

Market design for altruistic supply: Evidence from the Lab

Robert Slonim
Carmen Wang

Working Paper 16-112



Market design for altruistic supply: Evidence from the Lab

Robert Slonim
The University of Sydney

Carmen Wang
Harvard Business School

Working Paper 16-112

Copyright © 2016 by Robert Slonim and Carmen Wang

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

Market design for altruistic supply: Evidence from the Labⁱ

Robert Slonim^a and Carmen Wang^b

March 2016

Abstract

Volunteer supply is widespread. Yet without a price, inefficiencies occur due to suppliers' inability to coordinate with each other and with demand. In these contexts, we propose a market clearinghouse mechanism that improves efficiency if supply is altruistically provided. The mechanism, a registry, combines aggregate demand information with suppliers' willingness to help, and invites volunteers to help only when excess demand occurs. We experimentally study three registries with stochastic high-stake demand and heterogeneous supplier costs. We find that all three registries improve efficiency dramatically; they eliminate unneeded costly help when demand is unexpectedly low and significantly increase supply during shortages.

JEL Codes: D47, D64, C93

Key Words: Market Design, Laboratory Experiments, Volunteering, Public Goods Provision

^a: Corresponding author: The School of Economics, The University of Sydney, Robert.slonim@sydney.edu.au, +61 2 9187 9036.

^b: Harvard Business School and Harvard Department of Economics, carmenwang@fas.harvard.edu

ⁱ: We are grateful for the generous financial support of the Australian Research Council, Discovery Project #1095222. We received excellent feedback during seminars at the University of New South Wales, New York University, the University of British Columbia, University of Sydney, and Harvard University and from participants at the Economic Science Association, North American meetings in Tucson AZ, the Conference on Economic Design in Istanbul and the Russell Sage Foundation Early Career Behavioral Economists conference in Chicago IL. We are grateful for the research assistance from Roxanne C. Dawidoski, Min-Taec Kim, Matthew Lilley, Michael Major, and Helia Marreiros.

1. Introduction

Economists are increasingly designing markets to improve efficiency. The designs usually focus on changing the incentives for self-interested agents.¹ This paper extends this literature to contexts where altruism and social preferences presumably drive suppliers' behavior.² We show in a lab study that a novel market design can improve efficiency in markets with altruistic supplies.³

Our design addresses inefficiencies due to information and coordination problems in contexts where there is no market price. In particular, markets with altruistic supply usually have no price because suppliers provide goods and services out of intrinsic motivation rather than for monetary compensation. These markets range widely from blood and bone marrow to animal shelters and soup kitchens. Without a price, volunteers do not know how much the market values their supply. This can lead volunteers to help when their cost to help exceeds the social benefit (e.g., when there is over supply) or not help when their cost is less than the social benefit (e.g., when there is under supply). Therefore, most contexts where supply is driven by altruism and social preferences lack effective ways to inform and coordinate volunteers to produce an amount that matches aggregate demand, resulting in too little or too much supply. Even when aggregate supply equals aggregate demand, without a market price a method is needed so that the lowest cost suppliers, *ceteris paribus*, will be the ones who help.

To highlight these inefficiencies, first consider the blood donation context that motivates this research. In the U.S., more than 16 million people donate whole blood every year. Since whole blood can be stored for only a maximum of 42

¹ Examples include labor market clearinghouses (Roth, 1984; Roth and Peranson, 1999), school choice systems (Abdulkadiroglu and Sönmez, 2003; Abdulkadiroglu et al, 2005), spectrum auctions (Milgrom, 2000) and kidney exchanges (Roth, Sönmez and Ünver, 2004, 2005a,b, 2007).

² We use the terms 'altruism' and 'social preferences' broadly to include altruism (Andreoni 1989, 1990) and outcome-based social preferences such as inequity aversion (Fehr and Schmidt 1999, Bolton and Ockenfels 2000) and efficiency maximization (Charness and Rabin 2002; Fisman, Kariv and Markovits 2007).

³ Approximately 26% of the US population volunteer annually with a value of \$173 billion (Independent Sector 2010).

days, and hospitals often demand blood no more than seven days old, the timing of blood supply is critical. Slonim, Wang and Garbarino (2014) observe that most developed countries experience shortages during the winter. However, these shortages are unlikely to be due to a lack of altruism, since large surpluses occur at other times, often after major disasters, when suppliers appear to incorrectly infer a large increase in demand. While this observation indicates that suppliers are responsive to perceived market demand, it also indicates that during normal times, people may be inactive due to incorrect beliefs of low demand or that others have provided the needed supply. If donors could be perfectly informed and coordinated, supply during normal times could be higher while over-supply during disasters could be eliminated.⁴ These inefficiencies can also occur in other volunteer contexts, such as an over-supply of volunteers at natural disaster locations, and even in monetary donations; for instance, after major disasters, people can donate too much money to a specific cause that constrains charitable organizations from spending donations on other lesser-known needs.⁵

We propose a new mechanism to address these inefficiencies. Our design improves efficiency to the extent that people have social preferences and are willing to donate the good or service based on needs.⁶ We introduce a registry in these contexts so that in addition to donating directly, an individual can register

⁴ When there is unmet demand, inefficiencies occur due to the gap between the value of each potential recipient's gain (e.g., saved life or improved health) and the lower costs of the un-provided supply (e.g., the time and discomfort to donate). When there is excess supply, inefficiencies occur due to the wasted costs associated with collecting the surplus supply (e.g., the value of donor's time donating and the costs to collect, store and destroy unneeded blood). One unit of unmet demand in this context (e.g., a lost life) is presumably much larger than one unit of excess supply (the wasted time and costs to collect an extra unit of unneeded blood). We designed the lab study with a similar imbalance in which a unit of excess demand results in a much greater efficiency loss than a unit of excess supply.

⁵For example, the Guardian: "Don't rush to Nepal to help. Read this first."

<http://www.theguardian.com/commentisfree/2015/apr/27/earthquake-nepal-dont-rush-help-volunteers-aid>.

For another example, Gross (2005) notes that donations to many international AID organizations dried up after the 2005 Tsunami. See also ProPublica.org: "How the Red Cross Raised Half a Billion Dollars for Haiti and Built Six Homes." <https://www.propublica.org/article/how-the-red-cross-raised-half-a-billion-dollars-for-haiti-and-built-6-homes>.

⁶ In contrast, Kessler and Roth (2012) assume self-interested agents to design a registry mechanism that gives priority to members if they need help. They show experimentally that priority improves welfare.

with a central coordinator and express their willingness to help. Registered donors will receive an invitation from the central coordinator if and only if their help is needed. In this way, receiving an invitation signals that there is unmet demand and importantly, other donors have not been invited to fill this particular demand; when not invited, it signals that there is no demand or others have been invited. Therefore, by following the invitations, registered donors can make sure to donate when needed and not to donate when there is no need. The central coordinator only needs to work out the optimum number of people to invite at any given time. If people only care about their own monetary payoffs, they will have no incentive to help with or without a registry. However, for people with social preferences, the registry informs the suppliers of unmet demand and increases their expected benefits of helping by increasing the likelihood that their help is needed.⁷ The registry thus increases efficiency by providing a mechanism to coordinate voluntary supply to match demand. Additionally, the registry can coordinate suppliers using their stated preferences so that those with the highest net utility to help (*ceteris paribus* lowest costs) are the ones to help.⁸

In the field, our mechanism only requires an organization (but not individual suppliers) to observe the aggregate demand for the volunteer good or service. This requirement seems reasonable in many contexts; for example, in most countries there is a central collection agency for blood products and organs and most communities have organizations for specific volunteer needs. A few examples further suggest a registry or equivalent mechanism increases welfare. Denmark, one of the few developed countries to have a national blood donor registry, has

⁷ Consistent with helping less as the risk of wasted help increases, Gneezy, Keenan and Gneezy (2014) find that donors give less money to fundraising campaigns the more their donations are used for overhead rather than for those in need.

⁸ Other approaches to address blood shortages include using material incentives (Goette and Stutzer, 2008, Lacetera, Macis and Slonim 2012, 2013a,b, 2014), reducing donors' time waiting (Craig et al. 2015) and using unsolicited gifts to generate reciprocity (Garbarino, Slonim and Wang, 2013). Another solution is to introduce a price for donating blood. However, this is currently prohibited with longstanding debate of ethical and practical considerations (Titmuss, 1970; Lacetera et al 2013b; Roth, 2007; Becker and Elias, 2007).

one of the world's highest per capita donation and transfusion rates.⁹ Slonim et al (2014) present evidence from a field experiment in Australia showing that blood donors are very likely to join a registry (73% of eligible donors who were invited joined) and that during shortages registry members are 40% more likely to donate compared to control subjects who were not invited to join the registry.

A few registries and proposed registry designs already exist in thin markets with the objective to effectively match suppliers and recipients. Examples include the bone marrow registry, where the likelihood of matching a donor and recipient can be less than 0.01% (Becker and Elias, 2007), and the blood donor registry prior to being able to store blood, when donors and recipients needed to be in the same place at the same time for transfusions (Slonim et al 2014).

This paper experimentally shows that our registry design improves efficiency when the market is thick by coordinating altruistic suppliers' actions with each other and with demand. While field examples suggest registries in thick volunteer markets can increase efficiency, it is difficult to assess the effects since supplier costs and recipient benefits are unobserved. In the lab, these parameters are directly observable. Moreover, in the lab we can study multiple registry designs in the same setting to identify which components of the design are important.¹⁰ Finally, the lab also lets us directly compare the registries to an alternative approach that provides suppliers with the aggregate demand information, as organizations and media may provide during (severe) shortages. If coordination among suppliers is the main source of market inefficiency, then providing demand information should improve efficiency less effectively than introducing a registry.

⁹ To our knowledge, the Netherlands and Finland also have variants of coordination systems equivalent to a registry design. For Denmark, see <https://bloddonor.dk/>.

¹⁰ Kagel and Roth (2000) present one of the earliest lab studies testing market design. They examined the stability of matching markets. The most closely related market design research using the lab is Kessler and Roth (2012, 2014). They examine whether giving priority to receive organ donations if someone joins the registry affects organ donations. See Roth (2012) for a survey on the use of experiments in market design.

We use the lab to examine three closely related registry designs. In our setup, a market is a 10-person group. In every round, a random draw divides group members into the demand side, those who need help, and the supply side, those who can provide help. In the **baseline condition** (without a registry), suppliers simultaneously choose to help or not help knowing only their own cost, but not knowing how many subjects need help, the cost to help among others who can also help, nor anyone else's decision to help or not help.¹¹ If a subject helps, he incurs his cost to help but never learns if his help was needed. This setup captures key aspects of many markets in which volunteers provide the supply. For instance, blood donors receive little information about the demand, other potential donors' actions or feedback on the usage of their own donations.

In the **registry conditions**, subjects are able to help directly in the same way as in the baseline, but can alternatively join a registry. If they join, they will be asked to state their willingness to help, which will help determine the order in which they are invited to help. Subsequently, registry members will only be asked to help if their help is needed. Thus, registry members know that if invited, their help will definitely save someone who needs help. The registry conditions thus provide a method to coordinate suppliers' decisions to avoid donating when other suppliers have already provided enough supply to meet demand, to indicate when other suppliers have not provided enough to meet demand, and to coordinate which suppliers should help if not all suppliers are needed.

