the univariate Kolmogorov-Smirnov statistics for bivariate samples (Adler and Brown, 1986).

- There are subjects who fail Corollary I (selfish but display different behaviors in the $\mathcal{R}$ and $\mathcal{M}$) and others who fail Corollary II.

- These subjects might have preferences over $\mathcal{L}$ that do not obey independence (or might not be consistent with utility maximization).

- Individual preferences are very heterogeneous, ranging from utilitarian to Rawlsian.

- Actual preferences "mix-and-match" behavior in ways that no extant theory would regard as justified.

Takeaways

- A positive account of preferences for both personal and social consumption in rich choice environments.

- Two methodological contributions:

  - The establishment of theoretical links between preferences in various environments.

  - An experimental technique that allows for the collection of richer data about preferences.

- The experimental platform and analytical techniques are applicable to many other types of individual choice problems.

A field environment (FKM, 2009)

- YLS employs a mandatory first-term curriculum: constitutional law, contracts, torts, and civil procedure.

- Students are randomly assigned to classes taught by different instructors.

- There is no designated syllabus, and the several instructors are free (and indeed encouraged) to design their syllabus as they see fit.

- Substantial variation in contents, in particular with respect to conceptions of the economic role of the law.

YLS instructors

- We focus on contracts and torts – the courses with more substantive economic content:

  Economists emphasize efficiency, humanists in various ways emphasize equity, and other instructors are eclectic in their views and so fall in between these poles.

- The practical effects of these differences on teaching can be substantial both generally and, in particular, with respect to distributive questions.

Contracts

- Commercial relations among firms and efficiency of contract design versus contractual relations that involve individuals and equity among contract participants.

Torts

- The role of torts in making possible the efficiency gains of the marketplace versus torts as elaborating an individual ethic of care and responsibility.

- These differences may be *objectively* identified by looking to instructors' educational backgrounds and professional affiliations:

  Faculty with Ph.D. degrees in economics or humanistic disciplines are assigned accordingly, and those with only J.D. degrees receive a neutral assignment or an assignment based on professional affiliations.

- In the relevant period, a total of 16 instructors taught contracts and torts (all but one of these taught students in our sample).

Instructors' information

	Instructor	Classification	Highest degree
Torts	G. Calabresi	Economics	J.D., M.A. (Politics, Philosophy and Economics)
	J. Coleman	Philosophy	Ph.D. (Philosophy)
	J. Donohue	Economics	J.D., Ph.D. (Economics)
	D. Kysar	Neutral	J.D.
	P. Schuck	Neutral	J.D.
Contracts	I. Ayres	Economics	J.D., Ph.D. (Economics)
	L. Brillmayer	Neutral	J.D.
	R. Brooks	Economics	J.D., Ph.D. (Economics)
	S. Carter	Neutral	J.D.
	A. Chua	Neutral	J.D.
	R. Gordon	History	J.D.
	H. Hansmann	Economics	J.D., Ph.D. (Economics)
	D. Markovits	Philosophy	J.D., Ph.D. (Philosophy)
	C. Rose	Neutral	J.D., Ph.D. (History)
	A. Schwartz	Economics	J.D.

YLS students

- An aggregate measure of relative exposure to economic and humanist ideologies:

  $\boldsymbol{econ} = 1$ *(*$\boldsymbol{econ} = 0$*) if the subject was taught by at least one economist and no humanist (at least one humanist and no economist), and* $\boldsymbol{econ} = 0.5$ *if the subject was taught by neither an economist nor a humanist, or by one of each.*

- The 67 subjects in the experiment were recruited from the entire YLS student body.

The correlation between *econ* and subjects' individual characteristics

	Coefficient	Mean
Economics	0.100 (0.135)	0.152 (0.361)
Only child	0.075 (0.169)	0.091 (0.290)
Religious	-0.072 (0.100)	0.614 (0.487)
Male	-0.122 (0.096)	0.520 (0.510)
Log(Age)	-0.0070 (0.100)	3.217 (0.087)

The lab environment

- We consider two experimental treatments. The first treatment is identical to the (linear) two-person experiment of FKM (2007):
  - Subjects see budget lines on a computer screen and make choices through a simple point-and-click.
  - This allows for the quick and efficient elicitation of many decisions per subject under a broad range of budget lines.

