Title
Essays on Experimental Evidence of Peer Influence

Author
Zhang, Ruizhi

Publication Date
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

ESSAYS ON EXPERIMENTAL EVIDENCE OF PEER INFLUENCE

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ECONOMICS

by

Ruizhi Zhang

September 2020

The Dissertation of Ruizhi Zhang
is approved:

Professor Daniel Friedman, Chair

Professor Kristian López Vargas

Professor Natalia Lazzati

Quentin Williams
Interim Vice Provost and Dean of Graduate Studies

Contents

List of Figures

List of Tables

Abstract

Essays on Experimental Evidence of Peer Influence

by

Ruizhi Zhang

This dissertation presents three experimental studies with an emphasis on peer influence in people's behavior. The first chapter studies a mechanism that facilitates team formation and cooperation among peers. The second and third chapters focus on peer influence in altruistic behavior such as donations to charities and street performers.

The first chapter implements a field experiment in a digital environment with a matching mechanism that facilitates user-team assortativity. Using machine learning techniques, our system identifies users with high expected values and provides them the option to join highly active social groups in terms of engagement and expenditure. We deploy this mechanism experimentally in a popular online game. We find that assortative matching significantly increases new user engagement and productivity, and improves the overall health of the community. New users who join more active communities do exhibit higher engagement and productivity, but do not spend more money. Revenue, however, does increase as existing members of high-activity teams react positively to the influx of new, better-quality users. Teams matched with low-quality new users are negatively impacted leading to a more segregated team environment. We discuss implications of these findings both from a profit-maximizing firm's as well as a broader societal perspective.

The second chapter provides laboratory evidence on how different levels of monetary rewards affect image motivations in the context of charitable donations. We specifically study the interaction effect of image and material motivations on donation decisions. We implement a within-subject design with two treatments: monetary incentive and donation visibility. We have three levels of monetary incentives: no incentive (0% rebate), a low incentive (10% rebate), and a high incentive (50% rebate). We also have three levels in donation visibility: private donation, public donation without costly non-disclosure (ND) option, and public donation with costly non-disclosure option. Our results show that a small reward in terms of a 10% donation rebate does not impose any significant donation behavior changes, neither in private nor in public. When a large reward is given (50% donation rebate), people's perception of the monetary reward determines whether the reward crowds out donations. For those who believe the reward makes their donations appear 'less generous', a 50% rebate significantly crowds out charitable donations in public. For those who do not associate monetary rewards with a negative image, a high rebate significantly increases their donations in public. We also find that males, in general, are more sensitive to their public images and significantly reduce their donations in public when offered a high reward, while the effects on females are non-significant.

The third chapter evaluates the impact of peers' giving behavior on people's willingness to donate. We implement a field experiment studying donations to street performers at stoplights on the streets of Lima, Perú. The treatment condition is defined as observing another vehicle making a donation. We use natural variation in donations

made by passing-by drivers, as well as experimentally-manipulated variation in donations made by hired drivers. We study how observing a donation from another vehicle affects the probability and magnitude of donations of treated vehicles. Our experiment results indicate a strong substitution effect of peer influence. When drivers observe another vehicle donate, they are significantly less likely to make a donation to the same performer and overall average donations are lower as well. Our study contributes to the literature by bringing evidence of peer influence in donations in a developing world context.

To myself,

one journey's end is another's beginning,

the best is yet to come!

Acknowledgments

My journey at UCSC is full of memories that I will cherish for the rest of my life. I have encountered so many people, without whom, this journey would have been impossible. I wish to express my gratitude towards all those that have supported and encouraged me in the pursuit of my Ph.D.

First, I would like to thank my advisor, professor Kristian López Vargas, for his expert knowledge and excellent guidance in exploring new ideas, experiment design, and analysis methodologies throughout the five years that we work together. This journal would have been so much more difficult without his ongoing patience, continual encouragement, and the hundreds of hours he devoted to our research projects. Besides, professor López Vargas was very generous to share his network and connections, and extremely supportive in helping me develop a new career after my graduation.

I would also like to extend my gratitude to professors Daniel Friedman and Natalia Lazzati for their honest, constructive, and timely feedback on my research work. Professors Friedman and Lazzati have continuously offered their expertise in economics to give valuable inputs and guidance throughout my Ph.D. studies. I am especially grateful to professor Friedman for establishing and leading the experimental economics lab at UCSC, where all the experiments that explore our curiosity and ideas in economics take place.

Moreover, I would like to thank professors Donald Wittman and David Lee for being members on my qualification committee, offering valuable advice on my papers and future research directions. I would also like to express my gratitude to professors Marco Castillo and Christine Exley for their constructive feedback which helps shape the second chapter of this dissertation. I am also grateful to Eli Pandolfo, who programmed an excellent online experiment for the second chapter, as well as all my peers in the experimental workshop for giving valuable suggestions on my research work.

Furthermore, I want to thank my co-author, Julian Runge, for introducing me to conducting research in the industry. Julian acted as a great mentor with his expertise in both academia and industry. Without his presence, it would have been impossible for us to successfully cooperate with our partnering company and implement our research ideas.

My time at UCSC is so delightful because of the many friends I met here. We went through the most difficult times together, encouraging each other to pass the qualification exams, discuss research ideas, and offer feedback on each other's work. We also went through the most beautiful times together, hanging out in offices and on the beach of Santa Cruz. The beautiful times are to be continued even after graduation as we become life-long friends.

Last but not least, I would like to thank my family (my parents, husband, and friends) for their continuous financial and emotional support throughout the five years. I am especially grateful to my parents, who have been my role model since I was very young, and raised me to be an honest, humble, hard-working, and rational person just like them. I am willing to constantly explore new opportunities and take on challenges today because I know my parents will always be there to back me up. Mom and Dad, I

am very lucky to be your child.

Chapter 1

The Impact of Assortative Matching in a Digital Environment [1]

1.1 Introduction

Humans are increasingly interacting and operating their daily lives through structured digital environments. It is estimated that they spend 20% of awake time on mobile devices. In 2018 alone, there were almost 200 billion mobile app downloads worldwide and spending in app stores amounted to $101 billion (AppAnnie, 2018).

The *freemium pricing* model is an important driver of this growth. Under such pricing, apps can be downloaded free of charge, and app providers generate revenue from in-app purchases or in-app advertisement. Ghose and Han (2014) point out that apps with in-app purchases result in significantly higher app downloads than those with in-app advertisement, providing more ex-ante utility to consumers. In a freemium app,

[1]The first chapter is a joint work with Kristian López Vargas and Julian Runge

upon download, the user needs to decide to continue using the app based on her first experiences in the environment. Most people use the app only for a few minutes and only for a couple of days. Indeed, 1-day and 30-day user retention rates tend to be below 40% and 10%, respectively (Sifa et al., 2018; GameAnalytics, 2019).

Socialization is crucial in many of the digital ecosystems created by these applications (e.g. in social media, messaging, dating and gaming apps (Alsen et al., 2016)), and together such applications account for most of the revenue and time spent on apps (AppAnnie, 2018). Speaking to this observation, existing literature has established that social experiences are crucial to revenue generation in online freemium settings (Oestreicher-Singer and Zalmanson, 2013; Bapna and Umyarov, 2015; Bapna et al., 2018). Important features of the mechanisms that shape social interaction in these environments, e.g. in regards to the formation of digital communities, are not yet fully understood however.

On this background, this paper studies the impact of implementing an assortative matching mechanism (Becker, 1973; Kremer, 1993) on behavioral responses at the individual level as well as the digital community level (henceforth teams).[2] In particular, we study a mechanism that facilitates positive assortative matching between new joiners and existing teams, where players of an online Role-Playing Game (RPG) are matched to teams of similar behavior to promote a stable and fun-maximizing social environment for users. We test whether this matching mechanism impacts early app experiences of

[2] A matching mechanism is said to be positive assortative if individuals are sorted by characteristics and those with similar features are matched with one another more frequently (Becker, 1973; Kremer, 1993).

new app downloaders, changing users' behavior, longer term engagement and revenue.

We do so by devising a machine learning-based matching system that classifies entities on the demand (new app downloaders) and supply (existing teams of users in the app) side into high and low types and matches them accordingly. On the demand side, we separate premium users from free users by predicting their future purchase behaviors with demographic and behavioral traces using a machine learning algorithm that has been suggested as well suited in previous literature (Sifa et al., 2015). On the supply side, we classify existing groups of users into highly active and premium user-based teams and separate them from less active, free user-based teams.

We evaluate and test our system in a large mobile gaming app through a quasi-field experiment, turning the system on and off in three day intervals over a six week period to evaluate the overall impact of our system and identify the causal impact of assortative matching on users' behavior and that of teams.

We find that assortative matching increases overall user engagement and productivity, leading to net positive impact on engagement and productivity metrics. In particular, we find that new users that join more active communities exhibit higher willingness to play and spend time with longer retention rates, but do not spend more money. Revenue, however, does increase due to existing members of high activity-teams reacting positively to the influx of new users with better quality. However, teams matched with low quality new users experience negative effects on engagement and productivity, leading to more separating paths between teams of different qualities than in the absence of positive assortative matching.[3]

The presented findings establish the merits of such an approach to generate engagement and revenue in freemium settings, and help understand features that shape social interaction and community formation in these environments. They also indicate that the use of assortative matching can have detrimental effects on marginal parts of the user population. We discuss implications of these findings both from a profit-maximizing firm's and a wider societal perspective.

The paper is organized as follows: Section 2 provides a review of related literature on social experiences and premium conversion in freemium environments and of literature on assortative matching theory and applications in economics. Section 3 introduces the online game environment where our matching mechanism is deployed. Section 4 describes the matching mechanism, specifically the machine learning-based separation of premium and free users and high and low activity-teams, as well as results from offline evaluation of the matching mechanism using observational data. Section 5 lays out study design and results of a six week deployment of the system in the game environment. Section 6 concludes with a concise discussion of broader implications of our findings.

[3] The authors are aware that the manipulation is on the timing of "on" and "off" periods and not directly in the assignment of new users to teams. However, arrival times of users can be assumed to be random as the timing of system "on" and "off" periods is unknown to users. Therefore, whether someone was matched using the assortative system or not is independent of user characteristics. We sometimes omit the word "quasi" when describing the experimental intervention.

1.2 Related Literature

Two branches of literature are relevant for our paper: (i) literature on assortative matching, and (ii) literature on social interactions and premium demand in freemium environments.

The first body of literature that is relevant to our paper is research on assortative matching, where agents of similar characteristics are matched together. The assortative matching theory has been popular in the marriage and labor market since the 1970s. Becker (1973) first introduced the concept of assortative matching in the marriage market, and shows that this matching method is efficient under a two-person partnering environment with fully transferable utility and complementarity within the groups. Later, researchers have studied theoretically sufficient conditions for assortative matching to happen in a broader range of settings (Shimer and Smith, 2000; Durlauf and Seshadri, 2003; Legros and Newman, 2007; Chai et al., 2016; Ahlin, 2017). For example, Besley and Ghatak (2005) provides theoretical evidence that principal-agent matching based on mission preferences increases organizational efficiency. In an environment where there exist profit-oriented and mission-oriented risk-neutral agents and principals, agents in profit-oriented sector must always be paid to exert effort, but agents in the mission-oriented sector are affected by both payment and the mission preference of their principal. Thus, given the same wage, well-matched principals and agents in the mission-oriented sector will have higher productivity than in profit-oriented sector.

Besides the theoretical foundation of assortative matching, Kremer (1993) utilizes it to study labor market conditions. He proposes a production function that derives a state of assortativity of skills between workers and firms: high-skilled workers will be matched together in equilibrium, and wages and outputs will rise steeply in skill. This production function is consistent with the large income difference between rich and poor countries, and the positive correlation between wages and occupations within industries. There is empirical evidence supporting this theory using observational employer-employee matching data, where higher productivity workers tend to match with higher productivity firms (Mendes et al., 2010; Torres et al., 2018). Moreover, assortative matching and production complementarity between firms and labor quality also contribute to thicker urban labor markets (Andersson et al., 2007). [4]

Experimentally, introducing assortative matching of players in laboratory experiments (public goods game and prisoner's dilemma) significantly improves average contribution in later rounds of the games (Page et al., 2005; Gunnthorsdottir et al., 2007; Rud and Rabanal, 2015). In a public goods game setting, subjects sorted into pairs of two by pre-committed investment levels are significantly more cooperative in later rounds of investment decisions than randomly paired subjects (Rud and Rabanal, 2015). Gunnthorsdottir et al. (2007) find similar higher contributions when matching subjects based on their current round of contribution decisions in a 10 round public goods game with 4 subjects matched into one group each time. When allowing subjects

[4] For a few empirical studies that observe no evidence of assortative matching between worker and firm effects (Goux and Maurin, 1999; Woodcock, 2007; Abowd et al., 2009; Gruetter and Lalive, 2009), those results are likely due to limited mobility bias (Andrews et al., 2012), not the lack of assortativity. Correlation between worker and firm contributions to wages is negative when fewer workers move between firms, but the correlation becomes significantly positive for larger samples of employer and employee matching data with higher inter-firm mobility.

to voluntarily form associations in a repeated public goods game, Page et al. (2005) find that this process generates sorting by contribution levels over a series of re-groupings.

In this paper, we apply the concept of assortative matching to a digital environment, where players of an online Role-Playing Game (RPG) are matched to teams of similar behavior to promote a stable and fun-maximizing social environment for users, to in turn increase in-app engagement and revenue for the firm producing the game. In a previous study, Choi et al. (2008) report that a matching between players and teams based on task-orientation or social-orientation improves players' retention. However, their study is based on observational data collected post-voluntary-matching between players and teams, so the effect cannot be considered causal. In our paper, we will provide causal estimates of the impact of assortative matching on user retention, engagement and revenue by means of a quasi-experiment and appropriate econometric analysis.

The second body of literature that is relevant to this paper covers the nexus of social experiences and revenue generation in freemium settings. Oestreicher-Singer and Zalmanson (2013) and Bapna and Umyarov (2015) show that social engagement and peer influence are major drivers of the conversion of users from free to premium. Extending these findings, Bapna et al. (2018) present evidence in the opposite direction, i.e., that conversion from free to premium on a music platform encourages users to generate more content for community sharing and engagement. It hence appears that social interaction and premium conversion are reinforcing each other to generate more individual-level and community-level engagement and revenue.