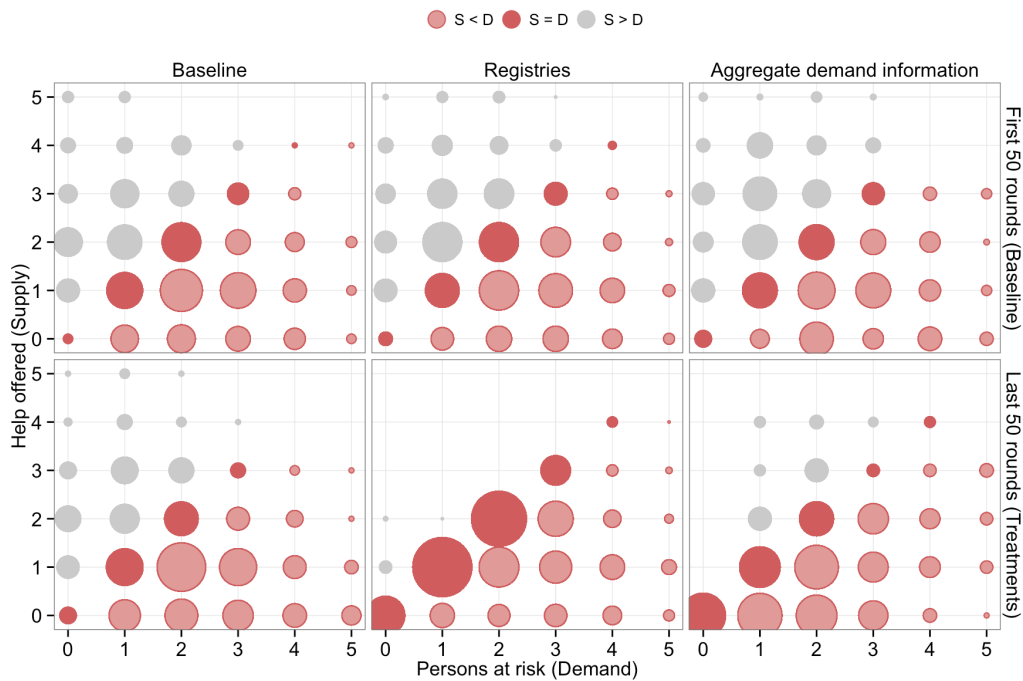
We also examine whether giving suppliers demand information improves efficiency. This '**aggregate demand information**' condition is identical to the baseline except that suppliers are informed of the exact demand before deciding whether to help. This condition captures what organizations do when they make announcements about their needs (e.g., announcing a blood shortage). However,

¹¹ The distributions of the number of demand and costs are common knowledge.

the coordination problem remains because suppliers do not know how many other suppliers will help nor which suppliers should help if not all suppliers are needed.

We implement a difference in difference design where all markets start in the baseline condition for the first 50 rounds. In the last 50 rounds, the market rules change according to the assigned treatment conditions: the baseline, the registry conditions, and the aggregate demand information condition.

Figure 1. Demand - Supply Distribution



The results show that welfare increases dramatically in the registry conditions compared to the baseline and aggregate demand information conditions. **Figure 1** displays weighted bubble plots of the distribution of demand (horizontal axis) and supply (vertical axis). The bubble size is the proportion of market level observations within each condition. The **top half of Figure 1** shows that without any market intervention, coordination failures of both over- and under-supply are prevalent; Supply (S) equals Demand (D) in only 20% of the market observations. The **bottom half of Figure 1** shows a dramatic reduction in coordination failures

when the registries are introduced; $S = D$ increases to 54% of the registry market observations but remains at only 18% in the baseline condition. Not only is supply higher in the registry than baseline conditions, but also 99% of the supply of help in registry conditions was used compared to only 67% in the baseline. With aggregate demand information, suppliers helped more the greater the demand, and efficiency improved when either extremely high demand or no demand was revealed, i.e., when help was almost certainly needed or not needed at all. However, when coordination among suppliers was needed, i.e. in normal demand situations, there was no increase in supply and we observe the same level of under- and over-supply occurred as in the baseline.

We examined three registries that produced the same market level results, but individual responses varied in expected directions based on the registry rules. All registries initially invited one registry member to help for each person needing help, prioritizing who was asked to help based on suppliers' choice of their willingness to help. If an invited registry member declined to help, the **'invitations-once'** and **'adaptive'** registries did not invite more members to help, whereas the **'sequential'** registry invited additional members until there was no more members to invite or all demand was met. Since the decision to help was less pivotal for sequential than invitations-once and adaptive registry members, we anticipated and found that suppliers were less likely to help when invited in the sequential registry, but more likely to join the sequential registry. The two effects balanced out, resulting in the same level of market efficiency. The **'adaptive'** registry adds to the invitations-once registry by giving priority to past helping behavior when determining which registry members to invite to help. However, we find no significant differences in individual level decisions between the invitations-once and adaptive registries, suggesting suppliers were not systematically joining and not helping.

Finally, the registries further decreased coordination failures by sorting help towards subjects with lower costs. We found that subjects effectively sorted themselves so that there was a significant negative correlation between suppliers' costs and those who were invited to help.

Overall, our study offers two major contributions. First, we show that by assuming people are at least partially motivated by social preferences, there are opportunities for novel market designs. Second, we show that redesigning the environment can have a major impact on volunteering without changing people's preferences or costs. For instance, a common attribution for blood product shortages is that people are not sufficiently altruistic, yet our experiment shows that more people help (and receive help) under a more effectively designed system, given the same distribution of social preferences and costs of helping.

The paper is organized as follows. Section 2 presents the experimental details, procedures, conjectures and efficiency measures, Section 3 presents the results and Section 4 concludes.

2. The Experiment

2.1 Baseline and Aggregate Information Conditions

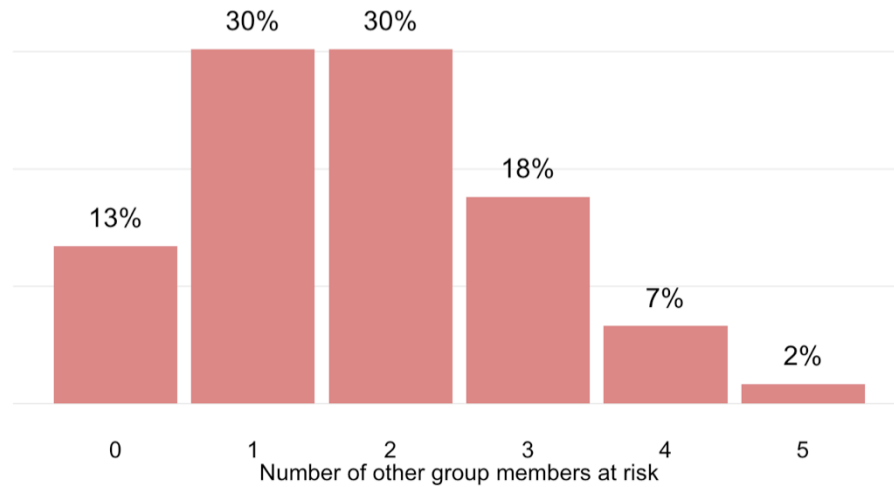
In each session, subjects were randomly and anonymously assigned to 10-person groups whom they participated with for the entire session. A session consisted of instructions and review questions for the baseline condition, 50 rounds of the baseline condition, further instructions and review questions for the treatment conditions, and then 50 rounds of the treatment conditions. In all conditions subjects knew the timing and structure of the session, but did not know the treatments in the last 50 rounds until after completing the first 50 rounds.

In each round each subject has a \$20 endowment. Every round in the baseline condition proceeded in three stages, and all procedures were common knowledge:

- 1. Determining demand and supply:** Each round began with an iid draw that determined who was ‘at risk’ (i.e., the demand for help) and who was ‘safe’ (i.e., the potential supply of help). For each subject, there was an 80% chance of being safe (20% chance of being at risk). Subjects who were safe were informed of their cost to help, c_i , which was iid on the uniform distribution from \$2 to \$16 in \$0.10 increments. Subjects were not informed of other subjects’ cost or how many other subjects were safe.
- 2. The supply decision:** Safe subjects had to privately decide to help or not help given their costs. If a subject chose not to help, he would earn his \$20 endowment. If a subject chose to help, he would earn his endowment minus his cost to help, $\$20 - c_i$. Subjects at risk did not make any decisions.
- 3. Determining who gets saved:** Let H and R be the total number of subjects who helped in stage 2 and who were at risk, respectively. If $H \geq R$, then all subjects at risk were saved. If $H < R$, then H of the R subjects at risk were saved, with each one having the same chance (equal to H/R). At risk subjects were informed individually whether they were saved; they received their \$20 endowment if they were saved or \$0 if they were not saved. Safe subjects who helped were not informed of whether their help saved anyone, and no subject was informed of anyone else’s decision or how many subjects were saved.

We used context-rich language in the instructions and on all decision screens. We referred to subjects as ‘safe’ and ‘at risk’ depending on their status. We referred to the choices that subjects had as ‘help’ and ‘not help’, and we referred to the outcome in which choosing to help could prevent an at risk subject from losing her endowment as ‘saving’ her. Experimental studies often avoid context-rich language; however, we are explicitly interested in studying volunteer contexts where people would naturally consider their actions as helping (or not helping) others, and would naturally identify with the roles of some people as being at risk (or not at risk).

Figure 2. Distribution of the Number of Other Group Members at Risk



The aggregate demand in each round was simply the number of subjects at risk. **Figure 2** displays the distribution of aggregate demand from the perspective of a potential supplier (i.e., a safe subject). For a safe subject, there are nine other subjects who each had an 80% chance of being safe and a 20% chance of being at risk. Therefore, there is a 13% chance that no subject is at risk (0.8^9), a 30% chance of exactly one of the other nine subjects being at risk ($9 * 0.8^8 * 0.2$), ..., and less than a 0.2% chance of more than 5 other subjects being at risk. We showed subjects **Figure 2** to not only provide them with a visual image to help them understand the distribution, but also so that it would be common knowledge that all subjects saw this display of the distribution.

2.2 Treatments

In the **Baseline** condition, subjects played the last 50 rounds following the identical rules used in the first 50 rounds. We included this condition to measure any potential changes in behavior that could be due to playing an additional 50 rounds independent of treatment effects; extensive experimental evidence shows that cooperation often declines with repetition in finitely repeated public goods games (e.g. Andreoni 1988).

2.2.1 Aggregate Demand Information (ADI)

The **Aggregate Demand Information (ADI)** condition was identical to the Baseline condition with one exception. In the first stage, safe subjects were also informed of the market demand realization, R (i.e., the total number of subjects at risk). This information provision was common knowledge. Thus, safe subjects knew the aggregate demand when choosing to help or not help in the stage 2 supply decision.

We included the ADI condition for two reasons. First, in many contexts some demand information is provided when shortages occur. For instance, blood collection agencies often publicly announce shortages when they occur. Thus, the ADI condition provides a benchmark to an approach commonly used in volunteer contexts. Second, the ADI condition will highlight the coordination challenge. In particular, there are two realizations of aggregate demand in which there is no coordination problem ($R = 0, R \geq 5$) and four realizations in which coordination issues remain ($1 \leq R \leq 4$). When no subjects are at risk, safe subjects know for sure that their help is not needed and when five or more subjects are at risk, safe subjects know for sure that their help will save someone, but when there are one to four subjects at risk, safe subjects will not know whether providing help will be needed. When these $R = 1$ to 4 realizations occur, which occurs 85 percent of the time (**Figure 2**), both under supply (lives not saved) and over supply (wasted help) are possible. In contrast, the registries provide a mechanism to coordinate supply for all realizations of demand.

2.2.2 The Registry Conditions

In all registry conditions, *safe* subjects were also given an option to join a registry in the stage 2 decision:

2R. The supply decision: Once subjects were shown their cost, they could help or not help (identical to the baseline condition) or they could join the registry and state their willingness to help from 3 (most willing), to 2 to 1 (least willing).

The registry works as follows: let H^d , J and R be the number of subjects who helped directly (i.e., helped without joining the registry), joined the registry, and the number at risk, respectively. The excess demand (R^E) after suppliers made their initial decision (not help, help directly or join the registry) was $R^E = \max\{0, R - H^d\}$, $H^d \geq 0$. The registry then invited registry members to help as follows:

- If $R^E = 0$, no registry member was invited to help.
- If $0 < R^E < J$, R^E of the J registry members were invited to help.
- If $R^E \geq J$ all J registry members were invited to help.

If a registry member was invited to help, the payoffs to help or not help were identical to helping or not helping directly (outside of the registry); if a member chose not to help, he would earn his \$20 endowment, and if a member chose to help, he would earn his endowment minus his cost to help, $\$20 - c_i$. However, in stark contrast to helping in the baseline and the ADI condition with $R < 5$, registry members knew for sure that if they helped they would save a subject at risk.

We examined three variations of the registry that differed in determining which members to invite to help when there were more members than excess demand ($J > R^E$). The registries operated identically if $R^E = 0$ (no member was invited to help) or if $R^E \geq J$ (all members were invited to help). In the **Invitations Once** registry, the subjects were ranked based solely on their stated willingness w ($w = 1, 2, 3$). Letting J_w be the number of members with willingness w , so $J_1 + J_2 + J_3 = J$, we used the following rule (and this was common knowledge):

If $J_3 \geq R^E$, randomly choose R^E members among those who stated $w = 3$.

If $J_2 + J_3 \geq R^E > J_3$, choose all members who stated $w = 3$ and randomly choose $R^E - J_3$ members among those who stated $w = 2$.

If $J_1 + J_2 + J_3 \geq R^E > J_2 + J_3$, choose all members who stated $w = 2$ and $w = 3$, and randomly choose $R^E - J_3 - J_2$ members among those who stated $w = 1$.

