Spring 5-15-2010

Essays in Experimental and Environmental Economics

Sarah Jacobson

Georgia State University

Recommended Citation

Jacobson, Sarah, "Essays in Experimental and Environmental Economics." Dissertation, Georgia State University, 2010.
https://scholarworks.gsu.edu/econ_diss/66

PERMISSION TO BORROW

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type.

I agree that permission to quote from, to copy from, or to publish this dissertation may be granted by the author or, in his or her absence, the professor under whose direction it was written or, in his or her absence, by the Dean of the Andrew Young School of Policy Studies. Such quoting, copying, or publishing must be solely for scholarly purposes and must not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential gain will not be allowed without written permission of the author.

_____
Signature of the Author

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University Library must be used only in accordance with the stipulations prescribed by the author in the preceding statement.

The author of this dissertation is:

Sarah A. Jacobson
5641 Rosser Court SW
Lilburn, GA 30047

The director of this dissertation is:

Ragan A. Petrie
Department of Economics and
Interdisciplinary Center for Economic Science (ICES)
George Mason University
4400 University Drive, MSN 1B2
Fairfax, VA 22030

Users of this dissertation not regularly enrolled as students at Georgia State University are required to attest acceptance of the preceding stipulations by signing below. Libraries borrowing this dissertation for the use of their patrons are required to see that each user records here the information requested.

			Type of use
Name of User	Address	Date	(Examination only or copying)

ESSAYS IN EXPERIMENTAL
AND ENVIRONMENTAL ECONOMICS

BY

SARAH ANDREA JACOBSON

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree
of
Doctor of Philosophy
in the
Andrew Young School of Policy Studies
of
Georgia State University

GEORGIA STATE UNIVERSITY
2010

ACCEPTANCE

This dissertation was prepared under the direction of the candidate's Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Andrew Young School of Policy Studies of Georgia State University.

Dissertation Chair: Ragan A. Petrie

Committee: H. Spencer Banzhaf
                        James C. Cox
                        Vjollca Sadiraj

Electronic Version Approved:

Robert E. Moore, Dean
Andrew Young School of Policy Studies
Georgia State University
May, 2010

ACKNOWLEDGEMENTS

CONTENTS

TABLES

FIGURES

ABSTRACT

ESSAYS IN EXPERIMENTAL AND ENVIRONMENTAL ECONOMICS

By

Sarah Andrea Jacobson

May, 2010

Committee Chair: Dr. Ragan Petrie

Major Department: Economics

The chapters of this dissertation explore complementary areas of applied

microeconomics, within the fields of experimental and environmental economics. In

each case, preferences and institutions interact in ways that enhance or subvert efficiency.

The first chapter, "The Girl Scout Cookie Phenomenon," uses a laboratory

experiment to study favor trading in a public goods setting. The ability to practice

targeted reciprocity increases contributions by 14%, which corresponds directly to

increased efficiency. Subjects discriminate by rewarding group members who have been

generous and withholding rewards from ungenerous group members. At least some

reciprocal behavior is rooted in other-regarding preferences. When someone is outside

the "circle of reciprocity," he gives less to the public good than in other settings. We find

no evidence of indirect reciprocity. We find two behavioral types in each treatment,

differing in baseline giving but not in tendency to reciprocate.

The second chapter, "The Effects of Conservation Reserve Program Participation

on Later Land Use," studies another public goods issue: conservation. The Conservation

Reserve Program (CRP) pays farmers to retire farmland. We use a treatment effect

framework to find that ex-CRP land is 21-28% more likely to be farmed than comparable

non-CRP land. This implies that the CRP improves low-quality land, making it more attractive to farm. This could demonstrate inefficiency, since farmers gain private benefit from a program meant to provide a public good. On the other hand, farmed ex-CRP land is more likely to adopt conservation practices, although this may not be caused by CRP participation.

The third chapter, "Learning from Mistakes," examines financial decisions by adult Rwandans in institutions inside and outside the lab. Over 50% of subjects make irrational choices over risk—choices that likely do not reflect their preferences, and are therefore likely inefficient—and these subjects share tendencies in their take-up of financial instruments. Risk-averse individuals are more likely to belong to a savings group and less likely to take out an informal loan. For those who make mistakes, however, as they become more risk averse, they are less likely to belong to a savings group and more likely to take up informal credit.

Chapter I: The Girl Scout Cookie Phenomenon
Peer Pressure in Grassroots Fundraising

Introduction

People seem to have a natural tendency to engage in reciprocal behavior. Their motives, however, are often ambiguous: the same actions can be explained by other-regarding preferences or by strategic motives. The distinction is important, because in many settings selfish motives alone cannot sustain cooperative behavior. The existence of other-regarding preferences has been documented in isolated laboratory experiments, but self-interest may dominate in richer field settings. Can institutions leverage other-regarding preferences while still harnessing the strategic motives of selfish people to promote pro-social behavior? One institution that may serve this purpose is grassroots fundraising. We report the results from a laboratory experiment that mimics elements of this institution to study the power of favor trading to support a public good. The design of our treatments "subtracts away" the possibility of each of a number of behavioral motives so that we can cleanly isolate the effects of direct and indirect reciprocity. We find evidence of direct reciprocity (rooted in other-regarding preferences), but a notable lack of indirect reciprocity.

Our methods and results are best understood through analogy to grassroots fundraising. Many nonprofit organizations fundraise by enlisting citizen boosters to solicit donations from their social networks. Such a fundraiser makes direct, personal solicitations to people with whom he has an ongoing relationship. Donations are revealed to the fundraiser and often to everyone else who is solicited. A classic example is the

Girl Scouts of the USA, an organization that works with American girls. The Girl Scouts' most prominent fundraiser is the annual drive in which member girls and their families sell cookies to friends, family, and neighbors. Klein (1990) argues that social favor trading played a role in the funding of turnpikes in early America. The same dynamic may occur when it is not money but time and effort being solicited, as may happen in a volunteer advisory board or a school's parent-teacher association.

This fundraising technique is so common that social networks are sometimes full of people with "pet causes" for which they are fundraising. Imagine that Joe and Frank work in an office together. Joe's son is in Boy Scouts, and Frank's daughter is in Girl Scouts. Each child's organization has an annual fundraiser: Joe's Boy Scout sells popcorn in January, and Frank's Girl Scout sells cookies in February. Each has a special interest in his child's fundraiser, because of the family's stake in the organization's success. As a result, Joe will buy popcorn and Frank will buy cookies, but in addition, Joe wants Frank to buy popcorn, and Frank wants Joe to buy cookies. Even if Frank is not interested in the Boy Scouts (or the popcorn), Frank may buy Joe's popcorn in January, hoping that Joe will reciprocate when Frank's cookie fundraiser comes around the next month. These peers use the promise of future reciprocation—in a favor trading exchange like many that arise in social situations—to pressure each other to perform the pro-social act of giving to charity.

We study whether and to what extent favor trading can increase pro-social behavior, and we look at the motives behind this behavior. Our research questions are as follows. Does the opportunity to practice targeted reciprocity increase contributions to a

public good? Is giving in this setting driven by direct reciprocity, indirect reciprocity, or only strategic self-interest? Is there important heterogeneity in behavior?

These questions, and the efficacy of the favor trading public goods institution in general, are difficult to study in non-experimental settings because selective use of fundraising techniques can cause endogeneity problems in the analysis. For example, Long (1976) found that the more "personal" a donor solicitation, the more contributions were solicited, and inferred that the reason was that a more personal solicitation allowed the fundraiser to exert peer pressure on the donor. However, this analysis is biased in favor of finding a relationship: personal solicitations are more costly to perform, so charities may limit their most personal solicitations to a "hot list" of likely donors.

This study fits into the literature examining social influences on cooperative behavior. We focus on the roles of immediate self-interest, altruism, strategic self-interest, direct reciprocity, and indirect reciprocity. Conditional cooperation, a general form of reciprocal behavior, may influence giving in public goods games as surveyed in Gächter (2007); this will be a background force in all of our treatments. While immediate self-interest and altruism are well-known concepts, the other motives require explanation. See Sobel (2005) for a review of the literature on interdependent preferences and reciprocity, including a helpful comparison of the terminology from various models. We use the phrase "strategic self-interest" to describe Sobel's "instrumental reciprocity," in which apparently reciprocal actions are undertaken to win future rewards (i.e. purely for selfish reasons). One characteristic of strategic self-interest is that, since strategic motivations depend on future rewards, the removal of those future rewards will destroy strategically-motivated cooperation. As a result, if actors are

strategically-motivated, finitely-repeated games that depend on cooperation tend to unravel from the end period. Strategic cooperation could look like unconditional cooperation, but it could also be conditional: a wise strategic person should "bribe" only people who have shown a tendency to reciprocate.

In contrast to strategic cooperation, we use the terms "direct reciprocity" and "indirect reciprocity" to describe conditional urges rooted in other-regarding preferences (Sobel's "intrinsic reciprocity"). Important models of this kind of reciprocity include Rabin (1993) and Cox, Friedman, and Sadiraj (2008). Most models of this type allow Joe's preferences over Frank's payoffs to depend on Frank's past actions. Because these kinds of reciprocity are rooted in other-regarding preferences, they can sustain cooperation even in cases where strategic cooperation would unravel. Biologists (e.g., Trivers 1971) have recognized the potential for reciprocity in general to promote pro-social behavior and to help pass on positive traits.

In the case of direct reciprocity, the reciprocator responds to acts that directly affected him. For example, if Joe is a direct reciprocator, he may reward Frank for Frank's past behavior toward Joe. Direct reciprocity is difficult to distinguish from strategic cooperation except that it implies continued reciprocation during end periods. Other researchers have examined direct reciprocity using one-shot games, often with some form of an investment game (e.g., Cox 2004). In our setting, we find direct reciprocity in the final periods of a repeated game.

Indirect reciprocity occurs when a disinterested third party rewards good behavior. For example, imagine that coworker Rita has no interest in Girl or Boy Scouts

and has no pet cause of her own. If Rita rewards Frank for patronizing Joe's charity, she's exhibiting indirect reciprocity. Some (e.g., Nowak and Sigmund 2005) argue that indirect reciprocity could be a powerful tool for boosting cooperation in a large, diffuse group.

We test for indirect reciprocity by exogenously assigning one subject to a position from which he can differentially reward others but he himself can never be rewarded. In this way, he benefits very little from establishing a norm of cooperation in his group. This is a strict test of indirect reciprocity. Other studies (for example, Engelmann and Fischbacher 2009; Seinen and Schram 2006) have found indirect reciprocity by having subjects play one-shot cooperation games in pairs, randomly re-matching them, and reporting to each subject's new partner a statistic (e.g., an "image score") that reflects this subject's past generosity. This definition of indirect reciprocity is less strict than our definition, in that it allows indirect reciprocators to benefit from a norm of cooperation in the group.

While laboratory experiments provide clean, controlled tests, critics have argued that other-regarding preferences inferred from such tests are artifacts of laboratory methods and have no relevance for behavior outside the lab. Dictator games and investment games, often used to study altruism and reciprocity, have received particular criticism. Some of this criticism focuses on the results of field experiments that imply that strategic (reputation) concerns are more important than other-regarding preferences. For example, List (2006) found that "gift exchange" behavior by sports card traders is only seen when there is an incentive to maintain a reputation, and Soetevent (2005) found that open church collection baskets (allowing churchgoers to see each others' donations)

increase donations. As discussed in Falk and Heckman (2009), incentivized behavior observed in the lab is just as "real" as behavior observed in field settings, and context-specificity is just as much of a problem for field results as it is for lab results. We take the view that apparent conflicts between lab and field findings simply show that social forces must be studied in a variety of controlled settings to learn in what ways they are sensitive to institutional features.

We contribute to the literature in a number of ways. First, our "Stakeholder" public goods game integrates a vehicle for targeted reciprocity into the public goods setting, to emulate the favor trading that occurs in social networks.[1] The design allows us to isolate reciprocal behavior and test for other-regarding preferences. Second, we include a "Bachelor" design as a clean test for indirect reciprocity. Third, using a within-subject design, we observe shifts in behavior across institutions and use an endogenous type classification method to study heterogeneity and sensitivity to institutions.

We find that average contributions increase by 14% when targeted reciprocal acts are possible. We confirm that reciprocal behavior is the cause for this increase, and that this is rooted (at least in part) in other-regarding preferences. Using our strict, clean test, we do not find evidence of indirect reciprocity. We find meaningful heterogeneity in giving behavior, and we find some stability in behavior across treatments.

The paper proceeds as follows. The next section describes the three experimental treatments. In the following section, we present results. In the final section, we conclude.

---

[1] The citizen booster as public good stakeholder is one lens through which to view asymmetric returns in a public goods game; others (e.g., Brandts, Cooper, and Fatas 2007) have examined the role of such a person as a potential leader.

Experimental Design

The experiment is a linear public goods game, with publicly revealed contributions and asymmetric payoffs. Subjects are assigned to five-person groups for a number of rounds. In each round, each person has an endowment of z tokens to allocate between a private investment with return a to himself and a public investment with some return to all group members. In each round, one group member is the Stakeholder: he has a bigger stake in the public good, in that he gets a higher return (relative to other members) from tokens invested there. The Stakeholder position rotates through group members. The public good return is b for non-Stakeholders and c for Stakeholders. Person i's contribution to the public good in round t is g_{it}, and $Stake_t$ is the index of the person who is Stakeholder in round t. Payoffs are:

$$\pi_{it} = \begin{cases} c\sum_j g_{jt} + a(z - g_{it}) & \text{if } Stake_t = i \\ b\sum_j g_{jt} + a(z - g_{it}) & \text{if } Stake_t \neq i \end{cases}$$

The parameters are such that $b < a < c$ and $a < 4b + c$, so the social optimum is achieved if everyone contributes fully. Since $(c - a)$ is positive, the Stakeholder maximizes profit by contributing as much as possible, so even a selfish Stakeholder always contributes to the public good. Since $(b - a)$ is negative, non-Stakeholders face a dilemma: they maximize profit by keeping all of their tokens, but this free riding is anti-social. Non-Stakeholder contributions will be the focus of our analysis.

We use $u_i(m, y_1, y_2, y_3, y_4)$ to represent utility for subject i in any given round from his earnings in this round m and the earnings of the other four members of his group y_1,

y_2, y_3, y_4 in this round. Designate the current Stakeholder as person k (and since non-Stakeholder contributions are of greatest interest, $k \neq i$).

We borrow a nonparametric model of reciprocity from Cox, Friedman, and Sadiraj (2008) to describe reciprocal preferences. Specifically, we note that i's marginal rate of substitution between his own payoff and that of another subject y_j may depend on j's past generosity toward i. If subject j was generous in the past, then i may have a lower marginal rate of substitution (a higher willingness to pay) and thus be willing to sacrifice his own payoff to increase j's. That is, as a non-Stakeholder he may sacrifice high earnings from the private investment so that he can (through a public good contribution) increase the earnings of a subject j who has previously increased i's earnings.

In any given round, a non-Stakeholder's (i's) contributions increase the payoff of the other three non-Stakeholders ($j \neq i,k$) and also the Stakeholder (k). Thus i's public good contributions may be intended to increase the payoffs of any subject (y_1, y_2, y_3, y_4). However, for each token contributed, the Stakeholder benefits by c and the non-Stakeholders only benefit by b (recall $b < a < c$). Similarly, his contributions may be a reaction to past contributions to the public good from any round. However, his earnings are increased most by contributions made when he was Stakeholder. Therefore, he is most likely to respond to generosity that occurred when he was Stakeholder.

A selfish person places no weight on others' payoffs y_1, y_2, y_3, y_4, regardless of the other subjects' previous generosity. An unconditional altruist has a constant willingness to pay to increase others' payoffs, and has the same willingness to pay for all four group members. A reciprocator has a willingness to pay for each other subjects' payoffs that

depends on that other subject's past generosity. (A person who is "practically selfish" for the purposes of this experiment is one whose willingness to pay to increase others' payoffs never increases above the opportunity cost of a public good contribution. A "practically selfish" person may be a mild altruist, a mild reciprocator, or a reciprocator who encounters only low contributors.)

Beliefs are important in determining the actions of strategic actors. If a subject is selfish and believes that everyone else is either selfish or an unconditional altruist, he will not contribute. If he is selfish but strategic, however, he may believe that he can influence his group members to have a greater willingness to pay for his benefit. That is, if he believes his group members are reciprocators, he may believe he can win future earnings (particularly in rounds when he will be Stakeholder) by contributing to the public good this period. Note that this future potential benefit becomes quite small once a subject has no future Stakeholder stints. Thus, a strategic but selfish person may cooperate in early rounds but cease to do so after he passes his last Stakeholder stint. His early-round cooperation may target people he expects to be reciprocators (no benefit can be gained from sacrificing for an unconditional altruist or a selfish person), and he may learn these types by observing other subjects' past behavior. As a result, this early-round strategic cooperation may depend on history in much the same way that reciprocal altruism may.[2]

We report three treatments: Private, Public, and Ineligible, all described in detail below. All three use endowment $z = 20$ tokens, private good return $a = \$0.02$, non-

---

[2] History could also affect current behavior through contagion: a person treated well (badly) in the past could react by behaving well (badly) not because they wish to reciprocate but because they have "caught" a good (bad) mood from their experience. This is outside of our model.

Stakeholder public good return b = $0.01, and Stakeholder public good return c = $0.03.

For non-Stakeholders, the "price of giving" is 2, while for Stakeholders it is 2/3. All

subjects participated in all three treatments, in a different group for each treatment, with

treatment order varied across sessions.

The experiment is computerized and proceeds as follows. Subjects enter the lab

and are given general instructions. (All instructions are in Appendix A.) They are told

that they will play three sets of rounds with three different groups. The first treatment

begins, with a set of specific instructions that explain the information conditions and the

number of rounds for this treatment. The subjects play through all of the rounds for that

treatment. When the first treatment is over, the groups are reshuffled. The second and

third treatments proceed in much the same way, each with treatment-specific instructions.

After all three treatments are complete, a questionnaire is administered and the subjects

are paid anonymously. Subjects' total earnings are the sum of their earnings in each

treatment, which in turn are the sum of earnings in each round.

In the software for all three treatments, subjects see first a decision screen and

then a review screen for each round. In both the decision and review screens, the central

feature is the contribution table. This table contains a row for each round in the

treatment. Columns contain information on the subject's contribution and the

contributions of others in his group, the group's total contributions, and the subject's own

earnings. Information is filled into this table after the decision stage of each round, and

remains visible for the rest of the treatment.

The Public treatment, which lasts ten rounds, follows the basic favor-trading public goods design outlined above. The Stakeholder position rotates through all five group members so everyone is Stakeholder twice. Contributions are publicly revealed and tracked individually, and Stakeholder assignments are common knowledge. Figure 1 shows the decision screen of the Public treatment. Each group member is randomly assigned a letter code (A, B, C, D, or E) and keeps the same letter code for all ten rounds. The contribution table shows in which rounds each subject will be the Stakeholder. Since contribution history is public and everyone knows when each group member will be Stakeholder, subjects can reward each other for past generosity. For example, if Joe is subject A and Frank is subject B, Joe can see how much Frank contributed in Round 1 when Joe was the Stakeholder (when he raised funds for the Boy Scouts). Joe can reward Frank with a large contribution when Frank is Stakeholder in Round 2 (when Frank fundraises for the Girl Scouts), or Joe may withhold that reward if he deems Frank's donation stingy.

You are now: **MAKING YOUR CONTRIBUTION FOR ROUND 1**
Your Letter Code: C   Stakeholder: A

CONTRIBUTIONS TO THE GROUP FUND

	Stakeholder		YOU			TOTAL TOKENS IN GROUP FUND	MY EARNINGS
	A	B	C	D	E		
Round 1	*****						
Round 2		*****					
Round 3			*****				
Round 4				*****			
Round 5					*****		
Round 6	*****						
Round 7		*****					
Round 8			*****				
Round 9				*****			
Round 10					*****		

***** indicates that you were/will be Stakeholder in the round indicated

DECISION PANEL

How much would you like to put in the **GROUP FUND**? (0-20) [　　]

Your **PERSONAL FUND** contribution will be **20 minus your GROUP FUND contribution**.

RETURNS:

Personal fund: **$0.02 per token** to you

Group fund: **$0.03 per token** to Stakeholder
             $0.01 per token to non-Stakeholders (including YOU)

CLICK TO SUBMIT

SUBMIT

Figure 1. Public Treatment Decision Screen

The Private treatment also lasts for ten rounds. The Stakeholder position still rotates through all group members so that everyone is Stakeholder twice. However, the information environment differs from the Public treatment. Each subject still sees the disaggregated (individual) contributions of his group members, but subjects are not assigned letter codes. It is no longer possible to track subjects' reputations. Figure 2 shows the review screen of the Private treatment. In each round's row, the contribution table reports the contributions of all group members in a randomly-ordered list, re-shuffled for each round. Further, even if a subject believes he can identify a group member as being particularly worthy (or unworthy) of reward, he still does not know

when that person will be Stakeholder. He only knows when he himself will be the

Stakeholder, so he can't target reciprocal acts toward any other particular subject. The

Private treatment is a good baseline against which to compare Public treatment behavior

because, as Sell and Wilson (1991) show, reporting disaggregated instead of mean or

total contributions in a public goods game may increase contributions.

Figure 2. Private Treatment Review Screen

Finally, the Ineligible treatment is very much like the Public treatment, with

randomly-assigned letter codes (not linked to the Public treatment letter codes), public

reputations, and public Stakeholder timing. However, one subject in each five-member

group (the "Bachelor") is ineligible to be the Stakeholder. To return to our guiding

analogy, the "Bachelor" is Rita, Joe and Frank's officemate who does not have a pet

cause for which to fundraise. Because only four subjects are eligible to be Stakeholder,

the Ineligible treatment lasts eight rounds, so each eligible subject is still Stakeholder

twice. The Bachelor is randomly chosen and remains the Bachelor for the entire

treatment. The computer screens for the Ineligible treatment (Figure 3) are like those for

the Public treatment, except that the Bachelor is indicated in the screen header and in the

contribution table as the "Ineligible" person. The Stakeholder position rotation skips the

Bachelor: if person D is the Bachelor, the Stakeholder is A, then B, then C, then E, etc.

	You are now: **REVIEWING RESULTS FOR ROUND 2**						
	Your Letter Code: E Stakeholder: B Ineligible: D						
	CONTRIBUTIONS TO THE GROUP FUND						
					YOU		
		Stakeholder		Ineligible			
	A	B	C	D	E	TOTAL TOKENS IN GROUP FUND	MY EARNINGS
Round 1	20	13	8	9	4	54	$0.86
Round 2	1	18	10	5	16	50	$0.58
Round 3			*****				
Round 4					*****		
Round 5	*****						
Round 6		*****					
Round 7			*****				
Round 8					*****		

***** indicates that you were/will be Stakeholder in the round indicated

DECISION PANEL

REVIEW RESULTS FROM ROUND 2 IN THE TABLE ABOVE.

YOUR EARNINGS WERE:
Your Personal Fund contribution (4) times $0.02 per token   = $0.08
PLUS: The total number of tokens in the Group Fund (50) times $0.01 per token   = $0.50
EQUALS: TOTAL   = $0.58

CLICK WHEN DONE
DONE

Figure 3. Ineligible Treatment Review Screen

The difference between the Public and Private treatments is that direct reciprocity, indirect reciprocity, and strategic self-interest cannot motivate giving in the Private treatment, because subjects don't have the information they would need to respond to each others' actions. (Unconditional altruism and conditional cooperation—response to overall group behavior—can affect giving in both treatments; actions targeted directly at another individual are only possible in the Public treatment.) In the Public treatment, all of these forces are in play because reputations and Stakeholder identities are public. Any difference between the Public and Private treatments must be due to direct reciprocity, indirect reciprocity, and/or strategic giving.

Within the Public treatment, we will be able to see whether subjects are responsive to past generosity: that is, whether they discriminate by giving larger contributions when the current Stakeholder is someone who was previously generous. In any given round, subjects may respond to the past behavior of both the current Stakeholder and the current non-Stakeholders. However, if we detect this kind of responsiveness with regard to the current Stakeholder's past actions, this is sufficient demonstrate reciprocal giving. We can identify direct reciprocity as responsiveness of this type that does not disappear as the end period approaches.

The Ineligible treatment allows us to investigate two additional questions. First, the Bachelor herself (Rita, in our analogy) is not subject to direct reciprocity or strategic self-interest. If the Bachelor *does* give in a way that responds to the Stakeholder's past generosity, this would be evidence of indirect reciprocity. Second, the presence of a Bachelor shrinks the "neighborhood of reciprocity" from five people to four people. This leaves the Bachelor to reap benefits from other subjects' increased generosity even if she

herself does not contribute. We can observe whether this change in the group dynamic affects non-Bachelor contributions. For example, others may be discouraged by supporting a Bachelor who freeloads off the public good.