Particularly, high-value premium users are essential for the profitability of freemium apps (Sifa et al., 2018). In freemium online games in particular, up to 50% of revenues can be contributed by less than 2% of premium users (Pei Chen et al., 2018). The identification of (high-value) premium users to provide them with personalized and good social experiences hence suggests itself as promising to produce lift in revenues from premium upgrades (Chica and Rand, 2016). We address this by proposing an assortative matching system that classifies users and teams into groups of similar characteristics.

On the user-side, high-value prospects can be identified based on their future demand which can be captures through customer lifetime value (CLV) [5] (Berger and Nasr, 1998; Reinartz and Kumar, 2003; Gupta et al., 2004). The prediction of CLV in existing literature is typically based on historic purchase behavior. Two seminal approaches are to group users based on the recency, frequency and monetary value (RFM) of their historic purchases (Verhoef et al., 2003; Fader et al., 2005), and stochastic models of consumer purchasing which e.g. use the Pareto distribution to determine retention and the negative binomial distribution to determine purchase frequencies (Fader and Hardie, 2005).

With the increased availability of data and computing power, more recent data science approaches have emerged where the prediction of CLV relies on supervised learning with Random Forest (RF) algorithms (Vanderveld et al., 2016; Chamberlain et al., 2017).[6] RFs have been successfully applied to the prediction of purchase decisions in

[5]Customer lifetime value (CLV) or lifetime value (LTV) is a prediction of the net profit attributed to the entire future lifetime of the customer. We use CLV and LTV interchangeably in our text.

[6]The Random Forest algorithm is an ensemble prediction method that builds a forest of random uncorrelated decision trees to solve classification and regression prediction problems.

freemium mobile apps (Sifa et al., 2015), and have been shown to perform well in the prediction of consumer behavior (Lemmens and Croux, 2006; Coussement et al., 2016).

In our paper, we adopt a data-driven approach that allows to flexibly combine purchase behavior information with other data such as demographic and usage behavior information to help predict future demand (Sifa et al., 2015, 2018). One key advantage of this approach is that CLV predictions can be obtained with no or limited past purchase behavior. In our freemium setting, the share of overall premium users is small, and many future premium users have not made a purchase at the time when we need predictions of future demand (when the matching decision is made as described in more detail below). User-level predictions for our matching system are obtained with a decision-tree-based XGBoost predictor [7] that is trained using synthetically oversampled data (Chawla et al., 2002). Team-level classification is derived from a geometric mean-based approach that combines crucial behavioral elements into one "activity score" that can be used for classification. We will provide more detail on the exact methods in a section dedicated to the description of our matching system.

1.3 Environment

The context of our study is an online role-playing game (RPG). This game is played using an application for mobile devices that is available in the iOS (Apple) and Android (Google) app stores. The main channel of user acquisition is through digital

[7] XGBoost stands for eXtreme Gradient Boosting. Similar to Random Forest, it is a decision-tree-based ensemble Machine Learning algorithm that uses a gradient boosting framework to solve classification and regressio prediction problems.

advertising. Once in the game, players need to identify possible color patterns and align objects to form such patterns and make points (similar to Candy Crush). The environment allows users to play solo campaigns but is designed to encourage players to form a team with other players to defeat powerful virtual opponents.[8]

The unit of user activity or participation is a campaign. A typical campaign has the goal of defeating an opponent and can comprise varying numbers of pattern finding puzzles. Upon succeeding in a campaign, the entire team that the winning player is a part of, collects in-game rewards. Typically, one player chooses five characters from a set of possible characters (up to 40) to participate in each campaign. The choice set of characters varies according to specific pre-defined settings of campaigns, and the goal is to achieve maximum damage in each strike against the virtual opponents by being proficient at pattern finding in puzzles.

After a user has installed the application, they are first allowed to play solo campaigns only to learn how the game works, but within a short period of time of playing, they rank up and reach eligibility to join a team. The corresponding 25th, 50th, and 75th percentiles of time taken to reach eligibility are 22, 33 and 217 minutes. The average time to eligibility is 474 minutes since the distribution is highly skewed to the right.

Once eligible, users see a list of teams that are available to join. All teams have names and some of them also have descriptions.[9] The set of teams that are offered in

---

[8]The industry collaborator chose to remain anonymous. For confidentiality reasons, we use general terms such as teams, characters, opponents, and campaigns/missions instead of game-specific terms throughout and do not identify the game.

[9]Here are some typical team names (description): *Midnight Elementals* (...ready to bring the power of

these lists is randomly chosen from all active teams [10]. There are a total of 9000 teams in the game environment, and around 6000 of them are actively involved in accepting new users with a typical size of 21 / 25 (average and median) members. The upper limit of team size is set to be 30 members by the game provider. A summary of the process in a timeline is the following:

- New install
- Play solo campaigns and rank up
- Reach rank 4 (eligible to join a team)
  - Typical time (50th percentile) to rank 4 is 33 minutes
  - Users see a list of teams they can join
    - Team 1: name, country, member counts
    - ...
    - Team 30: name, country, member counts
- After joining a team
  - Play event campaigns with team members
  - Exchange messages and gifts within team

Once users join a team, they are still able to play solo campaigns, but now they can also start team campaigns with team mates and share rewards (see below) if the campaign

the elements to destroy the competition...); *nflesh* (no description); *Bald dwarfs* (... a team determined to be one of the greatest...)

[10] If all members of a team have not logged into the game for consecutive 14 days, the team will be considered inactive and removed from the team pool

succeeds. In terms of in-game rewards, for a typical player, it is substantially more efficient to work in a team than alone.

There are several important resources and reward mechanisms in this environment.

In-game currency (IGC): These are objects that can be purchased with USD in the store of the game at a fixed price or be earned as rewards after campaign successes. Examples of IGC include gems and gold.[11] In the game, IGC can be used to purchase other resources and can be exchanged with other players as gifts as well. Around 6% of users become payers (purchase IGC in an in-app purchase) within 37 days after joining the game.

IGC-purchased capital goods: Once the player acquires IGC in in-app purchases, she is able to make in-game purchases. The in-game store includes capital goods (character cards and synergists) that make users' characters more powerful and give them higher chances to complete difficult solo campaigns and achieve higher contributions to team campaigns. These goods have different levels of durability. Some are targeted for specific campaigns, others can be used across campaigns and do not expire.

Within the team, users can write messages with and make gifts of IGCs and other in-game objects (character cards, synergists, keys) to other team members or to the team as a whole. Message space is not restricted, i.e., any written message in

[11] In the actual environment, IGCs have different names.

natural language is allowed. Messaging typically involves explaining game concepts and strategies to the new users (learning) and exchanging capital goods (trading) to equip characters before campaigns.

The environment is built to encourage participation in teams. The main type of player activity, called the *weekly event*, is designed so that players' strengths are highly complementary within the team. These weekly events are themed campaigns that run on a weekly basis and designed so that users can experience new tasks and challenges every week. These events also require a team to equip players in complementary characteristics and skill advantages to maximize striking power against virtual enemies. Consequently, these events can vary in quality and popularity, as well as in generating engagement, but consistently generate the majority of user engagement and revenue.

With a variety of solo and team campaigns, the environment can serve a rich variety of players with different skills and different levels of willingness to spend real money. In-app purchases, overall, provide a more advanced experience to premium users compared to that of free users. Similarly, teams that spend more tend to exhibit more engagement, perform better in their campaigns, and receive higher rewards when completing campaigns.

Despite the strong complementarity of skills and willingness to spend inherent in this setup, the environment does not facilitate or actively promote that players of higher skill or willingness to purchase form teams among themselves (seek positive assortative matching). Although players can move to other teams that fit them better, information on team characteristics is limited and scattered. Teams can only be searched by team name, thus for a new user who needs to make a team choice for the first time, search is uninformed. For existing users who have participated in joint campaigns and have access to the team rankings (displayed on a leaderboard visible to all players that joined a team), it is also hard to find a specific team name from thousands of teams on the leaderboard. The majority of users hence relies on the team recommendation page provided by the game provider when joining a new team instead of performing an active search for a team that matches their preferences. Essentially, new users' selection of teams of different quality, e.g. high or low activity, occurs at random.

On this background, this paper studies the effects of a positive assortative matching of new users and teams by surfacing high (low) quality teams – in terms of activity and future premium demand – on high (low) quality users' – in terms of future premium demand – team recommendation page.

1.4 Matching Mechanism

Given the skill complementarity in the environment, players with higher engagement potential benefit from joining a more skilled team. However, as mentioned before, at the moment of the design of this study, the game environment did not have relevant channels and mechanisms to facilitate users to seek a better matching. In that context, we designed a matching mechanism and implemented an experimental study to evaluate its effect on participants' behavior.

Since engagement and revenue are highly correlated, our matching mechanism is based on targeting potential high-value users and introducing them into high-value teams. Individuals with higher potential are given a list of high-activity teams (in terms of engagement and revenue) to choose, as opposed to the random set they are given normally. The user is ultimately the one that decides which team to join.

The model that sustains the mechanism can be decomposed into two classification problems: one for the users and one for the teams. Users are classified between likely payers and likely non-payers.[12] Teams are classified as high-activity or low-activity. With these two components (predicting the probability of new users to become "payers" in the future, and classifying teams by their engagement), the system is able to promote those with high paying probabilities to enter more engaged teams.

1.4.1 Team Classification

For implementing team classification, we leveraged institutional expertise from experienced product managers to determined what constitutes a highly active user community. Jointly with these experts, we devised an *activity score* at the team level. As inputs to this score, we use the number of active players per day, number of campaigns played together, the volume of gifts and messages exchanged between team members, and revenue generated. We calculated a 14-day moving average for each input feature, and then aggregated those into the "activity score" using a geometric mean.

We implemented a simple test of our team classification. We sorted all teams

[12] We use the term "likely" (as opposed to "actual" on the user side) for the following reasons. First, team matching happens at a very early stage of players' lifetime, usually 30 minutes after install. At this point in time, players usually have made zero in-game purchases. Second, users that actually convert to payers 30 days after install are still a small percentage of all users.

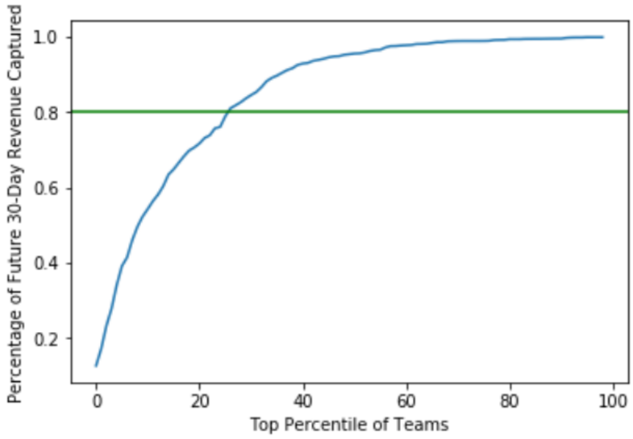


Figure 1.1: Team Activity Score and Future 30-day Revenue Captured.
The vertical axis represents the percentage of future 30-day revenue captured, and the horizontal axis indicates the percentile of teams (from high to low) sorted by team activity score

in an arbitrary day 0, and measured what share of revenues of the next 30 days came from the high-activity teams. We found that the top 30% of teams in terms of the score captured at least 80% of revenue (Figure 1.1). Therefore, we use the 30th percentile to set our "activity" threshold for the team classification model.

This method to calculate the activity score is more preferable to using only team revenue, because it not only ensures that the high-activity teams capture majority of game revenue, but also incorporates other metrics of engagement and experience that provide a more complete picture of highly engaged and successful teams.

1.4.2 User Classification

For user classification, we use supervised learning techniques, which use behavioral traces commonly collected by mobile application developers as input features to predict user lifetime values. There are several technical challenges associated with this exercise. First, since team recommendation typically occurs relatively early in a user's in-app experience, the behavioral traces collected are limited and there is almost no purchase history. Also, there is a strong class imbalance problem by which only a small share of users actually become payers in the future (this class imbalance problem has been documented in several previous studies (Weiss, 2004; Nieborg, 2015, e.g.)).

In this context, we use purchases in dollars (CLV or LTV) as the target variable to be predicted by our classification model. Specifically, we add all purchases in the next 30 days after the user becomes eligible to join a team. Input features of the user classification model includes both demographic variables (user's language, type of device and country) as well as data characterizing early behavior in the game, collected by the game provider from app install up to the moment of team eligibility. Behavioral data includes the time (in minutes) users take to reach each rank before team eligibility, number of sessions and solo campaigns played and succeeded, highest rank achieved and revenue generated before eligibility. Table 1.1 summarizes the user behavioral characteristics. The training data set for our user prediction model includes one month of new installs who reached team eligibility within 7 days after joining the game, which is around 52,000 individuals.

In order to solve the class imbalance problem, we apply a synthetic oversampling technique (SMOTE) to generate synthetic samples from the class of actual payers (Chawla et al., 2002). During training, we perform a five-fold cross-validation (CV) to avoid model over-fitting. On each training set during the CV process, we conduct SMOTE oversampling of the minority class and generate 2 times the number of actual

Table 1.1: User Behavioral Statistics Before Team Eligibility

	Count	Mean	std	Min	25%	50%	75%	Max
Rank2 Mins	51672.0	130.69	669.21	0.0	6.0	7.0	11.0	9995.0
Rank3 Mins	51675.0	220.28	843.46	3.0	10.0	14.0	25.0	10043.0
Rank4 Mins	51675.0	474.09	1239.28	4.0	22.0	33.0	217.0	10076.0
Sessions	51675.0	1.86	1.26	1.0	1.0	1.0	2.0	31.0
Campaign Start	51675.0	4.11	0.56	2.0	4.0	4.0	4.0	38.0
Campaign Win	51675.0	3.96	0.25	0.0	4.0	4.0	4.0	6.0
Max Rank	51675.0	3.81	0.48	0.0	4.0	4.0	4.0	17.0
Revenue	51675.0	0.11	1.78	0.0	0.0	0.0	0.0	173.62

Basic behavioral statistics used as input features for predicting user LTV: minutes taken to reach each rank, number of sessions played, number of campaigns played and succeeded, maximum rank achieved, and revenue generated.

payers, and perform feature selection and hyper-parameter tuning to minimize root mean squared errors (RMSE) to achieve best model performances.