The registry thus let subjects sort on their preferences to provide help. The randomly determined costs proxy for unobserved preferences (similar to Kessler and Roth 2012).¹² In our study, *ceteris paribus*, letting subjects state their willingness provides a mechanism to sort into being more likely to be invited to help the lower their costs are, and consequently for the help to be provided by those with the greatest preference to help.

The registries differed in what happened when a registry member who was invited to help chose not to help. In the **Invitations Once** registry, no more members were invited to help (even if there were members who had not been asked). In the **Sequential** registry, registry members who had not been invited to help initially would be invited next according to the same invitation rules above.¹³ This procedure would continue until either everyone at risk was saved or there were no more registry members to invite.

We included the **Invitations Once** and **Sequential** registries to measure the impact of whether the help decision was or was not pivotal, respectively. In the Invitations Once condition, the decision to not help would prevent anyone else from helping and would thus guarantee that a subject would not get saved, while

¹² Kessler and Roth 2012 use costs in a similar manner to proxy for unobserved preferences in their lab study of bone marrow registries examining the effects of providing priority

¹³ In the Sequential registry, the registry never indicated whether anyone else had been asked (and said no) before a member received his invitation. The timing of decisions was often extremely quick after a few rounds had been played, and delays before receiving a registry invitation could be attributed to other subjects taking longer to decide to join in the initial supply decision, thus it would be unclear to subjects whether they had received an initial invitation or an invitation after some other member had declined to help. This setup matches how registries operate outside the lab; someone invited to help would not know whether someone else had been asked previously.

in the Sequential condition, someone else could still help a subject in need. We included the Sequential registry not only since it mimics some existing registries (e.g., bone marrow registries), but importantly because we anticipate distinct behavior in the Sequential and Invitations Once conditions. We anticipated that Invitations Once registry members will be more likely to help if invited than those in the Sequential registry given their greater pivotal impact. We further anticipated that subjects in the Invitations Once condition will recognize the greater consequence if they join and are subsequently invited to help, and will thus be less likely to join the registry than subjects in the Sequential condition.

The **Adaptive** registry was identical to the **Invitations Once** registry, except that the Adaptive registry augments which registry members are invited to help when $J > R^E$ to take into account past behavior. In particular, the Adaptive registry gives each subject a status for their past behavior, and invites members with the highest status, then second highest status, etc. until it has identified R^E members to help. Among those who tied with the same status, the Adaptive registry uses the willingness rules used in the Invitations Once condition to determine who to ask to help. All subjects began with a status of 1000. The status s_{it} of each subject i in round t was updated each period as follows:

$$\begin{aligned}
 S_{i(t+1)} &= s_{it} - 10 && \text{if } i \text{ joined the registry, was invited to help, but chose not to help} \\
 S_{i(t+1)} &= 1,000 && \text{if } i \text{ helped directly or joined the registry, was invited and helped} \\
 S_{i(t+1)} &= s_{it} && \text{if } i \text{ chose not to help directly or joined the registry but was not invited to help}
 \end{aligned}$$

Thus, a subject's status fell if he joined the registry but did not help when invited, and was restored to its initial level if he helped. To the extent that there might be subjects who would join a registry but not help if invited, the Adaptive registry

would improve efficiency over the Invitations Once registry by sorting against inviting these subjects.

All registry procedures were common knowledge except that in the Adaptive condition we did not explain exactly how subject's past choices would affect the likelihood that the registry would ask them to help. We only told subjects that if they joined the registry and were invited to help that, "... if you do not help, that may reduce your chance to be invited in the future, and if you help, that may help your chance to be invited in the future." We designed it to mimic organizations that use past behavior to alter rules but often do not explicitly state how they use the past behavior.

2.3 Experimental Procedures

Subjects were recruited from a student population who had volunteered to receive email invitations regarding economic experiments using ORSEE (Greiner, 2015). The study was advertised as 'economic decision-making with others', and indicated sessions would take up to two hours. The experiment was programmed in zTree (Fischbacher, 2007).

The instructions and review questions for all of the conditions and the survey are in **Appendix B**. When subjects arrived they were randomly assigned seats and randomly and anonymously assigned to a 10-person group to play all 100 rounds with (which was common knowledge). The initial instructions informed all subjects that they would play 50 rounds in the baseline condition, receive further instructions, and play 50 additional rounds with the same group, but they were not told anything further about the last 50 rounds. After completing the first 50 rounds, all groups received further instructions and review questions for the condition they were randomly assigned to: 1) Baseline, 2) ADI, 3) Sequential registry, 4) Invitations Once registry and 5) Adaptive registry.

Subjects were given a hard copy of the instructions for the first 50 rounds that they could review at any time. The experimenter read these instructions aloud while the subjects could follow along, and their computers would show examples of the decision screens and how their payoffs would be calculated. The review questions were then given on their computers. After the first 50 rounds, hard copies of the instructions for the last 50 rounds were distributed, the experimenter again read these instructions aloud, and new review questions were given on their computers.¹⁴ The same experimenter read the instructions in every session.

At the end of the 100 rounds and before the final survey, an experimenter rolled a large dice in front of all subjects to randomly select two rounds that determined subjects' payoffs, with one round from the first 50 rounds and another round from the last 50 rounds. The payoffs were stated directly in Australian dollars, e.g. \$20 endowment (one Australian dollar was approximately 1.03 US dollars at the time of the experiment).

Subjects received payment based on the outcome of the two randomly selected rounds plus a \$10 show up fee and up to \$5 for answering review questions correctly. We incentivized the review questions to encourage subjects to pay close attention to the instructions. We randomly selected 2 review questions, one from the first 50 rounds, worth \$3 if answered correctly, and one from the last 50 rounds, worth \$2 if answered correctly. We did not reveal which questions were selected until all rounds were completed to avoid potential wealth effects. On average, subjects answered over 90% of the review questions correctly. The average earning was \$49.69 with subjects earning \$15 in a few cases (when the subjects were at risk and were not saved in either round chosen) to \$55. Subjects were paid in cash at the end of each session.

¹⁴ In order to parallel the treatment conditions that included three pages of new instructions and review questions, in the control condition for the last 50 rounds we included instructions as well. These instructions reminded subjects of the rules and the review questions were different than those asked in the first 50 rounds.

A total of 580 subjects participated in the experiment with each subject participating exactly once. There were 11 groups in each condition except the Sequential condition, which had 14 groups. Each condition had three sessions with three groups (except Sequential which had four sessions with three groups) and one session with two groups. All groups in a session were in the same condition. We ran all 21 sessions in two consecutive weeks during Apr-May 2012 at the University of Sydney Economics Decision Lab. We balanced the conditions across the day of week and the time of day.

2.4 Outcome Measures and Efficiency Benchmarks

To assess the effectiveness of the registries, we define the outcome measures and overall efficiency achieved for each group based on the group's realized payoffs compared to (1) a population that never helps, (2) a population that maximizes the group's *ex ante* total expected payoff in the baseline condition, (3) a population that maximizes the group's *ex ante* total expected payoff in the ADI condition, and (4) the *ex post* maximum possible group payoff, where the maximum possible payoff occurs when the number of subjects that help equals the number of subjects at risk (or all help if $R > 5$) and those who help have the lowest costs among those who are safe. As we describe below, the lab registry conditions with three levels of willingness are sufficient to allow a population that wants to maximize the group's payoff obtain over 99% of the maximum possible payoff.

We first define the outcome measures. Let r_{gt} be the number of persons at risk (total demand) for group g and round t , and $h_{igt} \in \{0,1\}$ be an indicator variable of a group member i 's decision to help with cost c_{igt} when i is not at risk ($h_{igt} = 0$ when i is at risk). Our main outcome measures are **total supply** h_{gt} , **persons saved** s_{gt} , **help wasted** h_{gt}^w , and **total group payoffs** π_{gt} for group g in round t :

$$h_{gt} = \sum_{i=1}^n h_{igt} \quad (1)$$

$$s_{gt} = \min(h_{gt}, r_{gt}) \quad (2)$$

$$h_{gt}^w = h_{gt} - s_{gt} \quad (3)$$

$$\pi_{gt} = (n - r_{gt} + s_{gt})e - \sum_{i=1}^n h_{igt}c_{igt} \quad (4)$$

where $n = 10$ persons per group, e is a constant of \$20 endowment for each individual.

We compare outcome changes for each group from the first to the last 50 rounds for each risk level $r = 1, 2, \dots$. Let $I_{grt} \in \{0,1\}$ be an indicator variable such that $I_{grt} = 1$ if group g has demand r in round t , and 0 otherwise. The change in the group payoff is $\Delta\pi_{gr}$:

$$\Delta\pi_{gr} = \frac{\sum_t \pi_{gt} I_{grt}}{\sum_t I_{grt}} \Big|_{t>50} - \frac{\sum_t \pi_{gt} I_{grt}}{\sum_t I_{grt}} \Big|_{t\leq 50} \quad (5)$$

And the change in **total supply** Δh_{gr} , **persons saved** Δs_{gr} , and **help wasted** Δh_{gr}^w are similarly defined. For changes over all risk levels, we have for group payoff $\Delta\pi_g$.¹⁵

$$\Delta\pi_g = \sum_r b(r; n, p) \Delta\pi_{gr} \quad (6)$$

where $b(r; n, p)$ is the binomial probability density function for risk level r , with $p = 0.2$ of being at risk and $n = 10$ persons in a group. Changes in **total supply** Δh_g , **persons saved** Δs_g , and **help wasted** Δh_g^w are defined similarly. *Each group* thus provides a single measure for the change in outcomes overall, and for each

¹⁵ We weight the overall efficiency by the theoretical distribution rather than the empirical distribution since using the empirical distribution has small differences in the frequency that groups realized different demand levels, and these small differences could bias the overall outcomes to the extent that the outcomes differ across demand levels. Using the empirical distribution would not change any of our results, however, most likely since we had enough observations that the empirical distribution was similar to the theoretical distribution.

level of demand, that we use as our independent units of observation. The first 50 periods provide a baseline measure that allows us to control for differences in preferences that could lead to different outcomes across the treatments. Given our large subject population (with 110 or more subjects per condition), the initial differences are relatively small and not significant.

Our critical tests for the registry effects are thus difference-in-difference analyses in which we compare the change in outcomes from the first to the last 50 rounds in the registry conditions to the change in outcomes in the control condition from the first to last 50 rounds (in which subjects remained in the baseline condition). Our core test for the effect of the registries is the change in overall payoffs $\Delta\pi_g$ in the registry conditions compared to the baseline condition.

We now consider four benchmark payoffs to assess the relative efficiency of the overall payoff for each group. Similar to *Equation 4*, let the benchmark payoff for the overall distribution of demand levels in a group be:

$$\Pi_{benchmark} = \sum_r b(r; n, p) [(n - r + s)e - \sum_i h_i c_i] \quad (7)$$

where the binomial distribution $b(r; n, p)$, and variables n , r , s , e , h_i and c_i are defined as above. First, if no subject helps, $h_i = 0$ for all i , it can easily be shown that a group's expected payoff Π_0 is \$160. Second, the maximum possible payoff Π_{max} for a population occurs when (1) the number of subjects who help exactly equals the number of subjects at risk (or all safe subjects help if more than half of the subjects are at risk) and (2) the subjects who help have the lowest costs among those who can help. When this occurs, the expected maximum payoff is \$188.98 (based on the average of one million simulation draws from our distribution). Although our registry design cannot fully achieve this expected payoff even with a population in which every subject wants to maximize the population's total payoff, simulations show that even three levels of willingness following a simple cutoff strategy that divides the cost range in thirds (join with the highest

willingness if costs are less than \$6.67, the second highest willingness if costs are between \$6.67 and \$11.33, and everyone else joins with the lowest willingness) results in a population average payoff within 0.5 percent (\$188.11) of the maximum possible. In other words, our registry design with three willingness levels is sufficient to allow close to the maximum possible payoffs. We thus define the group level change in efficiency from the first to the last 50 rounds due to the treatment effect as follows:

$$\Delta E_g = \frac{\Delta \pi_g - \Pi_0}{\Pi_{max} - \Pi_0} \quad (8)$$

where $\Delta \pi_g$ is defined in *Equation 6*. Each group thus provides a single measure for the change in efficiency from the first to the last 50 rounds relative to the maximum possible and a group in which no one helps.