The experiments were run in the Experimental Economics Center (E*x*CEN) at Georgia State University in six separate 20-subject sessions, for a total of 120 subjects. The software was written in z-Tree (Fischbacher 2007). The protocol was double anonymous (subject decisions were anonymous even from the perspective of the experimenter). Of the 120 subjects, 75 (62.5%) were female and the average age was 21.8. The experiment lasted about 90 minutes, and subjects earned on average $24.33 (standard deviation $2.67).

Results

As described above, each subject participated in all three treatments. The three treatments were run in all six possible orders. We do not observe order effects, so we pool the data across sessions. Contribution data are shown in Figure 4. Stakeholder contributions (the dashed lines) are close to the endowment, which is expected because the Stakeholder's return from the public good is greater than his return from the private good.[3] Non-Stakeholder contributions (the solid lines) for each treatment are lower but significantly positive in all rounds. These contributions show the downward trend usually seen in public goods games. Bachelor contributions in the Ineligible treatment do not show this trend, but they are well below contributions in the other treatments.

---

[3] We can reject the hypothesis that contributions are strictly 100% (t-test *p*-value 0.000 for the Public, 0.002 for the Private, and 0.000 for the Ineligible treatment). This could be explained by subject error, since error can only be made in the negative direction, or myopic inequity aversion.
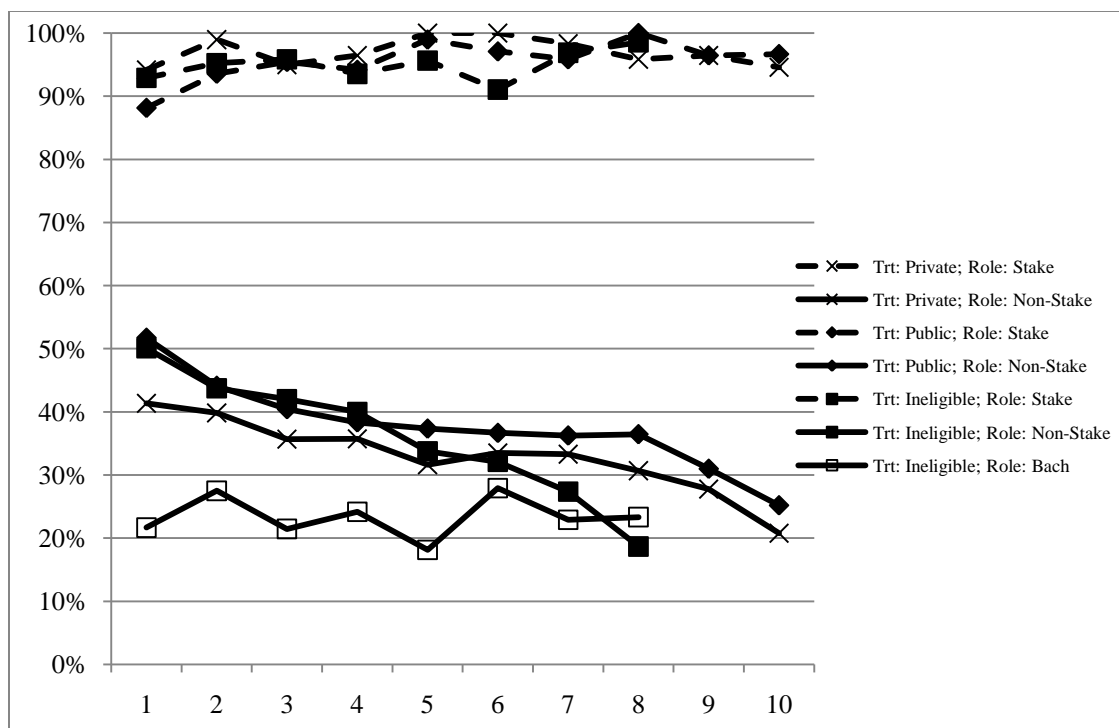
Figure 4. Contributions by Treatment and Role across Rounds (in Percent of Endowment)

Non-Stakeholder contributions in the Private treatment are comparable to contributions in other linear public goods games in the literature with similar (symmetric) "prices of giving." This was not a foregone conclusion, because the presence of a Stakeholder changes the non-Stakeholders' incentives as compared to the incentives in a symmetric public goods game. The Private treatment non-Stakeholder contributions start at 41% in round 1 and end at 21% in round 10, averaging 33% across all rounds. In the final round, 47% of non-Stakeholders make positive contributions. In Castillo and Petrie (2010), contributions decline from 41.6% to 23.6%, averaging 32.8%; Andreoni (1988) saw an overall average of 33.2%; Croson (1996) saw an overall average of 35.7%; relevant surveyed results in Cox and Sadiraj (2007) show positive contributions in the final period from 27-44% of subjects. Results from asymmetric-return public goods

games are hard to compare because of differences in the payoff structure, but subjects

with lower marginal per capita returns in somewhat similar treatments gave 20% on

average in Goeree, Holt, and Laury (2002) and 18% in Glöckner et al. (2009). Our game

has some similarities to sequential-play public goods games, and although contributions

are hard to compare because of differences in game structure, a similar sequential game

(Güth et al. 2007) finds average contributions of 47.68% of endowment.

Treatment Effects[4]

The difference between the Private and Public treatments is that the Public

treatment opens the door to targeted direct reciprocity, indirect reciprocity, and strategic

self-interest. Thus, we can answer our first research question (whether these forces can

increase the provision of a public good) simply by determining whether non-Stakeholder

contributions are higher in the Public treatment than they are in the Private treatment.

Figure 4 shows that average non-Stakeholder contributions in the Public treatment

exceed those in the Private treatment in all rounds. These differences are only statistically

significant in a few rounds, but are significant when pooled across rounds. Table 1

shows average Stakeholder and non-Stakeholder contributions for each treatment.

---

[4] Results that follow are mostly based on non-parametric tests; for the results of some selected parametric tests, including various regressions, see Appendix B. All results hold for these parametric specifications.

Table 1. Average Contributions by Treatment and Role (Percent of Endowment)

	Non-Stakeholder	Stakeholder
Private Treatment	33.02 (25.98) N=120	96.98 (10.30) N=120
Public Treatment	37.75 (25.93) N=120	95.63 (10.36) N=120
Ineligible Treatment, Non-Bachelors	35.96 (22.26) N=96	94.95 (11.74) N=96
Ineligible Treatment, Bachelors	23.39 (28.41) N=24	N/A

Standard deviations in parentheses

As shown in Table 1, the average non-Stakeholder contribution is 14.4% larger in the Public (37.75% of endowment) than in the Private treatment (33.02 %). This difference is statistically significant (Wilcoxon signed-rank test $p=0.051$). Thus, these social forces increase cooperation by a modest but economically significant amount.

Table 2. Average Contributions by Treatment and Bachelor Status (Percent of Endowment)

	Non-Bachelor	Bachelor
Private Treatment	33.80 (25.83) N=96	29.90 (26.89) N=24
Public Treatment	37.79 (24.86) N=96	37.60 (28.41) N=24
Ineligible Treatment, Non-Bachelors	35.96 (22.26) N=96	N/A
Ineligible Treatment, Bachelors	N/A	23.39 (28.41) N=24

Standard deviations in parentheses

As shown in Table 2, Bachelors in the Ineligible treatment contribute significantly less (23.39% of endowment) than they did in the Private (29.90%) or Public (37.60%) treatments (Wilcoxon signed-rank test $p=0.043$ and $p=0.020$, respectively). We cannot precisely identify the reason that Bachelors give less than they gave as Non-Stakeholders in the Private treatment. However, we conjecture that Bachelors feel a weakened urge to conform to the norm of contribution since they cannot be Stakeholder, or that their reduced earnings potential causes Bachelors to be less willing to trade off their payoff to benefit others (because of an income effect or spite). There is no difference between non-Stakeholders' behavior in the Ineligible treatment (35.96% of endowment) and their behavior in the Private (33.80%) and Public (37.79%) treatments (Wilcoxon signed-rank test $p=0.233$ and $p=0.410$, respectively). Thus, though the game has changed from the perspective of non-Bachelors (shrinking the "circle of reciprocity" to four people, and adding a public good beneficiary who is outside that circle), the effect of these changes is indeterminate and small.

Figure 5 shows the distribution of non-Stakeholder contributions, pooled across all rounds of all sessions. All treatments show a peak at zero tokens, a possible peak at 6-10 tokens, a dip between 10 and 20, and a peak at 20 tokens (full endowment). While qualitative differences appear between the treatments, many of these are not statistically significant. However, Bachelors give zero tokens more often in the Ineligible than in the Private (Wilcoxon signed-rank test $p=0.007$) and the Public ($p=0.014$) treatments. Also, in the Public treatment, subjects give 20 tokens more often than in the Private ($p=0.021$) and Ineligible ($p=0.098$ for Bachelors, $p=0.041$ for non-Bachelors) treatments.

Bins: 0; 1-5; 6-10; 11-15; 16-19; 20.
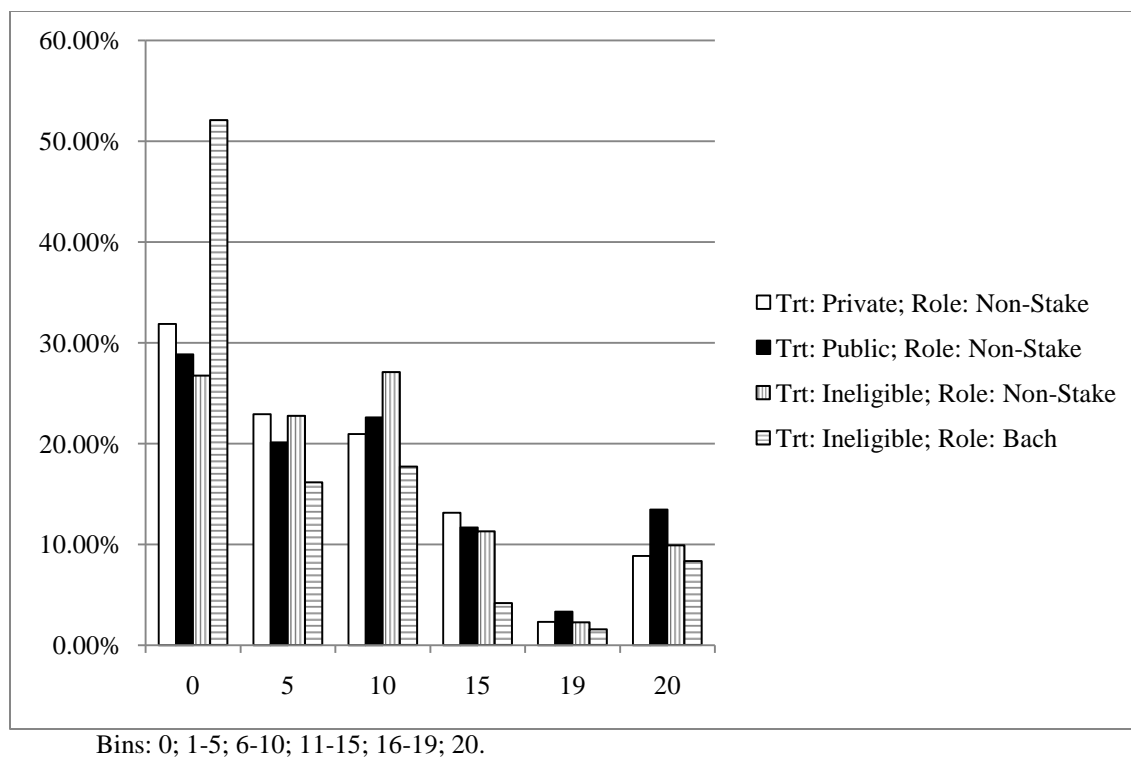
Figure 5. Distribution of Non-Stakeholder Contribution Amounts, Pooled across Rounds

Direct Reciprocity and Other-Regarding Preferences

We wish to detect reciprocity, i.e. responsiveness to the current Stakeholder's past behavior. To do this, we use nonparametric within-subject tests of aggregate statistics. For each person, we want to see whether (as non-Stakeholder) he gave more to the public good in rounds in which the current Stakeholder was previously generous to him, as compared to rounds in which the current Stakeholder was previously ungenerous. Stakeholder past generosity is determined by the current Stakeholder's average contribution to the public good in rounds wherein this subject was the Stakeholder. For example, Joe is Stakeholder in rounds 1 and 6, and Frank is Stakeholder in rounds 2 and 7. In round 2, Joe will remember how generous Frank was in round 1: "Stakeholder past generosity" will be Frank's contribution in round 1. In round 7, when Frank is

Stakeholder, our measure of "Stakeholder past generosity" for Joe would be the average of Frank's contributions in rounds 1 and 6 (when Joe was Stakeholder).

We define a "generosity threshold" such that contributions greater than this amount are considered generous. For each subject, we calculate his average contribution when facing a Stakeholder whose "past generosity" measure meets this threshold, and his average contribution when facing a Stakeholder whose past generosity does not. We tried many thresholds, including the "endogenous threshold" of the group's cumulative average contribution, and results were robust. We report results from a threshold of 10 tokens (50% of endowment). To see how this works, imagine that a group contains only Joe, Frank, and Mary. Frank gave 15 tokens every time Joe was Stakeholder, and Mary always gave 2 tokens. Joe's average contribution to a generous Stakeholder is his average contribution when Frank was Stakeholder, and his average contribution to an ungenerous Stakeholder is his average contribution when Mary was Stakeholder.

A subject displays reciprocal behavior if he gives more when facing a previously-generous Stakeholder than when facing a previously-ungenerous Stakeholder. Table 3 presents averages of these measures. In the Private treatment, the two statistics are not significantly different. This is expected (and reassuring) since in that treatment subjects can't tell who is Stakeholder or what the current Stakeholder did in the past. In both the Public and Ineligible treatments, subjects give significantly more—over 50% more—to previously-generous Stakeholders than they give to previously-ungenerous Stakeholders.[5] We use a difference-in-difference test to compare responsiveness (the difference between

---

[5] Recall that subjects could respond to past actions of the current Stakeholder *and* the current non-Stakeholders, although their contributions benefit the former three times as much as the latter. If subjects were responding to non-Stakeholders in this way, this would attenuate our within-subject test of responsiveness to Stakeholder history.

the amount given to a generous Stakeholder and the amount given to an ungenerous

Stakeholder) between treatments. Responsiveness is greater in both the Public (14.83%

of endowment) and Ineligible (17.3% of endowment) treatments than in the Private

(-1.59% of endowment) treatment (Wilcoxon signed-rank test $p=0.000$ and $p=0.000$,

respectively). This is evidence of directly reciprocal behavior. Responsiveness does not

differ between the Public and Ineligible treatments ($p=0.966$).

Table 3. Average Non-Stakeholder Contributions across All Rounds by Stakeholder's
Past Generosity toward Subject (in Percent of Endowment)

	Private Treatment	Public Treatment	Ineligible Treatment (excl. Bachelors)
Stakeholder gave ≥ 10 on average in past rounds in which subject was Stakeholder	27.57 (26.62)	41.92 (32.06)	41.22 (29.11)
Stakeholder gave < 10 on average in past rounds in which subject was Stakeholder	29.16 (29.59)	27.09 (27.63)	23.92 (24.08)
N	82	95	75
P-value (Wilcoxon signed-rank test)	0.773	0.000	0.000

Standard deviations in parentheses

N's are less than 120 because some subjects did not face both a generous and an ungenerous stakeholder. For this reason, we dropped 38 of 120 subjects in Private, 25 of 120 in Public, and 21 of 96 subjects in Ineligible.

Do subjects discriminate because of other-regarding preferences or because they

are strategic? We can dispose of strategic concerns by looking for reciprocal giving after

a person has passed his last Stakeholder stint. For example, again assume that Joe was

Stakeholder in rounds 1 and 6, and Frank was Stakeholder in rounds 2 and 7, and assume

that no further fundraising rounds follow. If Frank was kind to Joe in rounds 1 and 6,

will Joe reciprocate in round 7? If he is purely strategic, Joe has very little to gain, so he

should not contribute and therefore not reciprocate.

We test for reciprocal behavior rooted in other-regarding preferences by constructing statistics of each subject's average contribution to previously generous and ungenerous Stakeholders in rounds after this subject's last Stakeholder stint. These results, shown in Table 4, imply that other-regarding preferences play a role. Subjects in the Public treatment continue to discriminate between generous Stakeholders (to whom they give 30% of endowment) and ungenerous Stakeholders (16.93%) even after they have no strategic motive to do so. In the Ineligible treatment, the sample size is greatly reduced because fewer subjects can be considered (only non-Bachelors) and the treatment is shorter (eight rather than ten rounds). Because of this reduced power, discrimination in the late rounds of the Ineligible treatment is not statistically significant, although the point estimate is positive (26.60% is given to generous Stakeholders, and only 19.30% is given to ungenerous Stakeholders).

Table 4. Average Non-Stakeholder Contributions by Stakeholder's Past Generosity toward Subject, after Last Stakeholder Stint (in Percent of Endowment)

	Private Treatment	Public Treatment	Ineligible Treatment (excl. Bachelors)
Stakeholder gave ≥ 10 on average in past rounds in which subject was Stakeholder	14.60 (25.70)	30.00 (30.99)	26.60 (31.98)
Stakeholder gave < 10 on average in past rounds in which subject was Stakeholder	18.60 (29.24)	16.93 (27.43)	19.30 (27.75)
N	31	41	25
P-value (Wilcoxon signed-rank test)	0.436	0.026	0.435

Standard deviations in parentheses

N's are less than 120 because 48 subjects had to be dropped from each treatment because there were less than two rounds remaining after their last Stakeholder stint; additionally, more subjects had to be dropped from each treatment (41 of 120 from Private, 31 of 120 from Public, and 23 of 96 from Ineligible) if they did not face both a generous and an ungenerous Stakeholder after their last Stakeholder stint.

Indirect Reciprocity

An indirect reciprocator is a disinterested party who rewards one subject for generosity toward another subject. A strict definition of indirect reciprocity cannot be tested by looking at the responsiveness of a subject who has the opportunity to be Stakeholder, because he is not disinterested: he can benefit from a norm of cooperation. Our clean test for indirect reciprocity is the behavior of Bachelors in the Ineligible treatment. These randomly-selected subjects will never be Stakeholder, so they can never receive the benefits of targeted reciprocity. Therefore, direct reciprocity and strategic self-interest are not strong motives for Bachelors.

We have shown that Bachelors give significantly less than non-Stakeholders in the Ineligible treatment and less than they themselves gave in the other treatments. The people who are Bachelors show evidence of reciprocity in the Public treatment (i.e. they have positive "responsiveness," as defined previously; results available on request), so they do behave reciprocally when they are part of the circle of reciprocity. However, Bachelors in the Ineligible treatment do not give more (Wilcoxon signed-rank p=0.823) when facing previously generous Stakeholders (when they give 21.91% of endowment on average) than when facing previously ungenerous Stakeholders (25.54% of endowment). Therefore, indirect reciprocity does not seem to be a motivator in this setting.

This result is particularly intriguing, given that other studies have found indirect reciprocity. However, the re-matching structure of those experiments allows subjects to have an interest in the group's overall cooperation. Therefore, subjects in those experiments are not wholly disinterested, as our Bachelors are. Both conditions are valid

settings in which to examine cooperation; however, we feel that our result hews more

closely to the "bystander" interpretation of indirect reciprocity.

Heterogeneity in the Population

In estimating population effects, we may be missing important heterogeneity in

individual behavior. Our within-subjects implementation allows us to characterize each

subject within a treatment and see how subjects change their behavior between

treatments. We use an endogenous classification method to characterize subjects by their

baseline giving and their responsiveness to other subjects' past behavior. This method

(based on the "estimation / classification" method of El-Gamal and Grether 1995)

specifies a finite mixture model that mixes several instances of the same model and then

uses maximum likelihood to simultaneously assign subjects to types and choose the

parameters (coefficients) for each type. This allows the data to generate the "rules" that

best describe the data. A standard finite mixture model generates population percentages

for each of the nested models. To summarize the data, we assign each subject to the type

that best describes his contributions; that is, for each subject, we round to 1 the

probability of the type that best represents him, and round the rest down to 0.

We index types by *m*, and model contributions as a function of a history variable

h_{ikt} (subject *k*'s past contributions when person *i* was Stakeholder as of period *t*) and other

variables $\mathbf{X}_{it}$. The model uses type-specific parameters, including a constant parameter

a_m which represents baseline tendency to give (altruism), and a "responsiveness"

parameter b_m. The empirical model for each type is:

$$g_{it} = a_m + b_m h_{ikt} + \mathbf{C}_m \mathbf{X}_{it} + \varepsilon_{it}$$

This analysis is similar to Bardsley and Moffatt (2007). However, their game is different (as it is a sequential public goods game using a "Conditional Information Lottery" to determine payments). Also, while they specify four types (a free-rider type, an altruist type, a reciprocator type, and a strategic type) and then estimate each type's parameters and mixing probabilities using maximum likelihood, we specify a number of types but do not restrict parameters of any of the types (e.g., $a=b=0$ for an free-rider).

We performed this analysis with two, three, and four types. In the results we report, $\mathbf{X}_{it}$ includes only the group cumulative average contribution (excluding the subject's own past contributions and the current Stakeholder's past contributions). Results were robust to the inclusion of round number and a dummy indicating whether the subject had passed his last Stakeholder stint. Results were also robust to the inclusion of subject demographics (gender and age), but as none were significant we do not report those results.

For simplicity, Table 5 presents results from the two-type analysis. In each treatment, Type 1 is large (comprising 61-73% of the population) and has a small a_1 (low altruism), whereas Type 2 has a large a_2 (high altruism, $a_2 > a_1$). Thus, we will call Type 1 "low type" and Type 2 "high type." The b_m parameter (responsiveness) is large, significant, and positive for both types in both the Public and Ineligible treatments.[6] Within each treatment, the types show similar responsiveness ($b_1 \approx b_2$, t-test p-value greater than 0.13 in all cases). That is, while the types differ in altruism, all types tend to

---

[6] As the table shows, the responsiveness coefficient is also significant (but very small) in the Private treatment, though in this treatment it is not possible to target reciprocal behavior at individuals. This spurious significance comes from the dynamic entanglement inherent in panel analysis of this kind of data. Responsiveness is much higher in the Public and the Ineligible treatments as compared to the Private treatment (t-test p-value=0.000 and 0.001, respectively).

reciprocate. This universal responsiveness holds as more types are added. Thus, the

most important heterogeneity is in the baseline tendency to give and not in the tendency

to reciprocate.

Table 5. Behavioral Types from Estimation / Classification Procedure

		Private Treatment	Public Treatment	Ineligible Treatment
	Number (percent) in type 1	76 (63%)	73 (61%)	70 (73%)
	Number (percent) in type 2	44 (37%)	47 (39%)	26 (27%)
Type 1	Constant	8.34***	-3.91	8.51**
		(2.60)	(3.34)	(3.71)
	Stakeholder past contributions to me	0.09**	0.33***	0.30***
		(0.04)	(0.04)	(0.05)
	Group cumulative non-Stakeholder contributions (excluding me)	0.24***	0.37***	0.12*
		(0.06)	(0.07)	(0.07)
Type 2	Constant	38.72***	23.96***	50.07***
		(4.29)	(4.25)	(8.22)
	Stakeholder past contributions to me	-0.01	0.43***	0.36***
		(0.078)	(0.06)	(0.10)
	Group cumulative non-Stakeholder contributions (excluding me)	0.32***	0.28***	-0.27*
		(0.10)	(0.09)	(0.07)
	# of observations	720	720	432
	Log likelihood	-3389.976	-3411.576	-1612.1664
	AIC	6791.952	6835.152	3236.3328
	BIC	6808.677	6851.877	3251.719

Standard errors in parentheses. Normal distribution for error; mean is linear function of covariates.

* significant at 10%; ** significant at 5%; *** significant at 1%

For comparison, Bardsley and Moffatt (2007) find approximately 39% of subjects

to be strategists, 25% to be free-riders, 30% to be reciprocators, and 6% to be altruists. A

striking similarity between their results and ours is that neither they nor we find

significant evidence of unconditional altruism.

How do subjects change across treatments? There is imperfect stability in types: 56% of subjects are the same type for all treatments, and most of these (42% of the population) are consistently low type. Subjects who are high type in one treatment give significantly more in the other treatments.[7] Type transitions are shown in Table 5. The low type is "stickier" in a loose sense: for each treatment, subjects that are low type in that treatment are very likely (78-85%) to be low type in the other treatments, while high-type subjects are less likely (32-56%) to become low type.

Table 6. Treatment-to-Treatment Type Transitions: Percent of Subjects Making Each Transition

	Low to Low	Low to High	High to Low	High to High
From Private To Public	49%	14%	12%	25%
From Private To Ineligible	54%	9%	19%	18%
From Public To Ineligible	50%	9%	23%	18%

Transitions to and from Ineligible treatment types use only 96 subjects, since Bachelors are excluded.