1.4.2.1 Model Evaluation and Implementation

Final performance of the model is then evaluated on a 1-month time-independent hold-out set.[13] We focus on assessing the hit rates of our algorithm, where we sort users by their predicted 30-day revenue and evaluate how many actual payers we correctly capture for different ratios of this ranking. XGBoost stands out to have the best hit rate performance compared to Linear Regression, RF (Random Forest), and GBM (Gradient Boosting Machine). Figure 1.2 reports the hit rate performance of our XGBoost model, where we are able to capture about 80% of total actual payers with our top 50 percentile sorted prediction rankings.

[13]Using a time-independent hold-out set provides an unbiased evaluation of our user classification model's prediction power. Similar to the training data set, the test data set uses 1-month of new installs, independent of the 1-month of data used in the training set, and contains around 100,000 individuals.

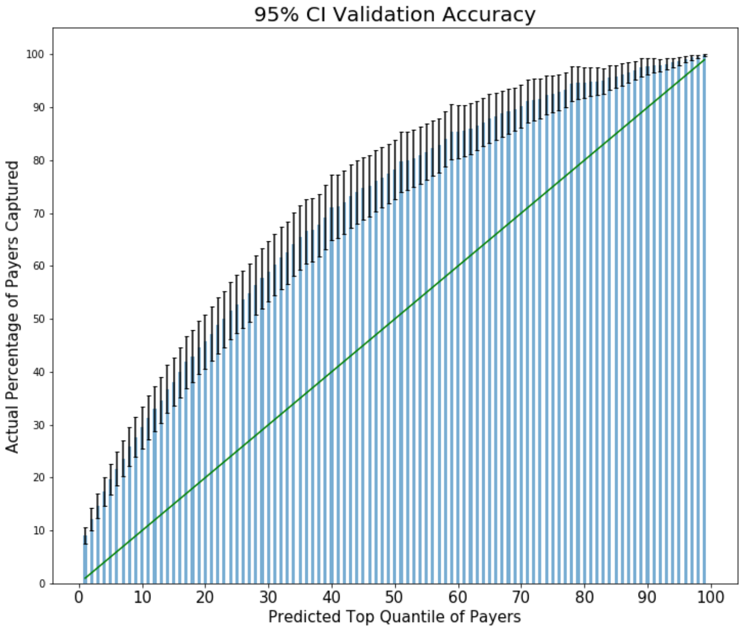


Figure 1.2: Hit Rate Performance of XGBoost Model.
The hit rate performance of the XGBoost Model is evaluated on the time-independent test set. The vertical axis represents the percentage of actual payers captured in our prediction model, and the horizontal axis indicates the different rankings (from high to low) after sorting all users by their predicted LTV.

The XGBoost model deployed to production continuously makes 30-day revenue predictions for new users when the user arrives at the team recommendation page. It re-trains itself daily on the newest training data available, and automatically performs feature selection and parameter tuning.

After obtaining a prediction for each new user entering the game, we classify users into two categories: likely payers and likely non-payer. Consistently on the team side, we have high-activity and low-activity teams. The number of available team spots in each category (i.e supply of the matching market) are fixed because we determine the number of high-activity teams up to the point that they capture at least 80% of future 30-day revenues. In order to equilibrate supply and demand in the matching market, the threshold that separates a likely payer vs. non-payer depends on the following rule:

- We sort all new users based on their predicted 30-day revenue in descending order.
- We direct as many users on the top of the ranked list to high-activity teams by pulling their recommended team lists from the high-activity team category only. Users will then have freedom to choose which specific team they join.
- We move the rest of the new users to be matched to the low-activity team category, and the threshold of separating likely payers and likely non-payers is determined by the lowest revenue projection of the "paying" group.

Thus the matching market will always be in equilibrium where demand (number of players looking for a team spot) is equal to supply (number of available team spots)[14].

[14] The number of available team spots in general is higher than the number of players looking for a team, so each player is guaranteed a team spot

The system that determines the threshold that separates high-activity teams and likely payers refreshes every night based on the most updated supply and demand information collected at the end of the day.

Since the implementation of the system, roughly 70% of new users are categorized as likely payers and the rest are classified as likely non-payers. For existing users who are looking to switch teams, the system will only recommend high-activity team environments for them, under the assumption that existing users actively switching teams are looking for a more engaged and active social environment. Figure 1.3 shows a time series of user predictions since the implementation of the system. Note that the system only makes a user prediction for new users obtaining team eligibility and existing users looking for a new team, and the majority of users are mapped to high-activity team environments.

1.4.3 Offline Evaluation of the Mechanism

Before deploying the system for experimentation, we performed offline evaluations to study the potential impact of our assortative matching system. We use historical observations of voluntary user-team matches, with a propensity score matching regression for causal analysis.

Propensity score is a technique that assigns treatment probability based on observed characteristics of the user: what is the probability for the user to join a high-activity team if the user possesses certain characteristics. The matching estimation then compares treatment effects between paired observations that share similar propensity

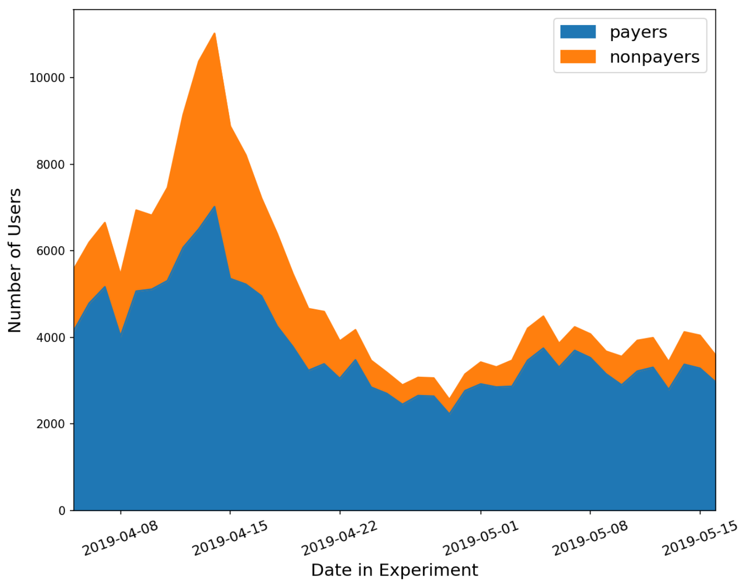


Figure 1.3: Time Series of System User Prediction Decisions.
This figure shows a time series of user predictions since the beginning of the experiment. The vertical axis represents total number of users classified as likely payers and likely non-payers, and the horizontal axis represents dates during the experiment

scores, which eliminates the confounding factors that predict probability of treatment (join high-activity teams) rather than treatment itself, and derives treatment effects on observational data that are more similar to randomized control experiments.

We first apply our team and user classification models on one month of historic users who joined between 60 to 90 days before we introduced our new matching system. Those users joined their first team based on a random list of teams that consists of both high- and low-activity teams. We then run propensity score matching regressions for both likely payers and likely non-payers, and report impact of joining a high-activity team on post-joining engagement and retention.

Figure 1.4 reports the coefficients of impact, as well as their 95% confidence intervals, of joining a high-activity team on likely payers and non-payers 14 days after joining. The figure illustrates that for likely payers, joining a high-activity team makes them significantly more engaged in sending messages and participating in weekly events. For likely non-payers, joining a high-activity team imposes non-significant impacts on them.

1.5 Experiment

1.5.1 Design

After obtaining positive results from our offline evaluations, we started our experimentation. However, due to engineering constraints, we were not able to conduct a fully randomized AB test for our matching system. Instead, we performed a time-split

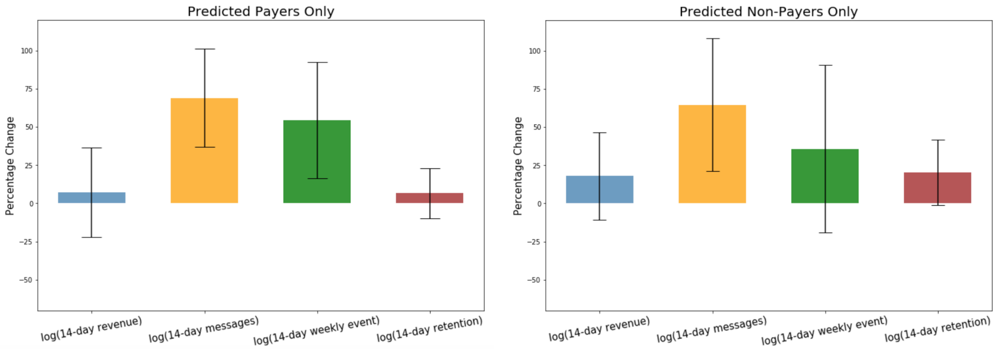


Figure 1.4: Propensity Score Analysis Coefficients.
The vertical axis represents percentage increases in target metrics due to joining a high-activity team compared to joining a low-activity team, and the horizontal axis is grouped by analysis metrics of interest

quasi-experiment, where the system was turned on and off on a 3-day interval over a 6-week period. During the "on" period, users and teams are matched based on assortative matching, and during the "off" period, users and teams are matched randomly.

The game environment where the system is implemented relies heavily on evolving weekly events for game performances. Weekly events can vary greatly in quality and popularity, as well as overall performance in terms of engagement and revenue. Running the experiment on a 3-day interval makes sure that each weekly event is split equally between treatment and control groups. The 6-week period ensures that the day of the week effect is equalized in both groups. Moreover, alternating between system "on" and "off" equally captures any potential time-series seasonality effects in the two groups over the 6-week period.

Figure 1.5 is an example of the monitoring dashboard we created for tracking system performance during the experiment. Users are grouped by whether they joined during experiment "on/off" on the x-axis, and the y-axis represents mean and 95%

confidence interval of the cumulative activity levels during day 0, 1, 3, and 7. The figures show that users who joined during the system "on" period are significantly more active in engagement (especially in sending messages) than users who joined during the system "off" period.

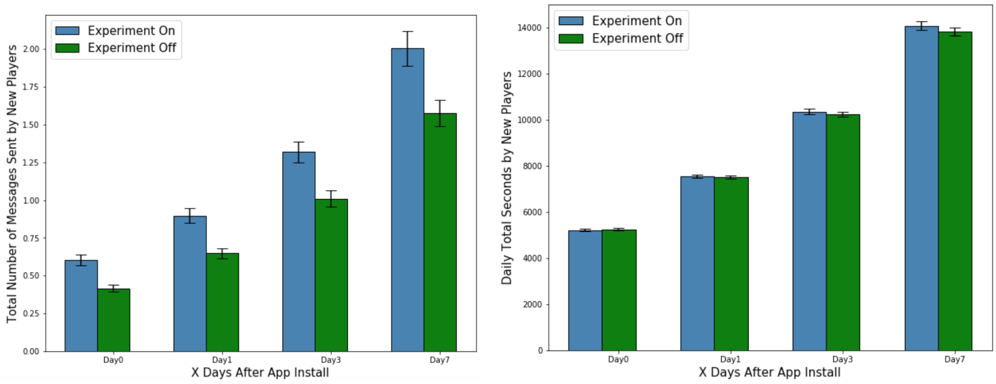


Figure 1.5: Monitoring Dashboard of Messages Sent and Daily Time (in seconds) Spent in Game by New Users.
The vertical axis represents total number of messages sent or daily seconds spent by the new users (averaged across all new users), and the horizontal axis is grouped by the number of days users have joined the game

1.5.2 Metrics and Hypotheses

We use the following metrics to evaluate the impact of our matching system. Metrics are used in individual-level analysis and also aggregated for the team-level analysis. For individual analysis, we consider 14 days of data starting at the moment of app install. For team-level analysis, we consider data for the last 14 days of the experiment, after all teams have received a considerable amount of experiment exposure. All metrics are in log terms. We also performed user-level analysis on metrics from 3 days to 30 days after the user installed the app, and team-level analysis from 14 days to 3 days towards the end of the experiment, and received consistent results on all metrics.

- **Revenue:** The cumulative amount of U.S. dollars the user(s) spent (for individual analysis), or the average among team members (for team-level).

- **Number of Message:** Total number of in-team messages sent by the user (for individual-level analysis) or the average team members (for team-level analysis).

- **Weekly Events:** Total number of weekly-event campaign starts by the user (for individual-level analysis) or the average team members (for team-level analysis).

- **Retention:** Indicator variable of whether the user is still active 14 days after app install.

- **In-Game Time Spent:** Time spent in the game in minutes, a measure of user effort. Total minutes for individual-level analysis or team average for team-level analysis

- **Campaign Success Rate:** Total campaign wins/total campaign starts, as measure of efficiency

We test the following hypotheses:

1. Matching likely payers into highly active teams will have positive impact on players' engagement and productivity (for both new as well as existing players).

2. Matching likely non-payers into low-activity teams will have negative impact on players' engagement and productivity (for both new as well as existing players).

3. The overall net effect of the assortative matching system will be positive in engagement and productivity.

1.5.3 Results

In the following two sections, we analyze effects of the proposed assortative matching system at the user- and team-level. Overall, results paint a positive picture of the system's impact on user engagement and monetization, suggesting that assortative matching approaches can contribute significantly to improve user experiences and firm profits in digital settings, and in freemium settings in particular. [15]

1.5.3.1 User-level Analysis

The user-level analysis focuses on users who joined during the experimental period. Our experimentation approach assumes that the quality of newly arriving users is randomly distributed across calendar days, and switches the matching system on and off in three-day intervals. As the treatment is not directly assigned to each individual, our approach is akin to a quasi-experiment, and we resort to an instrumental variables (IV) approach for causal inference. The instrumental variables approach enables us to evaluate the more direct effect of joining a team that matches the user's classification (i.e. the impacts of taking up the intervention) on outcomes, rather than the impact of the intention to treat. A flag indicating if a user joined during an "on" or "off" period serves as an instrument to predict the activity level of the team that the player

[15] Results on additional outcome variables are reported in the appendix.

joined, and we then regress user engagement and productivity metrics on predicted team activity with multiple user-level and weekly-event-level control variables. Concretely, we use the following two-stage instrumental variable specification:

Stage 1:

$$TActivity_i = \alpha_0 + \alpha_1 sysON_i + \alpha_2 User_i + \alpha_3 Event_j + \epsilon \tag{1.1}$$

Stage 2:

$$Metric_i = \beta_0 + \beta_1 TActivity_{pred} + \beta_2 User_i + \beta_3 Event_j + \epsilon \tag{1.2}$$

Where $TActivity_i$ is a binary variable that equals 1 for high-activity teams and 0 otherwise, $sysON_i$ indicates whether the matching system was on ($sysON_i = 1$) when the user first joined the game, $User_i$ and $Event_j$ are user- and weekly-event level control variables [16], $TActivity_{pred}$ is the predicted team health score from stage 1, and $Metric_i$ are the users' various engagement and productivity metrics as measured during a 14-day window after joining the game.