The last benchmarks of efficiency are for the baseline and the ADI conditions. In the baseline condition, we consider a benchmark payoff in which subjects maximize the population's *ex ante* expected payoff over the demand r and cost c_i distributions (since subjects do not know the realized demand or anyone else's cost). This requires subjects to help if and only if their cost is less than **\$6.36**.¹⁶ Based on this cost cutoff, the population's expected payoff is **\$177.43**. In the ADI condition, maximizing the population's expected payoff depends on the number of subjects at risk (since subjects are informed of the demand): with no one at risk, no one helps, if 1, 2, 3 or 4 subjects are at risk, then the optimal cutoffs to help are \$4.41, \$6.29, \$8.44, and \$11.18, respectively, and for five or more at risk, everyone safe will help. The expected population payoff overall in this case is **\$182.99**. Overall, using the optimal cutoff costs to maximize the population's payoffs in the baseline and information conditions result in **60.1%** $((177.43 - \Pi_0)/(\Pi_{max} - \Pi_0))$ and

¹⁶ It is also possible to show that this cutoff is an equilibrium given several assumptions.

79.3% $((182.99 - \Pi_0)/(\Pi_{max} - \Pi_0))$, respectively, of the maximum possible. In both the baseline and ADI condition, unlike the registry conditions, both wasted help (for 0 to 4 subjects at risk) and lives not saved (for 1 or more subjects at risk) remain possible even if subjects follow the optimal cutoff costs to maximize the population's payoffs. Thus, the registry conditions offer the potential for higher population payoffs by coordinating suppliers so that supply matches demand, and by sorting suppliers towards those with the lowest costs.

2.5 Hypotheses

If no subjects receive utility from helping others, no help will be provided and average payoffs (\$160), lives saved (0), wasted help (0) and efficiency (0%) will be identical across conditions. If subjects attempt to maximize the population's payoff, we have:

H1 (main hypothesis): Average payoffs and lives saved will be higher in the ADI condition than the baseline, and even higher in all of the registry conditions than in either the baseline or ADI conditions, and wasted help will be lower in the registry condition for all demand levels, and for $R = 0$ in the ADI condition.

H2: Subjects will be more likely to help, ceteris paribus, in the ADI condition the more subjects are at risk, and no subjects help if $R = 0$. When $R = 2$, the total supply, lives saved, wasted help and average payoffs will be almost identical to the baseline since the optimal cutoff cost (\$6.29) to maximize payoffs is almost identical to the baseline optimal cutoff (\$6.36). As the number of subjects at risk is further from $R = 2$, the expected payoffs will be increasingly greater in the ADI than baseline condition, with more subjects helping and saving more lives as R increases, and with fewer subjects helping and less wasted help as R decreases.

Hypothesis H1 and H2 follow directly from the above discussion assuming all subjects prefer to maximize the group's total payoff. If some but not all subjects prefer to maximize the population's payoff, we further hypothesis that:

H3: Registry members will be more likely to help, conditional on joining the registry, in the Invitations Once and Adaptive registries than in the Sequential registry. Note, if all subjects prefer to maximize the group's payoff, then there would be no difference and all members would help if invited in the registry. However, we anticipate that some members may decide not to help once in the registry, and we conjecture that this behavior will occur more in the Sequential registry since there is some chance for another member to help. We further anticipate that if this occurs, more subjects will join the sequential than the other two registries since they will be less concerned about having to help if invited.

H4: Registry members will sort into willingness based at least in part on their costs; subjects with higher costs will choose lower willingness levels.

3. Results

Throughout this discussion we focus on how behavior and outcomes changed from the first 50 rounds when all subjects participated in the baseline condition to the last 50 rounds when subjects either repeated the baseline condition, were given aggregate demand information or were in one of the registry conditions. **Section 3.1** compares the change in the total supply of help, then lives saved, wasted help and payoffs as a function of the number of subjects at risk, and finally the change in the efficiency. We first present the result graphically to highlight the key results, then present regressions to show the statistically significant effects. **Section 3.2** examines how individual decisions between the baseline and treatments and between the three registries differed.

For brevity here, we describe the subject characteristics (**Appendix Table A2.1**) and the realization of the random draws (number of subjects at risk and costs of the safe subjects for the first and last 50 rounds) by treatment (**Appendix Table A2.2**). Given more than 100 subjects per treatment, there are only small differences between treatments. Further, given our focus on outcomes based on changes from the first to last 50 rounds, any minor differences in characteristics between treatments net out. Also, given 100 rounds and 11 or more groups of 10

subjects, resulting in more than 10,000 draws per treatment for subjects at risk and more than 8,000 draws for costs, we also find only very minor differences in the distribution of demand or costs between treatments.

3.1 Market outcome and efficiency

3.1.1 Market outcome

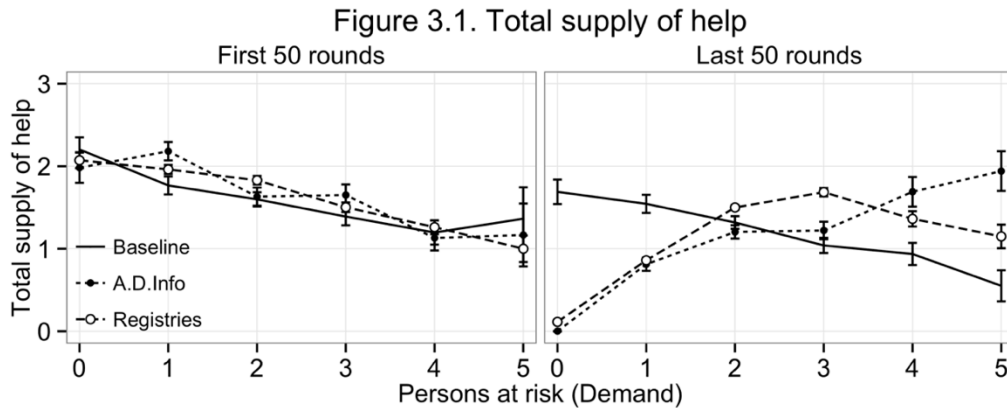


Fig. 3.1 shows the mean supply (with standard error bars) for each level of demand, using one observation per group. It shows that the total supply does not increase with the needed demand in the first 50 rounds. In all conditions, market supplies actually decrease as demand increases. This decrease follows logically given that subjects do not know the number of subjects at risk when deciding to help. If subjects follow a cutoff rule to determine if they offer help, then on average the likelihood of help h_i offered for each subject i in a group will be independent of r , and hence constant, across demand levels. However, since there are fewer subjects safe as r increases, the average amount of help offered per group will decline as r increases. In the last 50 rounds, our registry and information treatments successfully make aggregate supply upward sloping on average with regard to aggregate demand.

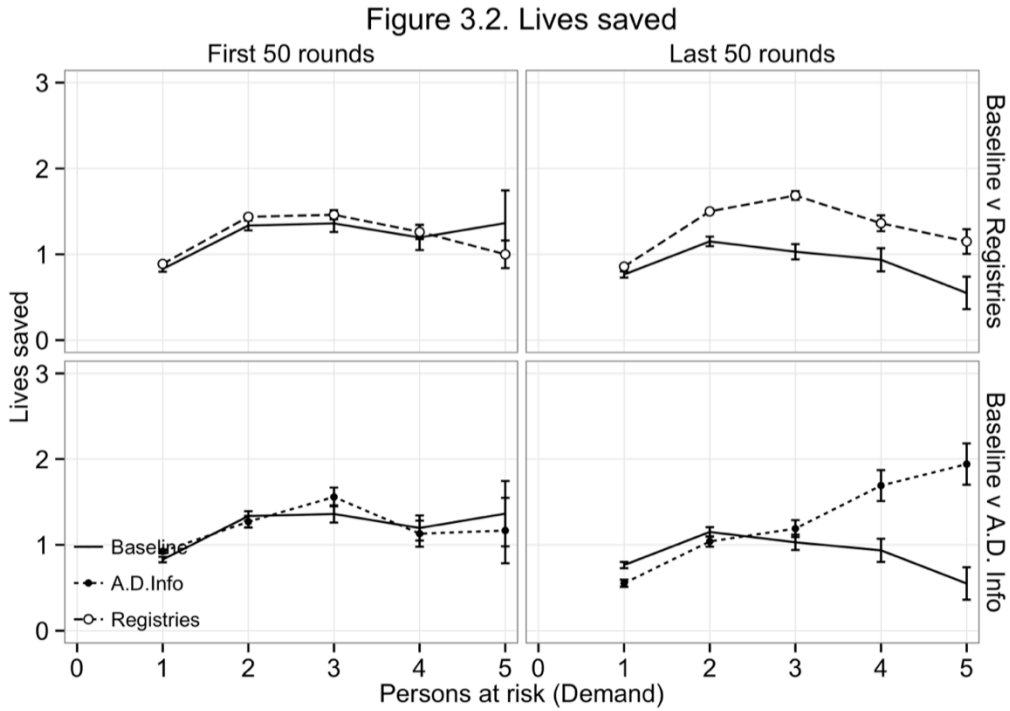


Fig. 3.2 shows that the average number of lives saved increased from $r = 1$ to 2, is similar for $r = 2$ and 3, and is slightly lower for $r = 4$ and $r = 5$ in the first 50 rounds for all conditions. As r increases there are opposing forces on the number of lives saved: there are more subjects at risk who can be saved, but there are also fewer subjects providing help. With our parameters, this led to subjects more likely to be saved from $r = 1$ to $r = 2$ (when the more-subjects-at-risk effect dominates the fewer-subjects-helping effect), and to less likely to be saved from $r = 3$ to 5 (when the fewer-subjects-helping effect dominates the more-subjects-at-risk effect). In the last 50 rounds, lives saved increases in the registries relative to the baseline for every level of aggregate demand greater than $r = 1$, while in the ADI condition, lives saved increased only for $r = 4$ and 5. When the demand is low, the information treatment is indistinguishable from the baseline when $r = 1-3$ and even slightly below baseline when $r = 1$. Comparing the information and registry treatments highlights the need for coordination, when there is little need for coordination ($r = 4$ and 5), information improves lives saved, but when

coordination is necessary ($r = 1-3$) the ADI condition performs similarly to the baseline and significantly worse than the registries.

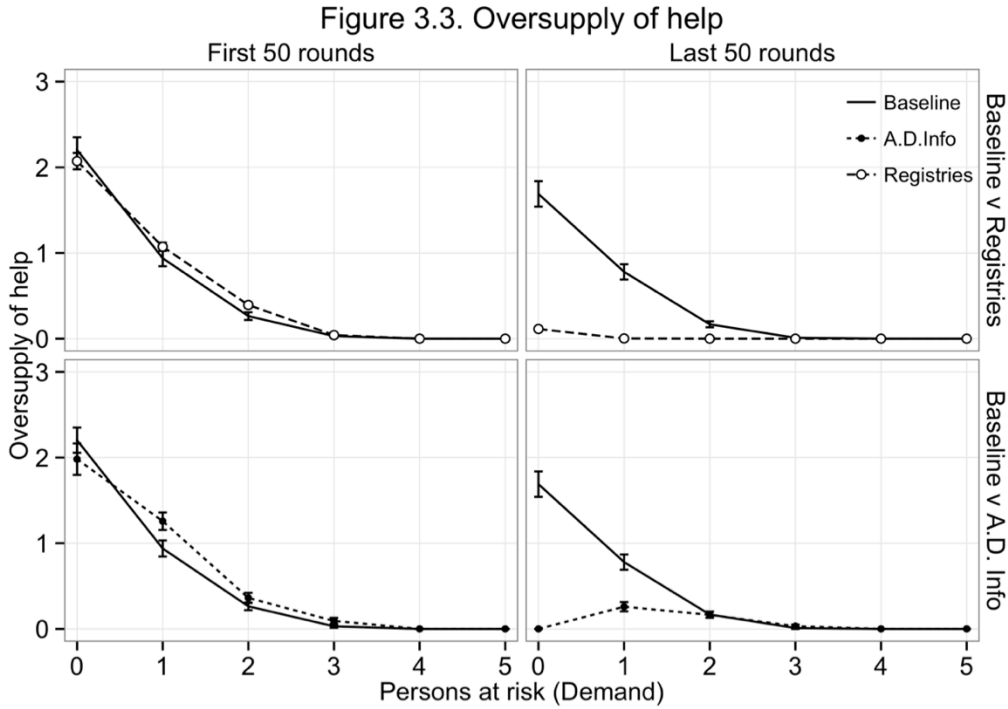


Fig. 3.3 shows that wasted help decreases from $r = 0$ to 3 and is almost nonexistent for $r > 3$ in the first 50 rounds for all conditions. The fall in wasted help follows logically as r increases, since fewer subjects are available to help and more subjects are at risk both reduce the wasted help. In the last 50 rounds, wasted help decreases by a (dramatically) larger amount in both the registry and ADI condition compared to the baseline when there were 0 or 1 subjects at risk. There is still slightly more wasted help in the information condition when $r = 1$ and 2 compared to the registry, even with the smaller number of lives saved compared to the registry conditions.

