Title
Three Essays on Experimental and Microeconomics

Author
Holder, Patrick James

Publication Date
2017

Peer reviewed|Thesis/dissertation

University of California

Santa Barbara

Three Essays on Experimental and Microeconomics

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy

in

Economics

by

Patrick James Holder

Committee in charge:

Professor Ryan Oprea, Chair
Professor Gary Charness
Professor Emanuel Vespa

December 2017

The Dissertation of Patrick James Holder is approved.

Professor Gary Charness

Professor Emanuel Vespa

Professor Ryan Oprea, Committee Chair

June 2017

Three Essays on Experimental and Microeconomics

by

Patrick James Holder

To my parents.

Acknowledgements

I'd like to acknowledge the assistance of Eduardo Zambrano, Sevgi Yuksel, Zachary Grossman, Robert Deacon, Cheng-Zhong Qin, KC Eames, Daniel Saunders, Heather Royer, Peter Rupert, the UC Santa Barbara Experimental Reading Group, and seminar participants at UC Santa Barbara, The Brattle Group, Colby College, and the 2016 North American Meeting of the Economic Sciences Association. Above all, I'd like to thank my dissertation committee – Ryan Oprea, Gary Charness, and Emanuel Vespa – for their guidance and encouragement.

Curriculum Vitæ

Patrick James Holder

Education

2017	Ph.D. in Economics (Expected), University of California, Santa Barbara.
2011	B.S. in Economics (emphasis in Quantitative Economics), California Polytechnic University, San Luis Obispo.

Publications

Kidney Coop: Proof of Concept (with KC Eames and Eduardo Zambrano)
Forthcoming in *The Journal of Health Economics*

Working Papers

Charity in the laboratory: Matching, competition, and group identity (with Gary Charness)
Revise and Resubmit in *Management Science*

Dynamic Investment and Preferences over the Resolution of Risk

Honors and Awards

UCSB Economics Department 2012 - 2013 Teaching Assistant of the Year

Ph.D. Pass With Distinction – Microeconomics Preliminary Exam

UCSB Economics Department Research Grant

Presentations and Workshops

2016 North American Meetings of the Economic Science Association; Tucson, Arizona

UCSB Experimental Reading Group (2014, 2015, 2016)

2015 IFREE Graduate Student Workshop

2013 Rady Spring School in Behavioral Economics

2013 UCE^3 Grad Camp

Abstract

Three Essays on Experimental and Microeconomics

by

Patrick James Holder

This dissertation contains three chapters on experimental economics and microeconomics. In the first chapter, *Dynamic Investment and Preferences over the Resolution of Risk*, I report the results of a laboratory experiment which attempts to explain the finding that individuals invest less in risky assets when risk is gradually resolved over time, rather than all at once. Though the literature has traditionally attributed this behavior to a cognitive error, Kőszegi and Rabin (2009) recently characterized this finding as the result of *non-standard preferences over the resolution of risk.* My results reject the traditional "cognitive errors" explanation in favor of Kőszegi and Rabin's "non-standard preferences" explanation. In the second chapter, *Kidney Co-operative: A Mechanism to Improve on Human Kidney Markets*, myself and coauthors propose a mechanism called the *kidney co-operative* which is designed to provide sufficient incentives to alleviate the human kidney shortage, while at the same time addressing the concerns regarding the potential losers to such a reform. We show that it is reasonable to expect that the number of transplants will be larger under the kidney co-operative mechanism than under either the status quo or the conventional market mechanism. In the third chapter, *Charity in the Laboratory: Matching, Competition, and Group Identity* myself and a coauthor study the effects of donation matching, competition, and group membership on charitable donations using a laboratory experiment. We find that providing matching donations to all subjects or having individuals compete for the privilege to have their donations matched (we match

the top half of donations in each session), raises donation levels modestly. However, arbitrarily assigning subjects to teams which competed for matching funds substantially raised donation levels. We appeal to the notions of group identity and team dynamics to explain our results.

Contents

Chapter 1

Dynamic Investment and Preferences over the Resolution of Risk

Patrick Holder

1.1 Introduction

Between 1889 and 1978, the average real return on the S&P 500 was 7% while the return on risk-free debt was less than 1%. Recreating such a large difference in returns within the standard Arrow-Debreu framework would require investors to have coefficients of relative risk aversion in excess of 30, whereas typical estimates suggest the coefficient is close to 1 (Mehra and Prescott, 1985). Like the S&P 500, many real-world assets have values which evolve slowly over time: stocks, bonds, and real estate, to name a few. Unlike standard "one-shot" gambles which resolve risk all at once, these investments provide feedback about their performance over the course of many periods.

A series of laboratory experiments suggest that it is this periodic feedback which is driving the large difference in returns between the S&P 500 and risk-free debt; researchers starting with Gneezy and Potters (1997)[1] have found that subjects invest less in risky assets as they evaluate financial outcomes more frequently. I call this finding *aversion to the gradual resolution of risk* (AGRR). As investors tend to monitor their stock portfolios closely (and thus receive frequent updates on financial outcomes), this large difference in returns has been attributed to AGRR[2].

Though AGRR has been widely regarded as a cognitive error, Kőszegi and Rabin (2009) recently introduced a preferences-based explanation for this empirical link between feedback frequency and willingness to invest in a risky asset. In this paper, I study whether or not Kőszegi-Rabin preferences can account for AGRR in laboratory subjects using two experimental tasks. The first task classifies subjects according to the preferences that Kőszegi and Rabin argue are responsible for the behavior observed by Gneezy and Potters (1997). The second task – a replication of Gneezy and Potter's experiment – is a simple dynamic investment problem, in which the frequency of feedback about the value of a risky asset is varied between treatments.

I exploit natural heterogeneity in subject preferences elicited in the first task to study whether these preferences can be used to predict investment levels in the second task. If the negative correlation between feedback frequency and investment is observed most often in those subjects who have non-standard Kőszegi-Rabin preferences, we can conclude that these preferences contribute to AGRR. However, if this negative correlation is present in all subjects – even those who have standard "expected utility" preferences in the first task – it would be likely that the cognitive error explanation is correct.

[1]Thaler et al. (1997) concurrently conducted similar research.
[2]Benartzi and Thaler (1995)

Benartzi and Thaler (1995) were the first researchers to suggest that individuals have AGRR. Their theory of *myopic loss aversion* relies on the interaction of two behavioral concepts: *loss aversion*[3], which asserts that individuals weigh losses more heavily than commensurate gains, and *narrow bracketing*, a process in which people consider the utility changes from multiple gambles in isolation, rather than considering the overall effect on utility of the resulting compound gamble.[4] As an example, consider an investor who faces the one-time opportunity to save for retirement by distributing a fixed wealth across two assets: a risky asset which pays higher average returns, and a safe asset which pays lower returns. This investor neither adds to nor subtracts from the portfolio (besides reinvesting all earnings) until the day he retires, at which point he consumes the entire value of the investment. He does, however, check the portfolio's performance once a year in the interim. Such an investor would maximize expected (discounted) utility by considering the expected utility from the distribution of returns that each portfolio induces on the day that he retires. A myopically loss averse investor, however, would choose his portfolio as if to maximize utility from wealth a year from now when he checks-in on his investment. Such behavior is called *narrow bracketing.* By assuming that people are loss averse over changes in wealth, Benartzi and Thaler show that narrow bracketing drives investors towards portfolios which mitigate potential short-term decreases in wealth, making them appear extremely risk averse.

Commentators like Tversky and Kahneman (1985), Kahneman and Lovallo (1993), and Read, Loewenstein, and Rabin (1999) regard narrow bracketing as a cognitive error – a mistake which inhibits utility maximization in environments with disaggregated risk. This cognitive error narrative is so ingrained that investment services regularly warn

[3]Kahneman and Tversky (1979), Tversky and Kahneman (1992)

[4]Benartzi and Thaler originally referred to the second factor as *mental accounting* (Thaler, 1985). I follow later authors like Read, Loewenstein, and Rabin (1999), by using the term *narrow bracketing.*

clients against the dangers of checking their portfolios too regularly, with some firms actively taking measures to prevent portfolio evaluations which they few as "too frequent".[5]

Kőszegi and Rabin (2009) provide an alternate, preferences-based, account of AGRR by presenting a model in which subjects have non-standard preferences over the resolution of risk. To provide an example of how these preferences might affect an agent's decisions, consider two gambles, presented to an agent in period t, each with identical probability distributions over earnings. Winnings in either gamble are received at a later time, period $t + 1$. The gambles differ only in when they resolve their risk: one gamble reveals its outcome in period t, while the other reveals its outcome more gradually, over the course of t and $t + 1$. If the agent doesn't make plans contingent on information learned in period t,[6] the standard economic model asserts that the agent must be indifferent between these gambles. An agent who preferred the one-shot gamble to the drawn-out gamble would have AGRR for purely preference-based reasons. If Kőszegi and Rabin are correct in their assertion that AGRR is driven entirely by preferences, investors may, in fact, choose their portfolios and information about financial outcomes "correctly" (i.e. to maximize utility). This would suggest that the push by investment firms to influence their clients' investment behavior may be unnecessary at best, and welfare decreasing at worst.

Research on preferences which allow agents to not be indifferent between the above gambles dates back to Kreps and Porteus (1978). These preferences have enjoyed significant attention from both theorists and empiricists studying a wide range of topics. For exam-

[5]Popular online investment services with such warnings include Betterment (https://www.betterment.com/resources/investment-strategy/behavioral-finance-investing-strategy/high-frequency-monitoring/) and Wealthfront (https://blog.wealthfront.com/often-check-portfolio/). The latter only allows clients to modify their portfolios once a month in the hopes that this will prevent those clients from checking their portfolios more frequently.

[6]That is, if this feedback is "noninstrumental".

ple, Epstein and Zin (1991) use these preferences to model intertemporal consumption choice, while Ely, Frankel and Kamenica (2015) use them to explain demand for entertainment goods like professional sports and mystery novels. The classification task doesn't only allow subjects to be classified as "expected utility" or "Kőszegi and Rabin" types; the experimental design uses a novel methodology to simultaneously compare several preference types at once. Authors like Ahlbrecht and Weber (1997), Lovallo and Kahneman (2000), and Falk and Zimmerman (2014) have previously performed similar classification exercises, however, the classification task contained herein allows for the comparison of a larger number of models than previous research has compared.

I find that there is significant heterogeneity in subjects' preferences over the resolution of risk, with both the standard expected utility model and Kőszegi and Rabin's preference model describing a large number of subjects in the classification task (26% and 36% of subjects, respectively). Moreover, I find that investments in the second task do appear to be driven by these preference differences, suggesting that AGRR is the result of Kőszegi-Rabin preferences rather than myopic loss aversion.

Section 2 reviews the experimental evidence in support of AGRR and describes a number of preference models which generate non-standard preferences over the resolution of risk, focusing on Kőszegi and Rabin's (2009) model. It concludes by presenting a simple framework which captures the predictions of these models. Section 3 describes an experimental design which consists of two tasks, a *classification task* and an *investment task*. It also presents a list of testable hypotheses which inform the experimental design. Results from this design are reported in section 4. Section 5 concludes.

1.2 Background & Theory

Benartzi and Thaler's (1995) theory of myopic loss aversion makes the testable prediction that as the frequency of feedback about the performance of a risky asset increases, the investment in that asset decreases. This assumption, which I call *aversion to the gradual resolution of risk* (AGRR) was first tested in a simple experiment developed by Gneezy and Potters (1997). Their subjects were given the option to invest in a risky asset which reveals its value over the course of nine periods. In one treatment, subjects receive feedback about the value of the asset after each of the nine periods, while in the other treatment, subjects only receive feedback once every three periods. As predicted by Benartzi and Thaler (1995), the researchers found that investments were significantly lower in the frequent feedback treatment, suggesting that laboratory subjects do, in fact demonstrate AGRR. Gneezy and Potters interpret this result as evidence in favor of myopic loss aversion.

This result was extended in two further studies. Haigh and List (2005) performed the same experiment with both standard subjects (undergraduate college students) and professional futures and options traders. They found that the professional traders exhibited AGRR to a greater extent (that is, experienced a larger treatment effect) than standard subjects, suggesting that exposure to risky assets does not attenuate AGRR. Additionallly, Bellemare et al. (2005) refined the original Gneezy and Potters result by showing that AGRR is driven by the frequency of *feedback*, rather than the frequency with which an investor can adjust his portfolio.

1.2.1 Preferences over the Resolution of Risk

Kőszegi and Rabin (2009) provide an alternative account of the AGRR observed by Gneezy and Potters and others. They argue that AGRR is purely a preference-based phenomenon. Their model makes two assumptions: first, that beliefs about future consumption affect individuals' utility, and second, that individuals are loss averse to changes in these beliefs. Specifically, Kőszegi and Rabin assume that agents in period t derive instantaneous utility from consumption, c_t, and beliefs about consumption in future periods, $F_{t,\tau}$ (where future periods are indexed by τ). Utility is assumed to take the following form[7]:

$$u_t = m(c_t) + \sum_{\tau=t}^{T} \gamma_{t,\tau} N(F_{t,\tau}|F_{t-1,\tau})$$

$m(c_t)$ is the standard reference-independent utility from consumption. $F_{t,\tau} \in \Delta(\mathbb{R}^+)$ [8] are the beliefs held in period t about period-τ consumption, and $N(F_{t,\tau}|F_{t-1,\tau})$ captures gain-loss utility generated by changes in beliefs from period $t-1$ to period t about these future consumptions. Agents make "ordered comparisons" between their past and current beliefs, comparing the highest percentile consumption in $F_{t,\tau}$ to the highest percentile consumption in $F_{t-1,\tau}$, the second-highest percentile consumption in $F_{t,\tau}$ to the second highest in $F_{t-1,\tau}$, etc. Agents are loss averse in all such comparisons.[9]

[7]This treatment of Kőszegi and Rabin's model makes some simplifying assumptions for the sake of exposition. Appendix A presents the full model, in addition to providing an illustrative example.

[8]Where $\Delta(\mathbb{R}^+)$ is the set of all probability distributions over $\mathbb{R}^+$

[9]Formally, for $p \in (0,1)$, if $c_F(p)$ is the consumption level at percentile p:

$$N(F_{t,\tau}|F_{t-1,\tau}) = \int_0^1 \mu[m(c_{F_{t,\tau}}(p)) - m(c_{F_{t-1,\tau}}(p))]dp$$

Where $\mu(\cdot)$ is the a standard piecewise loss-aversion function (with $\lambda > 1$):

$$\mu(x) = \begin{cases} x & x \geq 0 \\ \lambda x & x < 0 \end{cases}$$

Next, note the $\gamma_{t,\tau}$ parameter in the above utility function. This parameter "determine[s] the relative importance of news as a function of how far in advance of consumption the news is received" (Kőszegi and Rabin, pg. 913). The authors assume that $\gamma_{\tau,\tau} \geq \gamma_{\tau-1,\tau} \geq \gamma_{\tau-2,\tau} \geq ... \geq \gamma_{0,\tau} \geq 0$, and they specifically discuss two cases which obey these inequalities. When $\tau - t$ is large – that is, when beliefs about consumption in a very distant period change – the effect on utility might reasonably be assumed to be small. As the consumption period draws nearer, changes in beliefs about consumption in that period might have a more profound effect on utility. In this case, $\gamma_{\tau,\tau} > \gamma_{\tau-1,\tau} > \gamma_{\tau-2,\tau} > ... > \gamma_{0,\tau} > 0$. However, if this "news utility" is invariant to the temporal distance between the change in the beliefs and the consumption actually taking place, $\gamma_{t,\tau} = 1 \; \forall t, \; \tau$.[10]

Kőszegi and Rabin's model makes two primary predictions. First, agents prefer information about future consumption all at once, rather than receiving a trickle of information over several periods. This is the result of agents' loss aversion towards changes in beliefs. Intuitively, beliefs over future consumption which bounce back-and-forth repeatedly without changing much on net – as might be induced by the frequent monitoring of a stock portfolio – will be utility decreasing in the aggregate, since small down-ticks hurt more than up-ticks of equal size are enjoyable.

Second, if $\gamma_{\tau,\tau} > \gamma_{\tau-1,\tau} > \gamma_{\tau-2,\tau} > ... > \gamma_{0,\tau} > 0$, agents prefer information to be received earlier rather than later. This is intuitive in the case where there is an equal chance of beliefs moving up or down. Since loss aversion makes this change in beliefs utility-decreasing in expectation, the agent would prefer to receive the news as far away as

[10]To complete their preference model, Kőszegi and Rabin assume that, in period t, agents maximize the non-discounted sum of their instantaneous utilities,

$$U^t \equiv \sum_{\tau=t}^{T} u_\tau$$

possible from the period in which consumption will take place.

Kőszegi and Rabin aren't the only authors to model preferences over when risk is resolved. Next, I briefly introduce other prominent models of these non-standard preferences. As Kőszegi and Rabin isn't the only alternative to the standard expected utility model, I develop a simple framework which allows subjects to be classified as one of the various preference types predicted by these models. This framework is introduced at the end of the section. The interested reader can find much greater detail about these models in appendix A, including an example which illustrates the mechanics behind each model.

Like Kőszegi and Rabin; Ely, Frankel, and Kamenica (2015) assume that agents derive utility from changes in beliefs. These authors assume that changes in beliefs generate "entertainment utility", which is assumed to always increase when beliefs change, even for beliefs which update towards a less favorable outcome. As in Kőszegi and Rabin, utility is assumed to change by more the larger is the update in beliefs, however, agents display diminishing sensitivity towards changes in these beliefs. Due to this diminishing sensitivity, agents would always prefer to have uncertainty resolved over more periods rather than fewer.

Kreps and Porteus (1978) were the first authors to consider preferences over the resolution of uncertainty which differ from the standard model. Instead of assuming that changes in beliefs enter directly in the utility function (as in the previous models), Kreps and Porteus generate non-standard preferences by assuming that agents apply a distortion function to the utilities of possible outcomes before calculating expected utilities. Depending on the shape of this distortion function, agents can display a preference for uncertainty to be resolved either as early as possible, or as late as possible.

Individuals' preferences can be thought of as spanning two dimensions. The first dimension, which I refer to as *concentration*, dictates a person's preference for resolving uncertainty all at once or spread-out across multiple periods. Kőszegi and Rabin's model predicts that subjects display a preference for risk to be resolved all at once, which I call *clumped* preferences. Ely, Frankel, and Kamenica's (2015) model, on the other hand, predicts that subjects display a preference for risk to be resolved over multiple periods; I call these *piecemeal* preferences[11].

The second dimension, which I refer to as *timing*, describes a subject's preference to resolve uncertainty earlier or later. The model of Kreps and Porteus (1978) can rationalize both a preference for risk to be resolved as early as possible – which I refer to as an *early* preference – or as late as possible – which I refer to as a *late* preference. Kőszegi and Rabin's model can also generate *early* preferences with $\gamma_{t,\tau}$'s which are decreasing in $t-\tau$ (see section 2.2.1). Individuals with Kőszegi-Rabin utility can thus display clumped preferences, early preferences, or preferences for both clumped and early resolution of risk. These models' predictions are summarized in table 1.

Table 1: Preference Types

		Timing		
		Early	Neither	Late
Concentration	Clumped	KR	KR	-
	Neither	KR/KP	EU	KP
	Piecemeal	-	EFK	-

KR denotes Kőszegi and Rabin's (2009) preference types, EFK denotes Ely, Frankel, and Kamenica's (2015) types, KP denotes Kreps and Porteus's (1978) types, and EU denotes standard expected utility types. Note that, under the standard model, agents are indifferent towards the timing of the resolution of risk.

The classification exercise categorizes subjects based on which of the nine cells in table 1

[11] I borrow the terms *clumped* and *piecemeal* from Falk and Zimmermann (2014).

subjects' preferences correspond with. Thus, some subjects will be assigned types which do not match the predictions of any of the models. For example, a mass of subjects are classified as having late *and* piecemeal preferences (the lower-righthand cell), though no one model generates those preferences. These subjects' choices could, however, be rationalized with a utility function which has features of both Kreps and Porteus's (1978) and Ely, Frankel, and Kamenica's (2015) models.