Table 1.2 reports IV regression results on likely payers and likely non-payers, respectively. New users who are classified as likely payers (high future premium demand) by our prediction model join high-activity teams during periods when the system is on, and are equally likely to join high-activity and low-activity teams during periods when the matching system is off.

The impact of joining a high-activity team is positive on all 14-day[17] game-

[16] Both user- and weekly-event level control variables are dummies to control for the fixed effects of differences associated with users and weekly-events.

[17] We observe user behavior cumulatively over different time windows after users join the game. We call these x-day metrics with x in [1, 3, 7, 14, 30]. This approach is akin to an intent-to-treat measurement

specific metrics. The coefficients on engagement metrics (team chat messages and weekly event participation) are significant and positive: Messages increase by 36% and weekly event engagement increases by 26%. In addition, 14-day retention increases significantly by 7.8%.

On the side of users who are classified as non-payers, and hence join low-activity teams during "on" periods (and a team of either type with equal chance during "off" periods), we observe no significant impact of the matching system on all 14-day metrics. The user-level experiment results are also consistent with our propensity score regression analysis on historical data presented as part of the offline evaluation.

On the user-level productivity measures, we observe significant and positive impact of joining active teams on time spent in game (+8.1%), and non-significant impact for joining a low-activity team. However, we do not observe increased individual efficiency in terms of campaign success rate for either type of user.

User-level net effects of the matching system are summarized in Table 1.3. The regression model used for analyzing the overall effect of our new system is the following:

$$\begin{aligned} Metric_i = & \beta_0 + \beta_1 sysON_i + \beta_2 NonPayer_i + \beta_3 sysON_i * NonPayer_i \\ & + \beta_5 User_i + \beta_6 Event_j + \epsilon \end{aligned} \tag{1.3}$$

Where $NonPayer_i$ is a binary variable that equals 1 if user is a likely future non-payer and 0 otherwise.

Based on the Ordinary Least Squares (OLS) regression with the exogenous

to account for endogenous non-compliance.

Table 1.2: User-Level IV Regression on Game-Specific Metrics and Productivity Metrics

	Game-Specific Metrics				**Productivity Metrics**	
	(a) Predicted Payers					
	Revenue	Message	Weekly Events	Retention	Time Spent	Success Rate
Intercept	-2.2**	-1.3	1.6	-2.3***	6.3***	1.0***
	(1.1)	(1.2)	(1.8)	(0.62)	(0.77)	(0.064)
$TActivity_{pred}$	0.062	0.36***	0.26***	0.078***	0.081**	0.0026
	(0.046)	(0.053)	(0.08)	(0.027)	(0.038)	(0.0052)
Campaign						-0.00083***
$Starts_{day14}$						(2.1e-05)
R-squared	0.04	0.02	0.04	0.02	0.015	0.42
Observations	17965	17965	17965	17972	17965	17965
	(b) Predicted Non-Payers					
	Revenue	Message	Weekly Events	Retention	Time Spent	Success Rate
Intercept	-2.1***	-2.1***	-1.1*	-2.1***	4.8***	0.98***
	(0.24)	(0.42)	(0.65)	(0.23)	(0.44)	(0.05)
$TActivity_{pred}$	-0.011	0.12	0.24	0.037	-0.073	-0.003
	(0.088)	(0.15)	(0.24)	(0.085)	(0.13)	(0.019)
Campaign						-0.00088***
$Starts_{day14}$						(2.5e-05)
R-squared	0.02	0.02	0.04	0.01	0.027	0.28
Observations	13645	13645	13645	13646	13645	13645

* p<.1, ** p<.05, ***p<.01 (standard errors in parentheses)
User-level IV regressions include user device group controls, acquisition channel controls, weekly event controls, and day of week controls.

Table 1.3: User-Level Net Effects on Game-Specific Metrics and Productivity Metrics

	Game-Specific Metrics				**Productivity Metrics**	
	Revenue	Message	Weekly Events	Retention	Time Spent	Success Rate
Intercept	-1.9***	-1.7***	-0.024	-2***	5.2***	0.98***
	(0.33)	(0.42)	(0.64)	(0.22)	(0.38)	(0.041)
$SystemON(1)$	0.025	0.25***	0.18***	0.046***	0.044*	-0.0005
	(0.023)	(0.029)	(0.044)	(0.015)	(0.023)	(0.0031)
$NonPayer(1)$	-0.24***	-0.2***	-0.35***	-0.056***	-0.16***	-0.012***
	(0.023)	(0.028)	(0.043)	(0.015)	(0.021)	(0.0028)
$ON*$	0.0045	-0.24***	-0.14***	-0.025	-0.038	0.0024
$NonPayer$	(0.028)	(0.035)	(0.054)	(0.019)	(0.027)	(0.0038)
$Campaign$						-0.00084***
$Starts_{day14}$						(1.7e-05)
R-squared	0.046	0.028	0.055	0.022	0.024	0.37
Observations	31610	31610	31610	31618	31610	31610
Net Effects T-Test	0.019	0.1***	0.086***	0.025***	0.019	0.0004
	(0.014)	(0.017)	(0.026)	(0.0091)	(0.013)	(0.839)

* p<.1, ** p<.05, ***p<.01 (standard errors in parentheses)
User-level net effect regressions include user device group controls, acquisition channel controls, weekly event controls, and day of week controls
T-test hypothesis: $0.7 * SystemOn + 0.3 * (ON * NonPayer) = 0$

variable of interest $sysON_i$, we can calculate the overall net effect of the matching system. Since the share of likely payers is roughly 70% of all new users, the net effect for each metric can be determined using a t-test with the hypothesis that $0.7 * SystemON + 0.3 * (ON * NonPayer) = 0$. The results from the t-tests are shown at the bottom of Table 1.3 and indicate user-level significant net positive impact of the matching system on messages, weekly event participation and retention. Additionally, there are non-significant net positive effects on revenue and productivity measures.

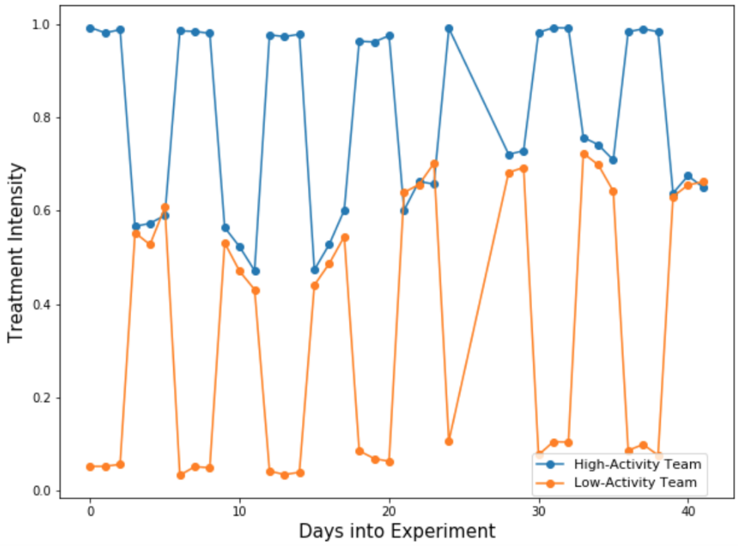
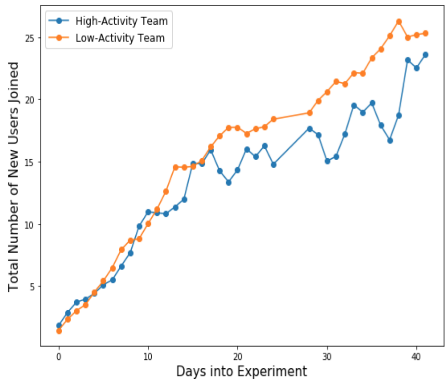


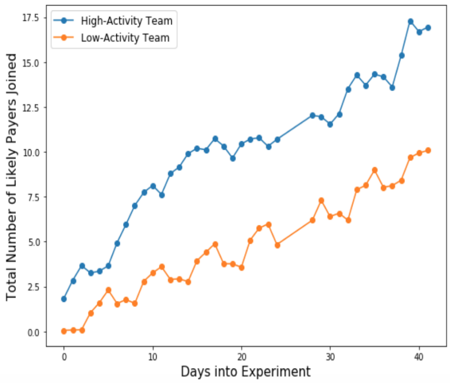
Figure 1.6: Treatment Intensity Over Experiment Time.
The vertical axis represents the proportion of likely payers joining two types of teams, and the horizontal axis indicates the number of days into the experiment with "on" and "off" treatment alternating in three day intervals, starting with the system being on. The orange lines indicate the proportion of likely payers joining a low-activity team, which is close to 0 during system "on" periods and around 60-70% during system "off" periods.

1.5.3.2 Team-level Analysis

Our experimental approach introduces exogenous variation in the likelihood of having predicted payers (versus predicted non-payers) join existing social groupings in the game. As shown in Figure 1.6, when the system is on, almost 100% of the the new joins to a high-activity team are likely payers. During periods when the matching system is off, the share of likely payers joining high-activity teams is around 30% which is consistent with approximately 30% of teams being classified as high-activity by the team classification mechanism. In essence, the experiment creates a separating path for high- versus low-activity teams. As the experiment progresses, high-activity teams will attract likely payers, while low-activity teams add non-payers (See Figure 1.7).



(a) Total Number of New Users Joined



(b) Total Likely Payers Joined

Figure 1.7: Teams taking separating path over experiment period.
The vertical axis represents the total number of new players joining a team, and the horizontal axis represents number of days into the experiment with "on"/"off" treatments alternating in three day intervals.

To analyze the effect of the system at the team-level, we run standard OLS regressions using engagement and productivity metrics from the last 14 days of the experiment to evaluate the impact of adding more potential payers on the high- and low-active teams. For engagement metrics, we use revenue, campaign starts in weekly events, and messaging. For productivity metrics, we use time spent and campaign success rate. For every dependent variable we test (except for the net effect analysis), we use team-level averages [18].

As regressors we include $ln(NPayers_g)$, which denotes the number of likely payers that have joined team g during the experiment, $TActivity_g$ is a dummy variable indicating high-activity teams ($TActivity_g = 1$), and the interaction term between the two. We also control for pre-experiment team-level characteristics, including 14-day cumulative campaign starts, messaging, gifts, revenue, and team age and size the day

[18] The average activity per user within a team.

before the experiment in $team_{g,pre}$, and the total number of new users ($NewJoins_g$) that joined during the experiment.

$$ln(Metric_g) = \beta_0 + \beta_1 ln(NPayers_g) + \beta_2 TActivity_g + \beta_3 ln(NPayer_g) * TActivity_g \\ + \beta_4 team_{g,pre} + \beta_5 NNewJoins_g + \epsilon_g \tag{1.4}$$

Table 1.4 reports team-level results using the average activity of all team users. A 1% increase in likely payers joining a high-activity team significantly increases average revenue (+0.19%), weekly event participation (+0.1%) and messages (+0.088%) exchanged in a team. While the matching system has no significant impact on revenue at the user-level (Table 1.2 and 1.3), it does have positive impact on revenue for high-activity teams. In essence, improving the overall quality of the team environment is positively affecting existing members of the teams, making them willing to spend and engage more.

If we exclude all newly joining users and focus on the existing members of high-activity teams, the effect (Table 1.5) becomes even stronger: Average revenue increases by 0.24%, weekly event participation improves by 0.12%, and messages are up 0.15% for every 1% increase in the number of likely payers joining. The addition of likely payers into the high-activity teams boosts the overall health of the team environment and encourages existing members to be more involved in spending and activities overall.

We observe a similar pattern of effects for time spent in the general productivity metrics for high-activity teams. However, the addition of high quality new users does

Table 1.4: Team-Level Analysis on Game-Specific Metrics and Productivity Metrics (Average of All Team Users)

	Game-Specific Metrics			**Productivity Metrics**	
	Revenue	Weekly Event	Message	Time Spent	Success Rate
Intercept	0.77***	4.1***	0.93***	5.5***	0.78***
	(0.16)	(0.16)	(0.12)	(0.16)	(0.018)
$T_{LowActive}$	-1.7***	-0.59***	-2***	-0.64***	-0.002
	(0.12)	(0.092)	(0.096)	(0.087)	(0.0094)
$ln(NPayer)$	0.19***	0.1***	0.088***	0.013	-0.0012
	(0.034)	(0.029)	(0.03)	(0.027)	(0.0027)
$ln(NPayer)*$	-0.089**	-0.15***	-0.071**	0.0096	0.0054
$T_{LowActive}$	(0.039)	(0.037)	(0.032)	(0.038)	(0.0039)
$NNewJoins$	-0.033***	-0.018***	-0.013***	-0.0028	0.00017
	(0.003)	(0.0027)	(0.0023)	(0.0024)	(0.00026)
Campaign					4.3e-06***
$Starts_{day14}$					(8.7e-07)
Team Controls	✓	✓	✓	✓	✓
R-squared	0.52	0.31	0.68	0.23	0.044
Observations	4635	4635	4635	4635	4635

* p<.1, ** p<.05, ***p<.01 (standard errors in parentheses)
Team control variables include pre-experiment 14-day cumulative campaign starts, messaging, gifts, revenue, and age and size the day before the experiment.

not improve the efficiency of campaigns, i.e. campaign success rate.

On the other hand, the addition of likely non-payers into low-activity teams leads to significantly negative effect on average team activity (i.e. revenues -0.089%, weekly events -0.15% and messages -0.071% for every 1% increase in the number of likely non-payers joining), leading to lower willingness to play and spend by existing low-activity team members in the game and a weaker environment for those teams.

There are multiple mechanisms through which these effects may realize. A better, more homogeneous team environment is a reasonable explanation for the increase in spending by existing players and the increase in overall team engagement and spending.