We estimate variations of the following model for the market level results:

$$y^* = \beta_0 + \beta_1 D_{Last50} + \beta_2 D_{ADI} + \beta_3 D_{Registries} + \beta_4 D_{Last50} D_{ADI} + \beta_5 D_{Last50} D_{Registries} + X'\gamma + \epsilon$$

where y^* is the outcome variable (lives saved in **Table 3.1** and wasted help in **Table 3.2**); D_{Last50} is a dummy for observations in the last 50 rounds; D_{ADI} and $D_{Registries}$ are dummies for observations in the aggregate demand information condition and the combined three registry conditions respectively; X include controls for rounds and differences in cost realizations in a group (see notes under the tables). Each group provides N_{grt} group-period level observations that depend on the number of periods t group g had r subjects at risk. We run Tobit regressions for each demand level censored between 0 and the maximum possible number of lives saved for **Table 3.1** and censored at 0 for wasted help in **Table 3.2**. If there is only one person at risk, we run a probit regression for lives saved (the one life was saved or was not saved).¹⁷

The group level regressions in **Table 3.1** show that the relative increase in lives saved in the registries compared to the baseline range from almost 0.5 (when $r=4$) to over 1.0 from the first to last 50 rounds and is significant for every level of aggregate demand for $r \geq 1$ ($p < .05$). **Table 3.2** show that the relative decrease in wasted help from the first to last 50 rounds is significantly greater in the registry than baseline condition for $r = 0$ or 1, and is not different otherwise. The regressions in **Table 3.1** and **Table 3.2** also show that the increase in lives saved from the first to last 50 rounds for $r > 3$ and decrease in wasted help for $r < 2$ are significantly different in the ADI than baseline condition. Appendix **Table A3.1b** and **Table A3.2b** further show that there are no significant differences between the three registries in terms of lives saved and help wasted, respectively.

¹⁷ Examining only the first 50 rounds, group level linear regressions presented in appendix **Tables A3.1a, A3.2a** for each level of risk that control for round (clustering s.e. at the group level) and the five lowest costs among the safe subjects robustly show that there are no statistical differences in lives saved or wasted help between conditions.

Table 3.1 Lives saved

	(1)	(2)	(3)	(4)	(5)
	Demand = 1	Demand = 2	Demand = 3	Demand = 4	Demand = 5
Help Wasted in First 50 Rounds in the Baseline Condition	.8288	1.3353	1.3600	1.1961	1.3636
Last 50 Rounds	-0.0641 (0.0424)	-0.0489*** (0.0155)	-0.165*** (0.0528)	-0.218*** (0.0652)	-0.545*** (0.170)
A.D. Info	0.0884** (0.0406)	-0.00596 (0.0459)	0.0274 (0.0946)	-0.115 (0.172)	0.0275 (0.292)
Registries	0.0576 (0.0421)	0.0373 (0.0290)	0.0629 (0.0826)	0.0741 (0.137)	-0.129 (0.215)
Last 50 Rounds * A.D. Info	-0.349*** (0.100)	-0.0213 (0.0359)	0.0527 (0.0687)	0.752*** (0.111)	1.155** (0.482)
Last 50 Rounds * Registries	0.0250 (0.0423)	0.0649** (0.0257)	0.250*** (0.0604)	0.249** (0.102)	0.606** (0.291)
Controls	Y	Y	Y	Y	Y
Observations	1,470	1,727	1,111	499	144
Log-Likelihood	-548.0	-1905	-1555	-730.0	-191.8
p values:					
Last 50 Rds*A.D.Info = Last 50 Rds*Regs	0.000***	0.0231**	0.000216***	0.000***	0.261

Marginal effects on group outcomes. Column (1) shows probit regression with $Y = 1$ if the one person at risk is saved. Columns (2)-(5) show Tobit regressions with $Y =$ the number of persons saved conditional on being at risk, censored between 0 and the number of persons at risk in a group in a round. The omitted category is the baseline condition. **Sample** consists of all observations in all treatments, grouped by each demand level from 1 to 5. Round 51 is excluded in all analysis due to a software error recording the data. **Controls:** Dummy variables for every 5 rounds, 5 cost variables for the 5 lowest costs in a group in a round.

Robust standard errors clustered on session level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3.2 Help wasted

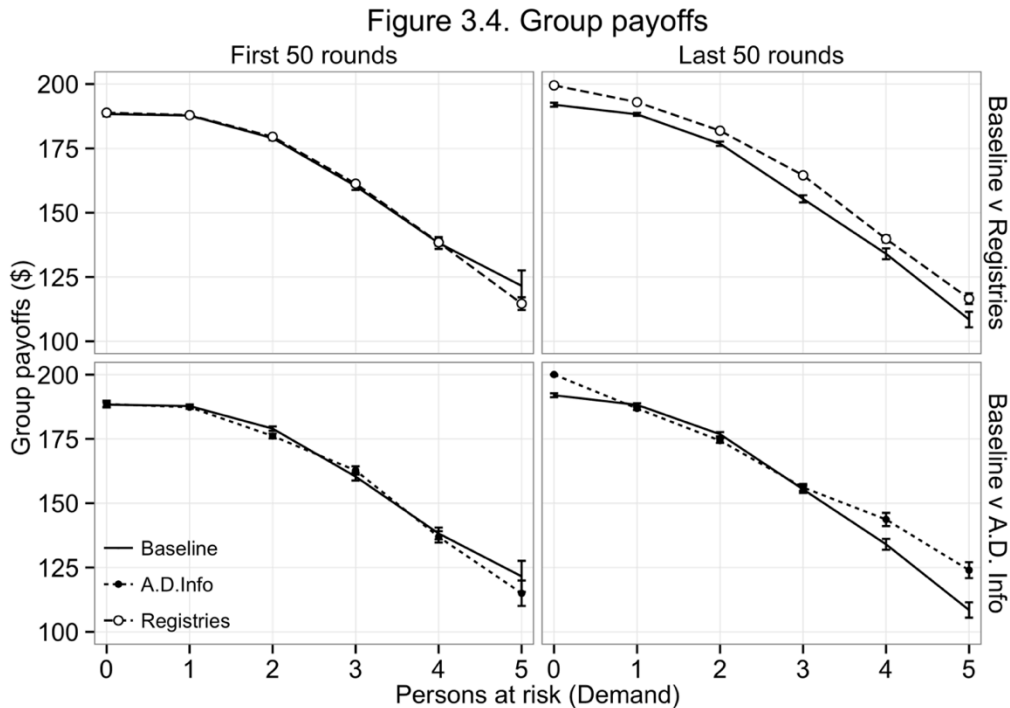
	(0)	(1)	(2)	(3)
	Demand = 0	Demand = 1	Demand = 2	Demand = 3
Help Wasted in First 50 Rounds in the Baseline Condition	2.203	.9383	.2635	.03
Last 50 Rounds	-0.194*** (0.0364)	-0.0860*** (0.0186)	-0.0423*** (0.0155)	-0.0447 (0.0397)
A.D. Info	-0.0787 (0.0678)	0.0807 (0.0880)	0.0414 (0.0277)	0.0362 (0.0251)
Registries	-0.0428 (0.0697)	0.0469 (0.0692)	0.0280 (0.0187)	0.0172 (0.0226)
Last 50 Rounds * A.D. Info	-1.240*** (0.0657)	-0.297*** (0.0347)	-0.0389 (0.0259)	0.0262 (0.0503)
Last 50 Rounds * Registries	-0.782*** (0.0509)	-0.859*** (0.0432)	-0.708*** (0.0341)	-0.382*** (0.0343)
Controls	Y	Y	Y	Y
Observations	608	1,470	1,727	1,111
Log-Likelihood	-654.4	-1310	-815.8	-119.2
p values:				
Last 50 Rds*A.D.Info = Last 50 Rds*Regs	0.000***	0.000***	0.000***	0.000***

Marginal effects on group outcomes. Tobit regressions with Y equal to the number of ‘help offers’ not used, censored above 0. The omitted category is the baseline condition. There were no ‘help offers’ not used (0) for 4 or more persons at risk. **Sample** consists of all observations in all treatments, grouped by each demand level from 0 to 3. Round 51 is excluded in all analysis due to a software error. **Controls:** Dummy variables for 5 rounds, 5 cost variables for the 5 lowest costs in a group in a round.

Robust standard errors clustered on session level in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Fig. 3.4 shows that average payoffs fall as the level of demand r increases in the first 50 rounds for all conditions. This follows immediately since as number of people at risk increases, both the total cost of saving those lives increase and the total cost of lives not saved increases because less people are available to help. In

the last 50 rounds, the payoffs are greater in all of the registry conditions than in the baseline condition for all levels of demand. The improved payoffs follow directly from (a) more lives being saved when two or more subjects were at risk (**Fig. 3.2**) and (b) less wasted help when less than two subjects were at risk (**Fig. 3.3**). **Fig. 3.4** also shows an increase in payoffs in the ADI compared to the baseline condition when there was no need to coordinate actions, either when there was no one at risk or when there were four or five subjects at risk. On the other hand, when there was need for coordination ($r = 1-3$), there is no improvement in group payoffs compared to the baseline.



The group level regressions in the **Table 3.3** indicate that the relative increase in payoffs in the registry than baseline conditions range from \$4.40 to \$9.77 and these differences are all significant. Regressions in Appendix **Table A3.3b** show that the relative increase in payoffs between the three registries are not significantly different for any level of demand. **Table 3.3** also shows that payoffs

increased significantly more from the first to last 50 rounds in ADI than baseline for $r = 0, 4$ and 5 .¹⁸

Table 3.3 Group payoffs

	(0)	(1)	(2)	(3)	(4)	(5)
	Demand = 0	Demand = 1	Demand = 2	Demand = 3	Demand = 4	Demand = 5
Payoff in first 50 Rds in the Baseline:	188.35	187.78	179.00	160.34	138.23	121.56
Last 50 Rounds	3.725*** (1.044)	0.300 (1.385)	-2.110** (0.797)	-3.952** (1.447)	-4.988** (1.951)	-9.506*** (2.795)
A.D. Info	0.117 (1.928)	-0.527 (1.429)	-2.213 (1.553)	-0.200 (2.370)	-2.125 (3.313)	-0.800 (5.495)
Registries	0.409 (1.619)	-0.0182 (0.959)	0.625 (1.051)	1.293 (2.036)	0.598 (2.522)	-4.576 (4.231)
Last 50 Rds * A.D. Info	8.051*** (1.615)	-0.827 (1.864)	0.225 (1.500)	1.132 (1.784)	12.69*** (2.544)	13.44** (5.569)
Last 50 Rds * Registries	6.812*** (1.269)	4.914*** (1.453)	4.399*** (1.121)	6.222*** (1.654)	5.351** (2.432)	9.770* (4.968)
Constant	184.7*** (2.113)	189.5*** (1.691)	196.3*** (1.552)	194.1*** (2.643)	170.0*** (4.734)	149.8*** (6.748)
Controls	Y	Y	Y	Y	Y	Y
Observations	608	1,470	1,727	1,111	499	144
R-squared	0.448	0.171	0.232	0.332	0.250	0.350
Log-Likelihood	-1956	-4865	-6349	-4360	-2023	-562.7
p values:						
A.D.Info*Treat =	0.425	<0.001***	0.0129**	<0.001***	0.0012***	0.501
Registries*Treat						

Coefficients of OLS regressions on group outcomes. Y equals the sum of individual payoffs in a group in a round. The omitted category is the baseline condition. **Sample** consists of all observations in all treatments, grouped by each demand level from 0 to 5. Round 51 is excluded in all analysis due to a software error in data collection. **Controls:** Dummy variables for every 5 rounds, 5 cost variables for the 5 lowest costs in a group in a round.

Robust standard errors clustered on session level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

¹⁸ Group level linear regressions presented in appendix **Table A3.3a** for each level of risk that control for round (clustering s.e. at the group level) and the five lowest costs among the safe subjects robustly show that there is no statistical difference in payoffs between conditions in the first 50 rounds.