A handful of authors have conducted classification experiments on either the concentration or timing dimensions in isolation, but this research is, to my knowledge, the first to classify subjects on both dimensions simultaneously. The concentration dimension has been studied by Zimmermann (2014) and Falk and Zimmermann (2014). Zimmermann (2014) elicits willingnesses to pay for gambles which reveal their outcomes in a clumped or piecemeal manner using a between-subjects design. He finds no difference in willingnesses to pay in the aggregate. Falk and Zimmermann (2014), on the otherhand, find that 90% of subjects prefer the clumped resolution of risk to piecemeal resolution when subjects are learning about whether or not they will receive a painful electric shock.

Chew and Ho (1994), Ahlbrecht and Weber (1997), and Lovallo and Kahneman (2000) find that, on the timing dimension, the majority of subjects prefer the early resolution of risk to the late resolution. These studies use exclusively hypothetical subject responses. Moreover, they all consider environments in which there is a planning benefit to resolving risk early (that is, these experiments dealt with instrumental information), which could result in a preference for early resolution even within the expected utility framework. Kocher, Krawczyk, and Van Winden (2014) find more preference heterogeneity in an experiment with incentivized subject responses, however, these authors also consider an environment in which information is potentially instrumental.

Kőszegi and Rabin's preferences – the central topic of this paper – span both the concentration and the timing dimensions. Thus, it is important that the preference elicitation task is able to distinguish across preferences on both dimensions simultaneously in order to detect the full set of subjects with preferences consistent with Kőszegi and Rabin's model. The preference elicitation task and the investment task are detailed in the following section.

1.3 Experimental Design & Hypotheses

This section first describes an experimental design motivated by the previous section. It then presents a series of hypotheses which address the underlying cause of AGRR.

1.3.1 The Classification Task

An experimental session consists of two back-to-back tasks. The first task, called the *classification task*, classifies subjects according to their preferences over the resolution of risk, while the second, the *investment task* is a replication of the dynamic investment problem from Gneezy and Potters (1997). Subjects were not aware of the second task during the first.

The first task serves to classify subjects as one of several types of agents identified in the literature. This classification is carried-out by eliciting from each subject a preference ordering over four gambles; these gambles differ only in when they resolve their risk, they are otherwise identical. Each pays $50 at the end of the experiment with probability 25.1% and pays nothing otherwise. Whether or not the $50 outcome obtains is determined by drawing two values, x and y, from the discrete uniform distribution $\{1, 2, 3, ...100\}$. If $x + y \geq 130$, the subject receives the \$50 payment. The gambles differ only in when x and y are revealed to the subjects. Section 1B of the classification task (see figure 1) is the hour long period over which x and y are revealed. The numbers can be revealed at three points over the course of section 1B: at the beginning, 30 minutes through the section, or at the end of the section.[12] Because subjects are in the laboratory for the duration of the experiment (without access to cell phones or other means of external communication), they cannot take actions contingent on the values of x and y before the experiment concludes, making any information portrayed by x and y noninstrumental.

Figure 1: The Timeline of a Session

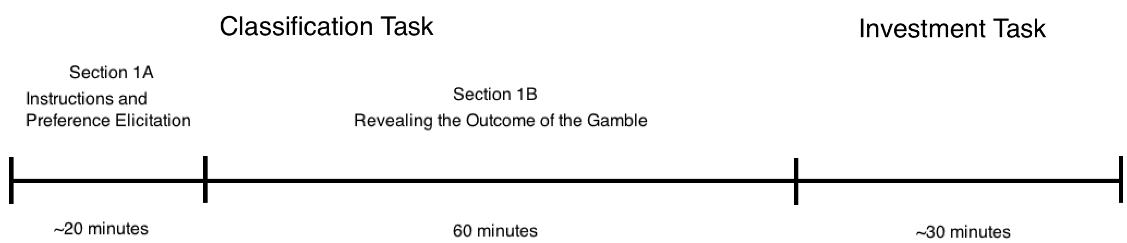


Table 2 details when each of the four gambles – A, B, C, and D – reveal x and y.[13] In section 1A, a full preference ordering over the gambles is elicited in a manner which is incentive compatible under all the models considered. This preference ordering allows subjects to be classified by their preferences over the resolution of risk.

[12] During the time between draws of x and y in section 1B, subjects performed unincentivized "filler tasks".

[13] Roughly half of the subjects saw the gambles named and arranged as they are presented here, while the other half of subjects were shown the gambles in the order C, D, A, B (with the gambles renamed appropriately).

Table 2: The Gambles

	The Beginning of Section 1B	30 Minutes Into Section 1B	The End of Section 1B
Gamble A	x, y	-	-
Gamble B	-	-	x, y
Gamble C	x	y	-
Gamble D	-	x	y

The preference ordering is elicited in a two-part procedure; subjects were informed about both parts before making any decisions. In the first part, subjects are instructed to rank the four gambles from favorite to least favorite. In order to incentivize subjects to reveal their true (weak) preference ordering, the gamble subjects actually faced is determined probabilistically, where:

- Subjects receive their favorite gamble with probability 42%
- Subjects receive their second favorite gamble with probability 33%
- Subjects receive their third favorite gamble with probability 17%
- Subjects receive their least favorite gamble with probability 8%

Though this procedure generates a weak preference ordering over the gambles, it doesn't allow for differentiation between strict preference and indifference. In order to achieve cardinal measures of preference, subjects were next presented with the opportunity to "buy"[14] probability distributions which were potentially more favorable.[15] Subjects were given three such opportunities. First, they were asked if they would like to pay c cents for a new probability distribution which moved 8% of the initial probability mass from

[14]Though the terms "buy" and "pay" are used throughout this section, subjects' instructions contain more neutral language. Wording which could imply ownership was avoided so as to minimize possible endowment effects (Kahneman, Knetsch, and Thaler; 1990).

[15]Allowing subjects the opportunity to buy weakly preferred distributions over bundles (in this case, the bundles are the gambles A, B, C, and D) is a design feature previously used by Toussaert (2016).

the second favorite to the first favorite gamble. Then, they were asked if they would like to pay c for the probability distribution which moved 8% probability mass from the third favorite to the second favorite gamble. Finally, they were asked if they would like to pay c for the probability distribution which moved 8% mass from the least favorite to the third favorite gamble[16] . Several values of c were presented to subjects for each of the three choices:

$$c \in \{1, 5, 10, 25, 50, 100, 200\}$$

For each session, one of the values of c was incentivized, while the rest were hypothetical. Roughly half of the subjects faced incentivized prices of $c = 10$, while the other half faced prices of $c = 25$.[17] Subjects' choices were not allowed to violate monotonicity in wealth, for example, a subject couldn't be unwilling to pay 10 cents to move 8% probability mass from their second to their first favorite gamble, but willing to pay 25 cents for the same offer.[18]

[16]The probabilities with which each gamble are initially chosen guarantee that subjects' first favorite gamble will always be most likely to be selected, their second favorite gamble will be second most likely to be selected, etc. If, for instance, the initial probabilities with which each gamble was chosen had been 40%, 30%, 20%, and 10%, a subject could have paid to move probability mass from their second to their first favorite gamble, and also paid to move probability mass from their third to their fourth favorite gamble. This would have resulted in a final probability distribution over gambles (from most to least favorite) of 48%, 22%, 28%, and 2%, making a subject's third most favorite gamble more likely to be selected than his second most favorite gamble. Probabilities were chosen to prevent such issues.

[17]A series of two-sample tests of proportion fail to reject the hypothesis that changing the incentivized value of c has an effect on subjects' decisions, thus, both groups are pooled in the following analysis.

[18]Subjects were first asked if they would be willing to pay the incentivized value of c (and they were told that their decision would be enforced). If they were willing to pay the incentivized level of c to buy a new probability distribution, they were then asked if they *would have* been willing to pay for the new distribution at the next highest level of c. For example, consider a subject for whom the incentivized c is 25 cents. He would first be asked whether or not he was willing to pay 25 cents to move 8% probability mass from his second favorite gamble to his favorite gamble. If he was willing to pay 25 cents, he would next be asked if he would have made the same decision for a "price" of 50 cents. This process would continue until the subject either declined to pay, or until he was asked about the maximum price, $c = 200$. A similar process ensured monotonicity in wealth for subjects who were not willing to pay for the incentivized level of c. The price was lowered until the subject either reported a desire to pay for the new probability distribution or until he was asked about the minimum price, $c = 1$.

1.3.2 The Investment Task

The investment task – a replication of the experiment developed in Gneezy and Potters (1997) – was revealed following the completion of the classification task. This task presents subjects with a simple dynamic investment problem which takes place over nine rounds. In each round, subjects are endowed with 150 cents, and are asked to decide what part of that endowment they would like to bet in a gamble. With probability $^1/_3$, the gamble returned any wagered money in addition to paying back 2.5 times the amount bet. With probability $^2/_3$, subjects lost any wagered money. Subjects could not bet any money won in previous rounds.

The experiment was conducted using physical randomization devices (as was the case in the original Gneezy and Potters experiment). Subjects are assigned a *win letter* – either A, B, or C – which remains constant throughout a session. In each round, three balls labeled A, B, and C are placed into a box, from which one of the three balls is drawn. If the ball matching a subject's win letter is drawn in a round, the subject wins the gamble, otherwise, the subject loses.

The task has two treatments which are *conducted in a between subjects design*: a frequent feedback treatment, F, and an infrequent feedback treatment, I. In the F treatment, subjects play rounds one-by-one, observing the outcome of each round before placing bets for the following round. In the I treatment, subjects play the rounds in blocks of three. For example, at the beginning of round 1, subjects in the I treatment are required to place their bets for rounds 1, 2, and 3. These bets were required to be equal. After subjects have decided upon their bets, they are simultaneously shown the results of all three rounds at once. They cannot assign a gain or loss to any particular round, but know in the aggregate how many rounds have been won and how many have been lost within the set of three rounds. For subjects in the I treatment, this procedure was repeated a second time for rounds 4, 5, and 6; and a third time for rounds 7, 8, and 9.

1.3.3 Hypotheses

This section presents the hypotheses that the above experimental tasks are designed to address. The classification task allows subjects to be classified as one of nine preference types described in the previous section. As agents in the standard expected utility model are indifferent towards the manner in which risk is resolved, this model predicts that subjects will be classified as EU types. On the other hand, Kőszegi and Rabin (2009), Ely, Frankel, and Kamenica (2015), and Kreps and Porteus (1978) all assume that individuals have non-standard preferences over the resolution of risk. Thus, the expected utility model acts as the null hypothesis for each of these alternative models.

Hypothesis 1 – Expected Utility

Subjects are indifferent towards when risk is resolved.

AGRR states that, as the frequency of feedback about the performance of a risky asset increases, investment in that asset decreases. Gneezy and Potters (1997) and others have shown that laboratory subjects have AGRR, a finding predicted by both Benartzi and Thaler's (1995) myopic loss aversion and the preferences model of Kőszegi and Rabin (2009). In my experimental design, subjects display AGRR if investments are lower among subjects in the F treatment than they are among those in the I treatment.

Hypothesis 2 – Aversion to the Gradual Resolution of Risk (AGRR)

> *Investments will be lower among subjects in the F treatment than among those in the I treatment.*

Kőszegi and Rabin's model assumes that individuals have preferences that are clumped, early, or early/clumped. Moreover, they assume that it is these preferences which cause AGRR. Thus, in the context of this experiment, Kőszegi and Rabin's preference model anticipates that clumped, early, and early/clumped types will invest less in the F treatment than in the I treatment, and that the overall difference between investment levels in the F and I treatments among all subjects will be driven by individuals with these three types.

Hypothesis 3 – Kőszegi-Rabin Preferences

> *Clumped, early, and early/clumped subjects will have lower investments in the F treatment than in the I treatment. Overall differences in investment between F and I treatments will be driven by these types.*

In Kőszegi and Rabin's framework, the stronger is an individual's preference for how risk is resolved, the more severe his AGRR becomes. Thus, subjects who have a greater willingness to pay to decide how risk is resolved would also be more sensitive to the frequency of feedback about the performance of a risky asset in deciding how much to invest. In the context of this design, subjects who are willing to pay more money for a preferred probability distribution in the classification task should have a larger difference in investments between the F and I treatments in the investment task. This result, of course, is only predicted among subjects with preferences consistent with Kőszegi and

Rabin's model: clumped, early, and early/clumped types.

Hypothesis 3.1 – AGRR Increasing in Strength of Preference

> *Among clumped, early, and early/clumped types, the difference between investments in the F and I treatments will increase as the total amount of money subjects are willing to pay in the classification task increases.*

Benartzi and Thaler's theory of myopic loss aversion claims that individuals have AGRR because they "narrowly bracket" risky prospects – that is, they fail to realize that current risk will be integrated with other risks. As myopic loss aversion does not rely on preferences over the resolution of risk, this model assumes that all subjects – even those with standard preferences over the resolution of risk (i.e. EU types) – will display AGRR.

Hypothesis 4 – Myopic Loss Aversion

> *Subjects of all preference types will have lower investments in the F treatment than in the I treatment.*

1.4 Results

This section presents data from 10 experimental sessions conducted at the Experimental and Behavioral Economics Laboratory at UC Santa Barbara. A total of 118 subjects[19] participated in sessions which lasted two hours. Including a $5 show-up fee, subjects earned an average of approximately $25. The first task was administered via computer

[19]This includes three subjects who left in the middle of the experiment, and are not included in the analysis.

and was programmed in z-Tree (Fischbacher, 2007), while the second task was conducted with pen and paper, using physical randomization devices when needed. Subjects were recruited with ORSEE (Greiner, 2004).

I first review the results of the classification task. Throughout, I assume that subjects who are willing to pay any amount (one cent or more) to move probability mass from a less preferred gamble to a more preferred gamble have a strict preference for the higher ranked gamble. I interpret an unwillingness to pay even one cent to move probability between a lower and higher ranked gamble as indifference between those gambles.

This preference ordering allows subjects to be individually classified as either one of the nine preference types in figure 1, or *unclassifiable*, otherwise. Table 3 summarizes the nine types, the preference orderings that are assigned to those types, and the predictions of each of the preference models discussed in section 2.2.

Table 3: How Subjects Are Classified by Preference Orderings over Gambles

		Timing		
		Early	Neither	Late
Concentration	Clumped	*KR* $A \succ C \succ D$ & $A \succ B$	*KR* $A \sim B \succ C \sim D$	- $B \succ D \succ C$ & $B \succ A$
	Neither	*KR/KP* $A \succ C \succ D \succ B$	*EU* $A \sim B \sim C \sim D$	*KP* $B \succ D \succ C \succ A$
	Piecemeal	- $C \succ D \succ B$ & $A \succ B$	*EFK* $C \sim D \succ A \sim B$	- $D \succ C \succ A$ & $B \succ A$

Subjects who have non-standard preferences on only one dimension are referred to by that preference only. For example, subjects with preferences $A \succ C \succ D \succ B$ are labeled early types. Subjects with preferences which cannot be rationalized using only the timing or concentration dimensions are labeled with a slash, for example, a subject with preferences $D \succ B \sim C \succ A$ would be labeled late/piecemeal.

Note that the set of subjects obeying some preference orderings are subsets of other preference groups. Specifically, subjects with the early type preference ordering also satisfy the preference orderings for early/clumped types and early/piecemeal types, since the ordering $A \succ C \succ D \succ B$ satisfies the conditions for both early, early/clumped, and early/piecemeal types. Similarly, late types also satisfy the preference orderings for late/clumped and late/piecemeal types. Therefore, when classifying subjects, early and late types take precedence over the other types.[20]

At this point is it also important to note that subjects labeled as EU are merely those who do not have non-standard preferences over the resolution of risk. Though myopic loss aversion is not a feature of the standard expected utility model, it is possible that subjects who are labeled as EU in the classification task may still be myopically loss averse. This is because the question of whether or not a subject displays myopic loss aversion is independent of his preferences over the resolution of risk.

In total, there are 192 possible preferences orderings that subjects can submit.[21] Early and late types correspond with only one of those preference orderings each. Any ordering in which subjects are indifferent between all four gambles corresponds with the expected utility type; there are 24 of these orderings. Clumped and piecemeal types correspond with 4 preference orderings each. The remaining four types – early/clumped, late/clumped, early/piecemeal, and late/piecemeal – correspond with 6 preference orderings each. The remaining 134 possible preference orderings cannot be rationalized in the present framework, and are thus considered unclassifiable.

[20]Thus, a subject will be classified as early/clumped if and only if he has preferences satisfying three conditions: (1) $A \succ C \succ D$, (2) $A \succ B$, and (3) *not* $A \succ C \succ D \succ B$.

[21]For the sake of this analysis, $A \succ B \succ C \sim D$ and $A \succ B \succ D \sim C$ would be considered different preference orderings.

Table 4: Distribution of Types

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="3">Timing</td><td rowspan="2">Unclassifiable</td></tr>
<tr><td>Early</td><td>Neither</td><td>Late</td></tr>
<tr><td rowspan="3">Concentration</td><td>Clumped</td><td>KR
16 (0.14)</td><td>KR
1 (0.01)</td><td>-
0</td><td rowspan="3">22 (0.19)</td></tr>
<tr><td>Neither</td><td>KR/KP
20 (0.17)</td><td>EU
26 (0.23)</td><td>KP
1 (.01)</td></tr>
<tr><td>Piecemeal</td><td>-
10 (0.09)</td><td>EFK
2 (0.02)</td><td>-
17 (0.15)</td></tr>
</table>

This table contains the number of subjects (out of 115) of each type. Proportions are given in parentheses.

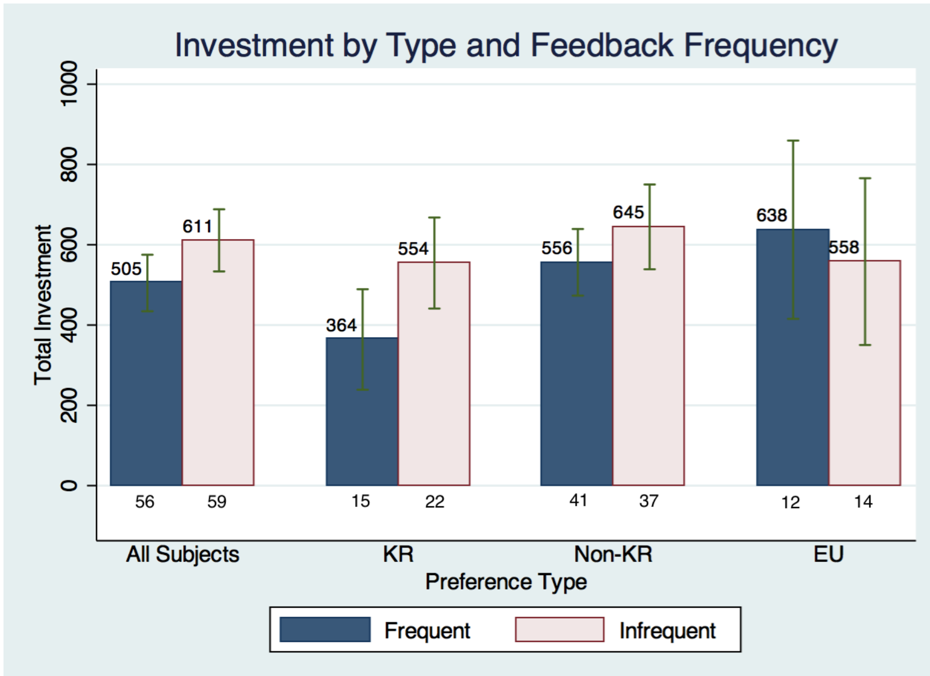
Table 4 shows the distribution of types resulting from the classification task, as well as the models which predict each type. If subjects were behaving randomly, we would expect the proportions of subjects of each type to correspond with the proportion of preference orderings which generate that type. For example, with random behavior, 24/192 or 12.5% of subjects would be EU, 134/192 or 69.8% of subjects would be unclassifiable, 1/192 or 0.5% would be early types, etc. A one-sample test of proportions strongly rejects this hypothesis, ($p < 0.001$), suggesting that subject behavior is not random. This finding is further supported by the fact that 17% (20/115) of subjects submitted the single preference ordering, $A \succ C \succ D \succ B$, that corresponds with the early type.