Table 1.5: Team-Level Analysis on Game-Specific Metrics and Productivity Metrics (Average of Existing Team Users)

	Game-Specific Metrics			Productivity Metrics	
	Revenue	Weekly Event	Message	Time Spent	Success Rate
$Intercept$	0.67***	4.1***	0.92***	5.4***	0.77***
	(0.16)	(0.18)	(0.13)	(0.17)	(0.02)
$T_{LowActive}$	-1.7***	-0.56***	-2***	-0.6***	-0.0019
	(0.12)	(0.098)	(0.1)	(0.091)	(0.0098)
$ln(NPayer)$	0.24***	0.12***	0.15***	0.058**	0.0029
	(0.037)	(0.032)	(0.032)	(0.03)	(0.0031)
$ln(NPayer)*$	-0.18***	-0.15***	-0.13***	-0.041	0.0017
$T_{LowActive}$	(0.041)	(0.04)	(0.034)	(0.039)	(0.0042)
$NNewJoins$	-0.03***	-0.01***	-0.011***	-0.001	-0.00012
	(0.0033)	(0.0032)	(0.0026)	(0.0028)	(0.00031)
$Campaign$					4.4e-06***
$Starts_{day14}$					(8.7e-07)
$TeamControls$	✓	✓	✓	✓	✓
R-squared	0.51	0.27	0.66	0.23	0.046
Observations	4635	4635	4635	4635	4635

* p<.1, ** p<.05, ***p<.01 (standard errors in parentheses)
Team control variables include pre-experiment 14-day cumulative campaign starts, messaging, gifts, revenue, and age and size the day before the experiment

A further mechanism, particularly explaining the increase in spending by existing players, may present itself in social competitive dynamics: Existing players may perceive and experience increased within-group competition when spending players and hence players with stronger card decks enter their team. They may therefore increase their effort comparatively, to preserve authority and/or to impress newly arriving team members.

To analyze the net effect of the system at the team-level, we run OLS regressions with team-level total activities (including all team members) as dependent variables to evaluate the impact of additional potential payers and non-payers on the overall performance of each type of teams. We then calculate a net effect aggregating the team-level impacts to address the overall effects of the matching system. As regressors we include $NPayer_g$ and $NNonPayer_g$ which denote how many likely payers and non-payers join during the entire experiment period. $TActivity_g$ is a dummy variable indicating high-activity teams, and the interaction term between the team activity dummy and the number of payers and non-payers joined. We also control for pre-experiment team-level characteristics, including 14-day cumulative campaign starts, messaging, gifts, revenue, and team age and size the day before experiment in $team_{g,pre}$.

$$\begin{aligned} ln(Metric_g) = \beta_0 + \beta_1 TActivity_g + \beta_2 NPayer_g * TActivity_g \\ + \beta_3 NNonPayer_g * TActivity_g + \beta_4 team_{g,pre} + \epsilon_g \end{aligned} \tag{1.5}$$

Table 1.6 reports the overall net effect of the matching system using team-level data. Given that the system classifies roughly 70% of new users as likely payers and 30% as likely non-payers, and that high-activity teams capture at least 80% of all future revenue and engagement in the game while low-activity teams only capture 20%, we perform the following t-tests with null hypothesis $0.7 * 0.7 * (0.8 * NPayer * T_{HighActive} - 0.2 * NPayer * T_{LowActive}) + *0.3 * 0.3 * (0.2 * NNonPayer * T_{LowActive} - 0.8 * NNonPayer * T_{HighActive})$. The results from the t-tests at the bottom of table 1.6 indicate significant net positive impact on revenue, weekly event participation and messages at the team-level. There further is a positive but non-significant effect on time spent in the game.

1.6 Conclusion

This paper provides experimental evidence of the impact of assortative matching between teams and players in a digital – specifically freemium – environment. We combine machine-learning classification of players and teams and test the matching system in a time-series quasi-experimental setting. Our premium player prediction model using behavioral and demographic features collected at an early stage of user lifetime in the app is 60% more likely to correctly capture actual payers than random guessing. Our team classification model is able to identify the top 30% of high-activity teams that contribute 80% of next month's revenue.

By matching users into teams of similar features and engagement levels we emphasize the complementarity of in-game experiences and encourage high-value new users to be more engaged in communication and campaign participation – leading to

Table 1.6: Team-Level Net Effects Analysis on Game-Specific Metrics and Productivity Metrics (All Team Users)

	Game-Specific Metrics			Productivity Metrics	
	Revenue	Weekly Event	Message	Time Spent	Success Rate
$Intercept$	2.3***	5.8***	2.4***	7.2***	0.78***
	(0.2)	(0.19)	(0.16)	(0.18)	(0.017)
$T_{LowActive}$	-3.1***	-1.5***	-3.2***	-1.6***	-0.011
	(0.15)	(0.11)	(0.12)	(0.11)	(0.0086)
$NPayer*$	0.022***	0.01**	0.018***	0.011***	1.2e-05
$T_{HighActive}$	(0.0052)	(0.0041)	(0.0044)	(0.0041)	(0.00025)
$NPayer*$	0.079***	0.0045	0.03***	0.043***	0.0018*
$T_{LowActive}$	(0.016)	(0.013)	(0.012)	(0.013)	(0.001)
$NNonPayer*$	-0.095***	-0.028***	-0.04***	-0.013	-0.0016***
$T_{HighActive}$	(0.017)	(0.01)	(0.012)	(0.0094)	(0.00056)
$NNonPayer*$	0.026	0.022*	0.031**	0.028***	0.002***
$T_{LowActive}$	(0.018)	(0.012)	(0.013)	(0.011)	(0.0007)
$Campaign$					4.3e-06***
$Starts_{day14}$					(8.6e-07)
$TeamControls$	✓	✓	✓	✓	✓
R-squared	0.58	0.45	0.68	0.42	0.045
Observations	4635	4635	4635	4635	4635
Net Effect T-Test	0.0081**	0.0059*	0.0076**	0.0016	-2.437e-05
	(0.004)	(0.003)	(0.003)	(0.003)	(0.000)

* p<.1, ** p<.05, ***p<.01 (standard errors in parentheses)
Team control variables include pre-experiment 14-day cumulative campaign starts, messaging, gifts, revenue, and age and size the day before the experiment
T-test hypothesis: $0.7*0.7*(0.8*NPayer*T_{HighActive}-0.2*NPayer*T_{LowActive})+$ $*0.3*0.3*(0.2*NNonPayer*T_{LowActive}-0.8*NNonPayer*T_{HighActive})$

higher retention rates and higher productivity.

Moreover, introducing new users of similar behavior as the existing team members creates a separating path between teams of different qualities. Highly engaged teams maintain their high-activity environment and existing high-quality team members spend and play more. However, the matching of low-quality players and teams has negative impact on the low-quality teams, further destabilizing the team environment and discouraging the development of these players and their communities. As, by definition, high-quality teams and players account for the majority of engagement and revenue, overall effects of the matching system are positive.

Our approach can be useful to improve firm-side outcomes and retain a larger number of valuable new users. However, from a wider societal perspective, this type of matching system can have detrimental effects on poorer, weaker or more marginal segments of the user population, leading to higher inequality in the digital environment. Such inequality can possibly adversely impact wider welfare, e.g. when considering online dating on platforms such as Tinder. Since an increasing, already sizable, portion of human life is being spent in digital environments, we believe that designers of social interaction systems (including policy makers) should be aware of the trade-offs between net private benefits that can be achieved by positive-assortative matching systems, and its potential costs on the weaker segments of users. We also believe that understanding these trade-offs of machine learning-powered matching systems should be a priority in the future research of fairness in machine learning.[19]

[19] Current literature is focused on other aspects of fairness in machine learning (e.g., Chouldechova and Roth (2018))

Chapter 2

The Effect of Monetary Rewards on Image Motivations in Charitable Giving [1]

2.1 Introduction

Prior literature has identified three motives for giving: intrinsic motivations from altruism (Andreoni, 1989, 1990), extrinsic motivations from monetary and symbolic rewards (Goette and Stutzer, 2008; Lacetera et al., 2012; Kosfeld and Neckermann, 2011; Ashraf et al., 2014), and image motivations from social approval (Gachter and Fehr, 1999; Rege and Telle, 2004; Andreoni and Petrie, 2004; Ariely et al., 2009; Filiz-Ozbay and Ozbay, 2014).

Concerns for what others think (image motivation) can drive substantial pro-social behavior. Many charities today publicize donations or donation brackets on their websites to encourage donors to give more. Others give out symbolic gifts (ribbons,

[1] The second chapter is a joint work with Kristian López Vargas

t-shirts, stickers, etc.) with charity logos that the donors or volunteers could show to others. Charities also host galas where the fund-raising process is public to all attendees. On the other hand, providing extrinsic motivations is also common in encouraging pro-social activities. In the U.S., there exist tax deductions to charitable donations which are equivalent to donation rebates that vary with the income brackets (IRS). Similarly, volunteer blood donors are often incentivized with money or small gifts for their activities.

Interestingly, while many charities provide both types of incentives, image and material, little is known about how these two motivations interact. For example, for those who act pro-socially due to image concerns, would offering monetary rewards affect their image-driven pro-social behavior? Benabou and Tirole (2006) suggest that, theoretically, monetary incentives may crowd out image motivations as volunteers who receive monetary rewards may want to avoid being seen as 'profit-seeking' by the public. Previous literature presents mixed evidence on this hypothesis. While some research supports the crowding-out theory (Ariely et al., 2009, e.g.), other studies document mixed evidence on the crowding-out effect when they consider other individual factors such as reputation and past altruistic behavior (Exley, 2018).[2]

Another question that still remains open is how different levels of material incentives might have qualitatively different crowding-out effects on image motivations. Small monetary rewards, even in the form of cash rewards, could be perceived as a

[2]In Exley (2018), if subjects established public volunteer reputation (i.e. participated in a public volunteer activity in the first round of the experiment), offering them monetary rewards in public for their volunteering in the second round does not crowd out as much volunteer willingness as those whose past volunteer reputation was private.

symbolic gift and be small enough not to cause any concerns of appearing ‘profit-seeking’, while large monetary rewards could crowd out image incentives in pro-social behavior. As illustrated by the comparative statics of the theoretical model we provide in this paper, people’s perception of the monetary reward determines whether the reward crowds out donations. When individuals dislike perceived ‘greediness’ more than perceived ‘altruism’,[3] offering a higher rebate will crowd out donations in public.

Analyzing the effect of different levels of monetary rewards has important real-world implications. First, volunteers are usually paid with small cash amounts as an appreciation gesture to cover small expenses like meals and transportation, instead of payments for services so that their activities are not perceived as ‘profit-seeking’.

Second, charitable contribution deductions in the United States is basically a rebate subsidy on donations, where the rebate varies with income brackets. The system can offer donation rebates up to 50.3% and can deduct up to 50% of the donor’s adjusted gross income.[4]

Last, rebate and matching subsidies are commonly used in charitable donation campaigns. It is important for charities who want to run public donation subsidy campaigns to learn what levels of monetary rewards would cause negative image concerns and crowd out the willingness to donate.

In this paper, we study the interaction of image and material motivations. We

[3]As ‘greediness’ and ‘altruism’ perceived by others.

[4]For example, consider a donation of $300 to the American Red Cross by some donor. The taxable income of this donor is reduced by $300. If the donor’s income level belongs to the top tax bracket (37% federal tax rate + 13.3% California state tax rate), the donor’s tax liability (amount of taxes owed to the government) is reduced by $150, which is equivalent to a 50.3% rebate on his/her donations

implement a within-subject design with two treatments: monetary incentive and donation visibility. We have three levels of monetary incentives: no incentive (0% rebate), a low incentive (10% rebate), and a high incentive (50% rebate). There are also three levels in donation visibility: private donation, public donation without costly non-disclosure (ND) option, and public donation with costly non-disclosure option. We will discuss the details of these treatments later in the experiment design section.

In our experiment, we use rebate subsidies as the monetary incentives. It is important to note that rebate subsidy and matching subsidy are mathematically equivalent. For example, a 10% donation rebate is equivalent to approximately an 11% matching subsidy, while a 50% donation rebate is equivalent to a 100% matching.[5]

Some previous experiments have compared matching and rebate subsidies and find considerably higher charity receipts under matching subsidies (Eckel and Grossman, 2003, 2005a,b, 2006; Davis et al., 2005). However, Davis (2006) indicates that such differences are less likely due to people's preference's for matching subsidies as suggested by Eckel and Grossman (2003), but to the *isolation effect*. In this context, this effect works similarly to *mental accounting bias*, where people focus on their own donation rather than on the total charity receipts. Given the same individual donation, a charity in matching subsidy receives more because the charity, instead of the individual, receives the third party's matching fund. When eliminating the isolation effect in the experiment, Davis (2006) finds no difference in donations under matching and rebate subsidies,

[5]The effective price of a $1 donation to charity in matching funds with matching rate S_m is $\frac{1}{1+S_m}$, while for rebates with rebate rate S_r is $1 - S_r$. Thus S_m and S_r induce the same effective price when $\frac{1}{1+S_m} = 1 - S_r$.

indicating that the two subsidies are equally preferred by subjects.

To bring realism to our image incentives, we use social media (Facebook) as the channel through which donations become public. This allows public donation decisions to occur in a more natural and realistic setting. This is an innovation of our paper with respect to previous studies, where the observability of donations was limited to lab participants (people sitting in the same room) or donation campaign solicitors.

Our experiment result shows that a small reward in terms of a 10% donation rebate does not cause any significant change in donation behavior, both in private and public. When a large reward is given (a 50% donation rebate) in private, it significantly increases people's donations with respect to the case without incentive.

When the reward is given in public the impact of the reward is heterogeneous. People's perception of the monetary reward determines whether the reward crowds out donations. For those who perceive the reward makes their donations appear 'less generous', a 50% rebate significantly crowds out charitable donations made in public (the interaction term of public donation and 50% rebate is negative); for those who do not associate monetary rewards with a negative image, but as a positive way to advertise, a high rebate significantly increases their donations made in public. The net effect for our sample is statistically zero. We also find that males, in general, are more sensitive to their public images and significantly reduce their donations in public when offered a high reward, while the effects on females' behavior are statistically insignificant.

The rest of the paper is organized as follows. Section 2 provides a review of related literature on monetary incentives and image motivation in pro-social behaviors.