3.1.2 Market efficiency

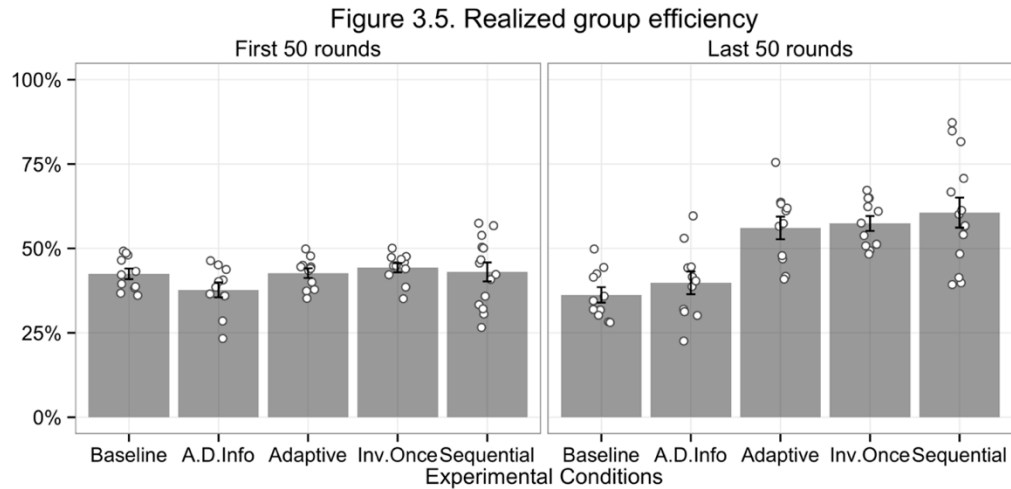


Fig. 3.5 shows the average percent of efficiency obtained relative to the maximum possible for each condition as well as for each group. In the first 50 rounds, groups were able to obtain only 42 percent of the maximum possible payoffs on average. The less than 50 percent efficiency obtained in all conditions reflects several factors including subjects with limited social preferences and the inability of subjects to coordinate on how many, and who, will help.¹⁹ Although the registry conditions do not alter preferences, they can address the coordination problems among those subjects with social preferences to provide a means to better coordinate to help when needed and to sort the subjects who help towards those with lower costs. The increase in efficiency from the first to the last 50 rounds is between 13 and 19 percentage points in the three registry conditions (on

¹⁹ Subjects in the first 50 rounds obtained 71 percent of the baseline benchmark payoffs of \$177.43 (without information, and if and only if subjects would have helped when their costs were less than \$6.36). If subjects followed this rule, then on average they would have helped 31.1% of the time they were safe. However, subjects over the first 50 rounds on average helped only 21.6% of the time they were safe, which at least partially explains the less than 100% efficiency. However, a lack of prosocial preferences resulting in less than the optimal amount of help is not all of the reason for the less than 100% efficiency. First, as **Fig. 3.3** shows, on average two units and one unit of help are wasted when there were 0 and 1 subjects at risk, thus resulting in too much help in these cases. Second, **Fig. 3.6** shows that relative to the registry conditions, subjects helped too often when costs are low (making it more likely to have wasteful oversupply) but not enough when costs are high (making it more likely to have unfulfilled demand). Thus, the inefficiency is a combination of too few subjects helping overall, wasted help when realized demand was low and subjects with higher costs not helping when they would have helped if they knew there was unmet demand.

average by 15 percentage points) while efficiency fell by 6 percentage points in the control condition. Thus, on average the registry resulted in a 21 percentage point relative increase in efficiency.

Table 3.4 Group efficiency

Each Mean Difference entry shows the Column condition minus the Row condition. E.g., the upper left cell indicates that efficiency increased 8.3 percentage points more in the ADI than Baseline condition from the first to last 50 rounds.

	<u>A.D. Info</u>	<u>Adaptive Reg</u>	<u>Inv.Once Reg</u>	<u>Inv.Seq Reg</u>
Baseline				
Mean Diff	+8.3%	+19.6%	+19.3%	+23.8%
p-value	0.0652*	0.0001***	0.0000***	0.0000***
Obs.	22	22	22	25
A.D. Info				
Mean Diff		+11.3%	+11.0%	+15.5%
p-value		0.0192**	0.0104**	0.0014***
Obs.		22	22	25
Adaptive Reg				
Mean Diff			-0.3%	+4.2%
p-value			0.8470	0.5719
Obs.			22	25
Inv.Once Reg				
Mean Diff				+4.5%
p-value				0.4030
Obs.				25

p-values from Wilcoxon Mann-Whitney test for comparisons between each pair of treatment conditions. **Sample** consists of group level observations in all treatments. There is one measure per group being the difference in efficiency from the first to last 50 rounds: *** p<0.01, ** p<0.05, * p<0.1

Group level non-parametric MW tests with one observation per group presented in **Table 3.4** show that the increased efficiency in each of the registry conditions is highly significantly different from the baseline ($p < .001$).²⁰ The MW tests also indicate no statistical difference in the change in efficiency between the three registry conditions and only marginally significant (and much smaller) difference in the change in the efficiency between the baseline and ADI

²⁰ **Table A3.4a** in the appendix shows that there is no statistical differences in efficiency between conditions except ADI which on average obtained 6.6 percentage points lower efficiency than the Invitation Once Registry condition in the first 50 rounds.

conditions. The reason the change in the ADI is small relative to the control condition is that, although subjects in the ADI condition received higher payoffs than subjects in the baseline for $r = 0$, $r = 4$ and $r = 5$, these realizations of demand only account for 22 percent of realized demand levels, and in ‘typical realization times’ when $r = 1-3$, there are directionally smaller payoffs in the ADI than control conditions (see **Fig. 3.4**).

3.2 Individual decisions and Registry design

In this section we examine individual level decisions that underpin the market level results and how these decisions respond to information and different registry rules.

3.2.1 Individual decisions to help

Fig. 3.6 presents decisions to help by costs (with cubic spline smoothing) with one standard error bands. It shows that as costs increase, the likelihood that a subject helps decreases. For instance, in the first 50 rounds, when costs are close to \$2, subjects helped over 70% of the time, whereas if costs are around \$4 they helped about 50% of the time, and less than 5% of the time if costs are over \$12.²¹ The percent of time subjects helped equals the number of times subjects helped divided by the number of times subjects were not at risk for each cost level. In the last 50 rounds in the registry conditions, this percent is deflated relative to an ‘intention-to-help’ metric since subjects who join the registries are not always invited to help. Comparing first to last 50 rounds, subjects are less likely to help overall in all conditions, consistent with most studies of finitely repeated public goods games that find cooperation falls across rounds (e.g. Andreoni 1988). **Fig 3.6** shows, however, different patterns of decrease in helping in the baseline

²¹ Subject level probit regressions presented in Appendix **Table A3.5a** that control for cost, cost-squared and round (clustering s.e. at the group level) robustly show that there is no statistical difference in the decision to help between conditions in the first 50 rounds, either excluding or including subject specific characteristics (those reported in Appendix **Table A2.1**).

and treatment conditions. In the baseline, the decrease in help is largest for the midrange of the costs (\$5-\$9), whereas in the registry conditions the decrease in help is largest for the lowest costs (\$2-\$5). Assuming the decrease in help in the baseline condition reflects a general reduction to help over time across all conditions, the *additional* decrease in help in the registries for the lowest costs (\$2-\$5) reflects the reduction in wasted help due to the registries not inviting subjects to help when help is not needed. Our regression analyses below will present evidence indicating how the registries achieved less wasted help, while it saved more lives and increased efficiency overall. In ADI condition, we also observe the largest decreases in help for the lowest costs (\$2-\$6), as well as slight increases in help for the highest costs (\$10 and above). This slight increase in help for the highest costs suggests that, similar to the registry conditions, subjects may also be responding to demand information.

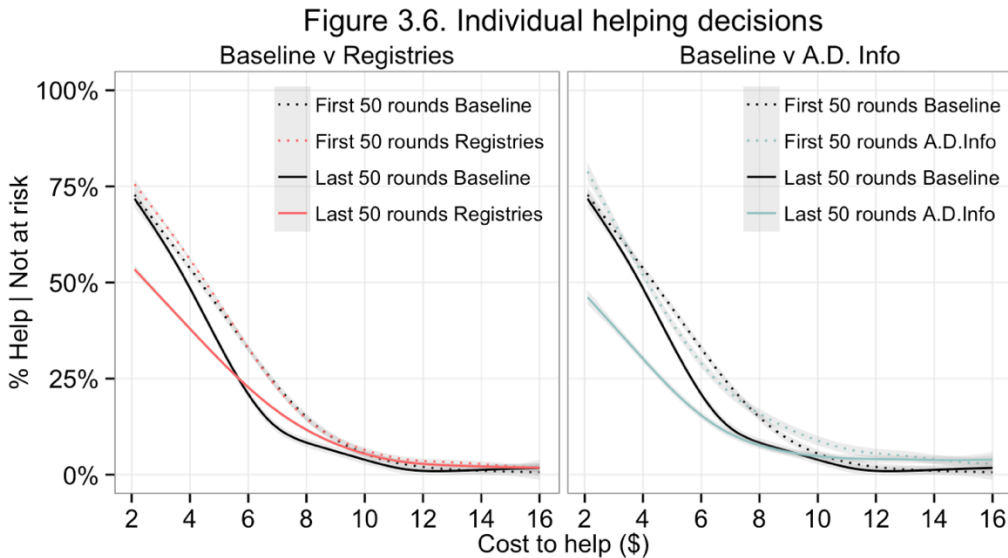


Fig. 3.7 presents changes in decision to help by costs in the ADI condition (with cubic spline smoothing). Overall, the change in the percent of help increased the more subjects were at risk. All help disappears when subjects are informed that no subjects are at risk. In contrast, when 4 or more subjects are at

risk (we aggregate for $r \geq 4$ otherwise there are two few group level observations), we observe that help increases the most, often by 15 percentage points or more, for costs up to \$14. This increase reflects that subjects knew for sure ($r \geq 5$) or almost for sure ($r = 4$) that if they help they will save someone. An interesting question, given this behavior, is why the registry conditions did not see an increase in help for higher costs (**Fig. 3.6**) since if subjects joined the registries with higher costs and were invited, they would have also known for sure that they could have saved someone at risk. The answer, as we show below, is that the registries allowed subjects to successfully sort so that the registries were more likely to invite subjects with lower costs, thus registry members with higher costs were rarely invited to help. Difference-in-difference-in-difference (last 50 rounds by ADI by demand level) regressions (Appendix **Table A3.5b**) at the group level indicate that the percentage help increased significantly as the level of aggregate demand increases for almost all costs.

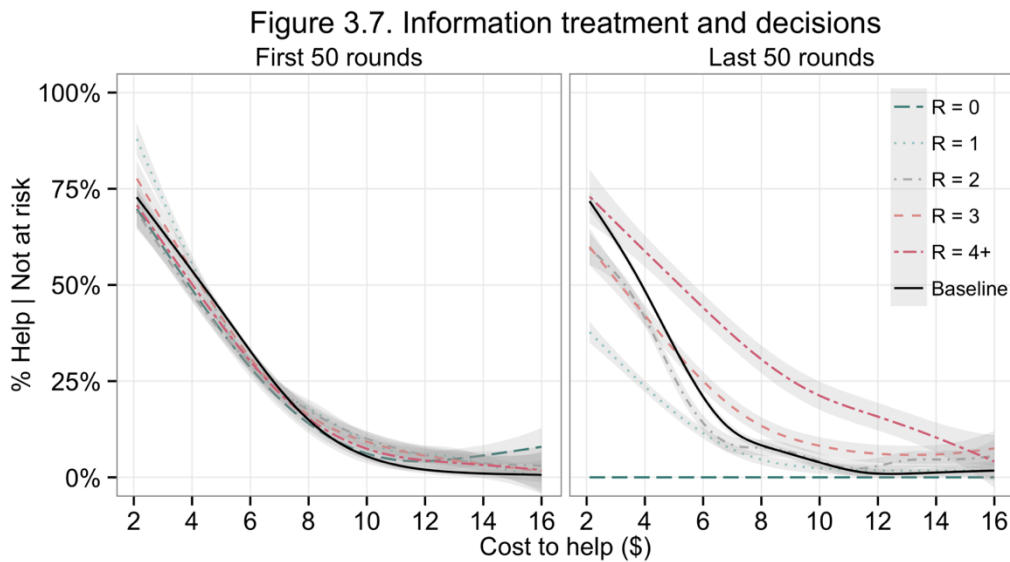


Table 3.5 presents individual level probit regression estimates on the likelihood to help over all costs and for three cost categories separately (\$2.00-\$5.00, \$5.10-\$10.00, \$10.10-\$16), specifically:

$$\Pr(h_{it} = 1|D, X) = \Phi(\beta_0 + \beta_1 D_{Last50} + \beta_2 D_{Registries} + \beta_3 D_{Last50} D_{Registries} + \beta_4 c_{it} + \beta_5 c_{it}^2 + X' \gamma)$$

where $h_{it} = 1$ if subject i helped in round t given she was not at risk, D_{Last50} is a dummy for observations in the last 50 rounds, $D_{Registries}$ is a dummy for observations in the three registry conditions, c_{it} and c_{it}^2 are the cost to help and the cost to help squared, respectively, and X includes controls for round and individual level information (see notes under **Table 3.5**). Robust standard errors are clustered at the group level. The key difference-in-difference estimator is the interaction Last 50 Rounds by All Registries ($D_{Last50} D_{Registries}$). The regressions collapse across the three registry conditions. Regression estimates presented in Appendix **Table A3.5c** show that the estimated difference-in-differences between the three registry conditions on helping are small and never significant. The estimates shown in **Table 3.5** indicate that the subjects in the three registry conditions over all costs helped 2.2 percent less often (column 1; $p < .05$), and this percentage decrease is driven entirely by when costs are lowest (Column's 2-4); when costs were less than \$5, help declined by nearly 16 percentage points more in the registry than baseline condition ($p < .001$), whereas we detect no significant difference when costs were greater than \$5.