Conclusion

Favor trading is a natural element of social networks. Grassroots fundraising harnesses people's tendency to trade favors and uses that reciprocal drive to increase the provision of public goods. Using an experimental design that allows different forms of reciprocity to be turned on and off, we explore the power of favor trading and the mechanisms through which this institution may work.

---

[7] Results available on request. For example, Private-high type subjects give more in the Public treatment than do Private-low type subjects (55.61% of endowment as compared to 27.41%, Wilcoxon rank-sum test p=0.000).

We demonstrate reciprocal giving rooted in other-regarding preferences, i.e. direct reciprocity. While some studies have robustly shown direct reciprocity in laboratory institutions like the investment game, others have criticized the abstraction and isolation of those institutions. The presence of direct reciprocity in our slightly richer setting may be the next step in building a bridge toward non-laboratory institutions that encourage the expression of direct reciprocity.

It is remarkable that in our strict, clean test, we do not find indirect reciprocity. Our definition of indirect reciprocity may be too strict, in a sense. While most experiments seeking indirect reciprocity allow actors to benefit from an increased tendency to cooperate within the group, we strip our Bachelor of incentives to foster cooperative norms. It is also notable that in our setting, Bachelors can never earn as much as other subjects. It is possible that this inequity causes some feeling of disaffection on the part of the Bachelors, and this suppresses the expression of their social preferences. This accords with the low level of contributions by Bachelors, and also with the remarkably flat profile of Bachelor contributions across the rounds of the treatment. Whatever the explanation, our results imply that indirect reciprocity is sensitive to institutional factors. The institutional factors essential to our test of indirect reciprocity—the disinterestedness of the indirect reciprocator and the potential for inequity—are factors that could certainly be present in many evolutionary settings, so the absence of indirect reciprocity in our results is quite interesting.

In our experimental setting, favor trading increases cooperation by 14.4%, an amount that is both statistically and economically significant. This should provide a lower bound for the level of efficiency that this kind of institution might achieve. In a

true social situation, rewards and sanctions are much stronger motivators than the

incentives offered in the lab, and social interactions can be an extremely long-term game.

These results support our intuition about the effectiveness of grassroots solicitation as a

fundraising tool, and suggest that favor trading could be successfully leveraged in other

settings as well. Our results also show the robustness of direct reciprocity, and the

fragility of indirect reciprocity, to small changes in the institution.

Chapter II: The Effects of Conservation Reserve Program Participation on Later Land Use

Introduction

The United States Conservation Reserve Program (CRP) pays farmers to retire land from agriculture. The program has several goals, including the preservation of environmental assets and the long-term improvement of the country's agricultural productivity. Balancing multiple goals is always difficult, and may be especially challenging with regard to the long-term effects of such a program. In the case of the CRP, the disposition of land that leaves the program deserves study because the program is extremely large and market and political factors can trigger policies that suddenly release quantities of CRP land. For example, the 2008 Farm Bill dictated a drop in the CRP's enrollment cap from 39.2 million acres to 32 million acres. If land that leaves the CRP persists in conservation, this long-term effect boosts the program's environmental benefits. If ex-CRP land returns to farming at a high rate, however, this works against the program's environmental goals, but may indicate that the CRP made the land more productive.

We examine land use of parcels that have been in the CRP to determine whether the CRP has an effect on the land's later use. We ask whether CRP participation causes land to be more or less likely to be farmed (or to take up another land use) and whether past CRP participation is correlated with the use of conservation practices on farmed land. We focus on land that exited the CRP between 1992 and 1997, the first period in which land left the program in great quantity. We use a treatment effect framework, with

regression and matching methods. This approach makes an explicit comparison between land that has been in the CRP ("treated" parcels) and land that has not been in the program ("control" parcels), a comparison we must make to attribute causality to CRP participation.

The CRP was established by the 1985 Food Security Act. A CRP contract binds a landowner to abstain from farming the land for a period (10-15 years) and to plant a conservation cover for that period. To join the program, farmers submit bids consisting of the rental payment they will accept, the land they would like to enroll, and the conservation cover they will plant. The "best" bids (according to criteria that evolved over time) are accepted. The USDA reimburses farmers for some of the cost of planting conservation cover. If a farmer exits before his contract expires, he pays a penalty.

The program has been very popular. Enrollment has usually been near its acreage cap (36.4 million to 39.2 million acres from 1992 to 2002), and competition has been keen for CRP contracts. The first contracts expired in 1996. Very few parcels left the program before that, and in 1996 (as in later years) most parcels with expiring contracts re-enrolled. There is some variability in a parcel's ability to exit or re-enroll in the program. For example, around 1996, some holders of unexpired contracts were allowed to remove their land from the program without penalty, and some expiring contracts were automatically extended for one year if the contract-holder wanted to do so. The CRP's eligibility criteria and bidding system also changed over the years to improve incentive-compatibility and to take into account an Environmental Benefits Index (EBI). As a result, some early enrollment waves accepted land for contracts that, when expired, were not renewable because the land did not meet the CRP's new criteria.

Post-CRP land disposition has been of interest for some time, but since CRP exits did not occur in quantity until the first contracts expired in 1996, studies before 1996 were performed using surveys and simulations. Researchers (for example, Cooper and Osborn 1998; Johnson, Misra, and Ervin 1997) surveyed farmers to gauge their intentions to remain in the program and their plans for the land if the CRP were eliminated. These surveys give useful qualitative results; for example, they indicate that not all CRP land would be farmed if it were not in CRP and that market prices drive land use decisions.

General equilibrium simulators estimate parameters or elasticities from observed land use and then use those estimates to predict the land use transitions that would occur under different market and policy conditions. These simulations account for the effects of CRP entry and exit on agricultural supply and therefore on price, which can feed back into other parcels' entry into and exit from agriculture.[8] Using the POLYSYS simulator, De La Torre Ugarte et al. (1995) estimate that 57% of CRP land would be farmed if the CRP was terminated, and De La Torre Ugarte and Helliwinckel (2006) estimate that 37% of CRP land would be farmed if the CRP were gradually eliminated. Secchi and Babcock (2007) use the EPIC simulator to show that crop prices have a very strong influence on the decision to un-retire CRP land. Lubowski, Plantinga, and Stavins (2008) find that the CRP accounted for a reduction in cropland of 29 million acres between 1982-1997 (given 32.8 million acres estimated enrolled in 1997, this implies that at least 88% of CRP land would have been farmed).

While general equilibrium simulation is essential for market-level analyses, it may not be well-suited to the analysis of parcel-level transitions, as discussed in Roberts

---

[8] One such phenomenon is "slippage," as discussed in Wu (2000) and a series of related papers.

and Lubowski (2007). General equilibrium simulators predict land use by assuming that land will enter its highest return use. These models' ability to predict in this context is limited by the fact that returns (and land characteristics determining those returns) are not fully observed. Therefore, micro-level analyses of land use changes are a useful complement to general equilibrium simulations.

One such analysis is Roberts and Lubowski (2007), who also examine post-CRP land use decisions. They use post-CRP land use in 1997 to parameterize a model of land use for ex-CRP land, and perform a partial equilibrium simulation to predict that 58% of CRP land would enter farming if the program were eliminated. Roberts and Lubowski use a Heckman two-stage model to control for selective exit from the CRP. However, a remaining concern is selection in CRP enrollment: land that enters the program must be have low opportunity costs, i.e. low returns from farming. To assign causality to CRP participation, this selection must be considered so that CRP parcels are compared to parcels that are similar in quality—that is, to land that is also marginal and therefore likely to leave farming anyway.

We compare ex-CRP parcels to parcels that have not been in the program to ask: did CRP participation cause changes in the later land use of enrolled parcels (compared to the use they would have entered had they not ever been in CRP)? In contrast, Roberts and Lubowski ask: given a known function that determines post-CRP land use, what will happen to currently-enrolled CRP land if the program disappears?

This paper contributes to the literature in a number of ways. We attribute causality to CRP experience by using plausible counterfactual (non-CRP) land parcels.

Our methodology is innovative in the use of a sample specification step that trims the treatment (CRP) and control (non-CRP) groups to the most comparable units. We also perform a multinomial logit analysis, which more thoroughly explores the land use decision. Finally, we study adoption of conservation practices on land that exits CRP and is later farmed.

We find that a naïve analysis, without careful specification of counterfactual land, shows that ex-CRP land is farmed at a lower rate than other parcels. However, when compared with the best counterfactual group, ex-CRP land is 21-28% more likely to be farmed than non-CRP land. Because this counterfactual group is very much like the ex-CRP group, we can infer that the increase in cultivation is caused by CRP participation. This result is novel in the literature, but is not unexpected, since the land should have improved while in the program. Thus, the CRP's long-term effects comport with some of the program's goals (such as agricultural efficiency) while working against others (environmental protection). We also show that cultivated ex-CRP land is more likely than similar land to adopt a conservation practice such as contour farming. However, we cannot infer whether the conservation practice result is caused by CRP participation.

The remainder of this chapter proceeds as follows. In the next section, we discuss a model of land use choice. Next, we describe the methods we will use to address the research questions. In the following section, we introduce the data and provide summary information. We give special attention to the potential "treatment" (CRP) and "control" (non-CRP) groups that will be used in the analysis. We present results in the following section. In the final section we conclude.

Theory

The essential insight behind the analysis is that a land parcel's use is determined by the relative returns of all possible uses for that parcel. This model has its roots in the ideas of Ricardo and von Thünen: land enters the use that provides the highest quasi-rents, and these quasi-rents depend on land quality and land location. In this model, we may abstract away from landowner preferences, because idiosyncratic values are reflected in quasi-rents if markets are competitive and complete.

We denote land use as $a \in A$, where A is the set of all possible land uses. We define $\mathbf{p}$ as the vector of input and output prices in the economy, $\mathbf{x}$ as the vector of land characteristics, and $\mathbf{t}$ as the vector of technology and policy factors. Each parcel has a quasi-rent corresponding to each potential use in set A. The quasi-rents are functions of $\mathbf{p}$, $\mathbf{x}$, and $\mathbf{t}$: that is, landowner rents from use a are $\pi(a \mid \mathbf{p}, \mathbf{x}, \mathbf{t})$. Observed land use is:

$$a^* = \arg\max_{a \in A} \pi\left(a \mid \mathbf{p}, \mathbf{x}, \mathbf{t}\right)$$

Some important points inform the analysis. First, we must be concerned with both observable and unobservable elements of the land characteristic vector $\mathbf{x}$. Land characteristics have static elements and elements that vary over time in ways that depend on the land's history. For example, soil can become more erodible as land is farmed. Notably, soil erodibility can improve as land is removed from agriculture and conserved, such as when it is planted with conservation cover, and CRP participation has been shown to improve land quality (e.g., Uri 2001). Low soil quality (e.g., high erodibility) makes the land less productive to farm, so parcel-level quasi-rents also depend on interactions between land characteristics and market prices.

Land use transitions should occur when the quasi-rent from the current land use is less than the quasi-rent that could be earned from some alternative land use. However, some parcels face transition costs when changing land use. For example, parcels in CRP contracts face penalties if they exit the program before the contract has expired. Even on contract expiration, land seeking to leave CRP and enter agriculture faces irreversible transition costs and uncertain future returns; these characteristics can delay or forestall transitions. For land leaving CRP, transition costs back into agriculture are affected by the type of conservation cover planted for CRP. Grasses are the most common cover (the others are forest and wildlife habitat) and carry the lowest transition cost. Because of transition costs, changes observed over short periods (such as the transitions we investigate) should *understate* the long-term transitions that will occur.

Transitions into the CRP are costly, but are subsidized by the government: the USDA reimburses some of the costs of planting conservation cover. These transitions are not possible for all parcels because of CRP eligibility requirements. As the CRP's criteria changed to emphasize environmental benefits, some parcels in the program ceased to be eligible. In these cases, if the contract-holder re-applied upon expiration of the contract, the application was rejected. Thus, when a parcel leaves the CRP, it could be because the landowner prefers to put the land into another use (i.e. because the return to another use is now higher) or because it is no longer eligible to stay in the program.[9]

When a parcel leaves the CRP voluntarily or is forced out, it should transition into the most profitable non-CRP land use. CRP contracts are only granted to land that was

---

[9] Parcels that are potentially eligible with respect to land characteristics may be rejected if the bid was too high. If the bidding mechanism is incentive-compatible, the bid reflects the parcel's opportunity cost. Such a rejected bid would mean that the return to another use (the opportunity cost) now exceeds the return to CRP participation.

cultivated cropland at the time of application, so farming may be the next most profitable activity for much of this land. However, as noted above, transition costs may delay or forestall desired transitions. Additionally, land characteristics and market conditions may have changed to make farming either more or less profitable for any given parcel.

Many factors described above would have influenced land use decisions even without the CRP. We seek to identify the change in land use decisions that occurs because of parcels' experience in the CRP. How can CRP tenure cause changes in post-program land use? The CRP experience could change landowner preferences for conservation to reduce the likelihood of farming. Recall, however, that the model assumes that preferences are reflected in quasi-rents, so the model does not allow identification of such an effect. Practically, however, these preferences could be imperfectly capitalized or not perfectly reflected in quasi-rent proxies. On the other hand, land improvement while in the program increases the returns to farming, and this should increase the likelihood of farming. Land that is retired from farming but does not enter CRP may also improve, but CRP's subsidization of conservation practices and CRP's enforced 10-year retirement should cause greater improvement. Thus, theory is ambiguous as to what effect CRP participation should have on later land use.

Methods

The fundamental problem of causal inference is that it is impossible to observe any unit as both a treated unit and a counterfactual (non-treated) unit. Here, end-of-period (1997) land use is the outcome and CRP participation is the treatment. To identify effects of CRP participation on land use, we must compare the land use outcomes of each CRP parcel to the outcomes that would have occurred had the parcel never entered CRP.

The fundamental problem of causal inference appears here in that we cannot know what would have happened to these CRP parcels had they not entered the program. These parcels are different from non-CRP parcels because of selectivity in CRP entrance, and therefore a simple comparison to the land use outcome of non-CRP parcels could produce biased results. We must choose non-CRP parcels that can act as good counterfactuals.

Controlling for land characteristics, the returns to various land uses, and the interactions between these would yield unbiased results if the treatment and control groups are sufficiently similar. For our first step, we perform a variety of "naïve" analyses using a full data set. We regress end period (1997) farming outcome on past CRP status, controlling for land characteristics, estimated returns to land uses, interaction terms, and region dummies, with appropriate spatial error clustering.

If the parcels in the control (non-CRP) group are very unlike parcels in the treatment (CRP) group, regression results can be biased, as demonstrated in Rosenbaum (2002). This is because treatment units are compared to inappropriate control units, and given the inevitable mis-specification of a linear model, differences in sample characteristics may drive apparent differences in outcomes. To try to counteract this bias, we perform a matching analysis on the same data set. The matching algorithm chooses the control (non-CRP) units that are "most like" the treated (CRP) units based on observable characteristics. All matching results reported in this paper use nearest neighbor matching, without calipers, using the Stata package psmatch2 (Leuven and Sianesi 2003). A propensity score is estimated (as a function of land characteristics, rent proxies, and region dummies) to determine a given parcel's likelihood of being in CRP. A Mahalanobis metric (a way of calculating the difference between points in a

multidimensional space) is used to do the matching, and this metric is calculated from the land characteristics, rent proxies, region dummies, and the propensity score. We calculate the average treatment effect on the treated (ATT) based on the matched sample.

Matching can only account for differences in observable characteristics. If treatment and control groups differ in unobservable characteristics that are correlated with the outcome of interest, matching results may still be biased. What unobservable characteristics are important in this analysis? Land characteristics are imperfectly observed, and transition costs are difficult to estimate, but both are crucially important in the determination of land use outcomes.

To address this, we perform an additional sample specification step. We trim the samples of CRP and non-CRP data to include only the parcels that are most appropriate to compare, based on theory and program characteristics. Conceptually, this is a population-level process much like the observation-level process of matching: the most comparable treatment and control group are selected based on observables to reduce bias. Comparability of the populations is difficult to ensure because of unobservable differences in the unobservable characteristics of land. However, inferences about these land characteristics can be drawn from the land use into which the parcel's owner has chosen to put it. This sample specification step, therefore, is a selection process targeting the CRP and non-CRP parcels that are most comparable based on the characteristics that can be inferred from their patterns of land use. The results comparing different subsamples used in the analysis in this paper can be thought of as ways of approaching the data with different research questions.

We specify two different treatment groups: CRP-Eligible and CRP-Exit. CRP-Eligible includes land that enrolled in the CRP in an early signup wave and was therefore probably eligible to exit by 1997. (Publicly-available data do not identify contract end dates, so exit eligibility is not known and must be guessed.) If the treatment group included land outside this group, many CRP parcels would be locked into contracts that are costly to exit, and the results would greatly understate the likelihood of transitions. However, much of the land that was eligible to exit actually stayed in the CRP, in part because of the automatic contract extension offered by the USDA in 1996. If a parcel stayed in the program, CRP participation may have changed the use this land would have adopted had it left the program, but it did not increase any of those returns enough to drive the land out of the program, or at least not by 1996. In this sense, it is not possible to fully observe the CRP's effect on land use for any parcels that stayed in the program.

We therefore create the group CRP-Exit, containing only parcels that actually exited. CRP-Exit contains some parcels that choose to leave (perhaps because they could earn a greater return outside the program), and some that are forced out (including parcels that do not meet new CRP eligibility rules), but it is not possible to determine why any given parcel left the program. Given the automatic contract extensions offered in 1996, it is likely that most CRP exits at this time were voluntary.

Using CRP-Eligible as the treatment group asks the question, "Given that staying in CRP is allowed, what land use is observed on land that has been in CRP?" Using CRP-Exit asks, "Given land that has exited the CRP, what land use is observed?" This is a narrower question, because it is likely that the parcels that exit CRP are better-quality than the parcels that remain in the program. Some parcels may have exited specifically

so that they can be farmed. The results from an analysis that uses CRP-Exit are informative, but caution should be exercising when extrapolating to other CRP parcels.

We also specify two different control (non-CRP) groups. Control-All includes all land that was farmed in 1982, the start of our analysis period. This restriction is important, because only farmed land can enter the CRP. Control-All includes productive farmland that is very unlike CRP land. We can control for observable characteristics, but these may understate the differences between CRP and Control-All parcels.

To address this difference in unobservables, we specify another non-CRP group called Control-Unfarmed. Control-Unfarmed parcels were farmed in 1982, so share a history of cultivation with the CRP parcels, but were not farmed in 1992. This ensures that non-CRP parcels face transition costs if they are to be cultivated, as CRP parcels do. It also restricts the sample to land of low enough quality that the landowner is willing to remove it from cultivation, which must also be true of CRP land. That is, by his land use choice, the landowner of a Control-Unfarmed parcel has revealed that his parcel's unobservable attributes are of low quality, just as the landowner of a CRP parcel has.

There are a few sources of possible differences between Control-Unfarmed and CRP land. First, Control-Unfarmed land may be of lower quality than ex-CRP land because Control-Unfarmed was willing to retire from farming without a subsidy, while some CRP land probably would not have retired without being paid. This would make Control-Unfarmed less likely to be farmed later. Second, and conversely, some Control-Unfarmed land may be higher quality because it left farming but did not choose to enter CRP (we cannot identify which land applied to CRP and was rejected and which did not

apply). These farmers may have wanted to keep the value of the option to farm during the coming ten years. This would make Control-Unfarmed more likely to be farmed.

Because of these factors, Control-Unfarmed land may be slightly better or slightly worse than CRP land. We acknowledge this potential difference. However, we expect that the resulting bias should be small, and we feel that Control-Unfarmed is still the best possible counterfactual group for CRP land.

Additionally, factors related to the timing of transitions may bias the results of comparison between CRP land and Control-Unfarmed. First, some Control-Unfarmed land may be in fallow cycles. The data set we use tries to classify land in a fallow cycle as cultivated cropland. Land classified as non-cultivated cropland (which is mostly hay) or pastureland is verified to have not been cultivated cropland for the last three years. However, if the fallow cycle is five years or longer, this land may be misclassified so that it ends up in the Control-Unfarmed category. Parcels that exit CRP may begin farming immediately because they will have just emerged from a 10-year fallow period (their CRP tenure), but only some non-CRP parcels in long fallow cycles will be ready to farm. This may introduce a bias making Control-Unfarmed less likely to be farmed in 1997 as compared to ex-CRP land. Second, and conversely, if transitions into farming take time our results may be biased to show Control-Unfarmed land more likely to be farmed. This is because, whereas non-CRP land can transition into farming in any year, CRP land can only transition after contracts end in 1996. Thus, ex-CRP land transitions may be under-recorded relative to non-CRP land transitions, thus biasing our results toward zero.

Using the pre-processed treatment and control groups, we perform the same regression and matching analyses described for the naïve sample. Comparing transitions

into cultivated cropland among CRP-Exit parcels to those observed among Control-Unfarmed parcels is akin to performing a difference-in-difference analysis. It is the difference between CRP and non-CRP land in the difference in land use between the earlier period (1982) and the end period (1997). Analyses on the broader data set, and on the other combinations of treatment and control groups, do not yield valid difference-in-difference results because in those cases, the samples are not comparable.

Data and Summary Statistics

CRP contract information is sensitive and therefore is not distributed at a level of detail that is useful for parcel-level analysis. We use data from a large nationwide land survey to collect parcel-level characteristics, and we match those data with county-level estimates of the returns to various land uses.

First, parcel-level data were obtained from the USDA's National Resource Inventory (NRI; see US Department of Agriculture 2001). The NRI is a panel survey of over 800,000 land parcel samples throughout the country. It provides data for 1982, 1987, 1992, and 1997. The NRI is a stratified survey, and the analysis that follows takes into account the NRI's sampling structure. Data are not available for Alaska or the US Virgin Islands, and we exclude parcels in Hawai'i and Puerto Rico, since land use decisions in those areas are likely to be quite different from land use decisions in the forty-eight contiguous states. Variables reflecting essential land quality data are not recorded for land that is urban, transportation, federal, or water. As a result, we exclude

parcels that entered one of these land uses in 1997. This exclusion is particularly

acceptable because this land's use may be idiosyncratic.[10]

The NRI land quality data of interest for this analysis are slope, erodibility, land

capability classification (defined in US Department of Agriculture 2009), and the prime

farmland indicator (*ibid.*). The NRI records land use in broad categories (e.g., cultivated

cropland, non-cultivated cropland (which is mostly hay), pasture, and CRP) and narrower

sub-categories. Throughout this paper, cultivated cropland will be referred to as

"farmed" land. This use includes all close and row crops, and will be an outcome of

particular interest because of farming's environmental impact. Parcel location

information is available but limited in precision because the NRI data are designed to

render precise sample location identification impossible.

The second source of data is a set of county-level proxies for the returns to

various land uses. These data consist of 1996 rent levels and 1986-1996 changes in rent

levels for various land use categories. These rent data are described in the appendix of

Roberts and Lubowski (2007). These county-level proxies are only available for 1600

counties, and the analysis considers only counties for which rent data are available.[11]

Table 7 shows the characteristics of land in subpopulations of interest. The first

column shows all land that was farmed in 1982, and the second column shows all land

that was in CRP in 1992. As expected, CRP land is worse than the broader sample of

originally-farmed land by all measures (less likely to be prime or have a good land

classification, more erodible, and more sloped). CRP land also tends to stay in CRP at a

---

[10] Results hold for specifications excluding land classification (the characteristic available for the fewest uses) from the control variable set and including the land from these 1997 uses (available upon request).
[11] All NRI counties with CRP parcels have rent data; only non-CRP parcels are thus excluded. Summary statistics including parcels without rent data are similar to those presented here (available upon request).

very high rate, and farmed land tends to continue to be farmed at a high rate. Among the specified subsamples, Control-All (the general non-CRP land sample) is of better quality than all of the other groups are, including Control-Unfarmed (the land that left farming but did not enter CRP). Both Control groups are of better quality than both CRP groups. CRP-Exit is of somewhat better quality by most measures than the CRP-Eligible group, which supports the intuition that some land leaves CRP specifically to be farmed. Interestingly, Control-Unfarmed is very slightly better than CRP-Exit on observable characteristics.