$p_s\pi_s + p_o\pi_o = 1$ where π_s and π_o correspond to the payoffs to persons $self$ and $other$, respectively, and $p = p_o/p_s$ is the *relative price of giving*.

- The budget line configuration allows to identify the equity-efficiency trade-offs that subjects make in their distributional preferences:
  - *decreasing* $p_o\pi_o$ when p *increases* indicates distributional preferences weighted towards efficiency (increasing total payoffs)
  - *increasing* $p_o\pi_o$ when p *increases* indicates distributional preferences weighted towards equity (reducing differences in payoffs).
- In contrast, *indexical selfishness* is the relative weight on the payoff for $self$ π_s.

- The second treatment was identical to the first with the exception that the computer identified three allocations consistent with maximizing utilitarian preferences, log utility, and Rawlsian preferences:
  - *This allocation always lies at the endpoint of the line segment that is farthest from the origin. This maximizes the sum of payouts.*
  - *This allocation always lies at the midpoint of the line segment. The allocation gives you and the other person each half of your maximum feasible payout.*
  - *This allocation always lies on the 45 degree line. The payouts are the same to yourself and to the other person.*

Related literature

- Prior studies have faced two primary obstacles, which our experiments are designed to overcome:
  - Self-selection into a discipline and the learning that education in this discipline provides.
  - Behaviors motivated by pure self-interest and by distributional preferences concerning efficiency versus equity.
- Perhaps due to such confounding factors, findings in this literature have been mixed – economists are born or made?

Summary of results

[1] We extend the conclusion of FKM (2007) that distributional preferences are highly heterogeneous and range from Rawlsian to utilitarian to perfectly selfish.

[2] Subjects exposed to economics instructors place a greater emphasis on efficiency relative to those exposed to humanist or neutral instructors, who emphasize equity.

[3] Subjects exposed to economics instructors also display greater levels of indexical selfishness relative to those exposed to humanist instructors; those exposed to neutral instructors exhibit intermediate selfishness.

[4] In the second treatment, the labeled (prototypical) allocations were chosen more often, but the correlation between economics exposure and distributional preferences was unaffected.

Econometric analysis

Let $(p^t_{s,i}, p^t_{o,i})$ denotes the t-th observation of the price vector and $(\pi^t_{s,i}, \pi^t_{o,i})$ denotes the associated allocation.

Let $\chi^{\boldsymbol{n}}$ and $\chi^{\boldsymbol{e}}$ be indicator variables for neutral ($\boldsymbol{econ} = 0.5$) and economics ($\boldsymbol{econ} = 1$) subjects, respectively.

The main econometric specification has the expenditure function of the form:

$$p^t_{o,i}\pi^t_{o,i} = \beta_1 + \beta_2\chi^{\boldsymbol{n}}_i + \beta_3\chi^{\boldsymbol{e}}_i + [\gamma_1 + \gamma_2\chi^{\boldsymbol{n}}_i + \gamma_3\chi^{\boldsymbol{e}}_i]\,\mathsf{log}(p^t_i) + \epsilon^t_i$$

where ϵ^t_i is assumed to be distributed normally with mean zero and variance σ^2_n.

β represents the *indexical* weight on $self$ versus $other$ payoffs, whereas the γ parameterizes attitudes towards the efficiency-equity tradeoff between $self$ and $other$:

- $\gamma < 0$ indicates distributional preferences weighted towards efficiency (increasing total payoffs).
- $\gamma > 0$ indicates distributional preferences weighted towards equity (reducing differences in payoffs).

We generate estimates of the β and γ coefficients using a Tobit model, and use robust standard errors that allow for clustering at the level of the individual subject i.

Econometric results

	(1)	(2)	(3)	(4)
χ^n	-0.124	-0.126	-0.165***	-0.167***
	(0.096)	(0.095)	(0.0607)	(0.0599)
χ^e	-0.212**	-0.222**	-0.179***	-0.189***
	(0.102)	(0.103)	(0.0677)	(0.0699)
$\log(p)$	-0.104***	-0.059	-0.105***	-0.0405
	(0.027)	(0.052)	(0.0311)	(0.0607)
$\chi^n \log(p)$		-0.007		-0.0366
		(0.061)		(0.0717)
$\chi^e \log(p)$		-0.143**		-0.176**
		(0.071)		(0.0820)
Selfish subjects	Yes	Yes	No	No
# of obs.	3300	3300	2200	2200