Table 3.5 Percent helped in Baseline and Registry conditions, Diff-in-diff

	(1)	(2)	(3)	(4)
Y=1 if Helped	All costs	Cost: \$2.1-5	\$5.1-10	\$10.1-16
Percent help in the Baseline in Rds 1-50:	.2034	.5877	.1964	.0170
<hr style="border-top: 1px dashed black;"/>				
Last 50 Rounds	-0.0377*** (0.00786)	-0.0438** (0.0212)	-0.0825*** (0.0160)	-0.00229 (0.00621)
All Registries	0.0156 (0.0188)	0.0257 (0.0514)	0.00981 (0.0282)	0.0115* (0.00655)
Last 50 * All Registries	-0.0215** (0.0101)	-0.158*** (0.0291)	0.0314 (0.0196)	-0.00274 (0.00696)
Cost to help	-0.0768*** (0.00570)	-0.0945*** (0.00857)	-0.0581*** (0.00363)	-0.00342*** (0.000721)
Cost to help ²	0.00207*** (0.000305)			
Controls	Y	Y	Y	Y
Observations	36,595	7,770	13,165	15,660
Log-Likelihood	-12,570	-5,046	-5,566	-1,738

Marginal effects of probit regressions on individual decisions. Y = 1 if an individual helped conditional on being safe in a round. The omitted category is the baseline treatment. **Samples:** Includes observations in the baseline and registry conditions. **Controls:** 10 dummies for every 5 rounds, frequency and amount of monetary donation last year, frequency and hours of volunteering last year, gender, ethnicity, English skills, academic major, university entrance exam performance, weekly work hours, weekly spending, family income.

Robust standard errors clustered on group level in parentheses; *** p<0.01, ** p<0.05, * p<0.1

3.2.2 Registry designs

We conclude by analyzing how each registry condition may have uniquely affected subjects' decision to join the registry and help when invited, even if they resulted in similar increases in payoffs (**Fig. 3.4**) and efficiency (**Fig. 3.5**). **Figs. 3.8 and 3.9** compare the decision to join the registry and the decision to help when invited for each registry condition by costs (with cubic spline smoothing). **Fig. 3.8** shows subjects in the Sequential registry are more likely to join the

registry for virtually all costs. **Fig. 3.9** shows the decision to help conditional on being a registry member. In contrast to the decision to join the registry, **Fig. 3.9** shows that Sequential registry members are less likely to help when invited. This result was anticipated since members of the Invitations Once and Adaptive registries face the decision to save someone or guarantee someone will not get saved, whereas Sequential registry members face a less certain consequential effect on a subject at risk if they do not help. Although the registries resulted in similar payoffs and efficiency in our lab study (**Fig. 3.4, Fig. 3.5**), if it is costly to enroll registry members, the non-sequential registries might be more efficient on the basis of having fewer people enroll (for instance, in the case of bone marrow registries where the cost to enroll can be non-trivial). Regressions presented in **Table 3.6** and **Table 3.7** confirm that these differences are significant.

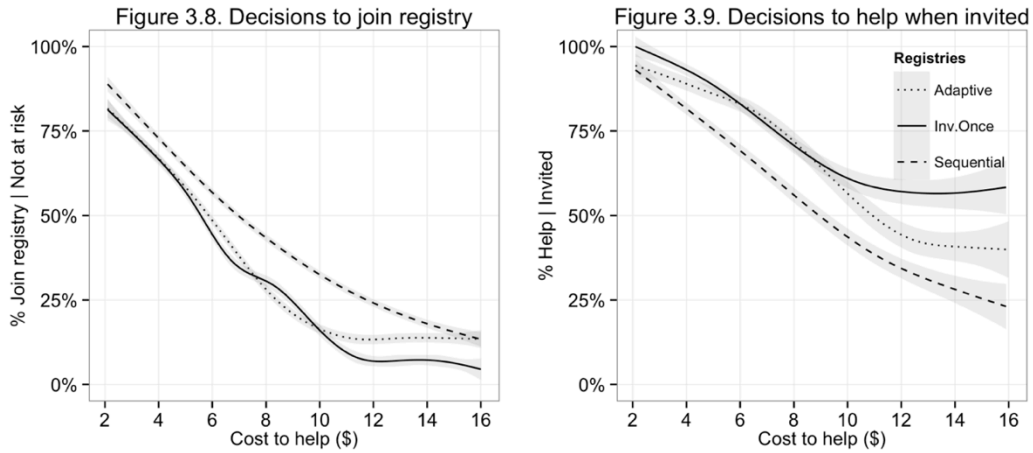


Table 3.6 show the regression results on the decision to either (a) help directly, (b) join the registry with willingness 3, (c) join the registry with willingness 2, (b) join the registry with willingness 1, or (e) not help nor join the registry. We ran ordered probit regressions estimating the following model:

$$h_{it}^* = \beta_0 + \beta_1 D_{Seq} + \beta_2 D_{Adapt} + \beta_3 D_{c \in [2.1,3]} + \beta_4 D_{c \in [3.1,4]} + \dots + \beta_9 D_{c \in [8.1,9]} + X'\gamma + \epsilon$$

$$h_{it} = \begin{cases} 0 \text{ (Help directly)}, & h_{it}^* \leq 0 \\ 1 \text{ (Join, Will = 3)}, & 0 < h_{it}^* \leq \mu_1 \\ 2 \text{ (Join, Will = 2)}, & \mu_1 < h_{it}^* \leq \mu_2 \\ 3 \text{ (Join, Will = 1)}, & \mu_2 < h_{it}^* \leq \mu_3 \\ 4 \text{ (Not join)}, & h_{it}^* > \mu_3 \end{cases}$$

where h_{it}^* is a latent variable for subject i 's propensity to help in round t , h_{it} is the observed help decision of subject i in round t , D_{Seq} and D_{Adapt} are dummy variables for the Sequential and Adaptive registries, respectively, $D_{c \in [l,m]}$ is a dummy for costs between l and m , with the omitted cost category is a cost between \$9.10 and \$16.00,²² and X includes subject specific variables (see Appendix **Table A2.1**) and dummy variables for every 5 rounds. We cluster standard errors at the group level. Each column indicates the marginal effect on the frequency for each possible choice from a single ordered probit regression. The estimates on the registry condition dummy variables compare their relative effects to the Invitations Once registry (the help levels for each choice for Invitations Once are shown on the first row), and the dummy variables on costs compare each of these cost ranges to costs between \$9.10 to \$16.00.

The estimates in **Table 3.6** show that the decision to join the registry, for every willingness level, is significantly higher in the Sequential than Invitations Once registry, whereas there is no difference between the Adaptive and Invitations Once registries. The regressions also show that for any cost less than \$9 subjects are more likely to both help directly and join the registry with any of the three willingness levels.

²² In other specifications we included dummy variables for higher cost categories, but there were never different from each other, we thus collapsed across these higher cost categories.

Table 3.6 Individual decisions to help immediately or join in the registry conditions

	(1)	(2)	(3)	(4)	(5)
Ordered probit Levels:	Help Directly	Join Reg Will = 3	Join Reg Will = 2	Join Reg Will = 1	Not Join & Not Help Dir
Invitations Once Help Percent	1.4%	3.4%	6.7%	18.4%	70.1%

Sequential Registry	0.00240* (0.00145)	0.0147* (0.00821)	0.0271* (0.0143)	0.0639** (0.0304)	-0.108** (0.0536)
Adaptive Registry	0.000411 (0.000813)	0.00262 (0.00549)	0.00499 (0.0102)	0.0123 (0.0250)	-0.0203 (0.0415)
Costs					
\$2.10-\$3.00	0.144*** (0.0280)	0.250*** (0.0262)	0.182*** (0.0161)	0.0521** (0.0249)	-0.628*** (0.0187)
\$3.10 - \$4.00	0.0905*** (0.0192)	0.201*** (0.0206)	0.175*** (0.0155)	0.0994*** (0.0215)	-0.566*** (0.0200)
\$4.10 - \$5.00	0.0602*** (0.0162)	0.161*** (0.0196)	0.160*** (0.0137)	0.129*** (0.0208)	-0.510*** (0.0203)
\$5.10 - \$6.00	0.0330*** (0.00999)	0.111*** (0.0157)	0.130*** (0.0114)	0.145*** (0.0156)	-0.418*** (0.0196)
\$6.10 -\$7.00	0.0177*** (0.00609)	0.0716*** (0.0113)	0.0966*** (0.0106)	0.137*** (0.0115)	-0.323*** (0.0224)
\$7.10 - \$8.00	0.0117*** (0.00405)	0.0525*** (0.00711)	0.0765*** (0.00964)	0.122*** (0.0121)	-0.263*** (0.0228)
\$8.10 - \$9.00	0.00562** (0.00238)	0.0291*** (0.00576)	0.0470*** (0.00897)	0.0893*** (0.0112)	-0.171*** (0.0243)
Controls	Y	Y	Y	Y	Y
Observations	13,443	13,443	13,443	13,443	13,443
Log likelihood	-11281	-11281	-11281	-11281	-11281

Marginal effects of Ordered probit regressions, with 5 levels if a subject helped immediately, joined the registry with willingness 3, 2 or 1, or did not join registry, conditional on not being at risk. The omitted category is the Inv. Once condition. **Sample** consists of the last 49 rounds of observations in the registry conditions. We exclude round 51 due to a software error that affected data in that round.

Controls: Dummies for every 5 rounds, frequency and amount of monetary donation last year, frequency and hours of volunteering last year, gender, ethnicity, English, academic major, university entrance exam performance, weekly work hours, weekly spending, family income.

Robust standard errors clustered on group level in parentheses; *** p<0.01, ** p<0.05, * p<0.1

The estimates in **Table 3.6** also indicate that subjects are sorting themselves on the basis of costs. These estimates are graphed in **Fig. 3.10**. Subjects increasingly chose to help directly (from 2 to 14 percentage points more) as costs fell from \$6.10-\$7.00 to the lowest costs (\$3.00 or less). Similarly, subjects increasingly chose to join the registry with the highest willingness (level 3), increasing this choice from 7 to 25 percentage points as costs fell from \$6.10-\$7.00 to the lowest costs. In contrast, as costs decreased over the same range, subjects decreasingly chose to join the registry with the lowest willingness level (level 1), decreasing this choice from 14 to 5 percentage points. **Table 3.6** and **Fig. 3.10** also show an increase in willingness level 2 as costs fell from \$9 to \$4, but little further change for costs from \$4 to \$2. One possible explanation for this lack of change in this lower cost range is that some subjects were switching from willingness level 1 to willingness level 2 while other subjects were switching from willingness level 2 to willingness level 3, and these effects roughly canceled each other out. In sum, the registries let subjects sort themselves on the basis of their preference to help, using costs as a proxy for these preferences; as costs fell, subjects increasingly joined the registry and switched to higher willingness levels and helping directly.