Comparing the preference models' predictions with the results in table 4 reveals some stark results. Clumped types, late types, and piecemeal types can all be rationalized by one of the preference models presented in section 2.2, yet there are not large numbers of subjects with any of these types in the data. There are however, moderate numbers of subjects with early/piecemeal and late/piecemeal preferences, which are not predicted by any one preference model. Thus, it appears that Ely, Frankel and Kamenica's (2015) model has some predictive power, though the majority of subjects with piecemeal preferences also have non-standard preferences on the timing dimension – a finding that these authors do not anticipate.

On the whole, Kőszegi and Rabin's model performs well, with 32% (37/115) of subjects' preferences falling under one of the three types predicted by this model. Kreps and Porteus's model also performs reasonably well, as it is able to rationalize 18% (21/115) of the data. Finally, the standard model can account for 23% (26/115) of subjects' preferences. As there appears to be significant heterogeneity in subjects' preferences over the resolution of risk, these results reject hypothesis 1, which states that subjects are predominantly EU types.

The remainder of the hypotheses involve data from the investment task, which are summarized in figure 2.

Figure 2: Investments by Feedback Frequency Within Preference Types



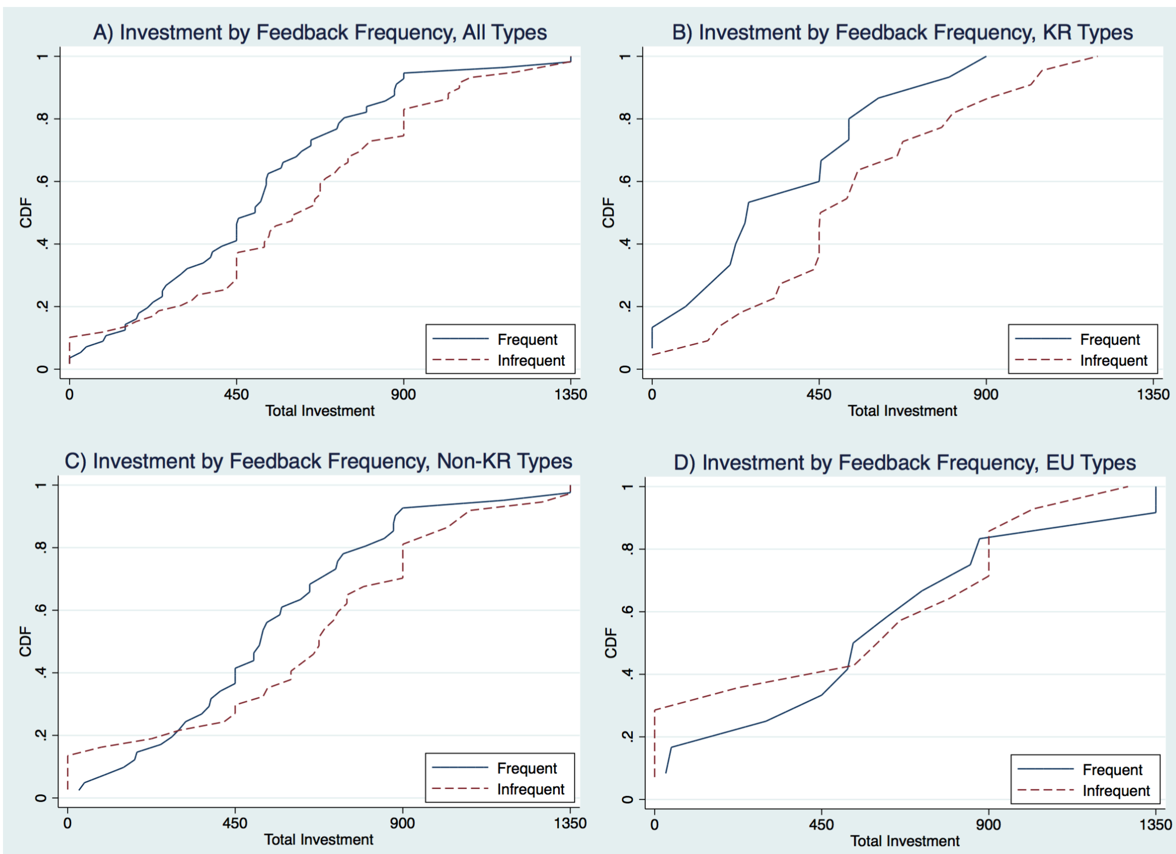
Error bars show 90% confidence intervals. Values above bars are average investments; values below bars are observations. The difference-in-differences (DID) between KR and Non-KR is not significant, though the DID between KR and EU is marginally significant ($p = 0.100$).

Hypothesis 2 states that, among all subjects, investments are lower in the F treatment than in the I treatment. Testing this hypothesis requires only a comparison of investment levels across treatments. The leftmost pair of bars in figure 2 shows the total investment over all nine rounds between treatments. As the figure makes clear, there is an overall treatment effect in the anticipated direction. A one-tailed Mann-Whitney U-test[22] confirms that this difference is, in fact, significant ($p = 0.033$). Panel A of figure 3 presents the CDFs for these same treatment groups. This graph shows that the treatment effect is roughly uniform across investment levels, with the exception of subjects with very low investments. Among these subjects, there does not appear to be a treatment effect.

[22] I follow Gneezy and Potters (1997) by reporting one-tailed significance levels, since AGRR makes a directional prediction (that investments in the F treatment will be lower than those in the I treatment).

The magnitude of the treatment effect – measured as the average investment in I divided by the average investment in F – is similar to that of Gneezy and Potter's original experiment (1.21 in the present study versus 1.33 in Gneezy and Potters).

Figure 3: CDFs of Investments by Feedback Frequency Within Preference Types



Hypotheses 3 and 4 make different predictions about the size of the treatment effect across preference types. Following hypothesis 3, Kőszegi and Rabin argue that AGRR is caused by preferences over the resolution of risk, so only early, early/clumped, and clumped types will invest less in the F treatment than in the I treatment. On the other hand, following hypothesis 4, myopic loss aversion assumes that AGRR is caused by a cognitive error, and therefore all preference types will invest less in the F treatment than in the I treatment.

The second and third set of bars in figure 2 address hypothesis 3. The second set of bars, *KR*, show the treatment effect for only those subjects who had non-standard preferences consistent with the predictions of the Kőszegi-Rabin model: early, early/clumped and clumped types. As the bar graph suggests, a one-sided Mann-Whitney U-test finds that the difference in investments across treatments is statistically significant for this subpopulation ($p = 0.045$). The third set of bars, *Non-KR*, shows the treatment effect among the complementary subpopulation, that is, the subjects who's preferences are *not* labeled *KR* in table 4. Though the treatment effect is in the direction consistent with AGRR, the difference in investments is only marginally significant (one-sided Mann-Whitney U-test, $p = 0.066$).

The treatment effect is considerably larger for subjects with Kőszegi-Rabin preference types (190 cents) than for subjects of other types (88 cents), though the difference-in-differences comparing the treatment effects between these groups is not significant at standard confidence levels ($p = 0.224$).[23] Panels B and C of figure 3 compare the CDFs of investments by treatment among KR and Non-KR types, respectively. These graphs lend further support to hypothesis 3, as I-treatment investments first-order stochastically dominate F-treatment investments among KR types, though this is not true among non-KR subjects.

Benartzi and Thaler's (1995) theory of myopic loss aversion is the motivation for hypothesis 4. These authors assume that individuals make a cognitive error in aggregating

[23] That is, for the dummy variables F_i (indicating whether or not subject i is in treatment F) and KR_i (indicating whether or not subject i is a Kőszegi-Rabin type, consider the regression:

$$y_i = \beta_0 + \beta_1 F_i + \beta_2 KR_i + \beta_3 F_i \times KR_i + \varepsilon_i$$

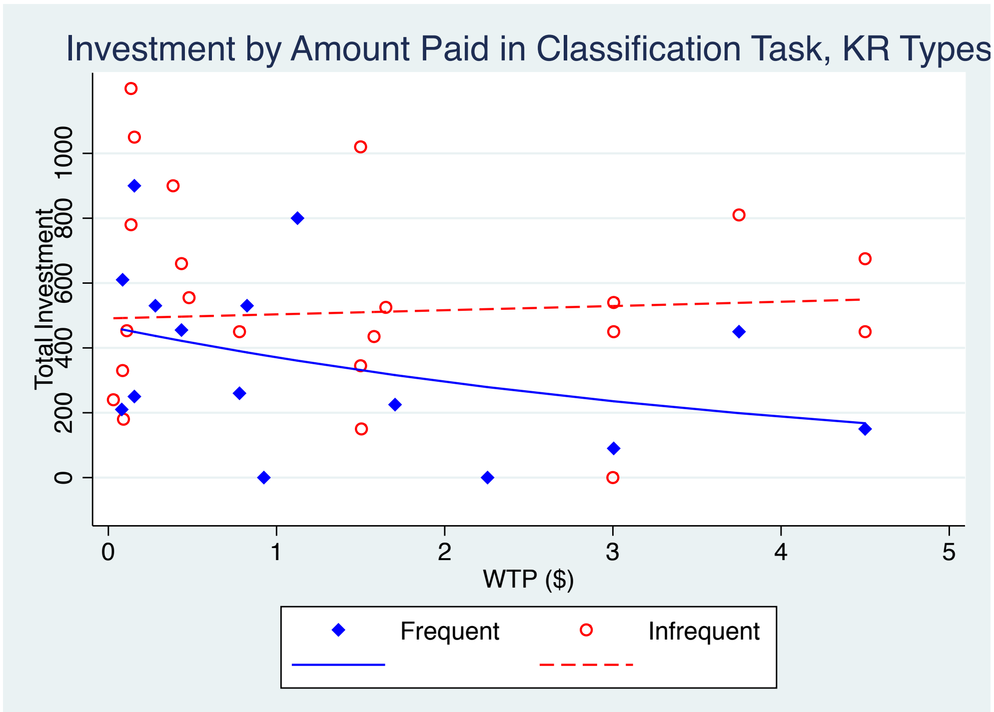
The p-value is that of the t-statistic associated with $\hat{\beta}_3$.

risk, and therefore, AGRR is independent of preference type. If hypothesis 4 is correct, investments should be lower in the F treatment than in the I treatment for all preference types. As the rightmost pair of bars in figure 2 shows, this is not the case. Among EU types, who constitute nearly a quarter of experimental subjects, investments in the F treatment are actually slightly higher than those in the I treatment, though the difference is far from from significant. Panel D of figure 3 emphasizes this point, as there appears to be no significant difference between treatments for any investment level. Moreover, the difference-in-differences comparing only KR and EU types is (marginally) significant (t-statistic, $p = 0.075$; see footnote 23), a result which myopic loss aversion cannot account for. Taken in sum, the previous results argue in favor of accepting hypothesis 3 and rejecting hypothesis 4. In light of these data, Kőszegi-Rabin preferences offer a more convincing explanation for AGRR than Benartzi and Thaler's theory of myopic loss aversion.

Next, I test hypothesis 3.1 by asking if, among KR types, AGRR increases as preferences become stronger. In order to measure the strength of a subject's cardinal preferences, I sum the total amount of money each subject stated they would be willing to pay in the classification task. I call this a subject's *WTP*. As this task involves subjects paying for (weakly) preferred probability distributions, I assume that a larger WTP corresponds with stronger cardinal preferences over the gambles A, B, C, and D. To examine whether stronger preferences correspond with more severe AGRR, I regress the natural log of the total amount bet in the investment task on WTP for subject i in treatment $t \in I, F$:

$$ln(\text{Total Investment}_i^t) = \beta_0^t + \beta^t \text{WTP}_i^t + \varepsilon_i^t$$

Figure 4: Investment versus WTP by Treatment in Investment Task, Kőszegi-Rabin Types Only



The results of these two regressions (one for each treatment) are plotted in figure 4. As this figure shows, KR types with small WTPs invest roughly equal amounts in the investment task, regardless of whether they are in the F or I treatment. That is, subjects with weak preferences tend to have less AGRR. However, as subjects' WTPs increase, the difference in investments across treatments increases as well. This is reflected in figure 4 by the widening gap between the dashed and solid lines. As the figure shows, the AGRR seems to be driven by subjects in the frequent feedback treatment. This is confirmed by the regression results, which find $\hat{\beta}^F = -0.227$ ($p = 0.041$). Interpreting this coefficient, I find that a marginal increase in WTP of \$1 is associated with a 22.7% decrease in total investment in the investment task for KR types in the F treatment. For comparison, I find $\hat{\beta}^I = 0.025$ ($p = 0.387$). This coefficient says that a marginal increase in WTP of \$1 is associated with a 2.5% *increase* in total investment among KR types in the I treatment (thought this increase is not statistically significant).

1.5 Conclusion

In this paper, I develop an experiment to identify the cause of *aversion to the gradual resolution of risk*, the empirical finding that individuals invest less in risky assets when provided with more frequent feedback about the value of those assets. This finding has been reproduced in the laboratory several times, and it is thought to be the driver of important macroeconomic phenomena, such as the large difference in average returns between risk-free debt and the S&P 500. The theory literature has produced two explanations for AGRR, each of which, if correct, would have important policy implications. Benartzi and Thaler's (1995) theory of myopic loss aversion argues that AGRR is the result of a cognitive error called "narrow bracketing", in which investors fail to consider the fact that the risk they currently face will be integrated with other future risk. By contrast, Kőszegi and Rabin (2009) propose a preferences-based explanation, asserting that AGRR is the result of non-standard preferences over the resolution of risk.

I test Kőszegi and Rabin's claim using two experimental tasks. In the first task, subjects are classified as one of nine preference types. These types not only include the predictions of Kőszegi and Rabin and the standard expected utility model, but also the types predicted by the work of Ely, Frankel, and Kamenica (2015) and Kreps and Porteus (1978). After the classification task, subjects make choices in a simple dynamic investment environment which is a replication of the experimental task of Gneezy and Potters (1997). In this second task, the frequency of feedback that subjects receive about the performance of their investments is experimentally varied in a between-subjects design. This task shows that experimental subjects do, in the aggregate, display AGRR. Using the results of the classification task, I show that this AGRR is driven by Kőszegi-Rabin preference types, suggesting that non-standard preferences over the resolution of risk are the underlying cause of the results in Gneezy and Potter's experiment. These findings suggest that the seemingly low amount of investment in dynamic assets like the S&P 500 may not be the result of a cognitive error, but instead the expression of preferences. An investor who, for example, chooses a conservative retirement portfolio and checks the performance of that portfolio on a regular basis, would, under the lens of Benartzi and Thaler's myopic loss aversion, be making a welfare-decreasing mistake. However, in Kőszegi and Rabin's framework, this investor may receive disutility by living with uncertainty about the future, and such a course of action may be that investor's utility-maximizing plan.

My results recommend a simple policy for financial institutions releasing information to investors about their portfolios' performance: do not *force* investors to view financial information, but make such information readily accessible for those to wish to have it. As preferences over the resolution of risk appear to be driving investment decisions, and my results show significant heterogeneity in these preferences, such a laissez-faire policy would allow investors to decide on feedback frequencies which maximize utility given their preferences.

1.6 Supplementary Details on Preference Models

This section provides additional detail on the preference models of Kőszegi and Rabin (2009); Ely, Frankel, and Kamenica (2015); and Kreps and Porteus (1978). It includes simple examples which illustrate the mechanics of each model.

1.6.1 Kőszegi and Rabin (2009)

There are two differences between Kőszegi and Rabin's full model, and the simplified version presented in section 2: first, the authors assume consumption is K-dimensional (instead of one-dimensional), and second, a more general loss aversion function is considered. In the more general model, Kőszegi and Rabin assume that agents have preferences described by the following function (Kőszegi and Rabin's equation 1):

$$u_t = m(\mathbf{c_t}) + \sum_{\tau=t}^{T} \gamma_{t,\tau} N(\mathbf{F}_{t,\tau}|\mathbf{F}_{t-1,\tau})$$

$m(\mathbf{c_t})$ is the standard reference-independent utility from consumption. $\mathbf{F}_{t,\tau} = (F^1_{t,\tau}, ..., F^K_{t,\tau})$ are the beliefs held in period t about period-τ consumption in each of the K dimensions. The terms $N(\mathbf{F}_{t,\tau}|\mathbf{F}_{t-1,\tau})$ capture gain-loss utility generated by changes in beliefs from period $t-1$ to period t about these future consumptions. $N(\mathbf{F}_{t,\tau}|\mathbf{F}_{t-1,\tau})$ is the sum of the gain-loss utilities in each dimension. That is, $N(\mathbf{F}_{t,\tau}|\mathbf{F}_{t-1,\tau}) = \sum_{k=1}^{K} N^k(F^k_{t,\tau}|F^k_{t-1,\tau})$, where the gain-loss utility on dimension k, $N^k(F^k_{t,\tau}|F^k_{t-1,\tau})$, is given by:

$$N^k(F^k_{t,\tau}|F^k_{t-1,\tau}) = \int_0^1 \mu[m^k(c_{F_{t,\tau}}(p)) - m^k(c_{F_{t-1,\tau}}(p))]dp$$

Where, for any distribution F over $\mathbb{R}$ and any $p \in (0,1)$, $c_F(p)$ is the consumption level at the percentile p, defined implicitly by the conditions that $F(c_F(p)) \geq p$ and $F(c) < p$, $\forall c < c_F(p)$. $\mu(\cdot)$ is a loss aversion value function[24]. Intuitively, the agent makes "ordered comparisons" between his current beliefs, $F^k_{t,\tau}$, and previous beliefs, $F^k_{t-1,\tau}$. He compares the best percentile of outcomes under $F^k_{t,\tau}$ to the best percentile of outcomes under $F^k_{t-1,\tau}$, the second-best percentile of outcomes under $F^k_{t,\tau}$ to the second-best under $F^k_{t-1,\tau}$, etc. He is loss averse in all such comparisons.

To illustrate the mechanics of Kőszegi and Rabin's model, I consider a simple example which appears in the authors' original text. Assume there are two periods, 1 and 2, and that no consumption takes place in period 1. There is only one consumption dimension in period 2, and $m(c_2) = c_2$. There are two equiprobable possible consumption levels, $c_2 = 0$ and $c_2 = 1$, and the agent has no control over his consumption. He may, however, receive a signal $s \in \{0,1\}$ in period 1 which is accurate with probability $q > 0.5$. For simplicity, assume the loss aversion function, $\mu(\cdot)$, is piecewise linear, with $\mu(x) = x$ for $x \geq 0$ and $\mu(x) = \lambda x$ for $x < 0$.[25] Define $\gamma \equiv \gamma_{t_1,t_2}$. Since consumption utility is independent of the arrival of information (as is true in the environment that experimental subjects face), I focus only on gain-loss utility.

[24]This value function has the properties originally suggested by Kahneman and Tversky (1979), which were later formalized by Bowman, Minehart, and Rabin (1999). Kőszegi and Rabin assume a value function which has properties corresponding with these authors' properties (A0) - (A4):

- $\mu(x)$ is continuous $\forall x$, twice differentiable for $x \neq 0$, and $\mu(0) = 0$.
- $\mu(x)$ is strictly increasing.
- If $y > x \geq 0$, then $\mu(y) + \mu(-y) < \mu(x) + \mu(-x)$.
- $\mu''(x) \leq 0$ for $x > 0$ and $\mu''(x) \geq 0$ for $x < 0$.
- ${}^{\mu'_-(0)}/_{\mu'_+(0)} \equiv \lambda > 1$, where $\mu'_+(0) \equiv \lim_{x\to 0} \mu'(|x|)$ and $\mu'_-(0) \equiv \lim_{x\to 0} \mu'(-|x|)$.

[25]$\lambda > 1$.

If the agent observes the signal, his expected utility is:

$$0.5\Big[\gamma(q-0.5)\Big]-0.5\Big[\gamma\lambda(q-0.5)\Big]+0.5\Big[q(1-q)-(1-q)q\lambda\Big]+0.5\Big[(1-q)q-q(1-q)\lambda\Big]$$

The first two terms capture the expected gain-loss utility in period 1. There is a 0.5 probability that the agent receives a signal of $q > 0.5$, leading to an increase of his belief that the good outcome will occur equal to $q - 0.5$. Likewise, with 0.5 probability, the agent receives a signal of $q < 0.5$, leading to an decrease of his belief that the good outcome will occur equal to $q - 0.5$. Because this news doesn't arrive in the same period that the consumption takes place, the effect of the changes in belief are "discounted" by γ.