Table 7. Characteristics of Land in Subsets of NRI Sample

	Farmed in 1982	CRP in 1992	CRP-Exit	CRP-Eligible	Control-All	Control-Unfarmed
Prime farmland? [a, d]	0.547	0.286	0.345	0.262	0.576	0.438
	(0.002)	(0.002)	(0.009)	(0.004)	(0.002)	(0.006)
Good land class? [a, c, d]	0.952	0.875	0.888	0.873	0.960	0.895
	(0.001)	(0.002)	(0.005)	(0.002)	(0.001)	(0.004)
Erodibility index [a, d]	7.330	12.990	12.209	13.759	7.176	10.579
	(0.026)	(0.971)	(0.221)	(0.108)	(0.030)	(0.221)
Slope [a, d]	2.759	4.260	4.841	4.270	2.783	4.166
	(0.009)	(0.018)	(0.010)	(0.025)	(0.011)	(0.046)
Return: crops [b]	93.365	80.739	96.169	81.074	94.863	98.626
	(0.150)	(0.097)	(0.754)	(0.275)	(0.168)	(0.820)
Return: government payments [b]	10.967	8.856	9.364	8.870	11.192	9.023
	(0.014)	(0.010)	(0.063)	(0.275)	(0.016)	(0.061)
Return: pasture [b]	29.050	27.052	30.967	28.109	29.267	31.005
	(0.064)	(0.024)	(0.181)	(0.104)	(0.072)	(0.308)
Return: range [b]	10.912	11.228	8.937	11.179	10.872	10.553
	(0.032)	(0.027)	(0.201)	(0.064)	(0.036)	(0.145)
Return: forest [b]	11.725	12.189	21.662	12.807	11.659	14.453
	(0.044)	(0.011)	(0.140)	(0.065)	(0.049)	(0.181)
Return: urban [b]	2331.954	2426.994	2450.342	2566.866	2318.095	2260.462
	(4.968)	(3.711)	(31.431)	(8.409)	(5.517)	(21.918)
Early CRP signup wave	-	0.617	0.821	1	-	-
		(0.002)	(0.005)	(0)		
CRP cover of grass in 1992	-	0.906	0.920	0.927	-	-
		(0.001)	(0.004)	(0.001)		
CRP in 1992	0.082	1	1	1	-	-
	(0.000)	(0)	(0)	(0)		
CRP in 1997	0.078	0.895	-	0.863	0.007	0.005
	(0.000)	(0.001)		(0.002)	(0.000)	(0.000)
Farmed in 1992	0.821	-	-	-	0.921	-
	(0.001)				(0.001)	
Farmed in 1997	0.790	0.053	0.528	0.068	0.891	0.271
	(0.001)	(0.001)	(0.007)	(0.002)	(0.001)	(0.005)
Conservation practice if farmed in 1997	0.233	0.251	0.258	0.271	0.232	0.181
	(0.001)	(0.013)	(0.014)	(0.016)	(0.001)	(0.008)
Hundreds of acres	3,755,742	340,400	31,417	188,208	2,699,633	211,329

Standard errors in parentheses

[a] 1997 land characteristic data

[b] 1996 land use returns

[c] "Good land class" is an indicator for land classification of 1, 2, 3, or 4, indicating few restrictions on use

[d] Land characteristic data only available for certain land uses, so these are means over available data: erodibility index not available for pasture; erodibility and slope not available for pasture, range, forest, other rural; erodibility, slope, prime, and land class not available for urban, water, and federal land.

Table 7 also shows that 61.7% of all land that was in the CRP in 1992 was part of an early signup wave, but only 82.1% of all land that exited CRP between 1992-1997 was in an early signup wave (probably because of contract releases that were granted at that time). Grass is the most common conservation cover planted on all land in CRP, and is not planted on CRP-Exit land at a higher rate (92%) than on CRP-Eligible land (92.7%). Consistent with intuition, most (86.3%) CRP-Eligible land stays in CRP in 1997, and most Control-All land was farmed in 1992 (92.1%) and 1997 (89.1%). Finally, CRP-Exit land is unconditionally more likely to be farmed than the Control-Unfarmed sample. This result will continue to hold throughout the analyses.

Results

The data and methods employed allow an examination of land use outcome and the adoption of conservation practices on farmed land. We will study land use outcome both as a binary choice (cultivated or not) and as a multinomial choice. Here we can infer causality on the part of the CRP because we have strong reason to believe that we have matched parcels with similar characteristics. In the conservation practice results, the sample is further restricted to land that is farmed in 1997. There are competing reasons why ex-CRP land may adopt conservation practices at a higher rate, so we cannot infer causality for this result.

Binary Land Use Outcome Results

To study land use, we begin with a simple binary choice model. The landowner chooses whether to put the land into "farming" (cultivated cropland) or into some other use. Other uses include non-cultivated cropland and pasture; these uses are obviously

important to agriculture, but cultivated cropland has a larger environmental impact so is of particular interest. Cultivated cropland includes row and close crops. The vast majority (95.58%) of 1997 non-cultivated cropland is used to grow hay.

We form a naïve sample of all NRI observations for which county-level rent data are available. Analysis of this broad sample provides interesting insights, as shown in Table 8. In three different OLS specifications,[12] we use mutually exclusive dummies to compare 1997 use of CRP land to 1997 use of other land. Specification I uses dummies to indicate 1992 CRP participation and 1992 farming. This specification shows that land that was in the CRP in 1992 is less likely to be farmed in 1997 than land that was farmed in 1992 and even than land that was neither farmed nor CRP in 1992. Using the same data set, we perform a matching analysis using 1992 CRP participation as the treatment variable and the same explanatory variables as elements of the Mahalanobis metric and propensity score. Using the matched sample, the estimate of average treatment effect on the treated indicates that CRP land is 73.8% (standard error 0.014) less likely to be farmed than all other land. This result was driven by land that was farmed in 1992 and 1997 and land that was in CRP in 1992 and 1997.

---

[12] Results are very similar for logit and probit specifications as well; results available upon request.

Table 8. Land Use Results with Naïve (Un-Pre-Processed) Sample

	Specification I	Specification II	Specification III
CRP dummy	-0.233***		
	(0.005)		
Early CRP dummy		-0.216***	
		(0.006)	
Late CRP dummy		-0.260***	
		(0.005)	
CRPEXIT dummy			0.401***
			(0.007)
FARM92 dummy	0.629***	0.629***	0.791***
	(0.005)	(0.005)	(0.003)
N (hundreds of acres)	3,006,841	3,006,841	3,006,841
F	15,619.34	19,859.11	3,311.24
R^2	0.616	0.616	0.612

* significant at 10%, ** significant at 5%; *** significant at 1%

Standard errors in parentheses. Regressions are OLS. All included land was cultivated in 1982, is in a county for which rent data is available, and does not enter 1997 land uses: urban/built-up, water, or federal. Specifications control for 1997 land characteristics (prime farmland indicator, good land class indicator), 1996 land use rents and changes, region code dummies, and land characteristic x rent interactions. Survey regressions performed with data points appropriately weighted and clustered errors.

Specification II in Table 8 replaces the CRP dummy with dummies to indicate early CRP signup and late CRP signup, with results similar to Specification I. Specification III in Table 8 hints at the need for pre-processing. A dummy is included to indicate CRP exit, and this dummy has a positive coefficient. This land is more likely to enter farming than the baseline group. The baseline group, however, is unintuitive: it includes land that was not farmed in 1992 and land that stayed in (did not exit) CRP.

For a more interpretable result, we must look to the re-specified subsamples. Results are shown in Table 9. The table shows the coefficient on (or the marginal effect of) the CRP dummy. Results are shown for OLS, logit, and probit regressions, and matching. Covariate balancing tables for matching analyses are available upon request.

Table 9. Land Use Results with Pre-Processed Samples

Specification:	I	II	III	IV
What CRP land included?	Early signup wave (CRP-Eligible)	Early signup wave (CRP-Eligible)	Exit 1992-7 (CRP-Exit)	Exit 1992-7 (CRP-Exit)
What non-CRP (previously farmed) land included?	All land (Control-All)	Not farmed 1992 (Control-Unfarmed)	All land (Control-All)	Not farmed 1992 (Control-Unfarmed)
OLS	-0.775***	-0.185***	-0.316***	0.231***
Logit (marginal effect)	-0.835***	-0.219***	-0.285***	0.275***
Probit (marginal effect)	-0.823***	-0.224***	-0.301***	0.266***
Matching (ATT)	-0.712***	-0.159***	-0.305***	0.207***
N (hundreds of acres)	2,887,997	399,719	2,731,206	242,928

* significant at 10%, ** significant at 5%; *** significant at 1%

All included land was cultivated in 1982, is in a county for which rent data is available, and does not enter 1997 land uses: urban/built-up, water, or federal. Specifications control for 1997 land characteristics (prime farmland indicator, good land class indicator), 1996 land use rents and changes, region code dummies, and land characteristic x rent interactions. Survey regressions performed with data points appropriately weighted and clustered errors. Matching uses Mahalanobis metric plus propensity score.

Each specification (column) in Table 9 represents a combination of a treated (CRP) sample and a control (non-CRP) sample. The results generate consistent estimates within each column, so the results are not driven by functional form. However, the results change greatly between columns. This is expected, since each treatment-control group pair represents a different question addressed to the data.

Specifications I and II in Table 9 show results using the CRP-Eligible sample (CRP parcels that were part of early CRP signup waves) as the treatment group. Regardless of the control (non-CRP) group chosen, the CRP-Eligible parcels are less likely to be farmed than the counterfactuals. Recall from Table 7 that 86.3% of CRP-Eligible stayed in CRP. As argued in the Methods section, the comparison of the CRP-Eligible sample with non-CRP parcels does not address the question of *post*-CRP land use since much of the eligible land stayed in the program.

Specifications I and III in Table 9 show results using Control-All (non-CRP parcels that were farmed in 1982) as the control group. In both cases, again, CRP parcels are less likely to be farmed. This result is again informed by Table 7. Most of the Control-All parcels are continuously farmed. When we compare CRP land to this land, we are asking whether CRP makes a parcel more likely to be farmed as compared to most of the nation's farmland, including the best land. This may not be a causal relationship, since it may be driven by unobservable differences in land quality and transition costs.

Specification IV in Table 9 is the most revealing. Here, CRP-Exit (land that exited the CRP program between 1992 and 1997) is the treatment group. It is compared to Control-Unfarmed—land that, though farmed in 1982, was in some other land use in the intervening years. Control-Unfarmed land has been revealed, by the land use chosen by its landowner, to be similarly low-quality as compared to CRP land and to face similar transition costs to enter farming. Therefore, the Control-Unfarmed land is the best counterfactual for CRP-Exit land for this research question. In these analyses, the CRP-Exit land is *more likely* to be farmed than the non-CRP land. This treatment effect is not just statistically significant and consistent across specifications, but it is also economically significant, at 21-28%.

To restate this result, the CRP seems to make this land 21-28% more likely to be farmed than it would have been had it never been in the CRP. This result is expected if the returns to agriculture have increased on this land, which in turn is expected because CRP improves land quality. (Control-Unfarmed land should have also improved while it was not farmed, but CRP's period of non-farming is probably longer, and conservation practices that cause land improvement are subsidized in CRP.)

Is this result driven by the selective exit of only the best CRP land? If CRP-Exit contained the upper envelope of CRP land, but Control-Unfarmed contained only very bad land, this difference in land quality would reduce the implications of the results. To check, we reduce the CRP and non-CRP land samples to only the best land using a specification with only land indicated as prime farmland, and a specification with only land that has a relatively good (unrestricted) land classification. Results are robust to these checks, with coefficients still significant and slightly increased in magnitude (available upon request). Of course, this check has limited power if observable characteristics understate the differences in land characteristics. It is still likely that the "best" CRP land (by unobservable characteristics) had a higher tendency to leave the program. If this is the case, the magnitude of the effect observed is only accurate for the parcels that actually exited; the effect could be different (most likely smaller) if parcels were ejected from the program randomly or if the program were ended. Even so, the tendency for CRP to cause at least some land to be more likely to be farmed is still compelling and interesting.

As discussed in the Methods section, these results may be biased toward zero if Control-Unfarmed land is unobservably better than CRP-Exit land or if there is a delay in transitioning into farming, and the results may be biased upward if Control-Unfarmed land is unobservably worse than CRP-Exit land or is in a long fallow cycle. This upward bias is less likely because Control-Unfarmed was shown to be (on average) better in observable qualities than CRP-Exit, and because the NRI tries to accurately record fallow land, but it is still possible.

We check the robustness of the matching result using Rosenbaum's recommended sensitivity test (Rosenbaum 2002). This test simulates conditions in which the treatment variable (1992 CRP participation) and the outcome (1997 farming of the land) are both driven by an unobservable factor. This test assumes that the unobservable factor determines with certainty whether land will be farmed in 1997. The correlation between this unobservable factor and the propensity to receive treatment (here, likelihood of being in CRP) is varied, and the test reports the range of unobservable values for which the matching result is still valid. When CRP-Exit land is compared to Control-All and Control-Unfarmed land, the result is robust up to an unobservable factor of 2.7 and 2.6 respectively. That is, even if some parcels are 2.7 (or 2.6) times as likely to enter CRP due to an unobservable factor perfectly correlated with later farming, the matching result showing that CRP-Exit land is more likely to be farmed than Control-Unfarmed (and less likely to be farmed than Control-All) would still indicate a positive causal effect.[13] These results do not tell us that there is such a factor; it simply indicates how strong a factor would have to be to create these results. Since this test assumes that this unobservable factor wholly determines farming outcome, we feel that a robustness of 2.6 is sufficiently strong for these results to be convincing.

Multinomial Land Use Outcome Results

Cultivated cropping is the land use of greatest interest. However, it is also interesting to learn how CRP participation affects the entire distribution of final land uses. Table 10 presents multinomial logit land use outcome results. The specifications

---

[13] When CRP-Eligible parcels are compared to either control group, the result is robust up to an unobservable factor of at least 5.

(columns) again correspond to pairs of treatment and control groups. Land use categories were: cultivated cropland, non-cultivated cropland, pasture, CRP, and other.[14] When CRP-Eligible is compared to either control group (Specifications I and II in Table 10), the results are mainly driven by the tendency of this CRP land, while probably eligible to exit, to stay in the program. CRP-Eligible land is much more likely to be in CRP in 1997, and less likely to be cultivated, than the control group. Also, when compared to the Control-Unfarmed group, CRP-Eligible land is less likely to become non-cultivated cropland or pasture, which again is because most of CRP-Eligible stayed in the CRP (and much of Control-Unfarmed became or remained non-cultivated cropland or pasture).

Table 10. Land Use Multinomial Logit Results

Specification:	I	II	III	IV
What CRP land included?	Early signup wave (CRP-Eligible)	Early signup wave (CRP-Eligible)	Exit 1992-7 (CRP-Exit)	Exit 1992-7 (CRP-Exit)
What non-CRP (previously farmed) land included?	All land (Control-All)	Not farmed 1992 (Control-Unfarmed)	All land (Control-All)[a]	Not farmed 1992 (Control-Unfarmed)[a]
Cultivated cropland	-0.832***	-0.252***	-0.274***	0.270***
Non-cultivated cropland	-0.020***	-0.251***	0.059***	-0.155***
Pasture	0.001	-0.245***	0.126***	-0.095***
CRP	0.848***	0.851***	N/A	N/A
Other	0.004***	-0.104***	0.090***	-0.019***
N (hundreds of acres)	2,887,997	399,719	2,712,792	241,650
F	3,169.99	722.44	213.20	51.66

* significant at 10%, ** significant at 5%; *** significant at 1%

Cells show marginal effects of 1992 CRP status dummy on each land use. All included land was cultivated in 1982, is in a county for which rent data is available, and does not enter 1997 land uses: urban/built-up, water, or federal. Specifications control for 1997 land characteristics (prime farmland indicator, good land class indicator), 1996 land use rents and changes, and region code dummies. Survey regressions performed with data points appropriately weighted and clustered errors.

[a] For these regressions only, both Control groups exclude land that becomes CRP in 1997 (0.67% of Control-All and 0.53% of Control-Unfarmed).

---

[14] "Other" includes rangeland, forest, and "other rural"; these uses make up small but significant elements of CRP-Eligible (2.19%), CRP-Exit (14.63%), Control-All (2.22%), and Control-Unfarmed (17.55%).

In Specifications III and IV of Table 10, in which land that exits CRP is compared to non-CRP land, the results again depend on the set of control parcels used. When CRP-Exit is compared to Control-All (Specification III), we see again that the ex-CRP land is less likely to be farmed, and that the ex-CRP land has a greater tendency to go into pasture, non-cultivated cropland (hay), and other uses. This is unsurprising: as discussed above, most land that exits CRP is planted with grasses as a conservation cover, so the conversion to pasture is trivial, and the conversion to hay or other non-cultivated uses may also be inexpensive. When CRP-Exit is compared to Control-Unfarmed in Specification IV, the ex-CRP land is 27% more likely to become cultivated cropland. This result is similar to the binary model estimates. The CRP-Exit parcels are less likely than Control-Unfarmed parcels to be non-cultivated cropland or pasture in 1997. This is probably because over 80% of Control-Unfarmed land had already been in non-cultivated cropland (47.54%) or pasture (36.25%) in 1992.

Both pasture and non-cultivated cropland convert to cultivated cropland at relatively high rates: 44.8% of 1992 non-cultivated cropland, and 28.2% of 1992 pasture, was farmed in 1997. The high rates of conversion to cropping over a five-year period demonstrated in these data may indicate some error in the NRI's verification of fallow cycles. We previously noted that (the issue of long fallows aside) CRP parcels have less opportunity to transition than do Control-Unfarmed parcels because CRP contracts end in 1996. This under-measurement of transition may be particularly large because CRP-Exit parcels that do not enter farming tend to enter non-cultivated cropland and pasture at a relatively high rate, and may therefore transition into farming in the future. However, again, this would bias CRP's tendency to enter cultivated cropland toward zero.

However, since the control (non-CRP) parcels of greatest interest were uncultivated in 1992 just as the CRP parcels were, transition costs are held constant between the treatment and control groups. Further, while the ex-CRP parcels exited the program in or around 1996, the control parcels were bound by no contract and could make any transition that was profitable at any time after the 1992 observation. This would bias the results to make the ex-CRP parcels look less likely to be farmed, and our result is robust to this direction of bias.

Conservation Practice Results

The 1997 NRI sample contains data on whether any of 22 different conservation practices was adopted on cultivated cropland, summarized in Table 7 for the populations of interest. Of all acreage cultivated in 1997, 22.77% was engaged in some conservation practice. The most popular practices were terraces (6.2% of land), contour farming (5.76%), grassed waterways (4.28%), and surface drainage (4.02%). Most of the conservation practices help to conserve water and/or reduce erosion. In some cases, these practices are strongly recommended or required for land that is very sensitive (e.g., highly erodible). There are government programs at various levels to promote and subsidize practices of this type.

Is ex-CRP land that becomes cultivated cropland more likely than other land to adopt these conservation practices? Table 7 shows that unconditionally, cultivated ex-CRP land appears more likely to adopt conservation practices than either cultivated Control-All or cultivated Control-Unfarmed land. However, this unconditional result is confounded by differences in land quality and other factors.

The conservation practice analysis can be performed in much the same way that the land use outcome analysis was performed. However, since some low-quality land is required to adopt conservation practices, the sample must first be reduced to include only land that is not highly erodible, with an erodibility index of 8 or less (results without the erodibility index restriction are similar, and available on request).

Even after this extra restriction on the samples, causality cannot be attributed to the CRP. While significant results could point to a causal role for the CRP, they could also indicate that the kinds of landowners who sign their land up for CRP are also the kinds of people who will adopt conservation practices (because of personal proclivity, better knowledge of government programs, or variation in local programs that promote conservation). The land use model abstracts from landowner characteristics in a way that is fairly plausible for land use outcomes, since market returns should be strong drivers of land use. However, in the decision to adopt conservation practices that assumption seems dubious for a number of reasons. For example, government programs are not always easy for potential participants to understand, and idiosyncratic landowner information may not make its way into market prices. Even without clear causality, however, the results are still interesting.

Table 11 shows the coefficient of (or marginal effect for) CRP participation on the adoption of some conservation practice in 1997. The columns again represent different combinations of ex-CRP and control (non-CRP) land, with data restricted to non-highly-erodible land that was cultivated in 1997. In Specification I, the control group includes the portion of Control-All (all land farmed in 1982) that was farmed in 1997. For this comparison, the CRP land is not significantly more nor less likely to adopt

a conservation practice. (Most of Control-All was farmed in 1992, but we cannot know which parcels adopted a conservation practice in 1992 because data on conservation practices for 1992 are not available.)

Table 11. Conservation Practice Results

	Specification I	Specification II
What CRP land included?	Exit 1992-7 (CRP-Exit)	Exit 1992-7 (CRP-Exit)
What non-CRP (previously farmed) land included?	All land (Control-All)	Not farmed 1992 (Control-Unfarmed)
OLS	0.018	0.045**
Logit	0.026	0.043**
Probit	0.026	0.043*
Matching	0.003	0.076**

* significant at 10%, ** significant at 5%; *** significant at 1%

Only non-highly-erodible land included

Land was in the categories described in the text but only non-highly erodible (erodibility index of 8 or lower). All included land was cultivated in 1982, is in a county for which rent data is available, and does not enter 1997 land uses: urban/built-up, water, or federal. Specifications control for 1997 land characteristics (prime farmland indicator, good land class indicator, slope, and erodibility index), 1996 land use rents and changes, region code dummies, and land characteristic x rent interactions. Survey regressions performed with data points appropriately weighted and clustered errors. Matching uses Mahalanobis metric plus propensity score.

For Specification II of Table 11, the control group includes the portion of Control-Unfarmed (land farmed in 1982 but not 1992) that was farmed in 1997. In this specification, both CRP and non-CRP parcels faced a transition back into farming. Adopting a new conservation practice for the first time requires some investment of time and money, particularly in planning and design, so transition costs into farming with conservation practices are greater than transition costs into farming without conservation practices. CRP parcels are 4-8% more likely to adopt a conservation practice than this very similar land. This is large relative to the baseline adoption rate of conservation practices of 18.1% for this land.

We perform another Rosenbaum test for sensitivity to unobservable factors, and the conservation practice result is very sensitive to unobservable factors that increase the CRP signup rate by as little as 40%. This leaves open the possibility that these results are driven by unobservable selection.

Conclusion

The Conservation Reserve Program uses selective retirement of agricultural land to protect the environment and to increase long-term agricultural productivity. There is a natural tension between this program's goals. The conflict between the agricultural and environmental goals of the CRP is particularly notable in land's disposition after it leaves the CRP. We seek causal effects of the CRP on later land use, and we find that CRP participation seems to cause land to be 21-28% more likely to be farmed than it would have been had it never been in the program. This result is congruent with findings that the CRP improves land quality, because that improvement would increase the returns to cropping. The CRP may act as a long, subsidized fallow period for some landowners.

The innovation of our analysis is the use of a sample re-specification step. This reduces bias by using the analyst's knowledge to restrict the treatment and control groups to the most appropriate parcels. Unobservable factors determine the potential returns to cropping and other land uses, so estimates of land use model parameters and elasticities may be biased if calculated using incomparable parcels. Comparable parcels can be identified by their use: land that is similarly low-quality and faces similar transition costs (i.e. land with similar unobservable characteristics) may drop out of regular cropping.

We find that treatment effect estimation is very sensitive to the use of inappropriate data samples. Once an appropriate sample is specified, however, results are

robust to different econometric specifications. This sensitivity to data set specification may also be applicable to the calibrations used, for example, in general equilibrium simulations. On a related note, we should be aware that the 21-28% increase in likelihood of farming is an average treatment effect on the treated (ATT) and that it is only applicable to parcels that exited CRP between 1992 and 1997. It would be incorrect to infer that the same rate of increase in cultivation would occur on CRP parcels that did not exit by 1997, or on parcels that may enter CRP in the future. Notably, it would be inappropriate to infer that if the CRP were dissolved, the end use of all ejected parcels would reflect a 21-28% increase in the likelihood of farming. However, the results are still interesting in indicating that *at least some* CRP participants take advantage of CRP-induced land improvements by farming their land more intensively.

Is the increase in cultivation on these ex-CRP parcels socially desirable? The welfare effects of this increased tendency to farm are ambiguous, so this question cannot be answered without use of a social welfare function that places explicit weight on the opposing environmental and agricultural benefits of the CRP. However, some of the environmental damage that could result from this increased tendency to farm is mitigated by farmed ex-CRP land's increased tendency to adopt conservation practices.

Finally, we might infer from these results that farmers are being paid too much to conserve their land since they are getting productivity gains from the improved land, and that government payments are therefore inefficiently high. On the other hand, if the bidding process were incentive-compatible, farmers' CRP bids would be reduced by their expectation of productivity benefits. If this were the case, inefficiency would be reduced.

Chapter III: Learning from Mistakes
What Do Inconsistent Choices Over Risk Tell Us?

Introduction

We all make mistakes. We take longer to complete tasks than we expected, or we make the wrong choice because we are in a hurry. Psychologists have long recognized this problem and regularly include multiple questions in surveys that ask the same thing in order to increase reliability. In experimental economics, the presence of inconsistent choices is common, particularly when experiments are taken to the field. The view in economics has been to either ignore subjects with inconsistent choices or to make specialized assumptions on the nature of mistakes and estimate the parameters of interest. We take the view that mistakes can be inherently interesting and informative with regard to subjects' decision-making rules and cognitive costs. We are particularly interested in how mistakes made in risk decisions can explain financial decisions. We find that measures of error are correlated with demographics and financial instrument choices. This paper shows that there is much to be learned from mistakes.