Section 3 provides a theoretical model with predictions on the effect of offering different levels of rebate on individual donation decisions in private and public. Section 4 presents the experimental design that tests predictions from section 3, and section 5 summarizes the experiment results. Section 6 concludes with a discussion of our findings and their implications.

2.2 Literature Review

A robust finding in the literature shows that image concerns drive a substantial portion of pro-social behavior. People tend to be more pro-social when their activities become more observable to lab or field experiment participants. In traditional economic games such as the dictator game and public goods game, revealing subjects' identity or their decisions substantially increases the splits by dictators (Haley and Fessler, 2005; Dana et al., 2006; Broberg et al., 2007; Ellingsen and Johannesson, 2007) and the voluntary contributions to public goods (Gachter and Fehr, 1999; Masclet et al., 2003; Rege and Telle, 2004; Andreoni and Petrie, 2004; SavikhinSamek and Sheremeta, 2014; Filiz-Ozbay and Ozbay, 2014).

Image incentives also lead to more generosity in charitable donations (Soetevent, 2005; Ekstrom, 2012) and fundraisings (Alpizar et al., 2008; DellaVigna et al., 2012), as well as volunteering activities such as blood donations (Lacetera and Macis, 2010b; Karlan and McConnell, 2014).

However, Bracha and Vesterlund (2017) report mixed signals of image motivations when donations are used to infer both generosity and income levels. They find that public donation might lead to lower contributions when income is invisible because donors want to be perceived as 'poor-and-generous' rather than 'rich-and-stingy'. When income is invisible but donations are public, donors will donate less to avoid being categorized as a high-income earner since higher incomes, relative to lower incomes, would make their donations appear 'less generous'.

Providing monetary incentives, on the other hand, may crowd out motivations for pro-social behavior, both intrinsically and image-driven. Benabou and Tirole (2003) and Meier and Stutzer (2008) show that when pro-social behavior (such as volunteering) are driven by intrinsic motivations, monetary rewards could decrease people's willingness to behave altruistically. Some empirical and experimental evidence support this theory (Rustichini and Gneezy, 2000; Mellstrom and Johannesson, 2008; Lacetera and Macis, 2010a) , while others (Goette and Stutzer, 2008; Lacetera and Macis, 2010b; Kosfeld and Neckermann, 2011; Lacetera et al., 2012; Ashraf et al., 2014; Olken et al., 2014) are against it, especially when the reward is small or symbolic (does not involve direct cash) (Goette and Stutzer, 2008; Lacetera and Macis, 2010b; Kosfeld and Neckermann, 2011; Ashraf et al., 2014).

In addition to the crowding-out effect on intrinsic motivations, Benabou and Tirole (2006) suggest that monetary incentives may crowd out image motivations as donors could be perceived as 'profit-seeking' instead of altruistic by the public. This theory is supported by Ariely et al. (2009): their experiment finds that introducing private monetary incentives crowds out efforts made in a public donation event. Without monetary incentives, subjects are significantly more motivated to put effort into the donation task in public than in private. With monetary incentives, where subjects will receive personal income that equals their donation amount, efforts in private increased significantly but declined insignificantly in public. This result implies that monetary incentives might dilute image motivations and crowd out the effort to behave pro-socially.

Moreover, Exley (2018) examines how individual reputation could play a role in alleviating the crowding-out effect of rewards on image incentives in volunteering activities. The author finds that the crowding-out effect is significantly lower for those with publicly known volunteering histories. When subjects' past volunteer decisions were made private (in the first round of the experiment), subjects become significantly less likely to volunteer in the second round of the experiment, where they will receive public monetary rewards for their volunteer activities. However, when subjects were able to establish public volunteer reputations in round 1, they are less concerned about appearing 'greedy' and the crowding-out effect in willingness to volunteer while receiving a monetary reward is less significant.

2.3 Theoretical Model

In this section, we present a simple theoretical model to derive comparative statics and hypothesis to bring to data, and there may be other models that would deliver similar comparative statics. We assume individuals maximize the following incomeconstrained utility function:

$$U_i = V_i[w - (1 - r)d] + \pi \times [\mu_a S_i(d) - \mu_y S_i(rd)] + H_i(d) \tag{2.1}$$

where V_i is the utility of direct consumption of material goods, after deducting the total out-of-pocket (OOP) donation $(1 - r)d$ from the individual's income w, where d is the donation decision and r is the rebate rate. Following the definition of Benabou and Tirole (2006), S_i is the perceived social status image (by others) from the individual's total donations and the rebate he/she receives. The sigh in front of μ_a and μ_y reflect the idea that individuals would like to appear as prosocial but not 'greedy', with $0 < \mu_a < 1$ and $0 < \mu_y < 1$. Individuals only gain social status utility when the probability of making their donation public, π, is greater than zero. H_i represents individual's intrinsic altruistic utility, which is a function of the total donation amount d received by the recipient.

All utility functions (V_i, S_i and H_i) are twice differentiable, increasing and concave. We further assume that the marginal impact of a \$1 OOP donation on personal consumption utility is greater than the utility gain from social status or altruism at any level of income or donation. Formally, we assume $S' < V'$, $H' < V'$, $|S''| < |V''|$, and $|H''| < |V''|$.

Since in private donations, $\pi = 0$, the objective function in 2.1 becomes: $U_i = V_i[w-(1-r)d]+H_i(d)$. Dropping the subscript i by convenience, the first order condition becomes:

$$\frac{\partial U}{\partial d} = H'(d^*_{pri}) - (1-r)V'[w-(1-r)d^*_{pri}] = 0 \tag{2.2}$$

Using implicit function theorem, we can calculate the impact of rebates on optimal donation.

$$\frac{\partial d}{\partial r}_{private} = -\frac{\frac{\partial^2 U_i}{\partial d \partial r}}{\frac{\partial^2 U_i}{\partial d^2}} = -\frac{V' - (1-r)dV''}{(1-r)^2 V'' + H''} > 0 \tag{2.3}$$

Comparative Static 1: when donating privately, individual's donation amount increases with rebate rate.

In public donations, $\pi = 1$, the objective function in 2.1 becomes: $U_i = V_i[w - (1-r)d] + \mu_a S_i(d) - \mu_y S_i(rd) + H_i(d)$. The first order condition becomes:

$$\frac{\partial U}{\partial d} = H'(d^*_{pub}) - (1-r)V'[w-(1-r)d^*_{pub}] + \mu_a S'(d^*_{pub}) - r\mu_y S'(rd^*_{pub}) = 0 \tag{2.4}$$

Using implicit function theorem, we get:

$$\frac{\partial d}{\partial r}_{public} = -\frac{\frac{\partial^2 U_i}{\partial d \partial r}}{\frac{\partial^2 U_i}{\partial d^2}} = -\frac{(V' - \mu_y S') - d[(1-r)V'' + \mu_y r S'']}{(1-r)^2 V'' + H'' + (\mu_a - r^2\mu_y)S''} \tag{2.5}$$

Equation 2.5 is strictly positive if $\mu_a - r^2\mu_y > 0$.

Comparative Static 2: when donating publicly, if rebate is small and individuals value perceived altruism slightly more than perceived 'greediness', donation amount *increases with rebate rate.*

To study the interaction effects of rebate and donation visibility, we consider the differences of how donation amount responds to rebate increases in private vs. public. We calculate the following equation: $\frac{\partial d}{\partial r}_{public} - \frac{\partial d}{\partial r}_{private}$, where the comparative statics are evaluated at the optimal donations d^*_{pri} and d^*_{pub}.

In order to arrive at a closed-form solution of the comparative statics, we assume that the utility functions are in log-forms. That is: $V_i = \alpha log[w - (1-r)d]$, $H_i = \beta log(d)$, $\mu_a S_i = \mu_a log(d)$, and $\mu_y S_i = \mu_y log(rd)$, where α, β, μ_a, and μ_y are between 0 and 1, and $\alpha + \beta + \mu_a - \mu_y > 0$.

Solving for the optimal donations and comparative statics in private and public, we have:

$$\begin{aligned}
d^*_{private} &= \frac{w}{1-r} \times \frac{\beta}{\alpha+\beta} \\
d^*_{public} &= \frac{w}{1-r} \times \frac{(\beta+\mu_a-\mu_y)}{\alpha+\beta+\mu_a-\mu_y} \\
\frac{\partial d^*}{\partial r}_{private} &= \frac{w}{(1-r)^2}\frac{\beta}{\alpha+\beta} \\
\frac{\partial d^*}{\partial r}_{public} &= \frac{w}{(1-r)^2} \times \frac{(\beta+\mu_a-\mu_y)}{\alpha+\beta+\mu_a-\mu_y}
\end{aligned} \tag{2.6}$$

Then we calculate the difference in changes in donations due to a rebate increase in public and private:

$$\begin{aligned}
\frac{\partial d^*}{\partial r}_{public} - \frac{\partial d^*}{\partial r}_{private} &= \frac{w}{(1-r)^2} \times \frac{(\beta+\mu_a-\mu_y)}{\alpha+\beta+\mu_a-\mu_y} - \frac{w}{(1-r)^2}\frac{\beta}{\alpha+\beta} \\
&= \frac{w}{(1-r)^2}\frac{\alpha(\mu_a-\mu_y)}{(\alpha+\beta+\mu_a-\mu_y)(\alpha+\beta)}
\end{aligned} \tag{2.7}$$

The comparative static in 2.7 is positive when $\mu_a > \mu_y$. When individuals value perceived altruism more than perceived 'greediness', individuals will increase donations more in public than in private in response to a higher rebate, and vice versa.

Comparative Static 3: When the individual's dislike perceived 'greediness' more than perceived altruism, offering a higher rebate generates less increase in donations in public than in private.

2.4 Experiment Design

This experiment has a within-subject design with two treatment arms. First, we use three levels of monetary Incentives. Second, we use three types of donation visibility.

Monetary Incentives:

1. No-incentive (control): no monetary reward on subjects' donations
2. Low-incentive: a 10% rebate on subjects' donations
3. High-incentive: a 50% rebate on subjects' donations

Donation Visibility:

1. Private donation (control): subjects are donating anonymously to the organization of their choice. Subjects receive the following thank-you message with their donation and rebate information on their computer screens (only visible to the subjects): 'We want to thank *subject's name* for donating *$X* to *charity Z*! *subject's*

name received a *Y%* rebate (rebate in dollars) from the UCSC LEEPS Lab.'

2. Public donation: subjects are donating anonymously to the organization of their choice. The experimenter will publicize their donation and rebate information in the following thank-you message on the experimenter's Facebook wall (visible to everyone): 'We want to thank *subject's name* for donating *$X* to *charity Z*! *subject's name* received a *Y%* rebate (rebate in dollars) from the UCSC LEEPS Lab.'

3. Public donation with a costly non-disclosure option: similar to public donation without non-disclosure, subjects' donation decisions will be public through a thank-you message posted on the experimenter's Facebook wall (visible to everyone). However, we offer subjects the opportunity to not disclose the rebate they receive in their donations, for a fee of $0.2. If the subject chooses to exercise this non-disclosure option, the thank-you message no longer reveals the rebate he/she receives: 'We want to thank *subject's name* for donating *$X* to *charity Z*!' This treatment is only interacted with the 10% and 50% rebate levels.

Figure 2.1 presents three sample Facebook posts for public donations with 0%, 10% and 50% rebate. Note that if the subjects choose not to donate in any of the decision rounds, their actions will not be visible to anyone.

We used block randomization of treatments to avoid order effects within subjects. The two main levels in *Donation Visibility* are the blocks (i.e. public or private donations), and within each block, the three levels of rebates are randomized, with a

Figure 2.1: Example Facebook Post for Public Donations: subjects' personal information is hidden

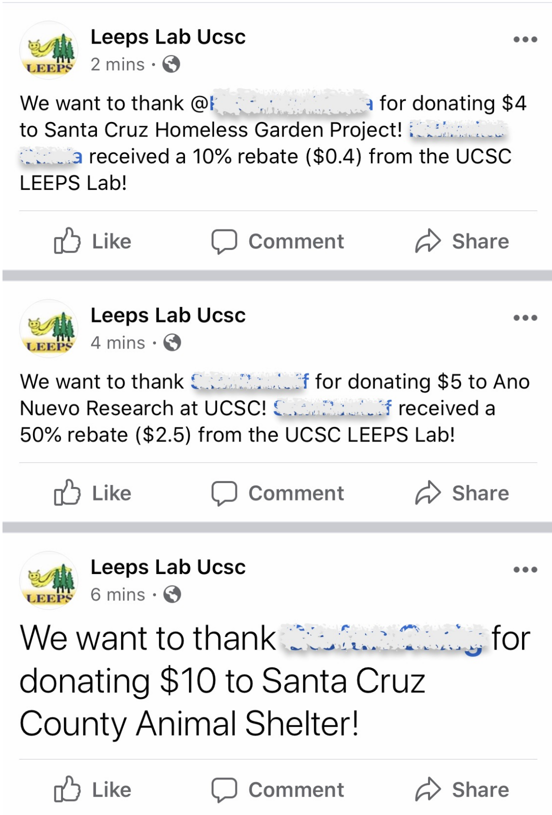


Table 2.1: A Sample of Treatment Permutations

Donation Visibility	Public	Public	Public	Private	Private	Private
Rebate	0 10 90 10 (ND) 90 (ND)	0 10 90 10 (ND) 90 (ND)	10 0 90 10 (ND) 90 (ND)	0 10 90	0 10 90	10 0 90
Donation Visibility	**Private**	**Private**	**Private**	**Public**	**Public**	**Public**
Rebate	0 10 90	0 90 10	0 10 90	0 10 90 10 (ND) 90 (ND)	0 90 10 90 (ND) 10 (ND)	0 10 90 10 (ND) 90 (ND)

This table presents a sample of block randomization on treatment combinations. Two levels in donation visibility are blocks and rebates are randomized within blocks.

total of 72 permutations (see Table 1 for an illustration).

At the end of the experiment, subjects take a survey that collects demographic information such as gender, age, major, their passion for the selected charity, past donation frequency, Facebook use frequency etc., and their beliefs about certain levels of monetary rewards [6].