The second two terms capture the expected gain-loss utility in period 2. With probability 0.5, the agent left period 1 believing he would receive the high consumption with probability q. If he does, in fact, receive the high consumption, his beliefs will further increase by $(1 - q)$. If not, they will decrease by q; this happens with probability $(1 - q)$. The last term is derived using a similar process for the case in which the agent leaves period 1 believing he will receive the high consumption with probability $(1 - q)$.

If the agent does not observe the signal, he will experience gain-loss utility only in period 2. In expectation, this results in a change in utility of:

$$0.5\Big[1-0.5\Big]-0.5\Big[\lambda(0-0.5)\Big]=0.25(1-\lambda)$$

Combining the previous two expressions, the signal generates higher utility when:

$$\gamma < 2(q - 0.5)$$

1.6.2 Ely, Frankel, and Kamenica (2015)

Ely, Frankel, and Kamenica consider an agent who does not know the state of the world, and who derives "suspense" or "surprise" utility from changes in beliefs about this state. Formally, the authors assume a finite state space Ω, with generic element ω. Typical beliefs are distributions over the state space, $\mu \in \Delta(\Omega)$. μ^ω designates the probability of ω, and the agent has a prior μ_0. Beliefs evolve over a finite number of periods $t \in \{1, 2, ... T\}$.

A *belief martingale* $\tilde{\mu}$ is a sequence $(\tilde{\mu})_{t=0}^{T}$ such that:

1. $\tilde{\mu}_t \in \Delta(\Delta(\Omega)), \ \forall t,$
2. $\tilde{\mu}_0$ is degenerate, and
3. $E[\tilde{\mu}_t | \mu_0, \mu_1, ... \mu_{t-1}] = \mu_{t-1}, \ \forall t.$

A realization of a belief martingale is a *belief path*, written $\eta = (\mu_t)_{t=0}^{T}$. Ely, Frankel, and Kamenica say that the agent has a preference for suspense if his utility function is:

$$U_{susp}(\eta, \tilde{\mu}) = \sum_{t=0}^{T-1} u\Big[\ E_t \sum_{\omega} (\tilde{\mu}_{t+1}^{\omega} - \mu_t^{\omega})^2\ \Big]$$

And the agent has a preference for surprise if his utility function is:

$$U_{surp}(\eta) = \sum_{t=1}^{T} u\Big[\sum_{\omega} (\mu_t^{\omega} - \mu_{t-1}^{\omega})^2\Big]$$

Where $u(\cdot)$ is an increasing and strictly concave function with $u(0) = 0$. Intuitively, suspense is induced by variance over the next period's beliefs and surprise is induced by a change in beliefs from the previous to the current period. Note that Ely, Frankel, and Kamenica consider preferences over beliefs about the state of the world, while Kőszegi and Rabin, and Kreps and Porteus model preferences over expected consumption. As Ely, Frankel, and Kamenica point-out, their model can accommodate an agent whose beliefs are about expected consumption (or wealth), which is the context in which I will consider their model. Both Ely, Frankel, and Kamenica's suspense and surprise models hinge on the idea that changes in beliefs are exciting, and the more beliefs change (or are expected to change), the higher will be the agent's utility.

Reconsider the example of the agent who lives for two periods that was discussed in the context of Kőszegi and Rabin's model above. The states of the world are $c_2 = 1$ and $c_2 = 0$. For simplicity, I will only consider the suspense and surprise utility generated by changes in beliefs regarding the $c_2 = 1$ state of the world. In cases where there are only two states of the world, as this example considers (and as will be true in the following experimental design), an increase in the belief of one state of the world occurring necessarily corresponds with a commensurate decrease in the belief of the other state of the world occurring. Thus, this assumption serves only to halve the argument of the $u(\cdot)$ function in the expressions for both suspense and surprise.

There are four possible belief paths:

$$\eta = (\mu_0, \mu_1, \mu_2) \in \Big\{[0.5,\ q,\ 1]; [0.5,\ q,\ 0]; [0.5,\ (1-q),\ 1]; [0.5,\ (1-q),\ 0]\Big\}$$

The first and last belief paths each have probability $0.5q$ of being realized, while the second and third belief paths are each realized with probability $0.5(1-q)$. Note that, for each belief path, the belief that the $c_2 = 1$ state obtains changes from a prior of $\mu_0 = 0.5$ to a period-one belief of $\mu_1 \in \{q, (1-q)\}$. Either way, beliefs will change by $(q - 0.5)$ between $t = 0$ and $t = 1$. In the second period, there is a q probability that beliefs change by $(1-q)$ and a $(1-q)$ probability that beliefs change by q.

Thus, the suspense induced by this belief martingale, for all belief path realizations η, is:

$$U_{susp}(\eta, \tilde{\mu}) = u\Big[(q-0.5)^2\Big] + u\Big[q(1-q)^2 + q^2(1-q)\Big]$$

$$= u\Big[(q-0.5)^2\Big] + u\Big[q - q^2\Big]$$

The surprise induced by this belief martingale depends on the realized belief path. Intuitively, the more surprising belief paths are those that are less likely to be realized: $(0.5,\ q,\ 0)$ and $(0.5,\ (1-q),\ 1)$. These generate surprise:

$$U_{surp}(\eta) = u\Big[(q-0.5)^2\Big] + u\Big[q^2\Big]$$

The less surprising – and more likely to be realized – paths are $(0.5,\ q,\ 1)$ and $(0.5,\ (1-$

q), 0). These generate surprise:

$$U_{surp}(\eta) = u\Big[(q - 0.5)^2\Big] + u\Big[(1 - q)^2\Big]$$

Ely, Frankel, and Kamenica note that the agent may have preferences over the state of the world, in which case suspense and surprise utility will augment the utility generated by the state. For example, a blackjack player may care primarily that they win the hand, but conditional on the outcome, the player derives more utility from hands that were laden with suspense or surprise. If the agent in the above example derives utility directly from period 2 consumption and from suspense, his total expected utility would be given by the following equation (where $m(\cdot)$ maps period 2 consumptions into period 0 utility):

$$U = 0.5m(c_2) + u\Big[(q - 0.5)^2\Big] + u\Big[q - q^2\Big]$$

Finally, note that experimental subjects choose belief martingales which they will face later in the experiment. Thus, I assume Ely, Frankel, and Kamenica-types maximize *expected* suspense or *expected* surprise.

1.6.3 Kreps and Porteus (1978)

Kreps and Porteus study environments with a finite, discrete number of time periods $t = 0, 1, 2, ... T$. For each period, the agent is in a state x_t and enjoys some immediate, period t payoff, z_t. The agent chooses an action to take in each period, and the state determines the set of actions available to the agent. The action taken, in turn, determines the probability distribution over (z_t, x_{t+1}). That is, the agent's action influences both her immediate payoffs, z_t, and the decision problem she will face in the following period, x_{t+1}.

The authors present preference axioms over (z_t, x_{t+1}) pairs, and show that these axioms are necessary and sufficient for the existence of a utility function. The utilities of (z_t, x_{t+1}) pairs in periods before the last period, T, are calculated recursively using a function $f_{KP}(z_t, U(x_{t+1}))$[26], where $U(x_{t+1})$ is the "continuation utility" of being at the state x_{t+1} in period $t+1$. Because neither the experimental design nor the following example allows for the possibility of "intermediate payoffs" (i.e. $z_t = 0, \quad \forall\ t < T$), the first argument is dropped from the f_{KP} function for simplicity. f_{KP} is a "utility distortion function" which "act[s] to 'convert' from the utility scale used at time $t+1$ to the scale used at time t" (Kreps and Porteus, pg. 191). As the following example makes clear, preferences for earlier or later resolution of uncertainty will be determined by the sign of the second derivative of f_{KP}.

Consider the two gambles from the introduction, g_1 and g_2. Each pays \$1 with probability 0.5 in period 2. The lotteries differ only in the period in which they resolve their uncertainty. Specifically, g_1 resolves its uncertainty in period 1, while g_2 resolves its uncertainty in period 2. For simplicity, suppose that the period 2 utility of winning either gamble is 1 util, the period 2 utility of losing is 0 utils, and that the outcome of the gamble is the only thing that affects the agent's utility[27]. Because the outcome of g_1 will have already been determined by the beginning of period 2, the expected utility of g_2 at this point is either 0 or 1.

[26] Assume f_{KP} is strictly increasing in both arguments.

[27] Note that, in period 2, the agent only derives utility from immediate payoffs, because the decision problem terminates after this period.

$$Eu_2^{g_1} \in \{0, 1\}$$

However, as the outcome of g_2 has not yet been determined at the beginning of period 2, the expected utility of g_2 is 0.5.

$$Eu_2^{g_2} = 0.5(1) + 0.5(0) = 0.5$$

Using the expected utility of the gambles in period 2, one can recursively calculate the period 1 expected utilities using $f_{KP}(U(x_{t+1}))$. For g_1, at the beginning of period 1, the agent realizes that there is 0.5 probability that his utility will be 1 at the beginning of next period and an 0.5 probability that it will be 0. Thus, his expected utility at the beginning of period 1 is:

$$Eu_1^{g_1} = 0.5 f_{KP}(1) + 0.5 f_{KP}(0)$$

Under g_2, the agent knows with certainty that his utility at the beginning of the next period will be 0.5, generating the following expected utility in period 1:

$$Eu_1^{g_2} = f_{KP}(.5)$$

As is common in the literature, assume f_{KP} takes the form $f_{KP}(\bar{u}) = \bar{u}^{\alpha}$ (for $\alpha > 0$). Then, $Eu_1^{g_1} = 0.5$ and $Eu_1^{g_2} = 0.5^{\alpha}$. Thus, for $\alpha > 1$, f_{KP} is convex and the agent strictly prefers g_1 (i.e. the early resolution of uncertainty). For $\alpha < 1$, f_{KP} is concave and the agent strictly prefers g_2 (i.e. the late resolution of uncertainty). For linear f_{KP}, the agent is indifferent between the gambles, and his preferences collapse to those of expected utility.

Chapter 2

Kidney Co-operative: A Mechanism to Improve on Human Kidney Markets

KC Eames, Patrick Holder, Eduardo Zambrano

2.1 Introduction

The kidney shortage is a serious problem. In the United States, 100,602 people were on the waiting list for new kidneys as of May 2014, while only about 16,500 kidney transplant operations take place every year.[1] Moreover, nearly 2,500 new patients are added, on average, to the waiting list each month. The consequences of the shortage are dire: 3,381 patients died in 2013 in the United States alone while waiting for a kidney transplant.[2]

[1] Organ Procurement Transportation Network (2014)
[2] National Kidney Foundation (2014)

At the core of this problem lies a basic, inescapable fact: one cannot compel potential donors to donate. The traditional solution to this problem is to allow a conventional market for live donor kidneys to develop, a solution advocated, for example, by Becker and Elias (2007). This solution has, however, been rejected for a variety of reasons by a large fraction of the population in the US and elsewhere.[3] For this reason, Roth, Sönmez, and Ünver (2004, 2005a, 2005b, 2007) began the important task of studying and promoting the institution of kidney paired donations as a solution to the kidney shortage. Unfortunately, the shortage is orders of magnitude larger than the number of kidney paired donations that are successfully achieved on an annual basis. It is therefore important to develop alternatives that would offer a large-scale solution to the kidney shortage.

As this paper demonstrates, the creation of a conventional market is not the only way to harness the power of incentives to motivate a large number of potential donors to donate. The solution we investigate, which we call a kidney co-operative, is designed to address the two main reasons why people consider the creation of a conventional market for live donor kidneys unacceptable: first, that such a market would completely exclude those patients unable to afford the kidney; and second, that donors will not understand the risks they're taking when donating, and that they will therefore not be properly compensated for taking that risk. Moreover, we show that it is reasonable to expect that the number of transplants will be larger under the kidney co-operative mechanism than under a conventional market mechanism.

Kidney co-operatives follow a simple set of principles. Patients in need of a kidney donate a set amount of money to the co-operative if they are able to do so. Those who

[3]For a broad discussion on people's perceptions regarding payment for organs see Leider and Roth (2010).

need a kidney but are unable to make the requisite monetary donation go on a "waiting list". Healthy patients donate kidneys to the co-operative, which first allocates kidneys to patients who donated money, then disburses any remaining organs to those on the waitlist. All revenue raised by the co-operative is split equally among the kidney donors, who also receive lifetime "kidney insurance".

Kidney co-operatives are not conventional markets, as some key conditions that define well functioning markets are not met. First, a kidney co-operative does not aim to maximize profits (either for itself or for its members). Second, the "law of one price" does not hold, as patients donate an amount per kidney that is larger than the cash payment that a donor receives for his or her donation. Third, the market "does not clear", that is, the quantity of kidneys supplied by the donors weakly exceeds the quantity demanded by paying patients. A kidney co-operative is, instead, a self-financing mechanism that, when designed efficiently, can maximize the number of kidney transplants that take place in the population at a minimum risk to the donors, while keeping all transactions voluntary. Furthermore, kidney co-operatives are (weakly) Pareto improving relative to the status quo.

2.2 Background

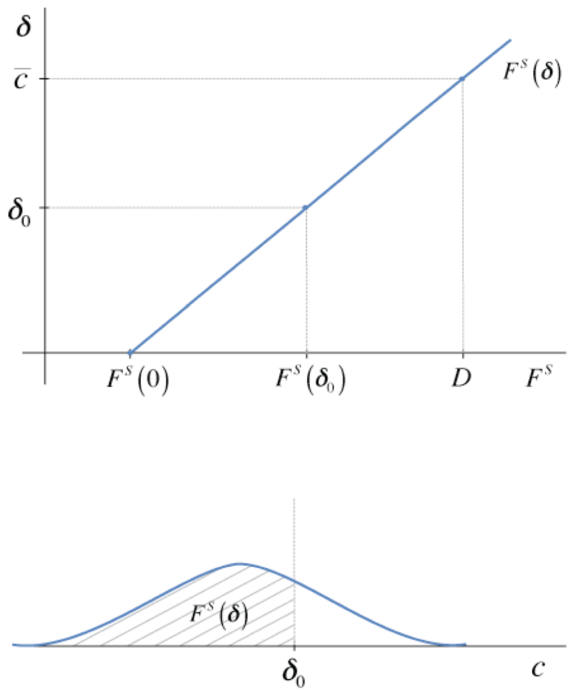
The simplest environment in which to consider a kidney co-operative is one in which the distributions of reservation prices for kidneys among both donors and patients is continuous and is known to the co-operative managers, when there is no risk to the donors from donating, and when patient/donor blood type incompatibilities are ignored. We relax these assumptions in section 2.4. Throughout the analysis, we treat all kidneys as being of equal medical quality, and all patients as being equally in medical need of a kidney. Though these assumptions can also be relaxed, we maintain them throughout.

The size of the population of patients is denoted by $D > 0$. Patients who contribute a set amount of money, p, designated by the co-operative, are called *contributing patients*. Patients who do not contribute are called *non-contributing patients*. The total amount of money collected by the co-operative divided by the number of kidneys donated is called the *co-operative dividend*, δ.

Each kidney donor forms an expectation of the co-operative dividend. Let the expectation for donor j be equal to δ_j. We postulate that, for each potential donor j there is a threshold $c_j \geq 0$ such that the potential donor donates if $\delta_j \geq c_j$ and does not donate otherwise. The distribution of reservation prices, c_j, for the potential donors is given by a probability distribution μ over the non-negative real numbers. Thus, if all potential donors expect a dividend equal to δ, the number of kidneys donated is given by $F^s(\delta) = \int \mathbf{1}_{\{c \leq \delta\}} d\mu(c)$, depicted in figure 1[4]. Let $\bar{c}$ be the smallest dividend level δ such that Anecdotal evidence suggests that $\mathrm{F}^s(\delta) = D$. Anecdotal evidence suggests that $\bar{c} \gg 0$.

[4]Where $\mathbf{1}_x$ is the indicator function. That is, $\mathbf{1}_x$ equals 1 when x is satisfied and 0 otherwise.

Figure 1: The Donors

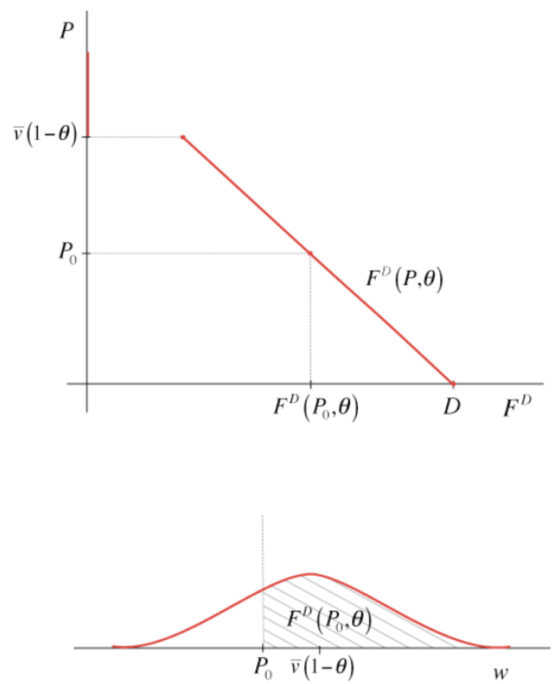


Kidneys are essential to life and all patients in need of one value them greatly. We thus assume that, without wealth constraints, every patient i is willing to contribute $v_i \geq \bar{c}$ to the co-operative to secure a kidney (for simplicity, we assume v_i is equal to $\bar{v}$ for all i). The alternative for each patient is to go on the waiting list and hope to be assigned a kidney this way, which, for patient i, happens with probability θ_i. Thus, without wealth constraints, patient i makes a contribution of size p to the co-operative when $\bar{v} - p \geq \bar{v}\theta_i$, that is, when $p \leq \bar{v}(1 - \theta_i)$, and does not contribute otherwise.

In practice, wealth constraints matter a great deal, as not all patients may be able to contribute an amount equal to p, even if they would like to. Therefore, in the presence of wealth constraints, patient i with wealth level w_i makes a contribution of size p to the co-operative when $p \leq \min\{\bar{v}(1 - \theta_i), w_i\}$, and does not contribute otherwise.

The wealth distribution for the patients is given by the probability distribution ρ over the non-negative real numbers. Thus, if all patients believe that the probability of obtaining a kidney by going on the waiting list is equal to θ, the number of contributing patients is given by $F^D(p,\theta) = \int \mathbf{1}_{p \leq \min\{\bar{v}(1-\theta_i), w} d\rho(w)$, depicted in figure 2. Note that $F^D(0,\theta) = D$ for any $\theta \in [0,1)$. We denote $F^D(p,0)$ simply as $F^D(p)$.

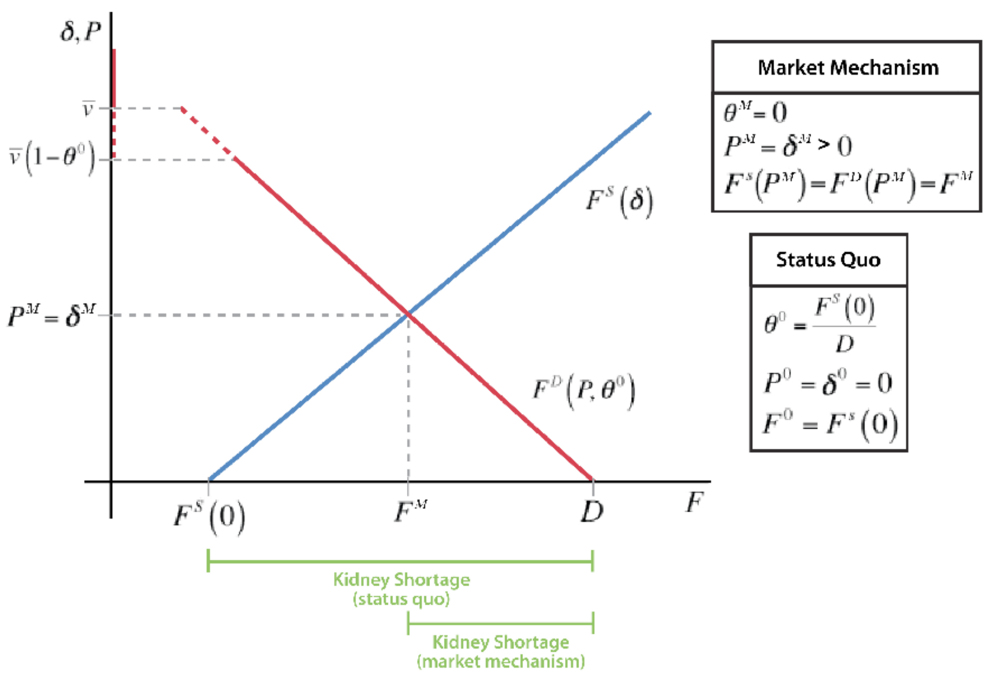
Figure 2: The Patients



The current situation in the United States can be described using the above notation. Under the status quo, $p^0 = \delta^0 = 0$ and $D - F^S(0) > 0$, thus, there is a kidney shortage. All patients go on the waiting list and have a positive probability of obtaining a kidney that is equal to $\theta^0 = \frac{F^S(0)}{D}$. Anecdotal evidence suggests that $F^S(0) \ll D$, and therefore that θ^0 is small.