Several researchers (Birnbaum and Schmidt 2008; Carbone and Hey 2000; Harless and Camerer 1994; Hey 2005; Hey and Orme 1994; Loomes, Moffatt, and Sugden 2002) have shown that assumptions on the source and nature of mistakes are not innocuous in the identification of behavior patterns. Much of this literature also shows, by collecting repeated observations on the same individuals, that the way subjects commit mistakes (or are inconsistent) is heterogeneous.

We use risk experiments that allow inconsistent choices to be made. We define a mistake as a choice that does not reflect an individual's true preferences. We use the results of these experiments to study mistakes and then link those mistakes to financial decisions made in the marketplace. Participants are drawn from a random sample of the adult population in Rwanda, thus providing important variation in demographics and economic outcomes. This setting is well-suited to our study because subjects face risky decisions regularly in their daily lives and have access to a variety of (mainly informal) financial instruments. The sample is typical of adults found in urban and rural areas in many developing countries. The experiment is designed so that with fairly unrestrictive assumptions on utility, choices inconsistent with most reasonable theories can be detected in the data. We find that risk aversion alone is a poor predictor of decisions, but mistakes and risk aversion correlate with the use of certain financial instruments, notably those that serve as safety nets.

Mistakes, or inconsistent choices, have been linked to non-cognitive abilities such as motivation, attention, and patience. Additionally, both cognitive and non-cognitive abilities have been linked to outcomes such as employment, wages, obesity, smoking, and saving decisions (e.g., Heckman, Stixrud, and Urzua 2006; Segal 2009). The use of economic experiments to look at cognitive and non-cognitive ability is more limited. Benjamin, Brown and Shapiro (2006) find that students with high standardized test scores are less likely to exhibit extreme risk aversion in small stakes gambles (which some consider irrational), and Sunde et al. (2010) find people with lower IQ's are more risk averse and impatient. Ashraf, Karlan, and Yin (2006) look at hypothetical survey questions on time preferences and find that those who make time inconsistent choices are

more likely to take up a commitment savings product in the Philippines. In this research, we ask whether inconsistent choices in risk experiments relate to patterns of financial decisions taken in the marketplace. We don't know what pattern of financial decisions is optimal for any person, but if we find that tendency to err is correlated with financial decision-making, this would imply either that these subjects are making errors in their financial lives or that they are hedging to protect themselves against their own errors.

Decision-making inconsistencies have been observed in several experimental studies, including decisions over risk (e.g., Holt and Laury 2002), health (Stockman 2006), and time preferences (Castillo et al. 2009). Prasad and Salmon (2007) find, using a risk experiment similar to ours, that subjects who make inconsistent choices over risk earn less than those who make consistent decisions in a principal-agent experiment.

In previous research, these inconsistencies have been frequently ignored, under the assumption that they are uninformative noise. Alternatively, choices are restricted so that inconsistent behavior cannot be observed. In risk experiments, this can be done by giving the subject only a single decision to make. For example, Binswanger (1980), one of the first to use such a method, gives a menu of lotteries over which the subject chooses one. Subjects can be asked to pick a point at which they switch from risky to safe lotteries (see Harrison and Rutstrom 2008; Tanaka, Camerer, and Nguyen 2010). Andersen et al. (2006) use an iterative procedure to hone in on the subject's switch point, and, consequentially, subjects make significantly fewer unexpected choices. The advantage of these methods is that they provide clear estimates of risk aversion by eliminating the possibility of mistakes.

We are interested in knowing the relationship between mistakes, lottery-measured risk aversion, and financial decisions. Results in the literature relating lottery-measured risk preferences to economic decisions have been mixed. Eckel et al. (2007) find that risk-averse individuals are more likely to take up experimentally-provided education financing. Bellemare and Shearer (2006) find that risk aversion is correlated with job sorting, and Dohmen et al. (2010) find that lottery measures of risk do not relate to risky behaviors but a general risk question does. Harrison, List, and Towe (2007) argue that the reason for these mixed results may be background risk.

Our results are surprising and may offer additional insight into the existing results relating risk aversion to important outcomes. While we find no correlation between risk aversion and outcomes, we do find one when we also control for mistakes. Absent mistakes, as we would expect, risk-averse subjects are more likely to be in a savings group and less likely to take an informal loan. Risk preferences interact with the tendency to make mistakes in significant and sensible ways. As risk aversion increases, those who are more likely to make mistakes become less likely to belong to a savings group and more likely to take out an informal loan. Those who make mistakes seem to be missing out on potentially beneficial opportunities or taking on risks when they should not. Making mistakes in one task may be linked to less than optimal behavior in another. Being able to observe these types of behavioral biases is essential to understanding financial decisions.

This chapter is organized as follows. Section 2 describes financial decisions people make in Rwanda. Section 3 describes the experiment and defines inconsistent choices. Section 4 describes the data and presents the results. Section 5 concludes.

Financial Instruments in Rwanda

The formal financial sector in Rwanda is limited and is consistently rated as one of the worst in the world (World Bank Group Doing Business Project 2007). Because of this, many Rwandans rely on informal channels for credit. We will look at three common types of informal credit (savings groups, insurance groups and informal loans) and formal credit offered through banks and credit unions.

Savings groups (*tontines*) are rotating credit associations that allow members to pool risk, keep precautionary savings, and have access to credit. These groups are common and popular in Rwanda. Members deposit a fixed amount of money at a fixed interval (typically monthly). Once every interval, one member of the group receives all the money deposited by the members. Members may leave the group without penalty once the cycle in which all group members receive the pool of money is complete. Groups vary in size and in interval length (e.g., monthly or every two months). It has been noted (e.g., Besley 1995) that tontines may serve risk-sharing functions for members who have negative shocks. We expect risk-averse individuals to be more likely to join this kind of group.

Insurance groups (*groupes d'entraide*) exist in two general forms. The first is a rotating work group for construction or agricultural work. Members help each other by exchanging labor. The second offers financial assistance in the case of a bad shock like death or illness. It functions as insurance. More generally, the group may offer moral support. Unfortunately, the data we have do not distinguish between these two general forms. These groups sometimes also have a religious component to them because many

are organized by churches. Members typically pay a monthly fee to belong. Since the risk-pooling nature of these groups is not necessarily monetary, we do not expect risk aversion over money to be as strongly correlated with membership.

Informal loans are widespread in Rwanda. They are usually short-term and small, and they are largely used for immediate consumption smoothing. Most are store credit or cash loans from family and friends. These loans are almost always interest free, but they are given with an expectation that the favor will be reciprocated. Default on these loans is risky: it may close doors to future borrowing and generally damage relationships with friends and family. Because of the obligation involved and the high cost of default, we expect risk-averse individuals to be less likely to take out these loans.

Formal credit is not widely accessible because of large financial barriers. To be eligible to apply for a formal loan, an individual must pay an application fee and maintain an account in the bank or credit union. Minimum deposits in banks are often very high relative to Rwandan incomes, and membership in a credit union requires paying a small fee. Credit unions are more accessible than banks for the poor since the membership fee is far lower than the minimum balance at a bank. Many poor people become credit union members precisely to get access to credit. Formal banks and credit unions are relatively stable; however, before 2000, credit unions had a reputation for making loans and not asking that they be paid back. Since 2000, regulation has made credit unions more accountable and stable. In this era, the country has been much more politically stable than in previous decades (after 1994's horrific Rwandan genocide). Trust in the government and in institutions in general is still weak, but there is some trust in the financial sector. Formal loans are primarily used for business and construction, rather

than consumption smoothing. It is not clear how risk aversion will affect the probability of taking out a formal loan because the barriers to entry clearly select a certain segment of the population. This population might be relatively risk averse or risk seeking.

Experiment

The experiment was done in conjunction with a 2002 World Council of Credit Unions survey on the economic activities and household characteristics of a random sample of credit union members and non-members in seven locations across Rwanda. In each location, fifty members and fifty non-members were interviewed, for a total of 700 respondents. Members were randomly selected from lists of active credit union members, and non-members were randomly selected from neighborhoods served by the credit union. Survey respondents were at least 18 years old and were asked questions about household demographics, economic activities of household members, and credit use. Interviews were conducted in Kinyarwanda, the primary Rwandan language, by Rwandan enumerators.[15] The enumerators were trained and tested by the experimenters.

At the end of the survey, each respondent was asked to complete two lottery experiments. One set of lotteries (which we will call the gain lotteries) had only positive earnings, and the other (the gain-loss lotteries) had positive and negative earnings. We chose these payment structures for the following reasons. The gain payment structure is similar to many experiments in the literature (e.g., Binswanger 1980). The gain-loss structure better mimics the type of risky outcomes an individual might face in his day-to-day life. We will see that the latter payment structure better explains economic decisions.

---

[15] For a complete description of the data and survey design, refer to Petrie (2002).

Of the 700 respondents, 15 received pilot treatments that were not designed to generate usable data, and another 62 were unable or unwilling to complete the full lottery experiment, so 623 individuals provided risk preference data. Of those, 442 received a treatment that presented a menu of five lotteries at once and asked the subject to choose one (similar to Binswanger 1980). This five-pair simultaneous presentation treatment, by design, did not allow the subjects to make mistakes. This will serve as a comparison treatment when we look at how risk measures predict financial decisions. The remaining 181 subjects received sequential binary-choice lotteries that did permit mistakes. Eighty-two subjects participated in a treatment with low payoff levels (55 with hypothetical and 27 with real payments) and 99 saw a treatment with high payoffs (all hypothetical).

The 181 participants of the sequential-choice lottery game are similar on observable characteristics to the larger survey population of 700 individuals. All 181 of these subjects lived either in the capital Kigali or in the towns of Gitarama or Butare in the south of Rwanda. The survey population and the sequential-choice lottery participants have similar gender ratios (39.0% female for the sequential-choice lottery participants and 39.4% for the survey population), average ages (36.6 and 37.2, respectively), and average monthly per capita incomes (30,897 RWF and 34,520 RWF, respectively). Like the survey population, 93% of the sequential-choice lottery participants are literate. Compared to the 2002 official Rwandan national census, the survey population is similar on many demographic dimensions. However, the survey population is slightly richer and more literate than the national average in Rwanda. This may be because credit union members, who made up 2% of the Rwandan population, were oversampled.

Experiment Design

In the sequential-choice experiment, subjects face a series of five pairs of lotteries, each with 50-50 odds, and are asked to choose one lottery (A or B) in each pair. The lottery pairs are shown in Table 12; the left half shows the low-payoff lotteries and the right half shows the high-payoff lotteries. The gain and gain-loss lotteries are increasing in expected payoff and variance, and in each pair, lottery B has a lower expected payoff and variance. Also, in each subsequent pair, option B has the same payoffs as option A in the previous pair. If choices are consistent, this lottery exercise is equivalent to presenting subjects six lotteries simultaneously and asking them to choose one.

Table 12. Lottery Treatment Payoffs (in Rwandan Francs, 500 RWF=$1 US*)

Low Stakes (Real and Hypothetical)			High Stakes (Hypothetical only)		
Gain Lotteries			**Gain Lotteries**		
	A	**B**		**A**	**B**
G1	(700, 400)	(500, 500)	**G1**	(1650, 1000)	(1250, 1250)
G2	(900, 300)	(700, 400)	**G2**	(2050, 750)	(1650, 1000)
G3	(1100, 200)	(900, 300)	**G3**	(2450, 500)	(2050, 750)
G4	(1300, 100)	(1100, 200)	**G4**	(2850, 250)	(2450, 500)
G5	(1500, 0)	(1300, 100)	**G5**	(3250, 0)	(2850, 250)
Gain-Loss Lotteries			**Gain-Loss Lotteries**		
	A	**B**		**A**	**B**
L1	(700, -100)	(500, 0)	**L1**	(1650, -200)	(1250, 0)
L2	(900, -200)	(700, -100)	**L2**	(2050, -400)	(1650, -200)
L3	(1100, -300)	(900, -200)	**L3**	(2450, -600)	(2050, -400)
L4	(1300, -400)	(1100, -300)	**L4**	(2850, -800)	(2450, -600)
L5	(1500, -500)	(1300, -400)	**L5**	(3250, -1000)	(2850, -800)

* At the time of this research, median per capita annual income in Rwanda was 118,000 RWF, according to the US Department of State, so 500 RWF was roughly equivalent to a day's wage. From our survey data, median monthly per capita income and expense measures were between 15,000 – 18,000 RWF, and this would imply a daily wage (based on 5 working-days a week) of 691- 830 RWF in our sample.

A risk-neutral or risk-loving individual would always choose A. An individual with a concave utility function would start with option A and switch to lottery B as expected payoffs and variance increase and continue to choose option B. Since the subject's "switching point" may occur above or below the wealth range of the lotteries presented to the subjects, strongly risk-averse subjects may always choose option B, while less risk-averse subjects may always choose option A.

The format of the experiment is similar to Holt and Laury (Holt and Laury 2002), hereafter HL, but with several key differences. HL keep payoffs constant and vary the probabilities of receiving the high and low outcomes. In our experiment, the probability is always 50-50, which is easier to administer and perhaps to understand, and the payoffs are varied. Also, HL present the lotteries all at once to the subjects. In our experiment, lottery pairs are presented sequentially. Finally, in addition to lotteries over gains, we also present lotteries over gains and losses (as in Laury and Holt 2005).

Experiment Implementation

Before they start making decisions, subjects are told that one of the five lotteries in each set will be randomly chosen for implementation by pulling a number between one and five from a hat and then a coin will be flipped to determine payment. After the procedures are explained, the subject is allowed to practice briefly with a sample lottery pair. Then all lottery pairs are presented one at a time and in the same order for all subjects. For example, in the gain lottery, subjects are first presented with the payoffs for G1 and are asked if they would prefer lottery A or lottery B. Next, they are presented the payoffs for G2 and asked to choose between A and B, and so on. They must choose one

of the lotteries in each case. They are not allowed to declare indifference. Once subjects have made decisions for the gain lotteries, they are presented with the gain-loss lotteries one by one and in the same order.

When all lotteries have been completed, one lottery is randomly chosen from the gain sequence and another from the gain-loss sequence. A coin is flipped for the chosen lottery in each sequence. If the coin turns up heads, the subject would earn the first number in the payoff pair for the chosen lottery. For example, if lottery G3 was randomly chosen in the low payoff treatment and the subject chose option A for G3, then if the coin flip turned up tails, the subject would earn 200 RWF.

Subjects in the real-payment treatment were given 500 RWF as a show-up fee and paid the outcomes of the coin tosses for the gain and the gain-loss lotteries. They were paid in cash. Subjects in the hypothetical treatments were not paid. After the two coin flips, they were told what they would have earned had they been paid.[16]

The majority of the subjects received a hypothetical-payment treatment because of limited funding. There is some debate in the literature as to whether real payoffs are necessary to incentivize choices in lottery experiments. Camerer and Hogarth (1999) note, in a review of 74 experiments, that the effect of real payoffs in risk experiments is unclear—in some cases there is no effect, and in some cases subjects appear more or less risk averse. Ortmann and Hertwig (2006), while noting the importance of financial incentives, emphasize the importance of a "do-it-both-ways" rule so that experimenters can compare results of financially motivated and non-motivated treatments. This is what

---

[16] While they were not paid, subjects in the hypothetical treatments evinced an interest in the outcome of the coin toss. That is, they seemed to care about their hypothetical earnings. This suggests they paid attention to their decisions.

we do in this paper. We test for a significant behavioral difference in subject choices

between the real and hypothetical treatments, and find none. Therefore, we pool the data.

Defining Anomalous Choices

Under relatively unrestrictive assumptions on preferences over risk, the pattern of

predicted choices in the sequential-choice experiment is as follows. A risk-neutral or

risk-seeking subject would always choose lottery A. A risk-averse subject may choose

one of six lottery patterns, AAAAA, AAAAB, AAABB, AABBB, ABBBB, or BBBBB,

in order of increasing risk aversion. If a subject ever chooses B (the safe lottery) in one

lottery pair and then subsequently chooses A (the risky lottery), his choice pattern is

inconsistent with most theories. Relative to a given lottery, he has now expressed a

preference for both a safer lottery and a riskier lottery. As shown in Appendix C, such a

pattern of choices is not predicted by Expected Utility Theory, the Dual Theory (which is

even more restrictive, only allowing the patterns AAAAA and BBBBB), or Expected

Value theory (which only allows AAAAA). As the Appendix also shows, with a few

additional (but reasonable) assumptions, Rank Dependent Theory, Prospect Theory, and

Cumulative Prospect Theory predict the same six lottery choice patterns. Thus, we argue

that choice patterns containing a switch from B to A are mistakes in the sense that they

are unlikely to reflect a subject's actual preferences over risk.

If a choice violating these patterns—that is, a choice pattern in which there is a

switch from B to A—represents true preferences and our assumptions are met, there are

three possible explanations. First, he could be truly indifferent between one or both pairs

of lotteries. This would require a flat utility function over wealth, and we consider this

rather unlikely. Second, he could be "practically" indifferent, in the sense that the

difference between his utility from option A is close to his utility from option B—so

close that he cannot detect the difference. This is quite possible, but we would consider

this a mistake because the decision does not reflect true preferences because cognitive

costs have (in a sense) interfered with the expression of those preferences. Third, he

could have a "humpy" indifference function, with an inflection point in the range of these

lotteries. Figure 6 uses expected utility is used to demonstrate an example in which a

sure thing of 500 is preferred to a lottery with equal-odds outcomes of 400 and 700, but a

lottery with equal-odds outcomes of 300 and 900 is also preferred to the 400-700 lottery.

A "humpy" utility function has been proposed (notably in Friedman and Savage 1948),

particularly for the case of consumption commitments (Chetty and Szeidl 2007), but it

seems unlikely that there should be a hump in just the right place to explain switching-

back behavior in our lotteries.
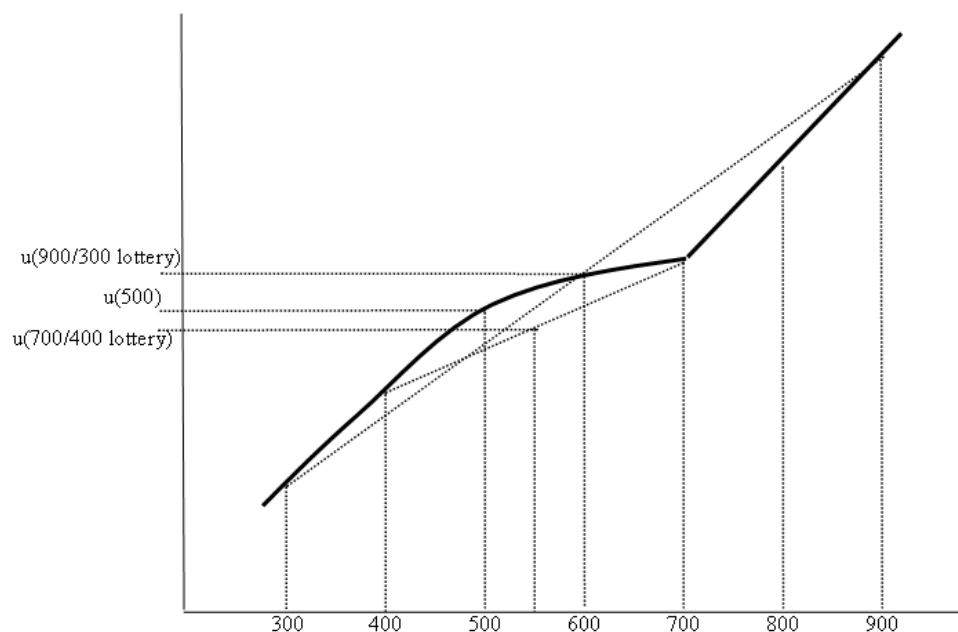
Figure 6. Preferences Over Wealth to Explain Inconsistency

There are three aspects of the experiment that could cause some unexpected behavior, but neither should cause inconsistent "switching-back." First, since two decisions will be implemented, subjects could make their decisions as if putting together a risk portfolio. This should make subjects choose riskier options, but consistently. Second, since one gain and one gain-loss lottery is implemented at random, errors may occur due to faulty compounding of lotteries. However, that should simply make people behave in a consistent, but more risk-averse, fashion (Holt 1986). Third, there may be order effects since the gain-loss lotteries are always performed after the gain lotteries. This is not a problem for this analysis because we do not try to compare choices over gain to choices over gain-loss.

As discussed briefly above, indifference between lottery pairs could generate some switches from safe to risky lotteries (switches that we consider inconsistent). For example, if a person is truly indifferent between a pair of lotteries, he may choose more or less randomly between the two. However, it is not clear how to interpret widespread indifference from the point of view of theory. If a person's preferences are such that lottery B, with a low expected value and a low variance, and lottery A, with a higher expected value and a higher preference, are exactly (or nearly) indifferent, this means that the increase in expected value gained with A perfectly offsets the disutility imposed by the increased variance. While this may explain a small number of inconsistent switches, it is unreasonable to expect that this would be the case in a large number of such switches as we see in this experiment, especially since some subjects exhibit "double switch-backs" (switch B to A twice). Alternately, a utility function that does not vary with wealth will generate indifference across all lotteries, which may cause people to choose

randomly and thus apparently make inconsistent switches. This also seems unlikely. It may be more realistic to interpret "indifference" as an inability to know one's own preference or an error in evaluating the lotteries; a person has real preferences between each lottery pair, but fails in some way to access those preferences (perhaps because of cognitive costs) and thus falls back on random choice or a rule of thumb for some lottery pairs. This kind of indifference can also be considered a mistake.

<div align="center">*Results*</div>

We look first at inconsistent choices. Then, we examine measurements of risk aversion and estimations of mistakes, and finally, we relate risk aversion measures and inconsistent choice to financial decisions.

Inconsistent Choices

Of the 181 people who completed the sequential choice lottery treatment, roughly 55% made at least one inconsistent switch over gains or losses. This is a similar percentage to that found with the same instrument in a random sample of adults in Peru.[17] Table 13 illustrates the distribution of lottery choices across predicted patterns and inconsistent switches. Because the real and hypothetical distributions for low payoffs are not significantly different, we pool these two treatments.[18] We can conclude that either

---

[17] Peru results are from a work in progress. The hypothetical choices from the Peru instrument (which allowed indifference, allowed subjects to review and change decisions, and included both hypothetical and paid lotteries) found 52% of choices over gains and 44% of choices over gain-loss were inconsistent. In other research, inconsistent choices made up varying percentages of choices: Holt and Laury (2002) found 13% in hypothetical choices with students, Stockman (2006) found 11% in hypothetical choices with adults, Castillo, Ferraro, Jordan and Petrie (2009) found 33% in paid choices with 13-year olds, and Prasad and Salmon (2007) found 30% in paid choices with students (presenting lotteries sequentially as we did).

[18] A Fisher's exact test between real and hypothetical stakes for equal distributions over gains has a *p*-value=0.431 and over gain-loss, *p*-value=0.439.

mistakes are not caused by lack of financial incentives, or the real lotteries are not
sufficiently incentivized. Since the pay from low stakes lotteries is on the order of one to
two days pay for our subjects, we believe the latter to be the case. However, mistakes are
the topic of this study, so our results are still interesting even if mistakes are due to a
failure of saliency (note that the difference in expected value of two lotteries in each pair
is actually rather low). There is, however, a significant difference between high and low
payoff treatments. High stakes may be easier to focus on or may be a cognitively easier
task. HL also found that higher-stakes lotteries resulted in a lower tendency to choose
inconsistently (5.5%).

Table 13. Distribution of Choices over Sequential-Choice Lottery Treatments (in percent)

| | Gain Lotteries | | | Gain-Loss Lotteries | | |
Choice Pattern	All	Low Payoffs*	High Payoffs	All	Low Payoffs*	High Payoffs
AAAAA	12.7	9.8	15.2	8.8	12.2	6.1
AAAAB	11.1	2.4	18.2	2.8	3.7	2.0
AAABB	5.5	4.9	6.1	0.6	0.0	1.0
AABBB	3.3	1.2	5.1	1.7	1.2	2.0
ABBBB	6.1	3.7	8.1	4.4	3.7	5.1
BBBBB	7.7	4.9	10.1	27.1	8.5	42.4
One anomalous switch	47.0	58.5	37.4	42.0	48.8	36.4
Two anomalous switches	6.6	14.6	0.0	12.7	22.0	5.1
Total	181	82	99	181	82	99

* Real and hypothetical payoff treatments are combined.

There are other possible categorizations that could capture people who are
consistent with some other decision mechanism. For example, a subject may use a "rule
of thumb" wherein he chooses all A's or all B's but can deviate once (to test the waters).
Using this categorization only helped explain 30 more choices (16.6%) over gains and 37

(20.4%) more choices over losses, leaving 67 (37.0%) choices over gains and 62 (34.3%) over losses still classified as inconsistent. No simple "rule of thumb" categorization explained a larger number of inconsistent choices, so all choices thus categorized remain identified as inconsistent.