2.4.1 Experiment Procedure

Prior to the experiment, subjects have to give authorization via email to the LEEPS lab for a possible post on the LEEPS Lab Facebook page about their decisions

[6]If subjects donated a positive amount ($X) in public with 50% rebate, they are asked whether they prefer: 1. Do not donate and have no message posted on the LEEPS Facebook Page; 2. Donate $X with the following message posted on the LEEPS Facebook Page: 'We want to thank *subject's name* for donating *$X* to *charity Z*! *subject's name* received a *Y%* rebate (rebate in dollars) from the UCSC LEEPS Lab.' Option 2 is consistent with their previous decisions in the experiment, while option 1 is inconsistent. Subjects who answered inconsistently are excluded from the sample.

during the experiment. Only those who have given authorization will attend the experiment. The experiment is conducted online. Upon arrival at the experiment's Zoom meeting room, subjects are assigned a random participant ID before they navigate to the experiment session page.

Subjects start the experiment by reading general instructions.[7] Then they participate in a square counting task to earn their incomes ($10). After the counting task, subjects are offered a list of charities that covers Covid-19 response funds, humanity care, natural preservation research, and animal shelters (see table 2.2 for a short description and Appendix for a detailed charity list). Subjects will select one charity for all subsequent donation decisions.

Subjects will make 8 independent donation decisions under different scenarios (see figure 2.2 for public donation decision pages without (top panel) and with (bottom panel) costly non-disclosure options. All rebates in the decisions are financed by the experimenter.

The experiment will conclude with a short survey. In order to incentivize subjects to respond truthfully during the experiment, the computer randomly chooses one round of decision for payment realization in the end. Subjects are paid via Venmo within 48 hours after the experiment.

[7]Detailed experiment instructions are included in the Appendix of the paper.

Table 2.2: List of Charities in Experiment

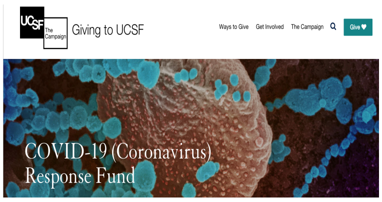
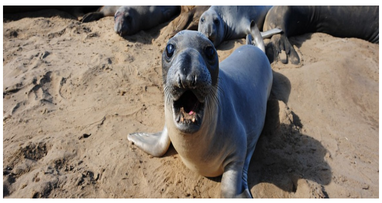
Charity Name	Logo	Descriptions
WHO Covid-19 Response Fund		The World Health Organization (WHO) is leading and coordinating the global effort with a range of partners, supporting countries to prevent, detect, and respond to the pandemic.
UCSC COVID-19 Slug Support Campaign		UCSC has created the COVID-19 Slug Support Campaign. Together, our community can help support students who are experiencing financial or personal crisis because of the infection and the drastic measures our state and country have had to take.
UCSF COVID-19 Response Fund		UCSF is among the top health sciences universities in the world. As the COVID-19 outbreak expands, teams throughout UCSF's hospitals, clinics, and research labs are actively monitoring and responding to the evolving situation.
Santa Cruz Homeless Garden Project		The Homeless Garden Project provides job training, transitional employment and support services to people who are experiencing homelessness. HGP's vibrant education and volunteer program for the broad community blends formal, experiential and service-learning.
Ano Nuevo Research at UCSC		Ano Nuevo Reserve is one of the University of California's 39 Natural Reserves. The close proximity of the Reserve to UC Santa Cruz makes it a hot spot for undergraduate experiential learning. Students obtain hands-on research experience while working on world-class scientific research projects.
Santa Cruz County Animal Shelter	SANTA CRUZ COUNTY ANIMAL SHELTER OPEN DOOR. OPEN HEART.	The Santa Cruz County Animal Shelter is a non-profit Joint Powers Authority that provides 24-hour animal rescue and is Santa Cruz County's only full service, open-admission animal shelter. We rescue around 5,000 stray, unwanted, abandoned and injured animals every year.

Figure 2.2: Donation Decision Page With/without costly non-disclosure option

Decision 1 - **Public** with **10%** Rebate

You have $10 in your cash account ($4 participation fee + $6 from counting task).

Total Cash Available	Would you like to donate any to the charity?	10% rebate cash back in your account	Total cash left in your account now (including rebate)
$10 ($4 participation fee + $6 from task)	**Yes**: choose between $1 to $10: 10 Your decision with rebate info will be posted on **LEEPS FB page**.	$1.00	$1.00
	No: $0. Your decision will **NOT** be posted.		

We will tag you in the following post on LEEPS FB page (**visible to everyone**): "We want to thank @[Your Name] for donating **$10** to UCSF COVID-19 Response Fund. @[Your Name] received a **10%** (**$1.00**) rebate from the LEEPS Lab!"

Submit Decision

Decision 4 - **Public** with **10%** Rebate

You have the option to **NOT** announce the rebate you received in the FB post

You have $10 in your cash account ($4 participation fee + $6 from counting task).

Total Cash Available	Would you like to donate any to the charity?	Would you like us to not announce the rebate you received in the FB post?	10% rebate cash back in your account	Total cash left in your account now (including rebate)
$10 ($4 participation fee + $6 from task)	**Yes**: choose between $1 to $10: 10 Your decision without rebate info will be posted on **LEEPS FB page**.	**Yes**: we will deduct $0.2 from your personal account	$1.00	$0.80
	No: $0. Your decision will **NOT** be posted.	**No**		

We will tag you in the following post on LEEPS FB page (**visible to everyone**): "We want to thank @[Your Name] for donating **$10** to WHO COVID-19 Response Fund!"

Submit Decision

Table 2.3: Subject Demographic Characteristics Summary Statistics

	Gender (M=1)	**Age**	**Work Now** (Yes=1)	**Donate Frequent** (Yes=1)	**Use FB** (Daily=0)	**Charity Passion** (High=10)	**Rebate Impact** (Negative =0)
Mean	0.37	21.46	0.4	0.53	1.24	7.76	0.50
SD	0.48	2.176	0.49	0.50	1.27	1.94	0.50

Summary statistics of subjects' demographic characteristics with a sample size of 140. Sample contains more females than males, with an average age of 21. 60% of the subjects are working, and 50% donated at least once in the last year. 41.4% of subjects use Facebook at least once a day, and 65% use Facebook at least once a week.

2.5 Experiment Results

We conducted 10 online experiment sessions with 140 undergraduate students from University of California, Santa Cruz. Table 2.3 summarizes subjects' demographic characteristics collected from the experiment's survey. We have slightly more females than males who participated in the experiment, with an average age of 21. 60% of the subjects are working part-time or full-time, and 50% donated at least once in the last year. 41.4% of subjects use Facebook at least once a day, and 65% use Facebook at least once a week.

Table 2.4 summarizes donation probability and mean donation for each treatment combination. When rebate increases in private, probability of donation increases, as well as average donation amount. When rebate increases in public, donation probability and mean also increase, but the magnitude of increase is smaller than that in private settings. Note that when subjects choose not to donate in the experiment, their decisions will not be posted anywhere (completely anonymous). We observe that within the 50% rebate, 13.6% (79.3%-65.7%) more subjects choose to not donate when the re-

Table 2.4: Summary Statistics of Donation Probability and Mean by Treatments

Donation Visibility/Rebate	0%	10%	50%	
Private	55.0% 3.09	60.7% 3.38	79.3% 5.41	% Donated Mean Donation
Public	46.4% 2.94	53.6% 3.08	65.7% 4.84	% Donated Mean Donation
Public (non-disclosure)		54.3% 3.18	70.7% 5.27	% Donated Mean Donation

Summary statistics of donation probability and mean by treatment combinations.

bate is offered in public without the non-disclosure option. The lost in the willingness to donate under a public 50% rebate is partially recovered when we offer the non-disclosure option to hide the rebate information in the Facebook post.

Moreover, comparing donation decisions when the non-disclosure option is offered to those where this option is not available, donation distributions shifted towards higher donations, especially under the 50% rebate (see figure 2.3).

To conduct a proper analysis of the average treatment effects of rebate and donation visibility on donation amount, we run the following regression specification with a Tobit model:

$$donation_i = \beta_0 + \beta_1 DonationVisibility_i + \beta_2 Rebate_i + \beta_3 DonationVisibility_i * Rebate_i + \beta_4 Rebate_i * NonDisclosure_i + \epsilon_i \tag{2.8}$$

where $DonationVisibility = 1$ if donations are made in public and 0 otherwise, and $NonDisclosure = 1$ if the public donation has a costly non-disclosure option. $donation_i$

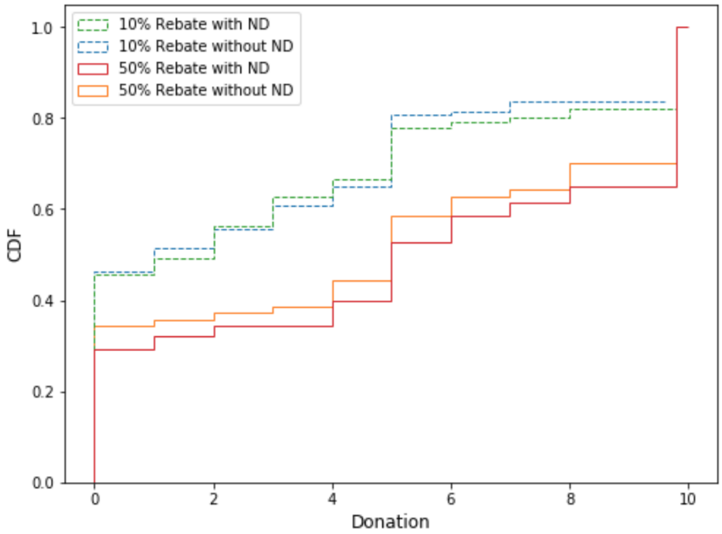


Figure 2.3: CDF of donations with/without costly non-disclosure (ND) option. Subjects donate more when the non-disclosure option is present, especially when rebate is high.

is censored at 0 and 10. Standard errors are clustered at the subject level to control for correlations of decisions within the same subject.

As we discussed in the theoretical model, we also study whether the impact of a public rebate is different for people who feel strongly about the negative image impact of the rebate vs. those who think rebate may bring lots of positive impact. The classification of the types of subjects is based on a survey question that elicited subjects' opinions on whether receiving a rebate would make their donations appear 'less generous' [8], we find that subjects are divided into two types. The first type (50% of subjects) believes that publicly announcing the rebate they received makes their donations appear

[8]What do you think is the impact of publicly announcing the rebate you received in the Facebook post?

1. The rebate makes my donations appear less generous
2. The rebate has an advertising effect that will attract more people to donate

‘less generous’, while the other subjects do not associate a negative image with receiving a rebate.

We present the Tobit regression results for all subjects, as well as for different types of subject in table 2.5. When we investigate the treatment effects on all subjects (column 1), we do not observe any significant crowding-out effect on donation decisions from receiving a rebate in public. However, when we analyze the two types of subjects separately, we find heterogeneous treatment effects of public rebates on donations.

For subjects who believe rebates make their donations appear ‘less generous’ (referred to as Type I subjects in column 2), offering a low (10%) rebate, whether in private or in public, does not have any significant impact on their donation decisions. However, when they receive a high rebate (50%) in private, Type I subjects significantly increase their donations (+7.112***). When the high rebate becomes public, however, it significantly crowds out donations (-3.712**) among these subjects.

Furthermore, when Type I subjects are offered a costly non-disclosure option in the 50% rebate scenario, they significantly increase their donations (+2.355**), alleviating the crowding-out effect caused by the negative image associated with receiving a rebate on donations. This is because a higher percentage of Type I subjects choose to exercise the non-disclosure option instead of not donating when the rebate is high (figure 2.4).

Column 3 of table 2.5 presents regression results for subjects who do not associate rebates with negative images (Type II subjects). Similar to Type I subjects, a small rebate (10%) does not have any significant impacts on their donation decisions,

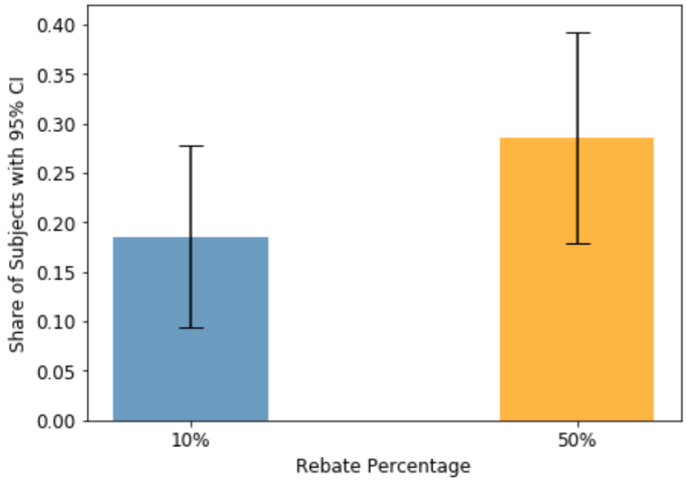


Figure 2.4: Share of subjects that chose the costly non-disclosure option.
More subjects chose the non-disclosure option when the rebate is high.

public or private. When offered a high rebate in private, Type II subjects significantly increase their donation amounts (+4.260***), and they are even willing to donate more when the donation and rebate becomes public (+1.619**). Since Type II subjects do not associate a negative image with rebate announcement, offering the costly non-disclosure option does not affect their donation decisions.

We also present the same Tobit regressions using Out-of-Pocket (OOP) donation ($OOP = Donation - Rebate$) as the outcome variable in table 2.6. We observe similar but smaller treatment effects (at the same significance level) from a public 50% rebate for type I and II subjects, indicating that the crowding-out/advertising effect of public rebate impacts not only donation amounts, but also donation net spending. However, the increase due to a private 50% rebate is no longer significant in the OOP donations. Without public image incentives, subjects are more inclined to keep their net spending constant, so larger donations under higher rebates are mostly financed by

Table 2.5: Tobit Regressions Across Subject Types on Donation Amount

	All Types Included	Type I: Rebate Negative Image	Type II: Rebate Non-Negative Image
Public Donation	-0.836	0.0803	-1.679**
	(0.593)	(0.942)	(0.751)
10% Rebate	0.828	1.053	0.616
	(0.522)	(0.990)	(0.442)
50% Rebate	5.650***	7.112***	4.260***
	(0.812)	(1.458)	(0.818)
Public*10% Rebate	-0.0828	-1.307	1.007
	(0.792)	(1.430)	(0.801)
Public*50% Rebate	-0.980	-3.712**	1.619**
	(0.912)	(1.717)	(0.759)
Public*10% Rebate* Non-Disclosure Available	0.269	0.288	0.264
	(0.483)	(0.842)	(0.538)
Public*50% Rebate* Non-Disclosure Available	1.197**	2.355**	0.0817
	(0.569)	(0.920)	(0.683)
Constant	0.684	-0.413	1.716
	(0.840)	(1.309)	(1.107)
Observations	1,120	560	560

* p<.1, ** p<.05, ***p<.01 (robust standard errors clustered at the subject level, displayed in parentheses)

3rd parties, not by subjects' own pocket. This finding of constant net spending under different rebates with private donations is consistent with the results from Davis (2006).