Under the conventional market mechanism, all patients buy kidneys from donors at a single price and markets clear, that is, $p^m = \delta^m > 0$ and $F^D(p^m, \theta^m) = F^S(p^m) = F^m$, implying that $\theta^m = 0$. Society exhibits a smaller kidney shortage, $D - F^S(p^m)$. Thus, under the market mechanism, there are more transplants than under the status quo. However, the outcome under this mechanism is not a Pareto improvement over the status quo, as a fraction of the patients (those with an inability to pay prices above p^m) go from having a positive probability of obtaining a kidney under the status quo to a zero probability of obtaining a kidney under the conventional market mechanism. Figure 3 contrasts the status quo with the conventional market mechanism.

Figure 3: The Status Quo Versus the Conventional Market Mechanism



In what follows, we sometimes assume that the equilibrium of the market mechanism occurs on the inelastic portion of the demand curve for kidneys; we call this *Assumption 1*. Algebraically, this assumption states that:

$$\left|\frac{\partial F^D(p^m, \theta^m)}{\partial p}\frac{p^m}{F^m}\right| < 1$$

Without wealth constraints, Pareto optimality (and basic humanitarian concerns) implies that all patients should receive a kidney.[5] This outcome is difficult to attain for two reasons. First, wealth constraints typically bind for many patients, and second, one cannot compel potential donors to donate. A kidney co-operative addresses both these concerns.

2.3 The Kidney Co-operative

The function of a kidney co-operative is simple: it collects a set amount of cash, p, from each contributing patient, and it distributes the collected cash evenly amongst all kidney donors. We identify two desirable traits that a co-operative can possess. The first is *viability* – a kidney co-operative is viable if all contributing patients receive a kidney. The second is *voluntariness*. A kidney co-operative is voluntary if all healthy individuals wishing to donate a kidney to the co-operative are able to do so, and if all healthy individuals wishing not to donate are free not to. Likewise, under a voluntary kidney co-operative, all contributing patients must prefer to contribute than otherwise. We call a co-operative *sustainable* if it is both viable and voluntary.

Formally, a sustainable kidney co-operative is characterized by a contribution level p^*, a supply of kidneys F^*, and a fraction of non-contributing members receiving kidneys θ^*,

[5]This is true when and the size of potential donors is greater than D. If either of these conditions do not hold, Pareto optimality requires that only donors for which $\bar{v} \geq \bar{c}$ donate, until either all D patients have received transplants, or all donors for whom $\bar{v} \geq \bar{c}$ have donated.

which satisfy:

1. $0 < F^D(p^*, \theta^*) \leq F^S(\delta^*) = F^*$

2. $\delta^* = \frac{p^* \cdot F^D(p^*, \theta^*)}{F^*}$

3. $\theta^* = \frac{F^* - F^D(p^*, \theta^*)}{D - F^D(p^*, \theta^*)}$

A kidney co-operative must be designed carefully for it to be sustainable. A low value for p^* is socially desirable because it would be affordable to a large number of patients, however, such a value of p^* may not generate enough revenue to incentivize a sufficiently large number of donations to make the co-operative viable. A large number of kidney donations, F^*, would also be desirable from an efficiency standpoint, but it may only be achievable with a level of revenue that the contributing patients are not able to provide. Lastly, a large probability of getting a kidney for those staying on the waiting list, θ^*, may also be desirable, but it might induce contributing patients to go on the waiting list instead of contributing. A sustainable kidney co-operative balances these tradeoffs and, when well designed, produces large values for F^* and θ^* nevertheless.

Sustainable kidney co-operatives always exist. This is illustrated by the fact that the equilibrium of the conventional market mechanism can always be implemented as the outcome of a kidney co-operative. If the co-operative's managers set $p^* = p^m$, the solution to equations 1 and 2 above generate a quantity of donations F^* equal to F^m, and a co-operative dividend δ^* equal to p^m. Thus, the same conditions that guarantee the existence of conventional market equilibria also guarantee the existence of sustainable kidney co-operatives. As a general rule, however, there are more sustainable kidney co-operatives than equilibria of the market mechanism, a fact that has important normative implications, as discussed below.

We now illustrate how to find values for (p, F, θ) that characterize a sustainable kidney co-operative. First, consider a value for p (say, p^*) no smaller than the conventional market equilibrium price p^m, and conjecture that $p^* \leq \bar{v}(1-\theta^*)$. Total co-operative revenue at such a price is given by $p^* \cdot F^D(p^*, \theta^*)$.

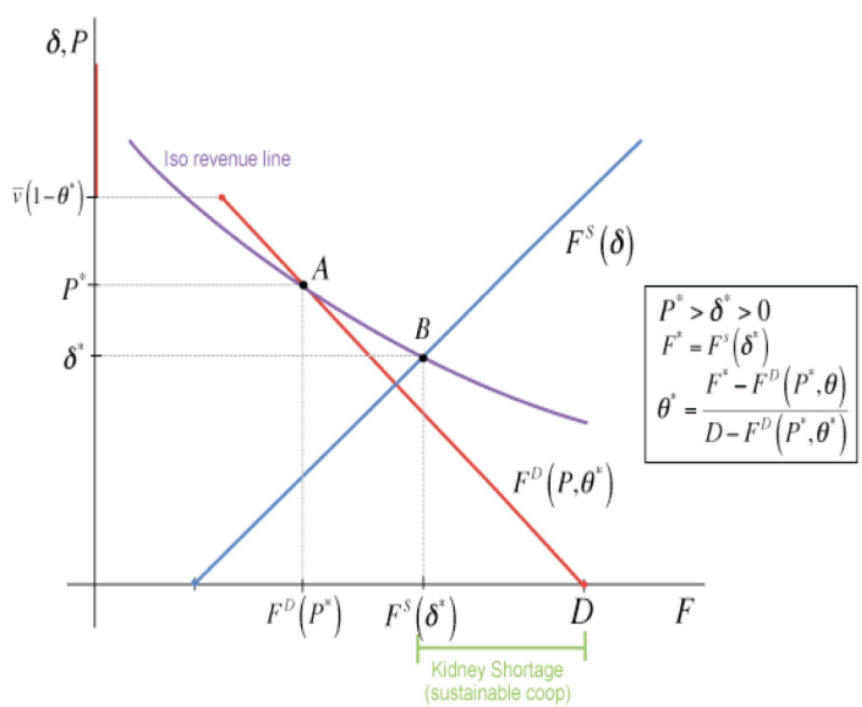
The sustainable number of kidneys donated, F^*, and the co-operative dividend, θ^*, jointly satisfy $\delta^* = \frac{p^* \cdot F^D(p^*,\theta^*)}{F^*}$ and $F^* = F^S(\delta^*)$.[6] Next, compute $\theta^* = \frac{F^* - F^D(p^*,\theta^*)}{D - F^D(p^*,\theta^*)}$ and verify that $p^* \leq \bar{v}(1-\theta^*)$. If so, we have found values for (p, F, θ) that characterize a sustainable kidney co-operative. If not, select a lower value for p^* and repeat this process.

A given patient contribution p induces at most one sustainable kidney co-operative. To see this, note that, for a well-behaved demand curve and some θ, a price p induces either 0 contributing patients (if $p > \bar{v}(1-\theta)$) or a weakly positive quantity, $F^D(p)$, of contributing patients, who are all the patients able to pay the contribution (if $p \leq \bar{v}(1-\theta)$). Then, depending on the value of θ, a generic p induces at most two possible total co-operative revenues ($p \cdot F^D(p)$), one of which is 0 (again, if $p > \bar{v}(1-\theta)$), and the other of which is positive (if $F^D(p)$ is positive). Then, for a smooth, well-behaved supply curve with $\frac{\mathrm{d}F^S(\delta)}{\mathrm{d}\delta} \geq 0$, each level of total co-operative revenue can purchase exactly one quantity of kidneys, F. Under the rules of the co-operative, all contributing patients receive a kidney, leaving $F - F^D(p)$ kidneys for non-contributing patients on the waitlist. Thus, each p^* is associated with at most one (positive) F^* and θ^*.

[6]To see this, notice that for values of F, say F_0, less than F^*, the co-operative dividend is $\delta_0 = \frac{p^* \cdot F^D(p^*,\theta^*)}{F_0} > \delta^*$. But that dividend level, δ_0 is high enough that it would entice more kidney donations than F_0, that is, $F^S(\delta_0) > F_0$. Therefore, the sustainable number of kidney transplants, for $p = p^*$, will be larger than F_0. For values of F, say F_1, greater than F^*, the co-operative dividend is $\delta_1 = \frac{p^* \cdot F^D(p^*,\theta^*)}{F_1}$. But that dividend level, δ_1 is too low; it would entice fewer kidney donations than F_1, that is, $F^S(\delta_1) < F_1$ Therefore, the sustainable number of kidney transplants, for $p = p^*$, will be smaller than F_1. This analysis is represented in figure 4, where expression 1 is illustrated as a (curved, downward sloping) purple line and expression 2 is illustrated as a (straight, upward sloping) blue line. The sustainable kidney co-operative is then represented by the pair (A, B).

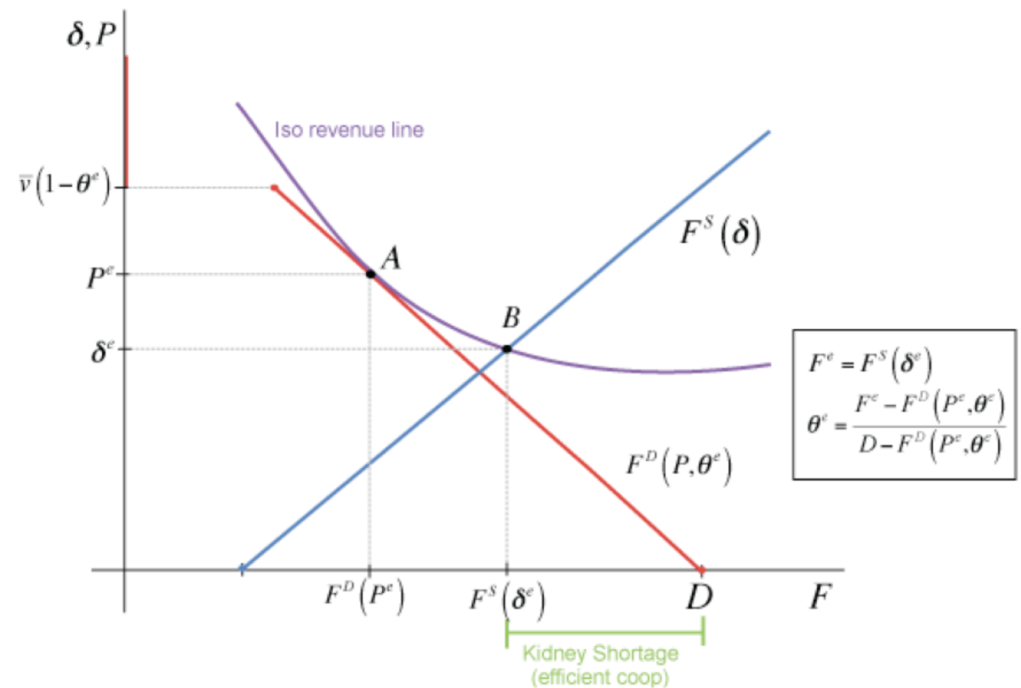
We say that a sustainable kidney co-operative is *(second best) efficient* if it maximizes the number of kidney transplants in the population among sustainable kidney co-operatives. We denote a generic efficient kidney co-operative with the values (p^e, F^e, θ^e). As discussed previously, an efficient kidney co-operative can never perform worse – in terms of the number of transplants – than the market mechanism. It can, however, perform better, under Assumption 1. To see this, notice that, under Assumption 1, a small increase in p, starting from the level prevailing at the equilibrium of the market mechanism, p^m, would simultaneously decrease the number of contributing patients and increase the co-operative's revenue, which has the effect of increasing the number of kidney transplants relative to the ones that obtain under the market mechanism. Figure 4 illustrates such a case.

Figure 4: A Sustainable Kidney Co-operative



Consider the kidney co-operative represented by the pair (A, B) in figure 4. This kidney co-operative can be improved on, from an efficiency standpoint, as there are price-quantity pairs, above and to the left of point A on the demand curve for kidneys, which yield higher co-operative revenue, and therefore more transplants. The kidney co-operative represented by the pair (A, B) in figure 5, however, cannot be improved on from an efficiency standpoint. It is thus an efficient kidney co-operative.

Figure 5: An Efficient Kidney Co-operative



One of the objections to the market mechanism is that it completely excludes those unable to pay for a kidney, as $\theta^m = 0$. Thus, as discussed previously, the market mechanism is never a Pareto improvement over the status quo. Efficient kidney co-operatives, on the other hand, have $\theta^e > 0$ (under Assumption I) and therefore have the potential for being a Pareto improvement over the status quo. This is the case as long as $F^S(0) \ll D$, that is, as long as the initial kidney shortage is sufficiently large,[7] which is consistent with anecdotal evidence.

A kidney co-operative is a Pareto improvement over the status quo whenever $\theta^* > \theta^0$, that is, when the probability of a non-contributing member receiving a kidney from the co-operative increases relative to the status quo. Contributing members are better off, as they pay an amount p^* for a kidney, which is less than the value $\bar{v}$ they assign to that kidney. Non-contributing members as better off, as the probability with which they receive a kidney has increased. Finally, donors are better off, as they receive δ^* for a kidney which they value (weakly) less than δ^* (in addition to full kidney insurance; see section 2.4).

Moreover, under Assumption I, an efficient co-operative not only maximizes the number of transplants among the sustainable co-operatives. It also maximizes, among the sustainable co-operatives, the probability θ of obtaining a kidney while on the waiting list. The intuition behind this result is simple: consider the number of transplants arising in the conventional market equilibrium, F^m. Under Assumption I, increases of p from p^m produce both an increase in revenue (hence an increase in transplants) and a decrease in contributing donors. Thus, the probability θ of obtaining a kidney while on the waiting list increases with the number of transplants.

An efficient co-operative is the most equitable among sustainable co-operatives in the sense that it makes θ as close as possible to the probability of getting a kidney enjoyed by the contributing patients in a kidney co-operative (that is, probability one). Note that complete equality between these probabilities is not viable. Such a rule would make the contributing patients want to go on the waiting list instead of contributing, making

[7] This essentially follows from the direct effect D has in the computation of θ^*

it impossible for the co-operative to raise any revenue with which to compensate donors.

2.4 Implementation and Extensions

Successfully implementing a kidney co-operative requires managers to address a number of concerns that have been heretofore ignored. This section addresses a number of these concerns.

2.4.1 Donation Risk

An important objection to the introduction of a conventional market mechanism is that donors will not understand the risks they're taking when donating and that they will therefore not be properly compensated for taking that risk.[8] Fortunately, the health risks from donating have been documented to be very small. Thus, kidney co-operatives will be able to afford an offer of "full kidney insurance" to the donors as part of their compensation for donating. Kidney insurance is very simple: in the unlikely event that a donor ever needs a kidney in the future, it will be provided to the donor free of charge. A kidney co-operative that offers kidney insurance thus allocates the kidneys it harvests among three groups of people, in this order: first, previous kidney donors who now happen to need a kidney; second, contributing patients; and third, non-contributing patients.

If α is the number of previous kidney donors who currently require a kidney, a sustainable kidney co-operative that offers kidney insurance is now characterized by (p^*, F^*, θ^*) that satisfy the following:

[8] Notice, however, that the same could be said about those who contribute kidneys under the status quo.

1. $0 < F^D(p^*, \theta^*) + \alpha \leq F^S(\delta^*) = F^*$

2. $\delta^* = \frac{p^* \cdot F^D(p^*, \theta^*)}{F^*}$

3. $\theta^* = \frac{F^* - F^D(p^*, \theta^* - \alpha)}{D - F^D(p^*, \theta^*)}$

A previous donor requiring a kidney later in life is so unlikely that the effect of providing kidney insurance on the operation of a kidney co-operative is likely to be minimal. For instances, only 11 out of 3,700 donors who donated a kidney between 1963 and 2007 later needed dialysis or a kidney transplant. This proportion is statistically identical to that arising in the overall population.[9]

Other safeguards already in place today would be exercised to ensure that the decision to donate is made conscientiously on the part of all kidney donors. A waiting period, alongside the provision of medical care to deal with any short- and long-term health consequences from donating, would also be instituted. A co-operative should, as much as possible, fully compensate the donors for any and all of the deterministic and probabilistic inconveniences that can accompany kidney donating.

2.4.2 Blood Type Compatibility

In a standard kidney transplant, recipients can only receive a kidney from donors with compatible blood types. Type O recipients can only be transplanted with kidneys from type O donors, type A recipients can be transplanted with kidneys from type O or A donors, and type B donors can be transplanted with kidneys from type O or B donors. Type AB recipients can be transplanted with kidneys from donors of any blood type.

[9] Ibrahim, H., et al. (2009).

A sustainable kidney co-operative ensures that anyone who wishes to donate a kidney may do so, that contributing patients are guaranteed a kidney, and that non-contributing patients each receive a kidney with equal probability. As no one is “turned away" in equilibrium, it is likely that there will not exist an assignment scheme that pairs each potential patient with a donor of a compatible blood type. This would occur if, for example, the quantity of type O patients were greater than the quantity of type O donors in equilibrium.

Until recently, a donor kidney of an incompatible blood type was almost certain to be rejected by the recipient's body. However, recent advances in kidney transplantation have not only made blood type incompatible transplants possible, such procedures have become quite common in some regions. For example, in Japan, over 30% of living-donor transplants are now between blood-type incompatible patient/donor pairs. Moreover, doctors specializing in donations between incompatible pairs now enjoy transplant outcomes that are comparable to those between blood-type compatible patient/donor pairs (see, for example, Takahashi and Saito, 2013). Despite these recent breakthroughs, blood type incompatible donations require the patient to undergo additional costly procedures prior to the transplant procedure itself. Stegall, et al. (2009) compares a number of these procedures.

A kidney co-operative will do the best it can to assign patients to donors of the same blood type. However, for the (almost certainly) small number of patients for whom there is not a donor kidney of a compatible blood type, the co-operative will cover the expenses associated with the additional procedures necessary for a blood type incompatible transplant. This wrinkle does little to affect the previous analysis. Consider a function $E(F)$, which, for some quantity of donations F, gives the expected minimum number of blood type incompatible transplants necessary. Then, if the additional cost associated with a blood type incompatible transplant is given by t, a sustainable kidney co-operative which accounts for potential blood type incompatibilities satisfies equations 1 and 3 above, while also satisfying $2'$:

$2'$. $\delta^* = \frac{p^* \cdot F^D(p^*, \theta^*) - t \cdot E(F^*)}{F^*}$

This equality ensures that the co-operative will have zero expected profits after accounting for the additional costs from blood type incompatible transplants.

2.4.3 Unknown Reservation Prices

The efficient design of a co-operative requires the co-operative manger to know μ and ρ which is, of course, unrealistic. The co-operative must be designed in such a way that it will provide patients and donors with the right incentives to reveal their reservation prices, so that the co-operative can operate effectively. This can be done using a mechanism that shares some properties with the VCG mechanism. The following provides a sketch of this mechanism.