How do subjects' choices change between gain and loss lotteries? We categorize the choice pattern AAAAA as "consistent risk loving," the patterns AAAAB, AAABB, AABBB, and ABBBB as "consistent risk averse," and the pattern BBBBB as "consistent strong risk averse." The shaded region of Table 14 shows the pattern of consistent and inconsistent choices over gains and losses. Of those choosing consistently, most are strongly risk-averse in the loss lottery. Conditioning on having made a consistent choice over gain and loss, 10% were more risk-averse over gain than over loss, 38% were equally risk-averse over gain and loss, and 52% were more risk-averse over loss than over gain. This means that of the subjects that made consistent choices, a little over half made choices consistent with a concave utility function over gains and losses. In terms of inconsistencies, roughly half of the subjects made the same number of inconsistent choices over gain as over losses. More people made two inconsistent switches over losses than over gains.

Table 14. Choices over Gain and Loss (Numbers of Subjects)

		Choices over loss					
		Consistent strong risk-averse	Consistent risk-averse	Consistent risk-loving	Inconsistent one-switch	Inconsistent two-switch	Total
Choices over gain	Consistent strong risk-averse	9	1	0	3	1	14
	Consistent risk-averse	14	6	4	21	2	47
	Consistent risk-loving	9	3	4	5	2	23
	Inconsistent one-switch	16	6	8	40	15	85
	Inconsistent two-switch	1	1	0	7	3	12
	Total	49	17	16	76	23	181

Risk Aversion

The overwhelming presence of inconsistent choices makes a risk measure based on a switching point from risky to safe lotteries in our data problematic, since many subjects have multiple switching points. It would be possible to use a structural model that allows for error to determine a risk aversion parameter and an error parameter either for the population or for subpopulations (for example, using a finite mixture model). However, we wish to obtain a measure of error and of risk aversion for each subject and to do so with the fewest possible assumptions on the form of preferences. As a result, we will generate a simple proxy index of risk aversion and (as discussed in the next section) estimate a probability of having made a mistake for each subject.

Our proxy for risk aversion (as in HL) is simply the number of B (safe) choices the person made. This gives us a risk-aversion measure over gain and one over gain-loss

for each subject. The higher this measure, the more risk-averse the person is. So that the risk aversion measure for subjects who participated in sequential-choice lotteries can be compared to that of subjects who participated in five-pair simultaneous presentation lotteries, both were normalized to fall in the range from 0 to 1.[19] This allows us to see if the two methods produce different measures of risk aversion.

The treatment that forces consistency (five-pair) yields risk measures that are significantly higher than in the treatment that allows for inconsistent choices (sequential choice). To compare similar payoffs, we restrict our discussion to the comparison of the five-pair treatment's results with those from the low payoff sequential treatment. Table 15 shows that risk measures are 0.12 points higher over gains, and 0.10 points higher over gain-loss. These differences are significant (t-test for difference in means: over gains p-value = 0.000, over gain-loss p-value = 0.076).

Table 15. Mean Risk Aversion Measures by Lottery Treatment

	Sequential-Choice Low Payoff*	Sequential-Choice High Payoff	Simultaneous Five-Pair
Risk aversion over gain	0.44	0.43	0.56
	(0.25)	(0.32)	(0.37)
Risk aversion over gain-loss	0.43	0.72	0.53
	(0.28)	(0.32)	(0.37)
N	82	99	442

* Real and hypothetical payoff treatments are combined.

Standard deviations in parentheses

[19] For the simultaneous five-pair lottery, subjects are shown five lotteries with a 50-50 chance of either payoff and asked to choose one. The five lotteries over gains are, (500,500), (800,400), (1100,300), (1400, 200), (1700,100). The lotteries over gain-loss are, (0,0), (300,-100), (600,-200), (900,-300), (1200,-400).

In the sequential-choice high payoff treatment, people are much more risk averse over losses (0.72) than any other treatments. Holt and Laury (2002) also saw an increase in risk aversion over high stakes.

Error Choice Model

While we can count the number of safe choices to derive a risk measure in the sequential choice lottery, this assumes that those who chose a predicted consistent pattern did not make a mistake. In a sense, choices we identified as inconsistent are only the most egregious mistakes; in addition to these errors, it is possible for a subject to mistakenly choose a consistent lottery choice pattern that is not best for him. To address this, and because we have few observations per individual, we use the error choice model proposed by Harless and Camerer (1994) to estimate the probability each individual made one or more mistake (given his lottery choice pattern) and the distribution of preferences over predicted patterns.

The intuition behind this approach is that we assume that every individual has a latent preferred pattern that reflects well-behaved preferences, and that some people make mistakes in one or more lottery decisions. We assume in each lottery decision, a subject chooses his preferred option with probability $(1-\varepsilon)$—that is, he errs with probability ε. If a subject actually prefers pattern AAAAA, for example, the probability he chooses that pattern is the joint probability that he will make no mistakes over the five choices: $(1-\varepsilon)^5$. If he makes exactly one mistake, he will instead choose BAAAA, ABAAA, AABAA, AAABA, or AAAAB, and the probability that he makes any of these choices is $\varepsilon\times(1-\varepsilon)^4$.

We assume that preferences are distributed through the population with proportions P(AAAAA), P(AAAAB), P(AAABB), P(AABBB), P(ABBBB), and P(BBBBB).

Say that a subject chooses a pattern BAAAA. That could reflect true preferences of AAAAA with one error, true preferences of AAAAB with two errors, true preferences of AAABB with three errors, true preferences of AABBB with four errors, true preferences of ABBBB with five errors, or true preferences of BBBBB with four errors. The likelihood of observing pattern BAAAA is the sum of the likelihoods of each of these possibilities: $P(AAAAA) \times \varepsilon \times (1-\varepsilon)^4 + P(AAAAB) \times \varepsilon^2 \times (1-\varepsilon)^3 + P(AAABB) \times \varepsilon^3 \times (1-\varepsilon)^2 + P(AABBB) \times \varepsilon^4 \times (1-\varepsilon) + P(ABBBB) \times \varepsilon^5 + P(BBBBB) \times \varepsilon^4 \times (1-\varepsilon)$.

Given individual likelihoods calculated as in this example, we maximize, with regard to ε and the population preference proportions (e.g., P(AAAAA)), the joint likelihood of observing the patterns in our data. The estimated ε and proportions allow us to estimate the probability that any pattern reflects at least one mistake. For example, a subject who chooses AAAAA may be making no mistake (may truly prefer AAAAA) or may be making one or more mistake. The probability AAAAA was chosen is the $P(AAAAA) \times (1-\varepsilon)^5 + P(AAAAB) \times \varepsilon \times (1-\varepsilon)^4 + P(AAABB) \times \varepsilon^2 \times (1-\varepsilon)^3 + P(AABBB) \times \varepsilon^3 \times (1-\varepsilon)^2 + P(ABBBB) \times \varepsilon^4 \times (1-\varepsilon) + P(BBBBB) \times \varepsilon^5$. Only the first term reflects the expression of true preferences; therefore, the likelihood that AAAAA reflects a mistake is reflected in all of the other terms—or, alternatively, one minus the probability of no mistakes. In this way, we calculate the likelihood that each choice pattern reflects at least one mistake (the likelihood that it does not reflect his true preferences). Thus, for each subject we have the probability he made a mistake over each set of lotteries (gain or gain-loss) based on his lottery choices and the parameters estimated by maximum likelihood.

Any subject who makes an egregiously inconsistent choice (i.e. a switch from B to A) has, using this method, a mistake probability of 1. If we simply used an indicator for these egregious mistakes, we would undercount mistakes by assigning all allowed patterns a mistake probability of 0 when some of those probably do not reflect the true preferences of the subject.

We estimated the distribution of true preferences given the observed choice patterns for all lotteries pooled together, for low-stakes only, and for high-stakes only. The results are shown in Table 16. The error rate, ε, is 0.222 and 0.217 for the pooled gain and gain-loss lotteries, and as expected, the error rate is higher at low stakes. The error rate estimates compare well to estimates for several experiments examined in Harless and Camerer (1994), where error rate estimates were 0.209 - 0.339 for Expected Utility Theory. Our gain-loss lotteries have a much higher proportion of people who choose all safe lotteries (BBBBB), especially for high payoffs.

Table 16. Estimated Population Preferences for Lottery Choice Patterns

Choice Pattern	All		Low Payoffs		High Payoffs	
	Gain	Gain-Loss	Gain	Gain-Loss	Gain	Gain-Loss
AAAAA	0.31	0.37	0.39	0.67	0.25	0.19
AAAAB	0.31	0.00	0.36	0.00	0.30	0.02
AAABB	0.02	0.00	0.00	0.00	0.03	0.00
AABBB	0.02	0.00	0.00	0.00	0.06	0.00
ABBBB	0.22	0.04	0.11	0.07	0.22	0.06
BBBBB	0.13	0.59	0.14	0.26	0.14	0.73
ε	0.22	0.22	0.34	0.31	0.13	0.15
Likelihood function	-3.18	-3.09	-3.39	-3.34	-2.85	-2.62

Inconsistent Choice and Financial Decisions

First, we look at who makes mistakes. Then, we examine how risk aversion and mistakes affect financial decisions.

Table 17 shows summary statistics for the entire sample and for those who completed the sequential choice lottery. The variables included are the percent of subjects who are female, the percent married, the subject's age in years, the number of years of education the subject has completed, the subject's household per capita monthly expenses in Rwandan francs, the number of children under 18 years old living in the subject's household, the number of adults 50 years and older in the household, the percent of subjects in a savings group, the percent in an insurance group, the percent having take an informal loan in the past year, and the percent having taken a formal loan in the past five years. We can see that the entire sample and those who completed the sequential choice lottery are similar on most variables, with the exception that those in the sequential choice lottery are not as rich and are more likely to have used credit.

Table 17. Summary Statistics

Variable	All Lottery Subjects		Sequential-Choice Subjects	
	Mean	Std Dev	Mean	Std Dev
Female	0.39	0.49	0.37	0.48
Married	0.64	0.48	0.65	0.48
Age (years)	36.58	11.23	38.52	10.94
Education (years)[a]	8.84	3.78	8.41	3.95
Per Capita monthly expenditures (1,000 Rwandan francs)[b]	30.62	61.97	21.15	43.68
Number of children (age < 18) in household[c]	2.99	2.15	3.02	1.95
Number of elderly (age ≥ 50) in household[c]	0.30	0.60	0.35	0.64
Member of savings group (tontine)	0.18	0.39	0.22	0.41
Member of insurance group	0.20	0.40	0.20	0.40
Have used formal credit	0.24	0.43	0.36	0.48
Have used informal credit	0.48	0.50	0.55	0.50
Number of Observations	623		181	

[a] There are 14 missing observations on education, so this is based on n=609 for the whole population, n=178 for the sequential-choice subjects.

[b] Household per-capita monthly expenses in Rwandan Francs divided by 1000, range was 0 to 1007.

[c] Ages were missing for some household members for some households, so these numbers are based on n=619 for the whole population, n=180 for the sequential-choice subjects.

Who tends to make mistakes? We regress the estimated probability of having made at least one mistake (from the error rate model) in the gain and gain-loss lotteries on demographic variables, as shown in Table 18.[20] Because we found no difference in choices between real and hypothetical treatments for low stakes, those treatments are pooled together. (If a dummy variable for the real payment treatment is included, it is not significant, and the results do not change.) As expected, subjects in the low-stakes treatment showed a greater tendency to make mistakes. The only other significant result is that over gain-loss women are more likely than men to make mistakes.

---

[20] A Tobit regression yielded the same results. OLS results are reported for ease of interpretation. Results are also generally similar if the dependent variable is an indicator of "switching back" rather than the probability a mistake was made.

Table 18. Probability of At Least One Inconsistent Choice (OLS Regression)

	Gain	Gain-Loss
Female	0.049	0.164**
	(0.063)	(0.071)
Married	-0.027	0.063
	(0.062)	(0.078)
Age	-0.000	-0.006
	(0.003)	(0.004)
Education (years)[a]	0.009	0.002
	(0.007)	(0.008)
Monthly expenses	-0.001	-0.001
	(0.001)	(0.001)
Number of children (age <18)	0.001	0.010
in household	(0.015)	(0.018)
Number of elderly (age $\geq$ 50)	-0.049	0.021
in household	(0.044)	(0.057)
Low stakes treatment	0.397***	0.305***
	(0.054)	(0.073)
Constant	0.442***	0.636***
	(0.140)	(0.182)
R^2	0.261	0.184
N	177	177

Standard errors reported in parentheses. *p-value<0.10, **p-value<0.05, ***p-value<0.01. Robust standard errors are used. All regressions include village-level fixed effects.

[a] There were missing values on education for three observations, and missing household ages for another, so those observations are dropped.

Next, we look at how risk aversion and tendency to make a mistake relate to financial decisions. Before considering mistakes, we want to see if our proxy risk aversion measure alone can explain decisions. Table 19 shows how membership in a savings group or insurance group and having taken either a formal or informal credit is affected by risk aversion, controlling for demographic covariates.[21] For completeness, we look at both the sequential choice lottery and the five-pair lottery (which does not allow inconsistent choice). To make these two sets of regressions comparable, we restrict the sample to the two survey locations where both types of lotteries were administered. (Assignment of lottery type to a subject within these sites was random.)

---

[21] Results are robust to alternative specifications, such as Logit and Dprobit.

Table 19. Financial Decisions without Inconsistency Measure (OLS Regressions)

	Five-Pair Lotteries							
	Savings Group		Insurance Group		Formal Credit		Informal Credit	
Risk Aversion (gain)	-0.025 (0.114)		-0.111 (0.132)		-0.067 (0.144)		-0.286* (0.164)	
Risk Aversion (gain-loss)		0.111 (0.101)		-0.071 (0.123)		0.193 (0.141)		-0.118 (0.177)
R^2	0.033	0.046	0.058	0.053	0.303	0.319	0.088	0.058
N	82	82	82	82	82	82	82	82
	Sequential-Choice Lotteries							
	Savings Group		Insurance Group		Formal Credit		Informal Credit	
Risk Aversion (gain)	0.281 (0.191)		0.189 (0.182)		-0.347 (0.216)		0.051 (0.189)	
Risk Aversion (gain-loss)		0.111 (0.154)		-0.056 (0.163)		0.260 (0.167)		-0.211 (0.184)
R^2	0.180	0.157	0.191	0.179	0.193	0.182	0.093	0.106
N	89	89	89	89	89	89	89	89

Standard errors reported in parentheses. *p-value<0.10, **p-value<0.05, ***p-value<0.01. Robust standard errors are used. All regressions include the control variables: gender, married, age, education, monthly per-capita expenses, number of children (<18 yrs), adults age 50 and older, and village-level fixed effects. There were missing values on education or on the ages of household members for 5 observations (3 from the Sequential-Choice Lotteries and 2 from the Five-Pair Lotteries), so those observations are dropped.

As the table shows, risk aversion is not significantly related to most financial instrument variables. The only weakly significant relationship is between risk aversion over gains in the five-pair lotteries and taking out an informal loan. Those who are more risk averse over gains are less likely to use an informal loan. Gender also does not seem to have any explanatory power.

Since risk aversion alone does not strongly correlate with using financial instrument use, we ask whether the tendency to make mistakes does. A person who makes inconsistent choices over risk may fail to take optimal advantage of the financial instruments available to him, or he may be less trusted by informal lenders and savings and insurance groups. On the other hand, if he is self-aware of his mistakes, he may

insure against this tendency and be more likely to select into these instruments. This is the argument made by Ashraf, Karlan and Yin (2006).

Table 20 shows our results of OLS regressions with bootstrapped standard errors.[22, 23] Being a member of a savings group or taking out an informal loan is significantly correlated with both risk aversion over gain-loss lotteries and tendency to make mistakes, but in opposite directions. As described earlier, savings groups serve as a risk pooling device, so we would expect risk-averse individuals to be more likely to be a member. They are. Also, borrowing money from informal sources (usually family and friends) involves the risk of losing face should one fail to repay the loan. These loans do not carry monetary interest, but they carry an expectation of reciprocation and default in this context can be socially costly. Therefore, we would expect risk-averse individuals to be less likely to take out an informal loan. They are.

---

[22] Because we use the estimated probability of making at least one mistake as an independent variable, we have a generated regressor. We therefore do not use OLS standard errors. Instead, the estimates were done by sampling with replacement 10,000 times to generate estimates and bootstrapped standard errors. The estimates are significant if they fall within the confidence interval specified in the table.

[23] Results are somewhat similar if an indicator of egregious mistakes (B-to-A switches) is used instead of the estimated mistake probability, but the results for informal loans, though reflecting the same sign, are no longer significant. This is not surprising, since the egregious mistake indicator understates true errors.

Table 20. Financial Decisions with Inconsistency Measures (OLS Regressions with Bootstrapped Errors)

Gain	Savings Group	Insurance Group	Formal Credit	Informal Credit
Risk Aversion	0.664*	-0.340	-0.583	0.150
	(0.421)	(0.411)	(0.437)	(0.478)
Est Prob of Mistake	0.465	-0.305	-0.108	0.127
	(0.283)	(0.273)	(0.312)	(0.336)
Risk aversion * Est Prob of Mistake	-0.820	1.014	0.443	-0.215
	(0.580)	(0.599)	(0.631)	(0.749)
Gain-Loss	**Savings Group**	**Insurance Group**	**Formal Credit**	**Informal Credit**
Risk Aversion	0.547***	0.096	0.290	-0.611***
	(0.252)	(0.247)	(0.271)	(0.218)
Est Prob of Mistake	0.515***	0.199	0.081	-0.606***
	(0.239)	(0.237)	(0.270)	(0.219)
Risk aversion * Est Prob of Mistake	-1.185***	-0.366	0.020	0.786***
	(0.393)	(0.361)	(0.443)	(0.427)
N	89 89	89 89	89 89	89 89

Standard errors reported in parentheses. The estimates are bootstrapped 10,000 times. *coefficient falls between the [0.05, 0.95] percentiles of the bootstrapped parameter distribution, ** coefficient falls between the [0.025, 0.975] percentiles of the bootstrapped parameter distribution, *** coefficient falls between the [0.01, 0.99] percentiles of the bootstrapped parameter distribution. All regressions include the following control variables: gender, married, age, education, monthly per-capita expenses, number of children (<18 yrs), adults age 50 and older, and village-level fixed effects. There were missing values on education or household member ages for 3 observations, so those observations are dropped.

More importantly, risk aversion and tendency to make mistakes interact in meaningful and significant ways. If we control only for tendency to make mistakes or risk and mistakes (not interacted), there is no significant effect. That is, mistakes alone, risk alone, and risk and mistakes do not correlate with financial decisions. It is the interaction of the two that is important. Conditional on being more likely to make mistakes, a risk-averse individual is less likely to be in a savings group and more likely to take out an informal loan. That is, those who make mistakes are less likely to choose as

we would expect. These results suggest that individuals who make mistakes are not self aware, but rather do not choose optimally or are excluded from groups or informal loans.

It is possible that the causation between our risk and mistake measures and the use of financial instruments runs in the opposite direction. People who are in savings groups may know that they have a safety net and therefore feel comfortable acting in a riskier fashion; similarly, people who know they have outstanding obligations in their social network due to informal credit usage may feel less comfortable acting in a risky fashion. These correlations are opposite from what our results show. However, it is conceivable that people who are in a savings group are less careful in the lab because they have a safety net (they make more errors) or that people who use informal credit are more careful because of their outstanding obligations (they make fewer errors). This is not precisely an endogeneity problem, but simply a different possible explanation for the result we observe.

Not all financial decisions are significantly correlated with risk aversion or mistakes. Being a member of an insurance group and taking out a formal loan are not correlated with risk aversion, tendency to make mistakes, or the interaction. This may be because insurance groups offer a variety of services, such as insurance, labor, and credit, so the correlation with risk and mistakes is obscured. Formal loans from banks may be given under a specific set of criteria that are unrelated (or only weakly related) to risk aversion and the tendency to make mistakes.

Conclusions

Can we learn anything from mistakes? This research suggests that we can. Using a lottery experiment designed to detect inconsistent choice in risk, we examine the correlation between risk aversion, mistakes, and financial activities that may serve as safety nets. We find that the tendency to make inconsistent choices over risk is widespread and women are more susceptible than men. When mistakes are ignored, risk aversion alone does not explain the use of financial instruments. When we control for the tendency to make inconsistent choices, however, risk and mistakes explain participation in savings groups and the use of informal credit. As we would expect, risk-averse individuals are more likely to belong to a savings group and less likely to take out an informal loan. Those who are more likely to make mistakes display the opposite behavior. Mistakes in the experiment are correlated with apparently suboptimal behavior in these financial decisions.

The policy advice implied by these results are ambiguous, because it's not clear whether financial choices are correlated with lab decisions because the financial choices are mistakes, because people are hedging against their own tendency to err, or because mistake-prone people are well-known in their communities and are therefore not offered credit. Financial literacy education that encourages proper usage of financial instruments may be fruitful in any of these cases. An interesting follow-up to this study would be to experimentally provide financial literacy education and see whether changes are observed later in either lottery choices or financial instrument take-up.

This paper makes two contributions to the literature. First, our results may help explain the inconclusive results that relate lottery-measured risk aversion to economic

outcomes. Like other researchers, we do not find a significant correlation between risk aversion and outcomes. While some have suggested that the lack of correlation may be caused by a failure to properly control for background risk (Harrison, List, and Towe 2007), our results suggest that it may also be due to not controlling for mistakes. Second, our results show that tests for consistent behavior in lab experiments may yield results that help explain choices in non-experimental settings. Previous work has mostly ignored inconsistent choices or designed experiments to not permit inconsistent behavior. This research suggests that important behavioral information may be embedded in mistakes. Interactions of mistakes and preferences clearly have power to predict behavior. Therefore, lab experiments that intentionally reveal mistakes may be important tools for understanding important economic behaviors in credit markets, and may be useful in a variety of other settings as well.

Appendix A: Instructions for the Girl Scout Cookie Experiment

This Appendix contains the instructions provided to subjects in the Girl Scout Cookie experiment. There are four components to the instructions: the General Instructions, which are read at the start of the experiment; and then the specific instructions for each of three treatments. Each treatment's instructions starts with the text, "Instructions for the [next] set of rounds." In this Appendix, the name of the treatment is noted on the same line of text, although the treatment name was not included in the instructions the subjects saw. Otherwise, except for details of pagination, these instructions are identical to those provided to the students.

General Instructions

Welcome to the experiment! This is a study of decision-making behavior. It will last about two hours.

Complete Privacy

This experiment is structured so that no one, including the experimenters and the other participants, will ever know the decisions or earnings of anyone in the experiment. You will collect your earnings, in a sealed envelope, from a numbered mailbox that only you will have the key for. Your privacy is guaranteed because neither your name nor your student ID number will appear on any form that records your decisions or your earnings. The only identifying mark that will be used is the identification number on the mailbox key that you will select. Each key opens a mailbox in the hallway adjacent to this room. After the experiment, you will each collect your payment from your mailbox alone and privately. The key and mailbox are labeled with the same identification number. You are the only person who will know your identification number.

Random Group Assignments and Anonymity

Each person will be randomly matched with 4 other people to form a 5-person group. No one will learn the identity of the people in his/her group. The composition of your group will change after a number of rounds, and you will be randomly matched in a new group of 5 people (you and 4 others). You will be notified when the groups have been rearranged. Once you are matched with a group, you will learn how many rounds you will play with this group. In total, you will play in 3 different 5-person groups.

Your Payment

In each round of the experiment, you will make a decision. What you earn in each round will depend on decisions that you and the other 4 people in your group make. Your payment for this experiment will be the sum of your earnings in all rounds.

The Task

For each decision, you will choose how to divide 20 tokens between two funds: the Personal Fund and the Group Fund. You are free to contribute some of your tokens to the Personal Fund and some to the Group Fund. Alternatively, you can contribute all of them to the Personal Fund, or all of them to the Group Fund. You will earn money from the tokens you put in each fund (below we describe how). Your total earnings for each round is the sum of your earnings from the Personal Fund plus your earnings from the Group Fund.

1. *The Personal Fund*: each token you contribute to the Personal Fund will earn you $0.02.

Example: Suppose you contribute 0 tokens to the Personal Fund. Then you would earn nothing from the Personal Fund.

Example: Suppose you contribute 10 tokens to the Personal Fund. Then you would earn $0.20 from the Personal Fund.

Example: Suppose you contribute 20 tokens to the Personal Fund. Then you would earn $0.40 from the Personal Fund.

2. *The Group Fund*: What you earn from the Group Fund will depend on the **total number of tokens** that you and the other 4 members of your group contribute to the Group Fund. Every token any member of your group puts in the Group Fund earns money for each member of the group, but not everyone will earn the same amount of money.

Before we explain more about the earnings from the Group Fund, you must learn about the **Stakeholder**. In each round, one member of your group will be the **Stakeholder**. This person earns more money (relative to the rest of the group) from all of the tokens in the Group Fund. You will always know whether you are the Stakeholder. If you **are not the Stakeholder**, you earn $0.01 for each token any member of your group puts in the Group Fund. If you **are the Stakeholder**, you earn $0.03 for every token any member of your group puts in the Group Fund.