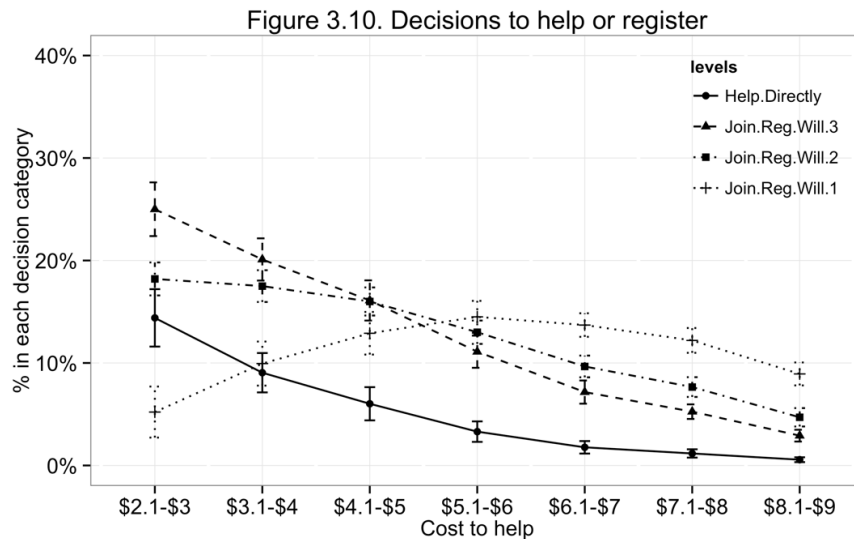


Table 3.7 presents probit regression estimates on the choice to help in the three registry conditions conditional on subjects joining the registry and being invited to help. We estimate versions of the following model:

$$\Pr(h_{it} = 1|X) = \Phi(b_0 + b_1D_{Seq} + b_2D_{SeqInvLater} + b_3D_{Adapt} + b_4c_{it} + b_5c_{it}^2 + bX_i)$$

where h_{it} , D_{Seq} , S_{Adapt} and X_i were defined above, and $cost_{it}$ is subject i 's cost to help in round t . Since the sequential registry continues to invite registry members until there are no subjects at risk or there are no more registry members, we estimate marginal effects of subjects who received invitations after the first set of invitations were made ($D_{SeqInvLater}$) since, if subjects were sorting towards the most likely to help being invited first, then these subsequent invitations could result in lower likelihood of help.

Table 3.7 shows that across all specifications the sequential registry members who receive the first set of invitations are about 13 percentage points less likely to help than Invitations Once registry members ($p < .05$). Sequential registry members who receive a subsequent invitation to help are an additional almost 17 percentage points less likely to help ($p < .01$). When we control for members' stated willingness to help (column 3), the registry members who are invited subsequent to the first set of invitations are only 8 percentage points less likely to help as willingness controls for some of the sorting reason for why these subjects are being invited later.

Table 3.7 also shows that there is no difference in the percent of time that registry members help comparing the Invitations Once and Adaptive registries. We had anticipated that a potential concern with the Invitations Once registry would be that some registry members may have a propensity to join but not help if invited. To address this concern, the Adaptive registry assigned subjects status based on their past behavior and used an algorithm to reduce the likelihood that

these members would get invited to help. However, the results show no discernable gain in the likelihood to help between the Adaptive and Invitations Once registries. The lack of greater success for the Adaptive than Invitations Once registry could be due to (a) a lack of subjects who systematically join but do not help or (b) our use of status did not successfully identify subjects who were joining but not helping.

Table 3.7 Registry member decisions to help when invited

	(1) All Registries	(2) All Registries	(3) All Registries	(4) Inv. Once & Adaptive Reg
Invitations Once Registry % help	84.1%	84.1%	84.1%	84.1%
Sequential Reg All Invitations	-0.133** (0.0653)	-0.125** (0.0522)	-0.127*** (0.0436)	
Sequential Reg Later Invitations	-0.167*** (0.0432)	-0.164*** (0.0478)	-0.0766* (0.0420)	
Adaptive Registry	-0.0533 (0.0451)	-0.0380 (0.0428)	-0.0407 (0.0410)	-0.0312 (0.0347)
Cost to help	-0.0903*** (0.0166)	-0.0912*** (0.0167)	-0.0727*** (0.0157)	-0.0565*** (0.0155)
Cost to help ²	0.00274*** (0.000857)	0.00276*** (0.000853)	0.00218*** (0.000800)	0.00160* (0.000947)
Willingness 1			-0.191*** (0.0179)	-0.137*** (0.0188)
Controls		Y	Y	Y
Observations	2,717	2,717	2,717	1,617
Log likelihood	-1234	-1154	-1094	-570.8

Marginal effects of probit regressions. Columns (1)-(3) compare the three registries: Y=1 if an individual helped immediately or helped when invited by the registry. The omitted category is the Inv. Once registry. Column (4) compares the Inv. Once and Adaptive registries and the omitted category is the Inv. Once registry. **Sample** consists of last 49 rounds of observations in the registry treatments including those who either helped immediately or was invited to help. We exclude round 51 in all analysis due to a software error that affected data in that round. **Controls:** Dummies for every 5 rounds, frequency and amount of monetary donation last year, frequency and hours of volunteering last year, gender, ethnicity, English skills, academic major, university entrance exam performance, weekly work hours, weekly spending, family income.

Robust standard errors clustered on group level in parentheses; *** p<0.01, ** p<0.05, * p<0.1

4. Conclusion

In our experiment, subjects faced uncertain demand to help subjects at risk. Although we did not alter subjects' monetary costs or benefits, our registry design improved efficiency by about 20 percentage points of the maximum possible payoff. Our design improved efficiency by (1) eliminating wasted help when aggregate demand was low, (2) increasing help and lives saved when aggregate demand was high, and (3) sorting subjects so that those with the lowest costs to help (*ceteris paribus*, greatest preference to help) were more likely to help. The sorting provided by the registries was especially pronounced, with subjects with the highest costs neither joining nor helping directly, subjects with intermediate costs joining with the intermediate willingness levels, and subjects with the lowest costs joining with the highest willingness or helping directly. Finally, although the three implementations of the registry resulted in similar market levels of lives saved, wasted help, payoffs and efficiency, we observed that subjects were more likely to join the Sequential registry but were less likely to help when invited in the Sequential than the other two registries.

Our registry design assumes a central organization that knows the aggregate demand. We found that when this aggregate demand information was provided publically, subjects were able to obtain higher payoffs, but only when the information indicated an extreme outcome that solved the coordination problem; in these cases, subjects either do not help (when no help was needed) or increased supply (when everyone's help was needed). However, when any level of aggregate demand occurred that required coordination, both wasted help and insufficient supply persisted. Overall, providing aggregate demand information did not significantly increase efficiency, since the extreme conditions in which the information solves the coordination problem are not common (in our setting, only 22% of the time). Outside the lab, solving the coordination problem with the

aggregate demand information is likely to be less effective since extreme events (such as no need) are rare in most contexts, and higher need periods could still be extremely difficult to coordinate supply. For instance, after disasters (e.g., 9/11 and the Australian Bushfires), aggregate demand for blood increased only slightly, but massive spikes in donations occurred that eventually resulted in wasted supply due to the inability to store blood more than a few weeks.

In sum, we showed that by assuming people are at least partially motivated by social preferences, there are opportunities for novel market designs. We also showed that it is possible to redesign the organization of an environment with supply provided by volunteers, without changing volunteer's costs or benefits, to increase efficiency.

In this paper, we did not delve into the specific nature of preferences to help (e.g., other regarding preferences, warm glow, pure altruism, etc.). Extensions to the current work could theoretically and experimentally examine designs that exploit the nuances of underlying preferences to help. Another extension could explore combining designs that solve the coordination problem (like the registries do) with incentives to increase donations among people with *a priori* lower intrinsic motivation to help.

References

- Abdulkadiroğlu, A., Pathak, P., Roth, A., & Sönmez, T. (2005). The Boston public school match. *The American Economic Review*, 95(2), 368-371.
- Abdulkadiroğlu, A., & Sönmez, T. (2003). School choice: A mechanism design approach. *The American Economic Review*, 93(3), 729-747.
- Andreoni, J. (1988). Why free ride?: Strategies and learning in public goods experiments. *Journal of Public Economics*, 37(3), 291-304.
- Andreoni, J. (1989). Giving with impure altruism: applications to charity and Ricardian equivalence. *The Journal of Political Economy*, 97(6), 1447-1458.
- Andreoni, J. (1990). Impure altruism and donations to public goods: a theory of warm-glow giving. *The Economic Journal*, 100(401), 464-477.
- Becker, G., & Elias, J. (2007). Introducing incentives in the market for live and cadaveric organ donations. *The Journal of Economic Perspectives*, 21(3), 3-24.

- Bolton, G., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *The American Economic Review*, 90(1), 166-193.
- Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117(3), 817-869.
- Craig, A., Garbarino, E., Heger, S., & Slonim, R. (2015). Waiting To Give. *Working paper*
- Fehr, E., & Schmidt, K. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3), 817-868.
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171-178.
- Fisman, R., Kariv, S., & Markovits, D. (2007). Individual preferences for giving. *The American Economic Review*, 97(5), 1858-1876.
- Garbarino, E., Slonim, R., & Wang, C. (2013). The multidimensional effects of a small gift: Evidence from a natural field experiment. *Economics Letters*, 120(1), 83-86.
- Gneezy, U., Keenan, E., & Gneezy, A. (2014). Avoiding overhead aversion in charity. *Science*, 346, 632-635.
- Goette, L., & Stutzer, A. (2008). Blood donations and incentives: Evidence from a field experiment. *IZA Discussion Paper No. 3580*
- Greiner, B. (2015). Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE. *Journal of the Economic Science Association*, 1(1), 114-125.
- Gross, D. (2005). Zero-sum charity: Does tsunami relief dry up other giving? *Slate Magazine*. Retrieved July 19, 2015 from <http://www.slate.com/id/2112485>.
- Independent Sector (2010). Value of Volunteer Time, www.independentsector.org/volunteer_time
- Kagel, J., & Roth, A. (2000). The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment. *Quarterly Journal of Economics*, 115(1), 201-235.
- Kessler, J., & Roth, A. (2012). Organ Allocation Policy and the Decision to Donate. *The American Economic Review*, 102(5), 2018-2047.
- Kessler, J. & Roth, A. (2014). Loopholes undermine donation: An experiment motivated by an organ donation priority loophole in Israel. *Journal of Public Economics*, 114, 19-28.
- Milgrom, P. (2000). Putting auction theory to work: The simultaneous ascending auction. *Journal of Political Economy*, 108(2), 245-272.
- Lacetera, N., Macis, M., & Slonim, R. (2012). Will there be blood? Incentives and displacement effects in pro-social behavior. *American Economic Journal: Economic Policy*, 4(1), 186-223.
- Lacetera, N., Macis, M., & Slonim, R. (2013a). Economic rewards to motivate blood donations. *Science*, 340(6135), 927-928.
- Lacetera, N., Macis, M., & Slonim, R. (2013b). In defense of WHO's blood donation policy--response. *Science (New York, NY)*, 342(6159), 692.
- Lacetera, N., Macis, M., & Slonim, R. (2014). Rewarding volunteers: a field experiment. *Management Science*, 60(5), 1107-1129.
- Roth, A. (1984). The evolution of the labor market for medical interns and residents: a case study in game theory. *The Journal of Political Economy*, 92(6), 991-1016.

- Roth, A. (2007). Repugnance as a Constraint on Markets. *The Journal of Economic Perspectives*, 21(3), 37-58.
- Roth, A. (2012). Experiments in Market Design. *Working paper (to appear in The Handbook of Experimental Economics 2, J. Kagel and A. Roth, eds.)*
- Roth, A., & Peranson, E. (1999). The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design. *American Economic Review*, 89(4), 748-782.
- Roth, A., Sönmez, T., & Ünver, M. (2004). Kidney Exchange. *The Quarterly Journal of Economics*, 119(2), 457-488.
- Roth, A., Sönmez, T., & Ünver, M. (2005a). Pairwise kidney exchange. *Journal of Economic Theory*, 125(2), 151-188.
- Roth, A., Sönmez, T., & Ünver, M. (2005b). A kidney exchange clearinghouse in New England. *The American Economic Review*, 95(2), 376-380.
- Roth, A., Sönmez, T., & Ünver, M. (2007). Efficient kidney exchange: Coincidence of wants in markets with compatibility-based preferences. *The American Economic Review*, 97(3), 828-851.
- Slonim, R., Wang, C., & Garbarino, E. (2014). The market for blood. *The Journal of Economic Perspectives*, 28(2), 177-196.
- Titmuss, R. (1970). *Gift Relationships*. London, George Allen.