Consider a finite but large population of potential donors and patients. Let I be the population of patients and J the population of potential donors. The distributions μ and ρ are redefined accordingly (i.e. they're now finite). Each patient and potential donor will be asked to report their private reservation price. Let w_i^r be patient i's reported wealth level and w_{-i}^r be the wealth level reported by all patients, except for patient i. Let c_i^r be potential donor j's reported reservation price and c_{-j}^r be the reservation prices reported by all potential donors, except for potential donor j. Such lists w_{-i}^r and c_{-j}^r define distributions ρ_{-i}^r and μ_{-i}^r over the characteristics of the populations $I \setminus i$ and $J \setminus j$. Let $w^r = (w_i^r)_{i \in I}$ and $c^r = (c_j^r)_{j \in J}$. Such lists in turn define the distributions ρ^r and μ^r.

For each $i \in I$ let p_{-i} and θ_{-i} be defined as in the efficient kidney co-operative given by the distributions ρ^r_{-i} and μ^r. For each $j \in J$ let δ_{-j} be defined as in the efficient kidney co-operative given by the distributions ρ^r and μ^r_{-j}.

Each patient is asked whether they want to contribute to the co-operative, and to report their wealth level. If they choose to contribute, they must contribute p^r_{-i} and if they don't contribute they will be put on the waiting list and face a probability of obtaining a kidney equal to θ_{-i}. Ask each potential kidney donor $j \in J$ whether they want to donate a kidney to the co-operative, and to report their reservation prices. If they choose to donate they collect δ_{-j}, plus full kidney insurance.

It is a routine matter to show that, in this setting, truthful reporting is a dominant strategy for all patients and potential donors. Moreover, it can be shown that, the dominant strategy implementation and the outcome from the efficient co-operative designed under the true distributions μ and ρ are very close in a "large co-operative" setting. Loosely, a large co-operative is characterized by the presence of lots of types of donors and patients, and a small number of people of each type, relative to the size of the co-operative. See the Online Appendix for details.[10]

2.4.4 Timing and Heterogeneity

Thus far, our analysis has exclusively taken place in a static framework. This is, of course, unrealistic, as the set of potential donors and patients is constantly in flux. How, then, should a kidney co-operative decide when and how often to "clear" the market?

It is tempting to execute a transplant as soon as a contributing patient/donor pair is

[10]Eames, Holder, and Zambrano (2016). See also Budish (2011).

found, or for the co-operative to purchase a kidney for a waitlisted patient as soon as the requisite funds become available. However, heterogeneity among recipients and donors makes this plan of action infeasible. The most obvious source of heterogeneity is location. Having the patient and/or donor travel long distances to the site of the operation is not only costly, it can be dangerous for an ill patient.

Another important source of heterogeneity is the tissue type, or human leukocyte antigens (HLA) type, of the potential recipient and patient; HLA type is determined by a combination of six proteins that reside in a person's body. Opelz (1997) shows that, as the number of proteins shared by the HLA types of the patient and donor decreases, so too does the likelihood of survival of the transplanted kidney. However, a "perfect" six-protein match between two people who are not blood relatives is exceedingly rare, and, indeed, many successful transplants occur between people who have few or no matching proteins.

Due to these important sources of heterogeneity, as well as the time required for co-operative managers to discover μ and ρ[11], the market should be cleared at discrete intervals, as in other centralized markets with significant heterogeneity (e.g. the National Resident Matching Program[12]). There are a number of factors to consider when determining the length of this time interval. Heterogeneity among patients and donors suggests that co-operative managers should wait long periods of time between clearing the market in order to accumulate a large pool of patients and donors from which good pairings can be chosen. On the other hand, living without a functioning kidney is associated with both a high mortality rate and a low quality of life. Furthermore, the longer a

[11] The process by which this can be done is described in section 2.4. Note that μ and ρ may evolve as time passes as, for example, potential donors with low c_j's donate kidneys and subsequently leave the market.

[12] National Resident Matching Program (2016)

patient has suffered from kidney failure, the higher is the probability the patient's body will reject the transplant.[13]

The fact that patients and donors are not immediately paired-up has a number of interesting implications. First, a kidney co-operative can deal only in living kidneys. Though cadaveric kidneys currently comprise over half of the kidneys transplanted each year,[14] surgery must be performed within a matter of hours of the donor's death for the kidney to be viable. Given the waiting periods between the co-operative "clearing" the market, cadaveric kidneys clearly cannot be offered through a kidney co-operative.

Furthermore, in many countries – including the United States – cadaveric kidneys are not the property of the heirs of the deceased, and thus private parties cannot legally offer cadaveric kidneys to the co-operative. A kidney co-operative would exist separate from, but in addition to, deceased donor transplantation programs which are currently in place. Of course, being on a waiting list for a cadaveric kidney would not exclude patients from participating in a kidney co-operative. The second consequence of such a market design is that the efficient number of kidney co-operatives is one. Though multiple kidney co-operatives could, of course, coexist, one co-operative would allow for the largest possible pool of potential donors and patients to accumulate in a fixed period of time, and would thus facilitate the best matches among donors and patients.

2.5 Conclusion

Kidney co-operatives are a form of implementing an excise tax levied on buyers – collected in units of the good being taxed – in an otherwise ordinary demand and supply

[13] Meier-Kriesche and Kaplan (2002)
[14] National Kidney Foundation (2014)

environment. In our context, the demand function is interpreted as the "ability to pay" rather than "willingness to pay", though the analysis does not depend on this interpretation. We have shown that kidney co-operatives have a number of desirable properties, such as the potential to generate a large number of transplants, a large probability of obtaining a kidney for those who can't afford to pay for one, and being self-financing. We thus believe kidney co-operatives are attractive institutions for alleviating the kidney shortage vis-à-vis the status quo and the conventional market mechanism.

The primary purpose of this paper is to explore a simple mechanism to solve the kidney shortage; this mechanism accounts for the objections that the public typically expresses towards conventional market-type solutions. Future research should involve seriously thinking about how the implementation of such a mechanism would be handled in practice, testing the mechanism experimentally, and carefully examining the consequences that might accompany the implementation of such a mechanism on a large scale.

Chapter 3

Charity in the Laboratory: Matching, Competition, and Group Identity

Gary Charness and Patrick Holder

3.1 Introduction

An issue of considerable economic importance is how to encourage philanthropic activity. In the U.S., $324 billion was donated to charitable organizations in 2015, with more than $2 billion spent annually in fundraising activities. Additional benefits from charitable contributions include the possibility of encouraging people to provide more help for those in need, as it would seem that a society in which people help each other out has more social capital and is much less likely to be divisive[1]. The policy implications of this

[1]Dictionary.com defines social capital as "the network of social connections that exist between people, and their shared values and norms of behavior, which enable and encourage mutually advantageous social cooperation". For a discussion, see Putnam (2000).

research program are clear. If society can devise effective mechanisms to induce more philanthropy, the social-welfare benefits would be quite substantial, even more so if such a program were at the same time able to induce a stronger sense of social cohesiveness.

Charitable giving has been found to be sensitive to a variety of conditions, so that behavioral mechanisms may very well affect increase charitable contributions. Yet the study of effective mechanisms for harnessing the potential underlying willingness to contribute is still in its early stages. One approach towards increasing contributions involves triggering social preferences, which are generally considered to involve a willingness to sacrifice some of one's material wealth in order to benefit other individuals or society at large. Prominent social-preference models (e.g. Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002) can explain charitable giving if the person considering making a donation believes that the recipient is poorer than the donor.

In addition, Charness and Rabin (2002) highlights the importance of social efficiency, whereby people are motivated to increase the total payoff received by the reference group. Experiments that involve providing external matching for the donations made by the participants show support for the influence of efficiency in charitable contributions. In field studies, Karlan and List (2007), Eckel and Grossman (2008), and Martin and Randal (2008) all find significant effects of matching on revenue, in the neighborhood of a 20-50% increase[2]. From the standpoint of a prospective donor, her donation is considerably more effective – for the same cost, more money is transferred from those who can afford it to less affluent recipients.

While charitable contributions have been examined largely in field experiments, there has

[2]However, Rondeau and List (2008) find that providing matching funds does not increase donations in the field, and Huck and Rasul (2011) find that providing matching funds when there is a lead donor in place is counter-productive.

also been research conducted in the laboratory. Perhaps the first such experiment was Eckel and Grossman (1996), which used laboratory donations to the American Red Cross as a treatment in a dictator-game experiment. Subjects were much more likely to donate to the charity than to an anonymous peer in another room. Eckel and Grossman (2003) test whether rebates or matching contributions are more effective for fund-raising in the laboratory. Each participant chose a charity from a list of 10. They varied whether a reduction in cost was due to a rebate of a portion of the contribution or a matching of the contribution, finding that “contributions are significantly higher with matching subsidies than with rebate subsidies". Eckel and Grossman (2004, 2006a, and 2006b) continue this investigation of the effectiveness of subsidies.

There has also been a recent boom in charity laboratory experiments, spurred on by funding from the Science of Philanthropy Institute. For example, the February, 2017 special issue of the Journal of Behavioral and Experimental Economics, while primarily focusing on field studies, does contain several field experiments. For example, Brown, Meer, and Williams (2017) test how “third-party ratings of charities impact charity choice and donative behavior". They do find that these ratings impact behavior. A surprising result is that there are no obvious preferences for local groups, much as we find in our experiment. Krieg and Samek (2017) conduct a laboratory experiment in which two public goods games are played at the same time, each involving a charity. They investigate the effects of recognition, a bonus for contributing, and sanctions on contributions, finding that bonuses increase contributions, but that recognition and sanctions are relatively ineffective.

A dimension that has received only limited attention is the importance of group identity in relation to charitable contributions. Experimental work (e.g. Charness, Rigotti, and Rustichini, 2007; Chen and Li, 2009; Benjamin, Choi, and Strickland, 2010) has shown that salient group membership can have strong effects on behavior. In relation to charitable contributions, the intuitive prediction is that people will donate more when the likely recipient belongs to the same group as the donor. Billig and Tajfel (1973) find that even with minimal groups (formed using a very modest sense of identity), group membership has strong effects on the allocations chosen for in-group and out-group members. Another form of group membership effect is related to peer effects. For example, Babcock, Bedard, Charness, Hartman, and Royer (2015) shows that participants provide more effort in a field experiment to avoid "letting down one's team". Can placing people in even an arbitrarily-formed competing team have a positive effect on charitable contributions?[3]

A final potential mechanism is that of competition, and we consider both individual and team competition for matching funds. Competition has generally been found to decrease social preferences. For example, Fershtman, Gneezy, and List (2012) show that social preferences in a standard laboratory dictator mini-game vanish when the scent of competition is in the air. Competition increases one's willingness to accept very small amounts in ultimatum-type games (see e.g. Roth, Prasnikar, Okuno-Fujiwara, and Zamir, 1991; Grosskopf, 2003). On the other hand, competition has been found to induce more contributions in a public-goods game where only the team with the higher level of contributions received payment (Rapoport and Bornstein, 1987), higher effort in a minimum-effort game where only the team with the higher minimum received payment (Bornstein, Gneezy, and Nagel, 2002), and more cooperation in a type of Prisoner's Dilemma (Bornstein and Ben-Yossef, 1994). Erev, Bornstein, and Galili (1993) used a "lifelike orange-picking task" and found that team competition led to higher productivity, particularly when the teams were similar in ability.[4]

[3]We use the terms "group" identity or membership and "team" competition throughout the paper.

[4]Note that our charitable-contribution environment differs from these others in that in our case one's decision affects only one's own payoff and not those of any other person in the experiment. For example, contributions in public-goods games with group competition affect the payoffs for both one's in-group

More recently, Ai, Chen, Chen, Mei, and Phillips (2015) conducted a field experimental study in which they worked with Kiva, a micro-finance organization, in trying to increase lender participation. They utilized existing lending teams, which were social groups that competed with each other in making loans; existing lenders received e-mail messages notifying them of the existence of teams and in six treatments recommending three teams to a lender. For those who joined a team, both the number of loans made and the amount loaned increased significantly. They conclude that the results support team competition as an effective mechanism for promoting pro-social behavior.

This study is close to ours in spirit, but there are some important differences. First, pro-social lending is not quite the same as charitable donations, since a lender expects (or at least hopes) to be repaid and earn a return on the loan, whereas a charitable contribution is freely given. Second, we randomly assign individuals to anonymous teams rather than offering them the opportunity to select in (perhaps without anonymity).[5] Third, it is quite likely that reputational considerations affect behavior in the field experiment, since there may very well be future interactions between various parties and many additional and extraneous features are also present; by contrast, the laboratory environment is controlled and anonymous. Fourth, the results of our competition are private while the results of the competition in the field study are public. Finally, we also consider matching environments and compare individual competition with team competition. We feel that these differences across the studies make them excellent complements.

We test the potential mechanisms in the laboratory, where it is possible to conduct clean tests of each potential mechanism.[6] To preview our results, we find that the level of

and out-group members.

[5]Babcock, Bedard, Charness, Hartman, and Royer (2015) provide very clean evidence that performance improves for those participants who have selected to be in a particular treatment rather than having been assigned to it.

[6]See Falk and Heckman (2009) for a discussion about the benefits and control available from labora-

contributions is relatively low in the baseline condition, when participants are simply asked to donate some portion of their endowment to charity; there is no significant difference depending on whether this is a local charity or one without local roots. However, providing matching funds has a substantial effect on donations. Relative to the baseline, this effect increases significantly when all contributions are matched or when there is individual competition for matching funds, and it increases dramatically with team competition for matching funds. We find that the most effective mechanism in the lab is to form teams (albeit anonymous and arbitrary ones) and to provide matching funds for which these teams must compete through their chosen level of donations. This effect of team competition ("letting down the team") being a very effective device for increasing charitable contributions is a main contribution of the paper. An additional main contribution is to show that while team competition is very effective, individual competition for matching funds is largely ineffective for increasing contributions relative to having one's donation automatically matched.

The remainder of this paper is organized as follows. Section 3.2 presents our experimental design and implementation, section 3.3 displays our experimental results, and section 3.4 concludes.

3.2 Experimental Design & Implementation

We conducted four treatments at UC Santa Barbara with an in-group charity (a UCSB charity) and one with an out-group charity (the United Way). In all treatments, we endowed each person with $21, which included a mandatory $5 show-up fee. Participants received decision forms, on which they designated whether they wished to contribute $0,

tory experiments.

\$4, \$8, \$12, or \$16 to the charity designated in the instructions (the instructions and decision sheets are shown in the section 3.5).[7] The relevant charity was described to the participants in every session.[8] In the baseline treatments (in-group and out-group charities, with no matching donations), participants simply chose a donation from the available set. In the all-match treatment, we matched each person's contribution on a 1:1 basis.[9] In the individual-competition treatment, we matched individual contributions if the individual was above the median contribution of the participants in the room. In the team-competition treatment, we formed teams of three at random from the participants (visual identification of other team members was not possible) and matched the contributions of the top half of these teams in the session. Each of the team-competition sessions had a multiple of six participants (either 12 or 18) participating. The in-group charity was used in each matching treatment. In all treatments, no feedback was given regarding the contributions of other participants; however, participants were informed that they would learn whether their contribution was matched.[10] Table 1 summarizes

[7]The choice to have a limited number of possible contributions was made because 1. we felt that decisions would be easier for participants to make, 2. we felt it would simplify the non-parametric analysis, and 3. it is common for solicitations for charitable contributions to offer a few boxes from which participants can make choices (although they typically do have a space for "Other"). For charities in the field, we suspect that point 3 reflects the notion in point 1 that people are more likely to contribute when decision costs are lower.

[8]Though we do not know our subjects' preconceptions about the in-group or out-group charities, we tried to select charities that subjects may have heard about previously (the in-group charity has been written about in student newsletters and the school newspaper, while the United Way is one of the largest national charity organizations). The charities were carefully selected to have similar mission statements; both charities emphasize leadership, education, and health. Furthermore, Michael D. Young, the namesake of the in-group charity, served a long tenure on the board of directors of the United Way's Santa Barbara chapter. Thus, any positive (or negative) feelings subjects have towards one charitable organization might be expected to carry-over to the other organization. The descriptions students were provided with regarding each organization were designed to be very similar (see supplementary material at the end of this chapter).

[9]We considered varying the matching rate, but settled on a 1:1 match, as this is most common in the field.

[10]We did not conduct a treatment with group identity but without competition. We felt that having more than one team would suggest to subjects that they are competing; having multiple teams without competition is unusual. In addition, evidence from a public-goods game in Eckel and Grossman (2005) shows that the mere existence of teams does not reduce free-riding, whereas team competition does.

our treatments.

Table 1: Treatment summary

Treatment	Charity	1:1 Donation Matching?
Baseline Treatments		
In-group no-match	UCSB Charity	Never
Out-group no-match	United Way	Never
Matching Treatments		
All-match	UCSB Charity	Always
Individual-competition	UCSB Charity	If individual contribution is greater than or equal to median
Team-competition	UCSB Charity	If group's total contribution is greater than or equal to median group contribution

As it was critical for each person to believe that the money that they sacrificed would go to the designated charity, participants were informed that a check would be written to the charity for the appropriate amount and that this check would be written in front of them and deposited in a stamped envelope addressed to the charity. We also told them that we would then hand them the envelope and they could mail these envelopes themselves. Of course, all of these statements were true.

Sessions were conducted at UC Santa Barbara, with students recruited through ORSEE (Greiner, 2015). The database includes a spread of different majors at UCSB, as the message inviting students to register for experiments was sent campus-wide. All in all, we had 144 participants in the sessions without matching funds and 130 participants in the sessions with matching funds. No one could participate in more than one session, so each observation is completely independent.

Given our design, what would theoretical models predict? The standard neoclassical model predicts that no contributions are made in any treatment. However, if one considers that people who are less well off will receive charitable donations and if these people are seen as being part of one's reference group, all of the prominent social-preference models discussed above allow for the possibility of charitable donations. If one is concerned about the effectiveness of one's donation (their "bang per buck"), all of these models predict an increase in donations when funds are matched; this increase should be larger when all donations are matched than when only half of the donations are matched. Finally, none of these models predict a difference in donations between the individual-match and group-match treatments; however, models of social identity (Akerlof and Kranton, 2000; Chen and Li, 2009) could potentially explain a difference across these treatments.

3.3 Results

Table 1 presents mean donations by treatment. In the no-matching treatments, there was an average contribution of 3.333 for the in-group charity and 4.111 for the out-group charity. This direction is not what was expected, although the difference is not significant (a Kolmogorov-Smirnov test of cumulative distributions gives $p = 0.411$, two-tailed test). However, in hindsight, we realize that the United Way recipients may well have been perceived as being poorer and more needy than recipients of the UCSB fund.[11] Also, if a donor wants to behave altruistically, donating to an external charity may be perceived as a more disinterested (and therefore more altruistic) decision.[12] These considerations may have been stronger than the sense of group identity we were able to induce.

[11] In this case, the full Charness and Rabin (2002) model (top of p. 852) would predict higher donations for the United Way, assuming that recipients are considered to be part of the reference group (which seems reasonable, given that participants are making positive donations). We thank Alex Imas for this point.

[12] Donating to the in-group charity could be perceived as more egotistic (showing interest in one's own organization instead of others. We thank Diego Pulido Lema for this idea.

Table 2: Average Total Donation by Treatment

	Avg. Contribution ($)	Number of Observations
No-Matching Treatments		
In-Group	3.333 (0.371)	72
Out-Group	4.111 (0.475)	72
Total No-Matching Treatments	3.722 (0.302)	144
Matching Treatments		
Match All	5.333 (0.718)	48
Individual Competition	5.200 (0.761)	40
Team Competition	7.524 (1.040)	42
Total Matching Treatments	6.000 (0.493)	130

Standard errors given in parentheses.