Example: Suppose you contribute 0 tokens to the Group Fund and no one else contributes any tokens to the Group Fund. Then you would earn nothing from the Group Fund. Everyone else in your group would also earn nothing from the Group Fund.

Example: Suppose you contribute 13 tokens to the Group Fund and no one else contributes any tokens to the Group Fund. Suppose that you are the Stakeholder. Then you would earn $0.39 from the Group Fund. Everyone else in your group would earn $0.13 each from the Group Fund.

Example: Suppose you contribute 20 tokens to the Group Fund and no one else contributes any tokens to the Group Fund. Suppose that you are not the Stakeholder. Then you would earn $0.20 from the Group Fund. The Stakeholder would earn $0.60 from the Group Fund. All of the other members of your group would earn $0.20 each from the Group Fund.

Everyone will earn money from every token contributed to the Group Fund, whether they contribute tokens to the Group Fund or not.

Decision Panel

For each decision, you will enter your Group Fund contribution in the Decision Panel. The Decision Panel is a box at the bottom of the screen that looks like the following:

Type the amount you want to put in the **Group Fund** in the box. This must be a number between 0 and 20. The amount that will go into your Personal Fund is 20 minus the amount you type. For example, if you type the number 0, 0 tokens will go into the Group Fund and 20 into the Personal Fund. If you type 6, 6 tokens will go into the Group Fund and 14 into the Personal Fund. If you type 13, 13 tokens will go into the Group Fund and 7 into the Personal Fund.

When you are satisfied with the allocation of tokens you entered, you will click the "SUBMIT" button to submit your allocation.

Contributions Table

The Contributions Table gives you information about what has happened in past rounds and what will happen in future rounds. The table will show all of the rounds you will play with your current group, and will indicate when you will be the Stakeholder. Also, after everyone has made their decisions for a round, the table will tell you what the Group Fund contributions were in that round, and what your earnings were.

The table will change slightly in different parts of the experiment. Before each part starts, you will receive specific instructions and you will see what the table will look like.

Each Round Has Two Steps

In each Round, you will proceed through two screens, one after the other. The first screen is the decision screen. In the decision screen, you will make your decision and click the "SUBMIT" button when you are done. After everyone has clicked "SUBMIT", everyone will proceed to the review screen.

The review screen will look very much like the decision screen, but you will have no decision to make. You will see information about the decisions that were made in that round and previous rounds, and you will learn how much money you earned in that round. In the Decision Panel of the review screen, you will see exactly how your earnings for that round were calculated. When you are done reviewing this information, you must click "DONE" to continue.

DECISION PANEL

REVIEW RESULTS FROM ROUND 2 IN THE TABLE ABOVE.

YOUR EARNINGS WERE:

Your Personal Fund contribution (5) times $0.02 per token   = $0.10

PLUS: The total number of tokens in the Group Fund (55) times $0.01 per token   = $0.55

EQUALS: **TOTAL**   **= $0.65**

CLICK WHEN DONE

DONE

The information in the Contributions Table will stay in the table through all rounds you play with the group, in both the decision and review screens

Questionnaire and Payment

After you have finished all of your decisions, you will complete a brief questionnaire. Then you will receive payment anonymously and privately, and the session will be over. You will collect your payment with no one watching, and no one else will learn your earnings.

QUIZ

Case 1: Suppose out of your 20 tokens you contributed 5 tokens to the Personal Fund and 15 tokens to the Group Fund. Suppose that the other 4 members of your group together contribute a total of 60 tokens to the Group Fund. Returns for the two Funds are:

Personal Fund: $0.02 per token to you

Group Fund: $0.03 per token to Stakeholder

                       $0.01 per token to non-Stakeholders

a) How much do you earn from the Personal Fund? _____

b) How many tokens in total were contributed by your group to the Group Fund? _____

c) If you ARE the Stakeholder, how much do you earn from the Group Fund? _____

d) If you ARE the Stakeholder, how much do each of the other members of your group earn from the Group Fund? _____

e) If you ARE the Stakeholder, how much do you earn altogether? _____

f) If you ARE NOT the Stakeholder, how much do you earn from the Group Fund? _____

g) If you ARE NOT the Stakeholder, how much does the group's Stakeholder earn from the Group Fund? _____

h) If you ARE NOT the Stakeholder, how much do each of the group's other non-Stakeholders earn from the Group Fund? _____

i) If you ARE NOT the Stakeholder, how much do you earn altogether? _____

Case 2: Suppose out of your 20 tokens you contributed 20 tokens to the Personal Fund and 0 tokens to the Group Fund.   Suppose that the other 4 members of your group together contribute a total of 8 tokens to the Group Fund. Returns for the two Funds are:

Personal Fund:　　　　$0.02 per token to you

Group Fund:　　　　　$0.03 per token to Stakeholder

　　　　　　　　　　$0.01 per token to non-Stakeholders

- How much do you earn from the Personal Fund? _____

- How many tokens in total were contributed by your group to the Group Fund? _____

- If you ARE the Stakeholder, how much do you earn from the Group Fund? _____

- If you ARE the Stakeholder, how much do each of the other members of your group earn from the Group Fund? _____

- If you ARE the Stakeholder, how much do you earn altogether? _____

- If you ARE NOT the Stakeholder, how much do you earn from the Group Fund? _____

- If you ARE NOT the Stakeholder, how much does the group's Stakeholder earn from the Group Fund? _____

- If you ARE NOT the Stakeholder, how much do each of the group's other non-Stakeholders earn from the Group Fund? _____

- If you ARE NOT the Stakeholder, how much do you earn altogether? _____

Instructions for the First Set of Rounds *(PUBLIC TREATMENT)*

- You will now be randomly assigned by the computer into groups of 5 people.

- You will be with the same group for 10 rounds.

- Each member of your group will be randomly assigned a *new, different* letter code that you will keep for these 10 rounds: A, B, C, D, or E.

- The Stakeholder position will rotate through all 5 members of the group, in alphabetical order (A, then B, then C, etc.). Each member of the group will be Stakeholder twice.

- You will always know who is currently Stakeholder, when you are going to be Stakeholder, and when each other member of your group (identified by their letter code) will be Stakeholder, because the Stakeholder position will be marked in the Contributions Table and in the status bar.

- In each round, you must decide how many of 20 tokens to put in the Personal Fund and how many in the Group Fund.

- When you have made your decision (by typing how many tokens (0 to 20) you want to put in the Group Fund), you must click "SUBMIT".

- The Contributions Table will show, for each past round:

    - what each member (identified by letter code) has contributed to the Group Fund; that is, there will be a column in the table corresponding to each group member, and that member's contributions for each round will be filled in after the round

    - the total number of tokens in the Group Fund

    - your earnings

- You will be able to tell what each group member contributed in each round; they will also be able to tell what you contributed in each round.

- The table will indicate in which rounds each person will be the Stakeholder by marking those rounds in each person's column with red, underlined stars (*********).

- You will see which column corresponds to you because your column will say the word **YOU** above it. Each person has **YOU** marked at the top of their own column.

- Also at the top of each table, in each round, the word **Stakeholder** will appear in red above the column of the person who is the Stakeholder in this round.

- After everyone has made their decisions for each round, you may examine the Contributions Table and the calculation of your earnings in the review screen. Click "DONE" when you are ready to move on to the next round.

When you are entering your contribution, the screen will look like this:

You are now: **MAKING YOUR CONTRIBUTION FOR ROUND 1**
Your Letter Code: C   Stakeholder: A

CONTRIBUTIONS TO THE GROUP FUND

			YOU				
	Stakeholder						
	A	B	C	D	E	TOTAL TOKENS IN GROUP FUND	MY EARNINGS
Round 1	*****						
Round 2		*****					
Round 3			*****				
Round 4				*****			
Round 5					*****		
Round 6	*****						
Round 7		*****					
Round 8			*****				
Round 9				*****			
Round 10					*****		

***** indicates that you were/will be Stakeholder in the round indicated

DECISION PANEL

How much would you like to put in the **GROUP FUND**? (0-20)

Your **PERSONAL FUND** contribution will be **20 minus your GROUP FUND contribution**.

RETURNS:

Personal fund: **$0.02 per token** to you

Group fund: **$0.03 per token** to Stakeholder
$0.01 per token to non-Stakeholders (including YOU)

CLICK TO SUBMIT

SUBMIT

When you are reviewing your results, the screen will look like this:

| You are now: **REVIEWING RESULTS FOR ROUND 2** |
| Your Letter Code: **E**   Stakeholder: **B** |

CONTRIBUTIONS TO THE GROUP FUND

| | | Stakeholder | | | YOU | | |
	A	B	C	D	E	TOTAL TOKENS IN GROUP FUND	MY EARNINGS
Round 1	10	6	2	14	18	50	$0.54
Round 2	19	11	3	7	15	55	$0.65
Round 3			*****				
Round 4				*****			
Round 5					*****		
Round 6	*****						
Round 7		*****					
Round 8			*****				
Round 9				*****			
Round 10					*****		

***** indicates that you were/will be Stakeholder in the round indicated

DECISION PANEL

REVIEW RESULTS FROM ROUND 2 IN THE TABLE ABOVE.

YOUR EARNINGS WERE:

Your Personal Fund contribution (5) times $0.02 per token   = $0.10

PLUS: The total number of tokens in the Group Fund (55) times $0.01 per token   = $0.55

EQUALS: TOTAL   = $0.65

CLICK WHEN DONE

DONE

Instructions for the Second Set of Rounds *(PRIVATE TREATMENT)*

- You will now be randomly assigned by the computer into groups of 5 people.

- You will be with the same group for 10 rounds.

- All 5 members of the group will be Stakeholder twice.

- You will only know when YOU are the Stakeholder—you will not know which group member is the Stakeholder in any given round if it is not you.

- In each round, you must decide how many of 20 tokens to put in the Personal Fund and how many in the Group Fund.

- When you have made your decision (by typing how many tokens (0 to 20) you want to put in the Group Fund), you must click "SUBMIT".

- The Contributions Table will show, for each past round:

    - what you have contributed to the Group Fund

    - a list of the Group Fund contributions of the other members of your group; these contributions will be listed in a **random order that is reshuffled** for each round

    - the total number of tokens in the Group Fund

    - your earnings

- No one will not be able to track any other individual group member's contributions.

- The table will indicate in which rounds you will be the Stakeholder by marking those rounds in your column with red, underlined stars (*********).

- After everyone has made their decisions for each round, you may examine the Contributions Table and the calculation of your earnings in the review screen. Click "DONE" when you are ready to move on to the next round.

When you are entering your contribution, the screen will look like this:

| You are now: **MAKING YOUR DECISION FOR ROUND 1** |
| **Stakeholder: Someone Else** |

CONTRIBUTIONS TO THE GROUP FUND

	YOU	OTHERS (Random Order)	TOTAL TOKENS IN GROUP FUND	MY EARNINGS
Round 1				
Round 2	*****			
Round 3				
Round 4				
Round 5				
Round 6				
Round 7				
Round 8	*****			
Round 9				
Round 10				

***** indicates that you were/will be Stakeholder in the round indicated

DECISION PANEL

How much would you like to put in the **GROUP FUND**? (0-20) ☐

Your **PERSONAL FUND** contribution will be **20 minus your GROUP FUND contribution**.

RETURNS:

Personal fund: **$0.02 per token** to you

Group fund: **$0.03 per token** to Stakeholder
$0.01 per token to non-Stakeholders (including YOU)

CLICK TO SUBMIT

SUBMIT

When you are reviewing your results, the screen will look like this:

You are now: **REVIEWING RESULTS FOR ROUND 2**

Stakeholder: Someone Else

CONTRIBUTIONS TO THE GROUP FUND

	YOU	OTHERS (Random Order)	TOTAL TOKENS IN GROUP FUND	MY EARNINGS
Round 1	9	1, 5, 17, 13	45	$1.57
Round 2	12	16, 4, 20, 8	60	$0.76
Round 3				
Round 4				
Round 5				
Round 6				
Round 7				
Round 8				
Round 9	*****			
Round 10				

***** indicates that you were/will be Stakeholder in the round indicated

DECISION PANEL

REVIEW RESULTS FROM ROUND 2 IN THE TABLE ABOVE.

YOUR EARNINGS WERE:

Your Personal Fund contribution (8) times $0.02 per token = $0.16

PLUS: The total number of tokens in the Group Fund (60) times $0.01 per token = $0.60

EQUALS: TOTAL = $0.76

CLICK WHEN DONE

DONE

Instructions for the Third Set of Rounds *(INELIGIBLE TREATMENT)*

- You will now be randomly assigned by the computer into groups of 5 people.

- You will be with the same group for 8 rounds.

- Each member of your group will be randomly assigned a *new, different* letter code that you will keep for these 8 rounds: A, B, C, D, or E.

- The Stakeholder position will rotate through 4 of the 5 members of the group.

- The one person in your group who will never be the Stakeholder is the **Ineligible Person**. This person is randomly selected at the start of this set of rounds, and will be Ineligible (will never be the Stakeholder) for all 8 rounds. The Ineligible Person is identified to all group members in the Contributions Table and in the text in the status bar.

- The Stakeholder position rotates through the four eligible members of the group for all 8 rounds, in alphabetical order. For example, if person C is randomly chosen as Ineligible, the first Stakeholder will be person A, the second will be B, the third will be D, etc.

- Everyone who is not Ineligible will be Stakeholder twice.

- You will always know who is currently Stakeholder, when you are going to be Stakeholder, and when each other member of your group (identified by their letter code) will be Stakeholder, because the Stakeholder position will be marked in the Contributions Table and in the status bar.

- In each round, you must decide how many of 20 tokens to put in the Personal Fund and how many in the Group Fund.

- When you have made your decision (by typing how many tokens (0 to 20) you want to put in the Group Fund), you must click "SUBMIT".

- The Contributions Table will show, for each past round:

    - what each member (identified by letter code) has contributed to the Group Fund; that is, there will be a column in the table corresponding to each

group member, and that member's contributions for each round will be
filled in after the round

- the total number of tokens in the Group Fund

- your earnings

- You will be able to tell what each group member contributed in each round; they
will also be able to tell what you contributed in each round.

- The table will indicate in which rounds each person will be the Stakeholder by
marking those rounds in each person's column with red, underlined stars (*********).

- You will see which column corresponds to you because your column will say the
word **YOU** above it. Each person has **YOU** marked at the top of their own
column.

- Also at the top of each table, in each round, the word **Stakeholder** will appear
above the column of the person who is the Stakeholder in this round.

- The word **Ineligible** will appear above the column of the Ineligible person.

- After everyone has made their decisions for each round, you may examine the
Contributions Table and the calculation of your earnings in the review screen.
Click "DONE" when you are ready to move on to the next round.

When you are entering your contribution, the screen will look like this:

You are now: **MAKING YOUR CONTRIBUTION FOR ROUND 1**

Your Letter Code: A   Stakeholder: YOU   Ineligible: D

CONTRIBUTIONS TO THE GROUP FUND

	YOU						
	Stakeholder			Ineligible			
	A	B	C	D	E	TOTAL TOKENS IN GROUP FUND	MY EARNINGS
Round 1	*****						
Round 2		*****					
Round 3			*****				
Round 4					*****		
Round 5	*****						
Round 6		*****					
Round 7			*****				
Round 8					*****		

***** indicates that you were/will be Stakeholder in the round indicated

DECISION PANEL

How much would you like to put in the **GROUP FUND**? (0-20) [ ]

Your **PERSONAL FUND** contribution will be **20 minus your GROUP FUND contribution**.

RETURNS:

Personal fund: **$0.02 per token** to you

Group fund: **$0.03 per token** to Stakeholder (YOU)
                         $0.01 per token to non-Stakeholders

CLICK TO SUBMIT

SUBMIT

When you are reviewing your results, the screen will look like this:

You are now: **REVIEWING RESULTS FOR ROUND 2**
Your Letter Code: E   Stakeholder: B   Ineligible: D

CONTRIBUTIONS TO THE GROUP FUND							
					YOU		
		Stakeholder		Ineligible			
	A	B	C	D	E	TOTAL TOKENS IN GROUP FUND	MY EARNINGS
Round 1	20	13	8	9	4	54	$0.86
Round 2	1	18	10	5	16	50	$0.58
Round 3			*****				
Round 4					*****		
Round 5	*****						
Round 6		*****					
Round 7			*****				
Round 8					*****		

***** indicates that you were/will be Stakeholder in the round indicated

DECISION PANEL

REVIEW RESULTS FROM ROUND 2 IN THE TABLE ABOVE.

YOUR EARNINGS WERE:

Your Personal Fund contribution (4) times $0.02 per token   = $0.08

PLUS: The total number of tokens in the Group Fund (50) times $0.01 per token   = $0.50

EQUALS: **TOTAL**   = $0.58

CLICK WHEN DONE

DONE

Appendix B: Parametric Tests of Reciprocity in the Girl Scout Cookie Experiment

In this Appendix, we use regression techniques to identify reciprocal behavior resulting from direct reciprocity and strategic self-interest. These results may be biased because of the endogeneity inherent in group dynamic behavior. We take care to limit the influence of this bias, but it is, to some extent, unavoidable.

We perform a panel regression for one treatment at a time, with fixed effects and errors clustered by group. Non-Stakeholder contributions in each round g_{it} are regressed on characteristics of that round, including h_{ikt} (a summary of the current Stakeholder k's past generosity toward subject i: here, the current Stakeholder's cumulative average contributions when subject i was Stakeholder) and $\mathbf{X}_{it}$ (other variables):

$$g_{it} = a + bh_{ikt} + \mathbf{CX}_{it} + \varepsilon_{it}$$

If direct reciprocity or strategic self-interest is important, b (the coefficient on h_{ikt}) should be positive in both the Public and Ineligible treatments. In the Private treatment, b should be zero, because in each round, no-one knows who the current Stakeholder is or what that person has done in the past. Group-level conditional cooperation could bias this coefficient upward, so we counteract that bias by including a control for group generosity in $\mathbf{X}_{it}$. Our control for group generosity is the group's cumulative average non-Stakeholder contribution in past rounds. For each subject in each round, this measure excludes his own past contributions and those of the current Stakeholder. In the Ineligible treatment, this group measure also excludes the Bachelor (although the same results obtain using a measure that includes the Bachelor, results available on request).

The panel regression also includes in $\mathbf{X}_{it}$ the current round number and an indicator for whether this subject has passed his last Stakeholder stint. If the coefficient on the round number is negative, there is a secular drop-off in cooperation. The coefficient on the post-last-Stakeholder stint dummy can be interpreted as the importance of strategic giving. If a subject gives strategically in the rounds before his last Stakeholder stint, there should be a discontinuity in contributions that should be reflected in a large, negative coefficient on this dummy.

Results are shown in Table 21. Directly reciprocal behavior is supported for the Public and Ineligible treatments. The coefficient on the post-last-Stakeholder stint dummy in the Public and Ineligible treatments is insignificant. This implies that strategic motives are not important. The same results obtain in an AR1 specification (results available upon request), except that in the Private treatment the post-last-Stakeholder stint dummy is no longer significant.

Table 21. Fixed Effects Panel Regression of Non-Stakeholder Contribution (in Percent of Endowment) on Period-Level Covariates

	Private Treatment	Public Treatment	Ineligible Treatment
Stakeholder average	0.05	0.24***	0.23**
past contributions to me	(0.06)	(0.05)	(0.09)
Group average	0.25	0.05	-0.08
contributions	(0.16)	(0.17)	(0.27)
Round number	0.03	-1.69	-2.86**
	(0.65)	(1.22)	(1.34)
Post-last Stakeholder	-9.66**	-1.66	-4.85
Stint? (dummy)	(3.86)	(5.05)	(5.40)
Constant	22.15***	34.65**	42.50**
	(7.48)	(13.68)	(16.18)
Observations (rounds)	720	720	432
Number of subjects	120	120	96
F	4.50	13.37	31.47
R^2 (overall)	0.078	0.182	0.156

Robust standard errors in parentheses; errors are clustered on groups; individual fixed effects
* significant at 10%; ** significant at 5%; *** significant at 1%

As in Chapter I, we can test for the importance of other-regarding preferences by restricting our attention to the rounds after a subject has passed his final Stakeholder stint. See results in Table 22. Due to the reduced population size, particularly for the Ineligible treatment, test power is significantly reduced and the Stakeholder past contribution coefficient is not statistically significant for the Ineligible treatment. However, the coefficient on Stakeholder past contribution is positive for the Public treatment. This implies that reciprocal behavior is at least partly rooted in other-regarding preferences.

Table 22. Fixed Effects Panel Regression of Non-Stakeholder Contribution (in Percent of Endowment) on Period-Level Covariates, Post-Last Stakeholder Stint

	Private Treatment	Public Treatment	Ineligible Treatment
Stakeholder average	-0.19	0.21*	0.17
past contributions to me	(0.11)	(0.11)	(0.22)
Group average	0.39	-0.09	0.30
contributions	(0.44)	(0.32)	(0.60)
Round number	-4.12**	-5.90***	-4.89
	(1.83)	(1.96)	(2.26)
Constant	48.39**	76.14***	42.89
	(22.93)	(25.80)	(25.08)
Observations (rounds)	240	240	144
Number of subjects	96	96	72
F	2.03	3.70	1.78
R^2 (overall)	0.036	0.106	0.083

Robust standard errors in parentheses; errors are clustered on groups; individual fixed effects
* significant at 10%; ** significant at 5%; *** significant at 1%

Finally, we build error clustering by group and fixed effects into a test using aggregate statistics, since this should eliminate the bias of our panel analysis while checking for robustness to intra-class correlation. We create two stacked observations per subject: one to represent the subject's behavior in rounds in which the current Stakeholder was previously generous (using the 10 token, or 50%-of-endowment,

threshold) toward this subject, and one to represent his behavior in rounds when the

Stakeholder was ungenerous toward this subject. A "nice dummy" differentiates between

the two observations for each subject. We perform a panel regression, with these two

observations per subject, in which we regress contributions on the "nice dummy." The

same results obtain if we use group dummies, individual random effects, and individual

fixed effects. In Table 23 we show results of the individual fixed effects regression

without group dummies. The "nice" dummy is significant and positive for the Public and

Ineligible treatments, but not for the Private treatment. The result that reciprocal

behavior continues after the last Stakeholder stint also persists in this specification

(results available upon request).

Table 23. Panel Stacked Regression of Average Contribution (in Percent of Endowment) on "Nice Dummy"

	Private Treatment	Public Treatment	Ineligible Treatment
"Nice dummy"	-1.59	14.83***	17.30***
	(2.72)	(2.68)	(2.72)
Constant	29.16***	27.09***	23.92***
	(1.36)	(1.34)	(1.36)
Subjects	82	95	75
R^2 (overall)	-0.001	0.058	0.096
F	0.34	30.54	40.37

Robust standard errors in parentheses; errors clustered on groups; individual fixed effects
* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix C: Predictions of Risk Preference Theories

In this Appendix, we describe in detail why the choices we call mistakes or inconsistent choices over risk are, in fact, inconsistent with standard theories of preferences over risk.

First, we note that the gain lotteries can be defined as follows. Lottery pair i (where $i = 1, …, 5$) offers two options:

Safe lottery = Lottery B = { $x + (i − 1)b$, $p=1/2$; $x − (i − 1)a$ }

Risky lottery = Lottery A = { $x + ib$, $p=1/2$; $x − ia$ }

Similarly, for the gain-loss lotteries, lottery pair i offers:

Safe lottery = Lottery B = { $x + (i − 1)b$, $p=1/2$; $0 − (i − 1)a$ }

Risky lottery = Lottery A = { $x + ib$, $p=1/2$; $0 − ia$ }

For low stakes lotteries, $x = 500$ RWF, $a = 100$ RWF, $b = 200$ RWF, and for high stakes lotteries, $x = 1250$ RWF, $a = 250$ RWF, $b = 400$ RWF. We will demonstrate the argument for gain lotteries; the same conditions hold for the gain-loss lotteries except as noted.

Expected Value

Proposition 1. An Expected Value maximizer will prefer Lottery A$_i$ for all $i > 0$.

Proof. For every lottery pair i, Lottery A has a lower expected value than Lottery B:

$$EV(A_i) = \frac{1}{2}\left[(x+ib)+(x-ia)\right]$$

This simplifies to:

$$EV(A_i) = x\frac{i}{2}[b-a]$$

$$EV(B_i) = \frac{1}{2}\Big[\big(x+(i-1)b\big)+\big(x-(i-1)a\big)\Big]$$

This simplifies to:

$$EV(B_i) = x\frac{i-1}{2}[b-a]$$

Therefore, choices governed by Expected Value will be A for all $i > 0$. Therefore, Expected Value only predicts the pattern AAAAA.

Expected Utility

Proposition 2. Assume that Expected Utility governs choices, and that the utility function is increasing and non-inflected in the range of the lotteries. Then if Lottery B is chosen for some lottery pair i, Lottery A may not be chosen for any following lottery $i+1$.

Proof. Define a utility function $u(x)$. Then Expected Utility from a lottery is $EU = \sum_i p_i u(x_i)$, so for the lotteries in this experiment where each outcome has probability ½:

$$EU = \frac{1}{2}\big(u(x_1)+u(x_2)\big) \propto u(x_1)+u(x_2)$$

Therefore, the expected utilities from the two lotteries can be defined as:

$$EU(A_i) = u(x+ib)+u(x-ia)$$

$$EU(B_i) = u\big(x+(i-1)b\big)+u\big(x-(i-1)a\big)$$

If the function $u(x)$ is convex or linear, then for all i, $EU(A_i) > EU(B_i)$ (note that A always has a higher expected value). In these cases, Lottery B will never be chosen, and therefore it would be impossible to switch from selecting B in pair i to A in pair $i+1$. Therefore, it suffices to show only that such a switch is impossible for strictly concave utility functions.

Figure 7 below demonstrates the prizes for lottery i and lottery ($i+1$) for a concave utility function over income.
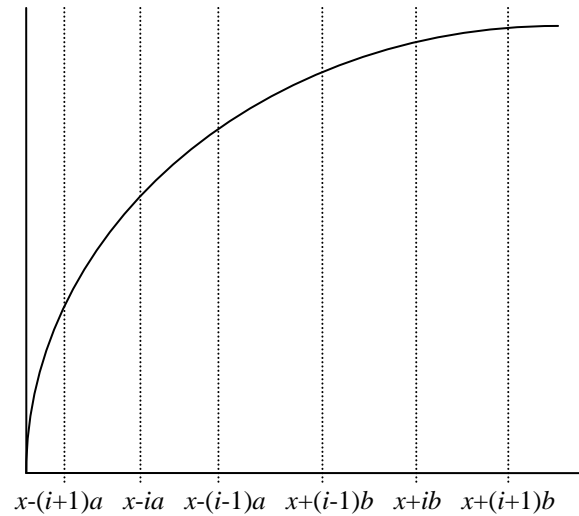

$$x\text{-}(i+1)a \quad x\text{-}ia \quad x\text{-}(i\text{-}1)a \quad x+(i\text{-}1)b \quad x+ib \quad x+(i+1)b$$

Figure 7: Concave Utility Functions over Lottery Prizes

We define:

$$udiff_1 = u\big(x-ia\big)-u\big(x-(i+1)a\big)$$

$$udiff_2 = u\big(x-(i-1)a\big)-u\big(x-ia\big)$$

$$udiff_3 = u\big(x+ib\big)-u\big(x+(i-1)b\big)$$

$$udiff_4 = u\big(x+(i+1)b\big)-u\big(x-ib\big)$$

A concave utility function gets flatter as wealth decreases, so for a concave utility function, $udiff_1 > udiff_2$ and $udiff_3 > udiff_4$.

In lottery pair i, choosing B means that $EU(B_i) \geq EU(A_i)$:

$$u\big(x+(i-1)b\big)+u\big(x-(i-1)a\big)\geq u\big(x+ib\big)+u\big(x-ia\big)$$

This simplifies to:

$$u\big(x-(i-1)a\big)-u\big(x-ia\big)\geq u\big(x+ib\big)-u\big(x+(i-1)b\big)$$

In other words, $udiff_2 \geq udiff_3$. In lottery pair i+1, choosing A means:

$$u\big(x+(i+1)b\big)+u\big(x-(i+1)a\big)\geq u\big(x+ib\big)+u\big(x-ia\big)$$

This simplifies to:

$$u\big(x-ia\big)-u\big(x-(i+1)a\big)\leq u\big(x+(i+1)b\big)-u\big(x+ib\big)$$

In other words, $udiff_1 \leq udiff_4$. However, it cannot simultaneously be true that $udiff_1 \leq udiff_4$ and $udiff_2 \geq udiff_3$ if the utility function is strictly concave throughout the region. This is because if it is strictly concave, then $udiff_1 > udiff_2$ and $udiff_3 > udiff_4$ and therefore $udiff_2 \geq udiff_3$ implies $udiff_1 > udiff_3$ and therefore $udiff_1 > udiff_4$, which implies that lottery B must be chosen in pair i+1. This is a contradiction.

Thus for a utility function that is concave through this region, an expected utility maximizer must never switch from B to A in a subsequent lottery pair. Therefore, Expected Utility Theory only predicts patterns that do not switch from B to A; namely, AAAAA, AAAAB, AAABB, AABBB, ABBBB, and BBBBB.

Rank Dependent Theory, Prospect Theory, and Cumulative Prospect Theory

Proposition 3. Assume that Rank Dependent Theory, Prospect Theory, or Cumulative Prospect Theory governs choices, and that the value function is increasing, concave, and

non-inflected in the range of the lotteries. Then if Lottery B is chosen for some lottery pair i, Lottery A may not be chosen for any following lottery $i+1$.

Proof. Rank Dependent Theory, Prospect Theory, and Cumulative Prospect Theory transform both payoffs as well as probabilities.

Although Prospect Theory allows an inflection point at zero, and Cumulative Prospect Theory allows inflection point at somewhat arbitrary reference points, the assumption of no inflection points in the range of the lotteries excludes the possibility of a reference points in the range of our lotteries. Note that this means that this Proposition cannot apply to Prospect Theory for the gain-loss lotteries.

Although the probabilities for the high and low outcomes are both ½, we allow the transformed probabilities to differ. We designate the transformed probabilities as constants k_1 and k_2. For our lotteries, these theories predict choice in lottery pair i based on a function of the form:

$$U_{RD/CPT}(A_i) = k_1 v(x - ia) + k_2 v(x + ib)$$

$$U_{RD/CPT}(B_i) = k_1 v(x - (i-1)a) + k_2 v(x + (i-1)b)$$

Analogous to the terminology in the Expected Utility demonstration, we define:

$$vdiff_1 = v(x - ia) - v(x - (i+1)a)$$

$$vdiff_2 = v(x - (i-1)a) - v(x - ia)$$

$$vdiff_3 = v(x + ib) - v(x + (i-1)b)$$

$$vdiff_4 = v(x + (i+1)b) - v(x - ib)$$

In these theories, if a subject prefers lottery B to A in pair i:

$$k_1 v\big(x-(i-1)a\big)+k_2 v\big(x+(i-1)b\big)\geq k_1 v\big(x-ia\big)+k_2 v\big(x+ib\big)$$

This simplifies to:

$$\frac{k_1}{k_2}\Big[v\big(x-(i-1)a\big)-v\big(x-ia\big)\Big]\geq v\big(x+ib\big)-v\big(x+(i-1)b\big)$$

In other words, $\frac{k_1}{k_2}vdiff_2 \geq vdiff_3$. In pair $i+1$, choosing A means:

$$k_1 v\big(x-(i+1)a\big)+k_2 v\big(x+(i+1)b\big)\geq k_1 v\big(x-ia\big)+k_2 v\big(x+ib\big)$$

This simplifies to:

$$\frac{k_1}{k_2}\Big[v\big(x-ia\big)-v\big(x-(i+1)a\big)\Big]\leq v\big(x+(i+1)b\big)-v\big(x+ib\big)$$

In other words, $\frac{k_1}{k_2}vdiff_1 \leq vdiff_4$.

If v is concave then $vdiff_1 > vdiff_2$ and $vdiff_3 > vdiff_4$. Therefore, if

$\frac{k_1}{k_2}vdiff_2 \geq vdiff_3$, then $\frac{k_1}{k_2}vdiff_1 > vdiff_3$ and therefore $\frac{k_1}{k_2}vdiff_1 > vdiff_4$. But that violates

the assumption that lottery A_{i+1} is preferred to B_{i+1}. This is therefore a contradiction, and

therefore given the assumptions imposed a person choosing according to Rank

Dependent Theory, Prospect Theory, or Cumulative Prospect Theory must never choose

lottery B in pair i and subsequently choose lottery A in pair $i+1$.

Given these assumptions, these theories only predict patterns that do not switch

from B to A: AAAAA, AAAAB, AAABB, AABBB, ABBBB, and BBBBB.

Discussion. Reference points create a particular challenge for Prospect Theory in the

area of 0 (i.e. for the gain-loss lotteries) and for Cumulative Prospect Theory for both sets

of lotteries. These reference points are inflection points, and given such an inflection point the prediction is indeterminate.

Discussion. If the value function is convex, it is possible for probability weighting to explain switching from lottery B in pair i to lottery A in pair $i+1$. If a value function is convex, then $vdiff_1 < vdiff_2 < vdiff_3 < vdiff_4$. Recall that both probabilities are ½, but they are transformed to weights k_1 and k_2. By logic like the above, if B_i is preferred to A_i, then $\frac{k_1}{k_2} vdiff_2 \geq vdiff_3$. If A_{i+1} is preferred to B_{i+1}, then $\frac{k_1}{k_2} vdiff_1 \leq vdiff_4$. These conditions can all be met if $\frac{vdiff_3}{vdiff_2} \leq \frac{k_1}{k_2} \leq \frac{vdiff_4}{vdiff_1}$.

Dual Theory

Proposition 4. Assume that the Dual Theory, which transforms probabilities but not payoffs, governs choices. Then if the subject strongly prefers A for some pair i, he must prefer A for all following pairs; similarly, if he strongly prefers B for some pair i, he must prefer B for all following pairs.

Proof. The dual theory does not transform payoffs but only transforms probabilities. We again use k_1 and k_2 to denote the transformed ½ probabilities for the high and low prizes respectively. Therefore, Dual Theory evaluations of the lotteries in pair i are:

$$U_{DT}(A_i) = k_1(x - ia) + k_2(x + ib)$$

$$U_{DT}(B_i) = k_1(x - (i-1)a) + k_2(x + (i-1)b)$$

This simplifies to:

$$U_{DT}(A_i) = (k_1 + k_2)x + i(k_2 b - k_1 a)$$

$$U_{DT}(B_i) = (k_1 + k_2)x + (i-1)(k_2 b - k_1 a)$$

Therefore, a subject who chooses according to the Dual Theory will choose A$_i$ if:

$$(k_1 + k_2)x + i(k_2 b - k_1 a) \geq (k_1 + k_2)x + (i-1)(k_2 b - k_1 a)$$

This simplifies to:

$$0 \geq k_2 b - k_1 a$$

Since this choice does not depend on i, it will either be true for all lottery pairs or for no lottery pairs. That is, a Dual Theory optimizer will always choose AAAAA or BBBBB if he has strong preferences.

Discussion. If a Dual Theory optimizer is indifferent for one lottery pair in the set, he will be indifferent for all lottery pairs in the set. True or "practical" indifference could produce patterns other than those predicted here.

References

Andersen, Steffen, Glenn W. Harrison, Morten Igel Lau, and E. Elisabet Rutstrom. 2006. Elicitation using multiple price list formats. *Experimental Economics* 9, no. 4: 383-405.

Andreoni, James. 1988. Why free ride? Strategies and learning in public goods experiments. *Journal of Public Economics* 37, no. 3: 291-304.

Ashraf, Nava, Dean Karlan, and Wesley Yin. 2006. Tying odysseus to the mast: Evidence from a commitment savings product in the philippines. *Quarterly Journal of Economics* 121, no. 2: 635-672.

Bardsley, Nicholas and Peter Moffatt. 2007. The experimetrics of public goods: Inferring motivations from contributions. *Theory and Decision* 62, no. 2: 161-193.

Bellemare, Charles and Bruce S. Shearer. 2006. *Sorting, incentives and risk preferences: Evidence from a field experiment*. Institute for the Study of Labor (IZA).

Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro. 2006. *Who is 'behavioral'? Cognitive ability and anomalous preferences*. Chicago, IL: University of Chicago.

Besley, Timothy. 1995. Nonmarket institutions for credit and risk sharing in low-income countries. *The Journal of Economic Perspectives* 9, no. 3: 115-127.

Binswanger, Hans P. 1980. Attitudes toward risk: Experimental measurement in rural india. *American Journal of Agricultural Economics* 62, no. 3: 395-407.

Birnbaum, Michael H. and Ulrich Schmidt. 2008. An experimental investigation of violations of transitivity in choice under uncertainty. *Journal of Risk and Uncertainty* 37, no. 1: 77-91.

Brandts, Jordi, David J. Cooper, and Enrique Fatas. 2007. Leadership and overcoming coordination failure with asymmetric costs. *Experimental Economics* 10, no. 3: 269-284.

Camerer, Colin F. and Robin M. Hogarth. 1999. The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty* 19, no. 1-3: 7-42.

Carbone, Enrica and John D. Hey. 2000. Which error story is best? *Journal of Risk and Uncertainty* 20, no. 2: 161-176.

Castillo, Marco, Paul J. Ferraro, Jeffrey Jordan, and Ragan Petrie. 2009. *The today and tomorrow of kids*.

Castillo, Marco and Ragan Petrie. 2010. Discrimination in the lab: Does information trump appearance? *Games and Economic Behavior* 68, no. 1: 50-59.

Chetty, Raj and Adam Szeidl. 2007. Consumption commitments and risk preferences. *Quarterly Journal of Economics* 122, no. 2: 831-877.

Cooper, Joseph C. and Tim C. Osborn. 1998. The effect of rental rates on the extension of conservation reserve program contracts. *American Journal of Agricultural Economics* 80, no. 1: 184-194.

Cox, James C. 2004. How to identify trust and reciprocity. *Games and Economic Behavior* 46, no. 2: 260-281.

Cox, James C., Daniel Friedman, and Vjollca Sadiraj. 2008. Revealed altruism. *Econometrica* 76, no. 1: 31-69.

Cox, James C. and Vjollca Sadiraj. 2007. On modeling voluntary contributions to public goods. *Public Finance Review* 35, no. 2: 311-332.

Croson, Rachel T. A. 1996. Partners and strangers revisited. *Economics Letters* 53, no. 1: 25-32.

De La Torre Ugarte, Daniel and Chad Helliwinckel. 2006. *Analysis of the economic impacts on the agricultural sector of the elimination of the conservation reserve program*. Knoxville, TN, USA: University of Tennessee.

De La Torre Ugarte, Daniel, Daryll E. Ray, Richard L. White, and Michael R. Dicks. 1995. *The conservation reserve program*. Agricultural Policy Analysis Center,

The University of Tennessee, and Great Plains Agricultural Policy Center, Oklahoma State University.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2010. Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey. *Journal of the European Economic Association* forthcoming.

Eckel, Catherine C., Cathleen Johnson, Claude Montmarquette, and Christian Rojas. 2007. Debt aversion and the demand for loans for postsecondary education. *Public Finance Review* 35, no. 2: 233-262.

El-Gamal, Mahmoud A. and David M. Grether. 1995. Are people bayesian? Uncovering behavioral strategies. *Journal of the American Statistical Association* 90, no. 432: 1137-1145.

Engelmann, Dirk and Urs Fischbacher. 2009. Indirect reciprocity and strategic reputation building in an experimental helping game. *Games and Economic Behavior* Forthcoming.

Falk, Armin and James J. Heckman. 2009. Lab experiments are a major source of knowledge in the social sciences. *Science* 326, no. 5952: 535-538.

Fischbacher, Urs. 2007. Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10, no. 2: 171-178.

Friedman, Milton and L. J. Savage. 1948. The utility analysis of choices involving risk. *The Journal of Political Economy* 56, no. 4: 279-304.

Gächter, Simon. 2007. Conditional cooperation: Behavioral regularities from the lab and the field and their policy implications. In *Economics and psychology. A promising new cross-disciplinary field*, ed. Bruno S. Frey and Alois Stutzer. Boston, MA: The MIT Press.

Glöckner, Andreas, Bernd Irlenbusch, Sebastian Kube, Andreas Nicklisch, and Hans T. Normann. 2009. *Leading with(out) sacrifice? A public-goods experiment with a super-additive player*. Bonn, Germany: Max Planck Institute for Research on Collective Goods.

Goeree, Jacob, Charles Holt, and Susan Laury. 2002. Private costs and public benefits: Unraveling the effects of altruism and noisy behavior. *Journal of Public Economics* 83, no. 2: 255-276.

Güth, Werner, M. Vittoria Levati, Matthias Sutter, and Eline van der Heijden. 2007. Leading by example with and without exclusion power in voluntary contribution experiments. *Journal of Public Economics* 91, no. 5-6: 1023-1042.

Harless, David W. and Colin F. Camerer. 1994. The predictive utility of generalized expected utility theories. *Econometrica* 62, no. 6: 1251-1289.

Harrison, Glenn W., John A. List, and Charles Towe. 2007. Naturally occurring preferences and exogenous laboratory experiments: A case study of risk aversion. *Econometrica* 75, no. 2: 433-458.

Harrison, Glenn W. and E. Elisabet Rutstrom. 2008. Risk aversion in the laboratory. In *Risk aversion in experiments*, ed. James C. Cox and Glenn W. Harrison, 12. Bingley, UK: Emerald.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics* 24, no. 3: 411-482.

Hey, John D. 2005. Why we should not be silent about noise. *Experimental Economics* 8, no. 4: 325-345.

Hey, John D. and Chris Orme. 1994. Investigating generalizations of expected utility theory using experimental data. *Econometrica* 62, no. 6: 1291-1326.

Holt, Charles A. 1986. Preference reversals and the independence axiom. *American Economic Review* 76, no. 3: 508-515.

Holt, Charles A. and Susan K. Laury. 2002. Risk aversion and incentive effects. *American Economic Review* 92, no. 5: 1644-1655.

Johnson, Philip N., Sukant K. Misra, and Terry R. Ervin. 1997. A qualitative choice analysis of factors influencing post-crp land use decisions. *Journal of Agricultural and Applied Economics* 29, no. 1: 163-173.

Klein, Daniel B. 1990. The voluntary provision of public goods? The turnpike companies of early america. *Economic Inquiry* 28, no. 4: 788-812.

Laury, Susan K. and Charles A. Holt. 2005. Further reflections on prospect theory. In *Experimental Economics Center Working Paper Series*. Atlanta, GA: Experimental Economics Center, Andrew Young School of Policy Studies, Georgia State University.

Leuven, Edwin and Barbara Sianesi. Psmatch2: Stata module to perform full mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing S432001. Boston College Department of Economics.

List, John A. 2006. The behavioralist meets the market: Measuring social preferences and reputation effects in actual transactions. *Journal of Political Economy* 114, no. 1: 1-37.

Long, Stephen H. 1976. Social pressure and contributions to health charities. *Public Choice* 28, no. Winter: 55-66.

Loomes, Graham, Peter G. Moffatt, and Robert Sugden. 2002. A microeconometric test of alternative stochastic theories of risky choice. *Journal of Risk and Uncertainty* 24, no. 2: 103-130.

Lubowski, Ruben N., Andrew J. Plantinga, and Robert N. Stavins. 2008. What drives land-use change in the united states? A national analysis of landowner decisions. *Land Economics* 84, no. 4: 529-550.

Nowak, Martin A. and Karl Sigmund. 2005. Evolution of indirect reciprocity. *Nature* 437, no. 7063: 1291-1298.

Ortmann, Andreas and Ralph Hertwig. 2006. Monetary incentives: Usually neither necessary nor sufficient? In *CERGE-EI Working Paper*. Prague, Czech Republic: Charles University.

Petrie, Ragan. 2002. *Rwanda credit unions member and non-member survey 2002*. World Council of Credit Unions.

Prasad, Kislaya and Timothy C. Salmon. 2007. *Self-selection and market power in risk sharing contracts*.

Rabin, Matthew. 1993. Incorporating fairness into game theory and economics. *American Economic Review* 83, no. 5: 1281-1302.

Roberts, Michael J. and Ruben N. Lubowski. 2007. Enduring impacts of land retirement policies: Evidence from the conservation reserve program. *Land Economics* 83, no. 4: 516-538.

Rosenbaum, Paul R. 2002. *Observational studies*. New York: Springer-Verlag.

Secchi, Silvia and Bruce A. Babcock. 2007. *Impact of high crop prices on environmental quality: A case of iowa and the conservation reserve program*. Ames, IA: Center for Agricultural and Rural Development, Iowa State University.

Segal, Carmit. 2009. *Misbehavior, education, and labor market outcomes*. Barcelona, Spain: Universitat Pompeu Fabra, Department of Economics and Business.

Seinen, Ingrid and Arthur Schram. 2006. Social status and group norms: Indirect reciprocity in a repeated helping experiment. *European Economic Review* 50, no. 3: 581-602.

Sell, Jane and Rick K. Wilson. 1991. Levels of information and contributions to public goods. *Social Forces* 70, no. 1: 107-124.

Sobel, Joel. 2005. Interdependent preferences and reciprocity. *Journal of Economic Literature* 43, no. 2: 392-436.

Soetevent, Adriaan R. 2005. Anonymity in giving in a natural context--a field experiment in 30 churches. *Journal of Public Economics* 89, no. 11-12: 2301-2323.

Stockman, Carol. 2006. *Differences and similarities in health risk preferences across two subject cohorts*. Pittsburgh, PA: University of Pittsburgh.

Sunde, Uwe, Thomas Dohmen, Armin Falk, and David Huffman. 2010. Are risk aversion and impatience related to cognitive ability? *American Economic Review* forthcoming.

Tanaka, Tomomi, Colin Camerer, and Quang Nguyen. 2010. Risk and time preferences: Linking experimental and household survey data from vietnam. *American Economic Review* forthcoming.

Trivers, Robert L. 1971. The evolution of reciprocal altruism. *The Quarterly Review of Biology* 46, no. 1: 35-57.

Uri, Noel D. 2001. A note on soil erosion and its environmental consequences in the united states. *Water, Air & Soil Pollution* 129, no. 1-4: 181-197.

US Department of Agriculture. 2001. *1997 national resources inventory (revised december 2000).* National Resources Conservation Service, Washington, DC, and Statistical Laboratory, Iowa State University, Ames IA.

US Department of Agriculture, Natural Resources Conservation Service. 2009. National soil survey handbook, title 430-vi. *Available online at: soils.usda.gov/technical/handbook, accessed October 1 2009.*

World Bank Group Doing Business Project. 2007. *Doing business 2008 rwanda.* Washington, DC, USA: World Bank Publications and the International Finance Corporation.

Wu, Junjie. 2000. Slippage effects of the conservation reserve program. *American Journal of Agricultural Economics* 82, no. 4: 979-992.

Vita

Sarah Andrea Jacobson was born on February 12, 1977 in Binghamton, New York. She received a Bachelor of Science degree in Engineering from Harvey Mudd College in 1998. She worked as an engineer for The Boeing Company from 1998 through 2000, and then for Scientific-Atlanta, Inc. from 2000 through 2005 (which became a part of Cisco during this time). In 2003, she returned to school to begin her graduate work. She was awarded a Master of Arts in Economics in 2005 from Georgia State University. She expects to receive her Doctor of Philosophy degree in Economics from Georgia State University in May, 2010. She has accepted a tenure-track faculty position as an Assistant Professor of Economics at Williams College in Williamstown, MA, to begin in July, 2010.

While at Georgia State University, Ms. Jacobson worked as a research assistant to Professor Ragan Petrie and to Professor James Cox. She served as a teaching assistant to Professor Ragan Petrie (for undergraduate Principles of Microeconomics and for two years of PhD-level Microeconomics) and Professor Paul Kagundu (for the undergraduate class "The Global Economy"). In Spring 2009, she taught "The Global Economy" as sole instructor.

During her time at Georgia State University, she received several awards recognizing her academic accomplishments: the Andrew Young School Excellence in Teaching Economics Award in 2010, the Third Year Paper Award in 2008, Carole Keels Endowed Scholarship in Economics in 2008, the Jack Blicksilver Scholarship in Economics in 2007, and the Master of Arts in Economics Award in 2005. She also

provided service as the Vice President of the Economics Graduate Student Association in 2008-2009, and as a member of the Dean Search Committee in 2008-2009.

Ms. Jacobson's research is in applied microeconomics. Her areas of specialization are experimental economics, environmental economics, development economics, and public economics. She has research programs in social preferences, public good provision, preferences under risk and uncertainty, and conservation. She published a paper titled "Learning from Mistakes: What Do Inconsistent Choices over Risk Tell Us?" (coauthored with Ragan Petrie) in the April 2009 issue of the *Journal of Risk and Uncertainty*. Additionally, she published "Using Laboratory Experiments in Public Economics" (coauthored with James Alm) in the March 2007 issue of the *National Tax Journal*. Grants from the National Science Foundation and Georgia State University funded the research project "The Girl Scout Cookie Phenomenon: Peer Pressure in Grassroots Fundraising."

Ms. Jacobson has presented her research at seminars at Georgia State University and at other institutions, at the Camp Resources workshop, and at conferences held by the American Economic Association, Southern Economic Association, and the Economic Science Association. She has served as a referee for the journals *Agricultural Economics* and *Southern Economic Journal*.

Ms. Jacobson's permanent address is 5641 Rosser Court SW, Lilburn, GA 30047.