Regarding the matching treatments, we find that matching contributions raises the contribution level in all cases. The level of contribution is almost the same when these are matched for all participants and when there is individual competition for matching funds ($p = 0.912$, two-tailed Kolmogorov-Smirnov test). The contribution levels in these competition treatments are 40-43% higher than in the pooled no-matching treatments.[13]

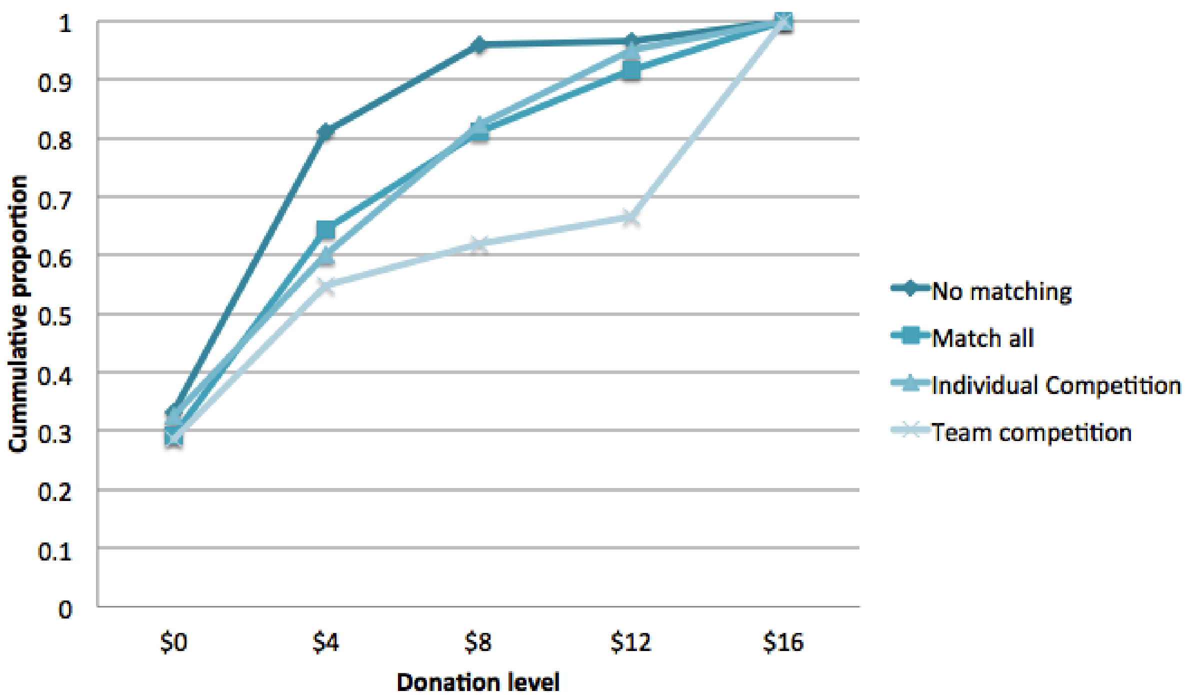
Table 3: Distribution of Contributions by Treatment

	$0	$4	$8	$12	$16
No matching	48	69	21	1	5
Match all	14	17	8	5	4
Individual competition	13	11	9	5	2
Team competition	12	11	3	2	14

For our statistical tests, we compare the distribution of contributions in the various cases, as seen in Table 3 and Figure 1. Given the lack of difference, we pool the in-group and out-group no-matching treatments for our main statistical tests.

[13]Overall donations in the matching treatments were significantly higher than overall donations in the no-matching treatments ($p = 0.003$, two-tailed Kolmogorov test).

Figure 1: Cumulative Distribution of Donations



Kolmogorov-Smirnov tests confirm the visual impression that the team-competition treatment has significantly more donations than any other treatment ($p = 0.045$ for the match-all treatment, $p = 0.025$ for the individual-competition treatment, and $p = 0.000$ for the pooled non-matching-funds treatments, all two-tailed tests). Neither the difference between the pooled non-matching treatments and the individual-competition treatment nor the difference between the pooled non-matching treatments and the match-all treatment is marginally-significant with two-tailed tests, although the first comparison is significant on a reasonable one-tailed test ($p = 0.030$), while the second comparison is marginally-significant with this test ($p = 0.065$).[14] We also note that the distribution

[14]We can also perform statistical tests comparing only the in-group charities; the smaller sample size leads to less significance. Nevertheless, the difference between the contributions in the team-competition treatment and in the in-group no-match treatment remains strongly significant ($p = 0.000$, one-tailed test). The difference between the in-group no-match treatment and the match-all treatment is significant at $p = 0.088$, while the difference between the in-group no-match treatment and the individual-competition treatment is significant at $p = 0.076$ (both one-tailed tests, in keeping with directional

for the team-competition treatment first-order stochastically dominates each of the other three treatments, and that the match-all and individual-competition treatments first-order stochastically dominate the pooled no-match treatments. This underscores these differences.

Thus far, the analysis has considered only donations from experimental subjects. However, for treatments in which donors may have their contributions matched, an alternative method of comparing treatments might consider the average total contributions, including any matched funds. This measure may be particularly of interest in an environment in which any matched funds are being provided by an outside party, rather than by the organization eliciting the donations. Appendix Table 1A gives average total contributions (including matched funds) for each treatment. The treatment effects described above are preserved when comparing average total contributions (including any matching funds) across treatments: donations in the team-competition treatment first-order stochastically dominate donations in other treatments, and Kolmogorov-Smirnov tests indicate that donation levels are significantly higher in the team-competition treatment than in the no-match, match-all, and individual-competition treatments ($p = 0.000$, $p = 0.083$, and $p = 0.048$, respectively, all two-tailed tests).

3.3.1 Structural estimation

In order to quantify the effects of the various experimental treatments, we estimate the coefficients in the following utility function:

predictions).

$$U_i^d = V_i^d + \varepsilon_i^d$$

$$\text{where} \quad V_i^d = \alpha^d + \beta_m^d x_{mi} + \beta_c^d x_{ci} + \beta_t^d x_{ti}$$

$d \, \epsilon \, \{0, 4, 8, 12, 16\}$ is the donation amount, i indexes the subject, m denotes the match-all treatment, c denotes the individual-competition treatment, and t denotes the team-competition treatment. x_m, x_c, and x_t are dummy variables which equal unity in the match-all treatment, individual-competition treatment, and team-competition treatment respectively, and 0 otherwise. U_i^d is the utility that individual i receives from donating an amount d to the charity. α^d is the utility associated with a donation of d dollars in the (baseline) no-matching treatment, and β_j^d is the additional gain or loss in utility from donating d in treatment j, where $j \, \epsilon \, \{m, c, t\}$. As we employ a multinomial logit model, unobservable errors, ε_i^d, are assumed to be distributed extreme value type I (Gumbel). We assume that, upon being presented with the donation task, subjects behave as if they calculate five utilities – $U_i^0, U_i^4, U_i^8, U_i^{12}, U_i^{16}$ – and choose the donation level associated with the highest utility.[15]

[15] In addition to the specification presented, we considered two other specifications. In the first, a random coefficient, γ_i^d, was introduced into the utility function. These terms were jointly normally distributed with mean 0, and were intended to capture "taste variation" that might be present in the population (see Train, 2003, Chapter 5 for details). This seems like a natural setting for introducing taste variation, as it is easy to imagine that the marginal utility from donating to charities varies widely from subject to subject. However, this alternative specification generated coefficients that are statistically indistinguishable from those in our main specification presented above. We also considered introducing self-reported demographic characteristics into the model. Specifically, we considered a specification in which $V_i^d = \alpha_d + \beta_m^d x_{mi} + \beta_c^d x_{ci} + \beta_t^d x_{ti} + X_i B^d$, where X_i is a matrix containing subject ages, in addition to dummy variables for gender, employment status, and degree of involvement with campus clubs. In total, this specification involved estimating 24 additional coefficients (captured by B), only 3 of which are significant at the 5% level. Again, including these additional terms has little effect on the β_t^d estimates presented in Table 3, and there are no notable patterns among the additional estimated coefficients. We thus find our results to be robust to alternative econometric specifications.

Table 4: Parameter Estimates in Multinomial Logit Model

Parameter	Estimate	$\alpha^d + \beta^d$	Wald test p-Value
α^4	0.363 (0.188)*	-	-
α^8	-0.827 (0.262)***	-	-
α^{12}	-3.871 (1.010)***	-	-
α^{16}	-2.262 (0.470)***	-	-
β_m^4	-0.169 (0.407)	0.194	0.591
β_m^8	0.267 (0.515)	-0.560	0.207
β_m^{12}	2.842 (1.137)**	-1.029	0.049**
β_m^{16}	1.009 (0.736)	-1.153	0.027**
β_c^4	-0.530 (0.451)	-0.167	0.683
β_c^8	0.459 (0.506)	-0.368	0.396
β_c^{12}	2.916 (1.139)***	-0.955	0.070*
β_c^{16}	0.390 (0.893)	-1.872	0.013**
β_t^4	-0.450 (0.458)	-0.087	0.835
β_t^8	-0.560 (0.700)	1.387	0.032**
β_t^{12}	2.079 (1.267)	-1.792	0.019**
β_t^{16}	2.416 (0.613)***	0.154	0.700

$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$. p-values in last column are for corresponding Wald tests. Observations by treatment: $N_{\text{no matching}} = 144$; $N_m = 48$; $N_c = 40$; $N_t = 42$. These parameter estimates are generated by pooling contributions to both the in-group and out-group charities in the no-match treatment. In Table 2A, located in the supplementary materials, we re-estimate the above parameters using only contributions to the in-group charity. However, because there were no donations of \$16 to the in-group charity in the no-match treatment, the $\alpha^{16}, \beta_m^{16}, \beta_c^{16}$, and β_t^{16} parameters cannot be estimated from this sample. For those parameters that can be estimated, the estimates above are similar to those in Table 2A.

Table 4 reports the results of our parameter estimation. To account for the scale invariance of utility, we set the non-random component of the utility from donating \$0 (V_i^0) to 0 in all treatments. To provide an example, a representative subject in the match-all treatment would first calculate the utility from the four positive donation levels; if they are all less than ε_0, the subject donates \$0. Otherwise, she donates an amount that generates the highest utility.[16]

The last column of Table 3 gives the p-values for the Wald test $\alpha^d + \beta_j^d = 0$. For example, 0.591 (in the fifth row of Table 3) is the p-value of the Wald test:

$$\begin{aligned} \alpha^4 + \beta_m^4 &= 0 \\ \text{or,} \quad V_m^4 &= V_m^0 \end{aligned}$$

The last column of Table 3 addresses the combined effect on utility of donating without any matching (α^d), plus the additional change in utility that the various matching treatments generate (β_j^d). Using a Wald test, the sum of these components is compared to 0, the utility of donating \$0. Note that, in this data, all statistically-significant sums are negative.

As the *estimate* column of the first four rows illustrates, (the non-random component of) utility is roughly decreasing in the amount donated for the no-matching treatments, with α^d significantly less than 0 at the one percent level for all $d >$ \$4. However, the last column of Table 3 shows that, in the m treatment, only donations of \$12 and \$16 generate utility levels significantly less than 0 at the 5% level, and in the c treatment, only donations of \$16 generate utility levels significantly less than 0 at the 5% level.

Finally, in the t treatment, donations of \$8 and \$12 generate utility levels significantly less than 0 (at the 5% level), but we fail to reject the hypothesis that there is a difference in utility generated by donating \$0 and donating \$16 in this treatment. This can easily be seen in the estimate column of Table 3; though α^{16} is considerably below 0, the additional utility boost from donating \$16 in the team competition treatment, β_t^{16}, is more than

[16] The corresponding utilities are $U^4 = 0.363 - 0.169 + \varepsilon^4, U^8 = -0.827 + 0.267 + \varepsilon^8, U^{12} = -3.871 + 2.842 + \varepsilon^{12}$, and $U^{16} = -2.262 + 1.009 + \varepsilon^{16}$.

large enough to make-up for the utility loss of giving away the maximum amount of \$16. Overall, the structural estimation squares well with the distribution of donations shown in Table 2.[17]

Though Table 2 presents the experimental data, there is so far no indication of whether or not the differences in donation frequencies within a specific treatment are statistically significant. The results from Table 3 help shed light on this issue. Pooling the donation frequencies in the baseline no-matching treatments, donations of \$0, \$4, \$8, \$12, and \$16 occurred 48, 69, 21, 1, and 5 times, respectively. The parameter estimates in Table 3 confirm what these donation frequencies suggest: for these treatments, $\alpha^0 > \alpha^d$ for $d\varepsilon\{8, 12, 16\}$, with $p < 0.01$.

The most dramatic change between the baseline no-matching treatments and the match-all treatment is the proportion of subjects donating \$12. In the baseline treatment, only one subject donated \$12, while more than 10% of subjects in the all-match treatment donated \$12 to the charity. This pattern is reflected by the parameter estimates in Table 3, as β_m^{12} is the only β_m^d estimate that differs significantly from 0 at $p < 0.05$. We observe a similar pattern in the individual-competition treatment, in which donations of \$0, \$4, \$8, \$12, and \$16 occurred 13, 11, 9, 5, and 2 times, respectively. Again, between the baseline no-matching treatments and this treatment, the most dramatic increase was in the proportion of subjects donating \$12; in this case, the proportion increased from 0.7% in the baseline to 12.5% in the individual-competition treatment. Once again, this

[17] In the baseline case we should see somewhat more donations of 0 than 4 (since the coefficient of 0.363 is marginally significant), substantially fewer donations of 8, and far fewer donations of 12 and 16; in the all-match condition, the coefficients suggest that we have slightly more donations of 4 than of 0, somewhat fewer donations of 8 than of 0, and a substantially lower rate for donations of 12 and 16. In the individual-competition condition, the coefficients suggest that we have slightly more donations of 0 than of 4 or 8, substantially fewer donations of 12 than of 0, and far fewer donations of 16 than 0; in the team-competition condition, the coefficients suggest that we have about the same number of donations of 0 and 4, considerably fewer donations of 8 and 12 than of 0, and slightly more donations of 16 than 0. All of these patterns are found in the raw data.

pattern is borne out in Table 3, in which $\beta_c^{12} is the only \beta_c^d$ estimate that differs significantly from 0, this time with $p < 0.01$.

This pattern doesn't hold in the team-competition treatment. In both the baseline and group competition treatments, a significant number of subjects donated \$0 or \$4, while only a few subjects donated \$8 or \$12. However, the proportion of subjects donating \$16 increased ten-fold, with only 3.5% of subjects donating the maximum amount in the baseline and 33.3% of subjects donating the same amount in the group-competition treatment. This dramatic increase is again reflected in Table 3, in which β_t^{16} is the only β_t^d term significantly greater than 0. Figure 2 summarizes the (non-random) total effects on utility from various contribution levels in our treatments, visually summarizing the parameter estimates in Table 3. For the no-match treatment, α^4, α^8, α^{12}, and α^{16} are plotted, while, for the matching treatments, the total utility associated with each contribution level, $\alpha_d + \beta_j^d$, is plotted (where d is the donation level and j is the matching treatment). There are two clear stylized facts: first, utility is decreasing in contribution level for all treatments *except group-competition*, and second, high donation levels (of \$12 or \$16) provide substantially less utility in no-match treatments than in matching treatments. While utility is roughly monotonically decreasing in the contribution level for the all-match and individual-matching treatments, results from the group-competition treatment suggest that there is a U-shaped relationship between contribution levels and utility, with contributions of \$0, \$4 and \$16 all generating roughly the same level of utility.

Figure 2: Parameter Estimates in Multinomial Logit Model

The structural model allows for a clear illustration of how a donor's attitude towards a particular donation level changes with the matching protocol, which may not be obvious from the raw data alone. For example, consider the question of how a donor's attitude towards intermediate donation levels of \$8 changes as the matching protocol changes. In the all-matching and individual-competition treatments, roughly 20% of subjects donated \$8. However, this proportion drops considerably to 7% in the group-competition treatment. There are two possible explanations for this decrease. The first potential explanation is that group-competition makes intermediate donation levels inherently unappealing, while an alternative explanation is that group competition makes another choice more appealing, and donors are simply substituting away from \$8 donations in favor of this other donation amount. As Figure 2 makes clear, the second explanation appears to be supported by the data. The utility associated with an \$8 donation remains relatively constant between the all-match, individual-competition and group-competition treatments, despite the large decrease in \$8 donations. This decrease appears to be due to group-competition making \$16 donations much more attractive than in other treatments.

Examining the proportion of subjects who donate \$0 across treatments suggests that there is a contingent of subjects that will pocket everything no matter what the treatment. This group of subjects – roughly a third of the population – is approximately the same across treatments.ă It appears that the effect we see in the matching treatments (and especially with team competition) isn't the result of influencing participants who were donating \$0 in the baseline treatment, but is instead from increasing the donations by participants who would have otherwise donated a smaller, but positive, amount.ă The proportion of high (\$12 or \$16) donations substantially lower in the no-matching treatment (6.2%) than in the match-all (26.5%) and individual-competition (25.9%) treatments, while the proportion was double in the team-competition treatment (53.3%) than in the other matching treatments. Thus, we see no change on the extensive margin, but a large shift on the intensive margin. While it seems quite difficult to affect the proportion of non-givers, one can in fact affect the level of generosity of those willing to donate.

3.4 Conclusion

Our laboratory design is geared towards identifying group-membership effects, matching considerations, and competition in relation to the charitable contributions made by participants in our sessions. To our surprise, donations to an in-group charity were not significantly greater than donations to a non-in-group charity; in fact, they were actually slightly lower. On the other hand, we find strong support for the influence of a different form of group membership ? belonging to a team in a competition.

We find that providing matching funds increases charitable contributions in the lab.

Automatic matching funds (without competition) increases donations by 10 percentile points (or 43%) from 23% to 33%. Competition for matching funds does not have a *per se* effect, since donations with individual competition for matching funds are actually very slightly lower (32 percent). However, team competition is quite successful at inducing a higher level of contribution, as here participants donate 47% of their endowment, more than twice the amount without any matching contributions. Our structural model lends support to our conclusions and matches up well with the observed distribution of donations. We find a considerable degree of heterogeneity in donations; the percentage of hardcore selfish participants is roughly constant across treatments.

We recruit participants who have not expressed any desire to contribute to charity (in fact, 32% do not make any donation) and we assign them randomly to a treatment; a targeted campaign could conceivably be more effective. Furthermore, as mentioned earlier, Babcock, Bedard, Charness, Hartman, and Royer (2015) find that performance improves substantially (by 27%) for participants who have chosen a treatment than who have been assigned to it.[18] One might therefore expect more of an impact from people who have voluntarily joined a team. The reluctance to “let down the team" seems to be a powerful motivation and we would expect this might well be stronger with endogenous group-formation.[19] Competition *per se* is not nearly as effective as when it is combined with being part of a team, as is shown by comparing the results from the individual-competition and team-competition treatments.

We note that the social-preference models mentioned in the introduction cannot explain the effect of team competition on donations, since the material payoff of another team member is unchanged by whether or not her donation is matched. However, an al-

[18]This is not driven by a selection effect, since 97% of the subjects given the choice made this choice.

[19]Gee and Schreck (2015) also find that beliefs about peers matter for the effectiveness of donation matching, in both the field and the laboratory.

ternative explanation is guilt aversion (Charness and Dufwenberg, 2006; Battigalli and Dufwenberg, 2007), whereby one experiences psychological disutility (feels badly) for disappointing the expectations of one's teammates regarding their donations being matched.

The complementary field study on micro-finance lending by Ai, Chen, Chen, Mei, and Phillips (2015) provides strong support for the notion that team identity and competition are forces that can be harnessed effectively for pro-social purposes; the confluence of a laboratory study and a large field study makes us more confident of this prediction. Thus, we feel that charities can potentially usefully employ the mechanism of creating (or encouraging people to form) competing teams.[20]

It is clear that considerably more research (and replication) on this topic is necessary before strong conclusions can be drawn. Nevertheless, our results indicate that there are indeed mechanisms that can increase charitable giving, even in the unembellished, artificial, and anonymous environment of the laboratory.

[20] In fact, evidence from a philanthropy program at Purdue University provides further support for the effectiveness of combining team identity with competition. The contributions of participating units (teams) were matched and these contributions were ranked on a leaderboard. While there is no control treatment for comparison, Purdue raised $14 million dollars on one day in 2015. See http://spihub.org/newsroom/blog/item/11-purdue-day-of-giving. We thank Anya Samek for bringing this to our attention.

3.5 Supplementary Material

3.5.1 Instructions

Introduction

Thank you for participating in this experiment. Please follow along carefully as the experimenter reads through these instructions. If you have questions at any point, please raise your hand. This is an experiment in the economics of decision-making. A research foundation has provided the funds necessary for conducting this study. For your participation, you will be paid at the end of this session. Your earnings will be paid in cash and/or a check made out to a charity, at your discretion. At a minimum, you will receive $5 for your participation. Instructions of how you will make decisions and earn more are provided below.

Decision Task

In addition to receiving the $5 payment for participation, you will be given a budget of $16 for this task. The experimenter will present you with informational materials regarding a specific charity. You will decide how much of your budget you would like to receive in the form of a check made out to the specified charity, and how much you would like to receive as a cash payment. The experimenter will provide envelopes and postage in addition to the check, but it will be up to you to place your donation in the mail.

Decisions will be limited to $4 increments. For example, you may choose to contribute $4 to the charity, while keeping the remaining $12 for yourself. However, you may not choose to contribute \$6, while keeping \$10 for yourself.

[THE FOLLOWING PARAGRAPH ONLY INCLUDED IN ALL-MATCHING SESSIONS]

Every dollar donated will be matched by the experimenter. For example, if you choose to donate \$12 of your \$16, you will receive the remainder – \$4 – and the charity will receive \$24 ($\$12 * 2$).

[THE FOLLOWING PARAGRAPH ONLY INCLUDED IN INDIVIDUAL COMPETITION SESSIONS]

You will compete with the other participants in the room. Participants will be ranked in terms of their donations, and the participants that are in the top half in terms of total donations will have their donations doubled. For example, if there were ten participants in the room, and you ranked fifth (or tied for fifth) in terms of total donations, your donation would be doubled. If you had chosen to donate \$12, the charity would receive \$24 instead of \$12. If you ranked sixth in terms of total donations, your donations would not be doubled, and the charity would receive \$12. In both cases, you would receive the remaining \$4.

[THE FOLLOWING PARAGRAPH ONLY INCLUDED IN TEAM COMPETITION SESSIONS]

There are two other participants who have the same colored slip as you ? this is your team. Your team will compete with the other teams in the room. The teams will be ranked based on the total donations of all team members, and the teams that are in the top half in terms of total donations will have their donations doubled. For example, if there were six teams in the room, and your team ranked third (or tied for third) in terms of total donations, your donation would be doubled. If you had chosen to donate \$12, the charity would receive \$24 instead of \$12. If your team ranked fourth in terms of total donations, your donations would not be doubled, and the charity would receive \$12. In both cases, you would receive the remaining \$4.

After the experimenter has received your responses, you will be provided with a receipt documenting your choice. This receipt is for our records. Once you have returned your receipt, you will receive payment and the task will be over.

Review

Please be sure to review these instructions, as well as the document provided. You will receive your \$5 participation fee in addition to the payments for the decision task. Once the experimenter has answered any questions you may have, we will begin.

3.5.2 Task Sheet

Charity Description

[THE FOLLOWING PARAGRAPH ONLY INCLUDED IN OUT-GROUP SESSIONS]

The United Way

The United Way's focus is to identify and resolve pressing community issues, and to make measurable changes in communities across America through partnerships with schools, government agencies, businesses, and other organizations. The agency strives to provide members of underserved communities with amble education, a livable income, and basic healthcare needs.

[THE FOLLOWING PARAGRAPH ONLY INCLUDED IN IN-GROUP SESSIONS]

The Michael D. Young Endowed Fund

The Michael D. Young Endowed Fund for Scholarship, Leadership, and Citizenshipăwill promote these three pillars of excellence advocated by Dr. Young during his 25-year career as the Vice Chancellor for Student Affairs at UC Santa Barbara. Funding priorities will be for Student Affairs services and programs that support low-income, underserved, and first-generation college students and initiatives that promote student mental health and wellness.

Payment Choice

Please indicate how much you would like to donate to the selected charity by circling one of the following options:

\$0	\$4	\$8	\$12	\$16

Survey Questionnaire

1. Please indicate your age ____________________

2. Please indicate your gender Female Male

3. Please indicate your major at UCSB ____________________

4. Please list any groups or clubs that you belong to ____________________

5. How regularly do you participate in these groups or clubs here at UCSB?

Regularly Occasionally Rarely

6. Have you ever given donations to charity in the past? Yes No

7. Have you ever volunteered your time in the past? Yes No

8. Do you currently hold a part-time or full-time job?

Part-time Full-time Neither

Table 1A: Average Total Donation by Treatment (Including Matched Funds)

	Avg. Contribution ($)	Number of Observations
No-Matching Treatments		
In-Group	3.33	72
Out-Group	4.111	72
Total No-Matching Treatments	3.722	144
Matching Treatments		
Match All	10.667	48
Individual Competition	10.200	40
Team Competition	12.667	42
Total Matching Treatments	10.783	130

Figure 1A: Donation Levels for In-group and Out-group Charities

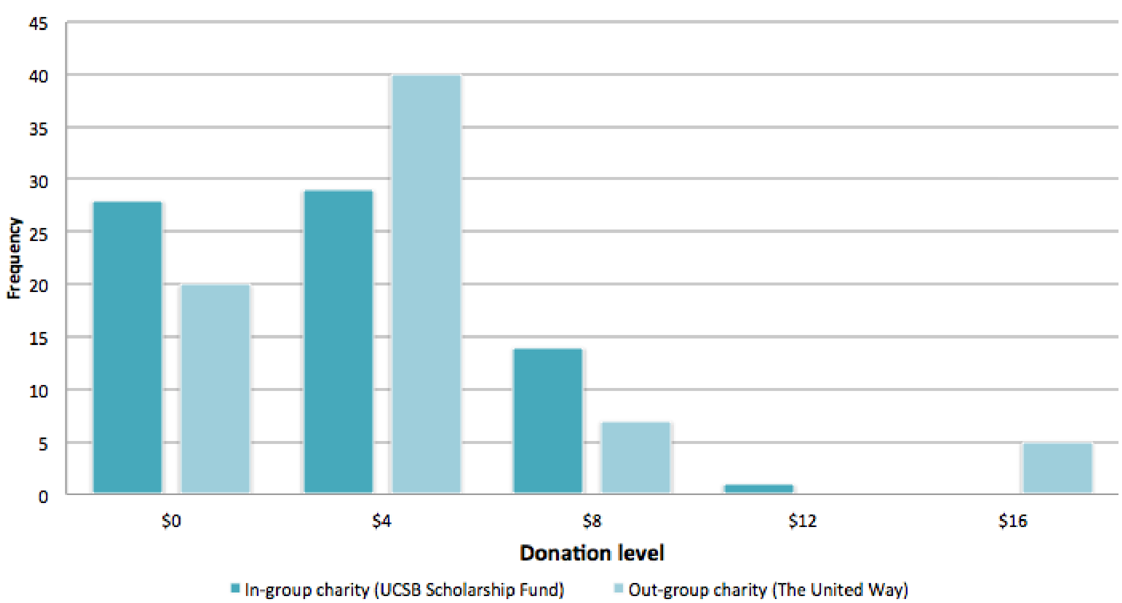


Table 2A: Parameter Estimates in Multinomial Logit Model (In-group Charity Only)

Parameter	Estimate	$\alpha^d + \beta^d$	Wald test p-Value
α^4	0.035 (0.265)	-	-
α^8	-0.693 (0.327)**	-	-
α^{12}	-3.331 (1.017)***	-	-
α^{16}	-	-	-
β_m^4	0.159 (0.448)	0.194	0.591
β_m^8	0.133 (0.551)	-0.560	0.207
β_m^{12}	2.302 (1.143)**	-1.030	0.049**
β_m^{16}	-	-	-
β_c^4	-0.202 (0.488)	-0.167	0.683
β_c^8	0.325 (0.543)	-0.368	0.396
β_c^{12}	2.376 (1.145)**	-0.955	0.070*
β_c^{16}	-	-	-
β_t^4	-0.122 (0.494)	-0.087	0.835
β_t^8	-0.694 (0.724)	1.387	0.032**
β_t^{12}	1.539 (1.272)	-1.792	0.019**
β_t^{16}	-	-	-

$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$. p-values in last column are for corresponding Wald tests. Observations by treatment: $N_{\text{no matching}} = 72$; $N_m = 48$; $N_c = 40$; $N_t = 42$

Bibliography

Ahlbrecht, Martin, and Martin Weber. "The resolution of uncertainty: An experimental study." *Journal of Institutional and Theoretical Economics* (JITE)/Zeitschrift für die gesamte Staatswissenschaft (1996): 593-607.

Ai, Wei, et al. "Recommending teams promotes prosocial lending in online microfinance." *Proceedings of the National Academy of Sciences* 113.52 (2016): 14944-14948.

Akerlof, George A., and Rachel E. Kranton. "Economics and identity." *The Quarterly Journal of Economics* 115.3 (2000): 715-753.

Babcock, Philip, et al. "Letting down the team? Social effects of team incentives." *Journal of the European Economic Association* 13.5 (2015): 841-870.

Battigalli, Pierpaolo, and Martin Dufwenberg. "Guilt in games." *The American Economic Review* 97.2 (2007): 170-176.

Becker, Gary S., and Julio Jorge Elias. "Introducing incentives in the market for live and cadaveric organ donations." *The Journal of Economic Perspectives* 21.3 (2007): 3-24.

Bellemare, Charles, et al. "Myopic loss aversion: Information feedback vs. investment flexibility." *Economics Letters* 87.3 (2005): 319-324.

Benartzi, Shlomo, and Richard H. Thaler. "Myopic loss aversion and the equity premium puzzle." *The Quarterly Journal of Economics* 110.1 (1995): 73-92.

Benjamin, D. J., J. J. Choi, and A. J. Strickland. "Social Identity and Preferences." *The American economic review* 100.4 (2010): 1913-1928.

Billig, Michael, and Henri Tajfel. "Social categorization and similarity in intergroup behaviour." *European Journal of Social Psychology* 3.1 (1973): 27-52.

Bolton, Gary E., and Axel Ockenfels. "ERC: A theory of equity, reciprocity, and competition." *The American Economic Review* (2000): 166-193.

Bornstein, Gary, and Meyrav Ben-Yossef. “Cooperation in intergroup and single-group social dilemmas." *Journal of Experimental Social Psychology* 30.1 (1994): 52-67.

Bornstein, Gary, Uri Gneezy, and Rosmarie Nagel. “The effect of intergroup competition on group coordination: An experimental study." *Games and Economic Behavior* 41.1 (2002): 1-25.

Bowman, David, Deborah Minehart, and Matthew Rabin. “Loss aversion in a consumption-savings model.” *Journal of Economic Behavior & Organization* 38.2 (1999): 155-178.

Brown, Alexander L., Jonathan Meer, and J. Forrest Williams. “Social distance and quality ratings in charity choice." *Journal of Behavioral and Experimental Economics* 66 (2017): 9-15.

Budish, Eric. “The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes." *Journal of Political Economy* 119.6 (2011): 1061-1103.

Charness, Gary, and Martin Dufwenberg. “Promises and partnership." *Econometrica: The Journal of the Econometric Society* 74.6 (2006): 1579-1601.

Charness, Gary, and Matthew Rabin. “Understanding social preferences with simple tests." *The Quarterly Journal of Economics* 117.3 (2002): 817-869.

Charness, Gary, Luca Rigotti, and Aldo Rustichini. “Individual behavior and group membership." *The American Economic Review* 97.4 (2007): 1340-1352.

Chen, Yan, and Sherry Xin Li. “Group identity and social preferences." *The American Economic Review* 99.1 (2009): 431-457.

Eames, KC, P. Holder, and E. Zambrano. “Kidney Co-operative: Proof of Concept. Online Appendix," http://ssrn.com/abstract=2390352, 2016.

Eckel, Catherine, and Philip Grossman. “Altruism in Anonymous Dictator Games." *Games and Economic Behavior* 16.2 (1996): 181-191.

Eckel, Catherine C., and Philip J. Grossman. “Rebate versus matching: does how we subsidize charitable contributions matter?" *Journal of Public Economics* 87.3 (2003): 681-701.

Eckel, Catherine C., and Philip J. Grossman. “Giving to secular causes by the religious and nonreligious: An experimental test of the responsiveness of giving to subsidies." *Nonprofit and Voluntary Sector Quarterly* 33.2 (2004): 271-289.

Eckel, C. C., and P. J. Grossman. “Do donors care about subsidy type? An experimental study" in D. Davis and M. Isaac eds., *Research in Experimental Economics*, vol. 11: Experiments Investigating Fundraising and Charitable Contributions." (2006).

Eckel, Catherine C., and Philip J. Grossman. “Subsidizing charitable giving with rebates or matching: Further laboratory evidence." *Southern Economic Journal* (2006): 794-807.

Eckel, Catherine C., and Philip J. Grossman. “Subsidizing charitable contributions: a natural field experiment comparing matching and rebate subsidies." *Experimental Economics* 11.3 (2008): 234-252.

Ely, Jeffrey, Alexander Frankel, and Emir Kamenica. “Suspense and surprise.” *Journal of Political Economy* 123.1 (2015): 215-260.

Epstein, Larry G., and Stanley E. Zin. “Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis.” *Journal of Political Economy* (1991): 263-286.

Erev, Ido, Gary Bornstein, and Rachely Galili. “Constructive intergroup competition as a solution to the free rider problem: A field experiment." *Journal of Experimental Social Psychology* 29.6 (1993): 463-478.

Falk, Armin, and James J. Heckman. “Lab experiments are a major source of knowledge in the social sciences." *Science* 326.5952 (2009): 535-538.

Falk, Armin, and Florian Zimmermann. “Beliefs and utility: Experimental evidence on preferences for information.” Working paper, 2014.

Fehr, Ernst, and Klaus M. Schmidt. “A theory of fairness, competition, and cooperation." *The Quarterly Journal of Economics* 114.3 (1999): 817-868.

Fershtman, Chaim, Uri Gneezy, and John A. List. “Equity aversion: Social norms and the desire to be ahead." *American Economic Journal: Microeconomics* 4.4 (2012): 131-144.

Fischbacher, Urs. “z-Tree: Zurich toolbox for ready-made economic experiments.” *Experimental Economics* 10.2 (2007): 171-178.

Gee, Laura K., and Michael Schreck. “Do Beliefs About Peers Matter for Donation Matching? Experiments in the Field and Laboratory." Unpublished manuscript. (2016).

Gollier, Christian, and Richard J. Zeckhauser. “Horizon length and portfolio risk.” *Journal of Risk and Uncertainty* 24.3 (2002): 195-212.

Gneezy, Uri, and Jan Potters. “An experiment on risk taking and evaluation periods.” *The Quarterly Journal of Economics* (1997): 631-645.

Greiner, Ben. “The online recruitment system ORSEE 2.0 - A guide for the organization of experiments in economics.” *University of Cologne, Working paper series in economics* 10.23 (2004): 63-104.

Grosskopf, Brit. "Reinforcement and directional learning in the ultimatum game with responder competition." *Experimental Economics* 6.2 (2003): 141-158.

Huck, Steffen, and Imran Rasul. "Matched fundraising: Evidence from a natural field experiment." *Journal of Public Economics* 95.5 (2011): 351-362.

Ibrahim, H., et al. "Long-Term Consequences of Kidney Donation," *New England Journal of Medicine* 360, 459-469, January 2009.

Kahneman, Daniel, and Amos Tversky. "Prospect theory: An analysis of decision under risk." *Econometrica: The Journal of the Econometric Society* (1979): 263-291.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. "Experimental tests of the endowment effect and the Coase theorem." *Journal of Political Economy* (1990): 1325-1348.

Karlan, Dean, and John A. List. "Does price matter in charitable giving? Evidence from a large-scale natural field experiment." *The American Economic Rview* 97.5 (2007): 1774-1793.

Kocher, Martin G., Michal Krawczyk, and Frans van Winden. "'Let me dream on!' Anticipatory emotions and preference for timing in lotteries." *Journal of Economic Behavior & Organization* 98 (2014): 29-40.

Kőszegi, Botond, and Matthew Rabin. "Reference-dependent consumption plans." *The American Economic Review* 99.3 (2009): 909-936.

Kreps, David M., and Evan L. Porteus. "Temporal resolution of uncertainty and dynamic choice theory." *Econometrica: The Journal of the Econometric Society* (1978): 185-200.

Krieg, Justin, and Anya Samek. "When charities compete: A laboratory experiment with simultaneous public goods." *Journal of Behavioral and Experimental Economics* 66 (2017): 40-57.

Leider, S. and A. Roth. "Kidneys for Sale: Who Disapproves and Why?" *American Journal of Transplantation* 10: 1221-1227, 2010.

Lovallo, Dan, and Daniel Kahneman. "Living with uncertainty: Attractiveness and resolution timing." *Journal of Behavioral Decision Making* 13.2 (2000): 179.

Martin, Richard, and John Randal. "How is donation behaviour affected by the donations of others?" *Journal of Economic Behavior & Organization* 67.1 (2008): 228-238.

Mehra, Rajnish, and Edward C. Prescott. "The equity premium: A puzzle." *Journal of Monetary Economics* 15.2 (1985): 145-161.

Meier-Kriesche, HU. and B. Kaplan. "Waiting time on dialysis as the strongest modifiable risk factor for renal transplant outcomes: A Paired Donor Kidney Analysis." *Transplantation* 74: 1377-1381, 2002.

National Kidney Foundation, http://www.kidney.org/news/newsroom/factsheets/Organ-Donation-and-Transplantation-Stats.cfm, retrieved on August 3rd, 2016.

National Resident Matching Program, http://www.nrmp.org/, retrieved on July 18th, 2016.

Opelz, G., "Impact of HLA compatibility on survival of kidney transplants from unrelated live donors." *Transplantation* 64: 1473-1475, 1997.

Organ Procurement Transportation Network, http://optn.transplant.hrsa.gov/latestData/rptData.asp, retrieved on August 3rd, 2016.

Putnam, Robert D. *Bowling alone: The collapse and revival of American community. Simon and Schuster*, 2001.

Rapoport, Amnon, and Gary Bornstein. "Intergroup competition for the provision of binary public goods." *Psychological Review* 94.3 (1987): 291.

Read, Daniel, George Loewenstein, and Matthew Rabin. "Choice bracketing." *Elicitation of Preferences* Springer Netherlands, 1999. 171-202.

Rondeau, Daniel, and John A. List. "Matching and challenge gifts to charity: evidence from laboratory and natural field experiments." *Experimental Economics* 11.3 (2008): 253-267.

Roth, Alvin E., Tayfun Sonmez, and M. Utku Unver. "Kidney Exchange." *Quarterly Journal of Economics* 119 (2), 457-488, May 2004.

Roth, Alvin E., Tayfun Sonmez , and M. Utku Unver. "Pairwise Kidney Exchange." *Journal of Economic Theory* 125 (2), 151-188, December 2005a.

Roth, Alvin E., et al. "Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An experimental study." *The American Economic Review* (1991): 1068-1095.

Roth, Alvin E., Tayfun Sonmez , and M. Utku Unver. "A Kidney Exchange Clearinghouse in New England" *American Economic Review Papers and Proceedings* 95 (2), 376-380, May 2005b.

Roth, Alvin E., Tayfun Sonmez , and M. Utku Unver "Efficient Kidney Exchange: Coincidence of Wants in Markets with Compatibility-Based Preferences." *The American Economic Review* 97 (3), 828-851, June 2007.

Stegall, M.D., et al. "A Comparison of Plasmapheresis Versus High-Dose IVIG Desensitization in Renal Allograft Recipients with High Levels of Donor Specific Alloantibody." *American Journal of Transplantation* 6, 346351, 2006.

Thaler, Richard. "Mental accounting and consumer choice." *Marketing Science* 4.3 (1985): 199-214.

Thaler, Richard H., et al. "The effect of myopia and loss aversion on risk taking: An experimental test." *The Quarterly Journal of Economics* (1997): 647-661.

Toussaert, Séverine. "Eliciting temptation and self-control through menu choices: a lab experiment" Working Paper, 2016.

Train, Kenneth. *Discrete choice methods with simulation*. Cambridge university press, 2003.

Tversky, Amos, and Daniel Kahneman. "Advances in prospect theory: Cumulative representation of uncertainty." *Journal of Risk and Uncertainty* 5.4 (1992): 297-323.

Tversky, Amos, and Daniel Kahneman. "The framing of decisions and the psychology of choice." *Environmental Impact Assessment, Technology Assessment, and Risk Analysis* Springer Berlin Heidelberg, 1985. 107-129.

Zimmermann, Florian. "Clumped or piecewise? Evidence on preferences for information." *Management Science* 61.4 (2014): 740-753.