Thus, it is important to identify subject types when analyzing the heterogeneous crowding-out effect of rebates on image incentives. With ex-ante information about subjects' perceptions of the public rebate, we further analyze the correlation of different observable characteristics with subjects' types (see table **??** in Appendix for

Table 2.6: Tobit Regressions Across Subject Types on Out-of-Pocket Donations

	All Types Included	Type I: Rebate Negative Image	Type II: Rebate Non-Negative Image
Public Donation	-0.477	0.145	-1.054**
	(0.395)	(0.644)	(0.468)
10% Rebate	-0.204	0.0611	-0.452
	(0.342)	(0.617)	(0.345)
50% Rebate	0.109	0.954	-0.681
	(0.457)	(0.725)	(0.575)
Public*10% Rebate	-0.119	-0.998	0.673
	(0.492)	(0.865)	(0.510)
Public*50% Rebate	-0.338	-1.675*	0.902**
	(0.473)	(0.871)	(0.395)
Public*10% Rebate*	0.135	0.194	0.0868
Non-Disclosure Available	(0.272)	(0.473)	(0.309)
Public*50% Rebate*	0.429**	0.909**	-0.00779
Non-Disclosure Available	(0.207)	(0.353)	(0.228)
Constant	1.917***	1.188	2.596***
	(0.571)	(0.847)	(0.780)
Observations	1,120	560	560

* p<.1, ** p<.05, ***p<.01 (robust standard errors clustered at the subject level, displayed in parentheses)

correlation regression). We find that males are more correlated with the negative image type, while subjects who are more passionate about the charity they chose tend to report more advertising effect of the public rebate. When the frequency of Facebook usage (announcement channel) decreases, subjects associate the public rebate less with the advertising effect.

We further run Tobit regressions for males and females separately, presented in table 2.7. In private donations, males react more to donation rebates, especially when rebate is high (+7.983*** vs. +4.645***). Moreover, offering a 50% rebate in public significantly crowds out donations among males (-3.955*), but nor for females (+0.305). However, the analysis using gender breakdown is only exploratory because our experiment is not designed to identify gender differences in the crowding-out effect, and we have gender imbalance (fewer males than females) in our experiment sample.

Table 2.7: Tobit Regressions Across Gender on Donation Amount

	Males	Females
Public Donation	0.142	-1.359**
	(1.303)	(0.677)
10% Rebate	1.311	0.532
	(1.134)	(0.590)
50% Rebate	7.983***	4.645***
	(1.870)	(0.884)
Public*10% Rebate	-2.304	1.012
	(2.126)	(0.728)
Public*50% Rebate	-3.955*	0.305
	(2.297)	(0.925)
Public*10% Rebate*	0.972	0.0152
Non-Disclosure Available	(0.950)	(0.590)
Public*50% Rebate*	1.500	1.160*
Non-Disclosure Available	(1.204)	(0.666)
Constant	-0.621	1.216
	(1.831)	(0.951)
Observations	416	672

* p<.1, ** p<.05, ***p<.01 (robust standard errors clustered at the subject level, displayed in parentheses)

2.6 Conclusion

This paper provides experimental evidence of the impact of different levels of monetary rewards on image incentives in charitable donations. Our experiment results show that a small reward does not impose any significant changes in donation behavior, both in private and public. When a large reward is given in private, it significantly increases people's donations. When the large reward is given in public, we observe heterogeneous treatment effects depending on people's perception of the monetary reward. For those who believe the reward makes their donations appear 'less generous' (negative image association), a 50% rebate significantly crowds out charitable donations made in public. When these people are offered a costly non-disclosure option to hide the rebate information in the announcement, the crowding-out effect is alleviated. For people who do not associate monetary rewards with a negative image, a high rebate significantly increases their donations made in public, and offering the non-disclosure option does not have any impact on their donation behavior. We also find that males are more sensitive to their public images and significantly reduce their donations when offered a high reward in public, while the effects on females are non-significant. Overall, the net effect from our sample is statistically zero.

Fundraising in public with rebate subsidies should be careful with the population group that are sensitive to image incentives. For those groups who associate receiving rebates with negative public images, offering rebates in private can better stimulate donations while not introducing concerns of negativity or crowding out willingness to donate. For future research, we should investigate which observable demographic characteristics are associated with the two types of subjects we identify in our experiment, so that charities can better target certain groups when running fundraising campaigns with subsidies.

Chapter 3

The Impact of Observing Peers' Giving Behavior: A Field Study [1]

3.1 Introduction

Is people's altruistic behavior in donations affected by peers' giving? In the literature, there are competing theories that predict opposite effects about peer's influence on people's pro-social behavior.

Many theories predict a complementary, positive effect, where individuals will become more generous if their peers are altruistic. This hypothesis argues that peers constitute 'socially appropriate' behavior in a given context, so if others behave in a certain way, say donate, one will want to conform with that behavior and donate as well. That is, this theory states that pro-social behavior is largely affected by pressure to comply with social norms (Bernheim, 1994). On the other hand, more standard theories predict

[1] The third chapter is a joint work with Kristian López Vargas and Angelo Rossi

a negative, substitution effect. In this theory, the individual utility depends mostly on the efficient allocation of resources among the recipients (Warr, 1982; Roberts, 1984). Therefore, people substitute away their willingness to behave pro-socially towards the same recipient that others have already helped. Put differently, seeing more donations from peers will crowd out an individual's donations to the same recipient. In this theory, a person can be purely or impurely 'altruistic', and that will determine whether the crowding out is close to be dollar-by-dollar or less than dollar-by-dollar, respectively. In this paper, we will shed light on the overall direction of peer influence on people's donation behavior, but not on the mechanism.

We study the effects of peer donations on people's giving behavior in a field context. We choose a donation setting that is common in the developing world: street performers at stoplights. In many cities of the developing world, it is common to see artists perform during the red light and then collect money from drivers before the light turns green. In this setting, we specifically study whether observing another vehicle donate would increase or decrease driver's willingness to donate to the same performer. Our field setting also contributes to the literature by bringing evidence of peer influence in giving from the developing world.

In this context, it is viable to study peer influence in a completely observational manner. In a large city like Lima, Perú, the set of vehicles that queue up at a red light is mostly random and anonymous amongst themselves. Thus, this donations environment already gives us a natural experiment. However, the natural rate of donation is rather low (roughly 5%). For this reason, we exogenously manipulate this rate upward. We deploy an experiment using hired drivers to increase the frequency of donations so that more people can be exposed to the treatment condition.

The unit of analysis in our experiment is the vehicle, and we are interested in studying the behavior of passing-by drivers and passengers in terms of how often and how much they donate to a street performer. The treatment variable is a binary identifying whether each vehicle has observed a donation from another one. We do not measure the attention points of drivers, but we construct a proxy based on whether a donation happened within their range of sight. A vehicle X is in the range of sight of the driver if vehicle X is positioned on the left, right, in front, or diagonally in front of the driver. We discuss the details and present graphics of how we construct our treatment variable in the experiment design section.

We conducted the experiment at three locations and dates on the streets of Lima, Perú in October 2019. We collected a total of 1,041 observations (excluding drivers we hire) and performed regression analysis to evaluate the effect of peer influence on probability and amount of giving. Results from our study indicate a strong substitution effect in the probability of donation and amount of donation. When drivers see other drivers donate to the street performer, they are significantly less likely (reduced by 4%) to donate and overall average donations are lower (reduced by 0.9 PEN) as well.

The rest of the paper is organized as follows: section 2 provides a discussion of the related literature on peer influence in pro-social behavior. Section 3 introduces the experiment design, treatment definition and outcomes. Section 4 presents our empirical results. Section 5 concludes and presents a brief discussion about future research.

3.2 Related Literature

The extent of people's pro-social behavior is sensitive to whether they observe their peers doing so. There are two competing theories that predict opposite effects of peer influence on pro-social behavior. One suggests complementary peer influence where people behave more pro-socially if their peers are altruistic. The other theory, however, predicts an opposite substitution effect, where people substitute away their willingness to act pro-socially towards the same recipient that others have already helped.

Most experimental evidence supports the positive peer influence prediction. In a dictator game, Cason and Mui (1998) show that if a dictator is informed about another dictator's allocation they tend to behave more pro-socially over time, relative to the control group where subjects become more self-regarding as rounds progress. Similarly, Krupka and Weber (2009) find that subjects tend to select more equitable allocation choices when they observe more peers behaving pro-socially in the game.

Conditional cooperation in public goods game is another example of positive peer impacts in the lab (Falk et al. (2013), Fischbacher et al. (2001) and Fischbacher and Gachter (2010)) and in the field where the setting resembles public goods provisions (Chen et al. (2010) and Shang and Croson (2009)).

Evidences of positive peer influence are also shown in people's donation decisions in the field. Frey and Meier (2004) show that charitable contribution increases when people know others donate, and the effect is strongest on people who are indifferent about donations.

Smith et al. (2015) suggest that online donation decisions respond to both very large and very small historic amounts and to changes in the mode. In their study, Smith et al. (2015) rely on the exogenous arrival time of donors to the same online fund-raising page, which creates different observed history of past donations. They find that very large (small) donations, twice (half) the size of the average donation, trigger similar-sized donations. Donors use past donation distribution as a benchmark to decide appropriate donation amounts.

One underlying mechanism behind positive peer influences is the theory of conformity as proposed by Bernheim (1994), where individual's utility not only depends on consumption, but also on conforming with social norm status. In this setting, peers constitute 'socially appropriate' behavior in a given context, and pro-social behavior are largely affected by pressure to comply with social norms. Krupka and Weber (2013), Smith et al. (2015), Krupka et al. (2017), and Reuben and Riedl (2013) show that enforcing contribution norms (either maximum contribution or equitable contribution) in public goods game among agents can overcome free-riding problems.

Other than conformity with social norms, equitable social distribution (Fehr and Schmidt (1999)) might better explain the positive peer influence, especially in settings where agents choose effort levels in response to wages given by a principal. In a three-person gift exchange setting,[2] the second agent observes the wage and effort level made by the first agent before making her choices. Gachter et al. (2013) find that the second agent's effort choice is consistent with payoff equality. When the second agent's

[2] the principal pays a wage to each agent, and agents make effort choices sequentially

effort choice does not affect payoff equality, positive peer effects disappear.[3]

In the substitution peer influence theory, most arguments come from 'pure' and 'impure altruism' models. The 'pure altruism' model predicts complete crowding-out when observing others' contributions (Warr (1982) and Roberts (1984)): the individual's utility depends on their own consumption and the consumption of others. As contribution of peers (mostly in the form of government transfers) to the same recipient increases, the agents' own contribution (in the form of private charities) would decrease dollar by dollar.

'Pure altruism' and dollar-by-dollar crowding out is rare in daily observations. Later Andreoni (1990) introduces 'impure altruism', which predicts a partial crowding-out effect. Other than the altruistic characteristic as proposed by Warr (1982) and Roberts (1984), 'impurely altruistic' individuals also generate utility from 'warm-glow' giving, thus the crowding-out from public transfer to private donations is less than 1-by-1.

Empirical evidence that supports the crowding-out effect mostly analyzes the substitution between government grants and private donations. Steinberg (1991) reviews 13 previous studies and concludes that the estimated crowding out effect in private donations ranges from 0.5% to 35% per unit of government funding. Payne (1998) finds that government grants offered to non-profit firms crowd out roughly 50% of private donations. Payne (2001) reports evidence that federal research grants significantly crowd

[3]When the game provider randomly selects one agent's choice to be realized, making the agent's personal choice not affecting the payoff equality in the realized state (the other agent's pay and effort are 0 if not selected, thus no payoff inequality exists), there exist no peer effects as predicted by the social norm model.

out private donations to non-research universities.

In lab experiments, Bolton and Katok (1998) find extensive but incomplete crowding-out in dictator allocations, while Eckel et al. (2005) find that the crowding out of private donations due to public spending is significant if public spending is framed as taxes from individual incomes.

3.3 Experiment Design

As we have seen in the literature review, there are theoretical reasons that predict both positive impact of others' giving behavior (e.g., social norms), as well as negative impact (e.g., substitution or crowding out). Most of the field experiments find evidence of positive impact. That is, due to norms or related motivations, people regard others' giving and their own as complements. To our knowledge, however, there has not been evidence of negative impact or crowding-out effects on giving in the field. With that consideration, we choose to study the effects of peers' donations in a field context that is common in the developing world: street performers at stoplights. We specifically study whether a person who observes another driver or vehicle donate before them, increases or decreases their probability and magnitude of giving.

This is an interesting and viable context for several reasons. In a large city, such as Lima, Perú, the set of vehicles that queue up at a red light is mostly random and anonymous amongst themselves. This means we could study the impact of peers' behavior in a purely observational manner. However, the rate of donation is very low in this natural setting (donation probability without our drivers is around 5.6%) and therefore very few drivers are exposed to these stimuli (observing another vehicle's donation). With this consideration we implement an experiment that adds artificial donations so more drivers can be under the condition of observing others giving.

For this purpose, we hire two drivers who go in circles around few blocks and repeatedly return to the same stoplight positioning themselves in the first row (so they can be seen). That is, the two hired drivers on our experiment will try to park in the first row in front of the traffic light, or as close to the traffic light as possible, in order to facilitate the visibility of their donations. They make donations every other light. However, it is not guaranteed that these drivers will always arrive in the front rows of the traffic during every round, thus in rounds when our drivers were not able to arrive at their designated spots, their were instructed not to make a donation.

We hire one performer whose task is to perform in the first 30 seconds of a one-minute-per-light stoplight. In the last 30 seconds upon completion of the performance, the performer will pass through vehicles to receive money with a hat. The performer memorizes the amount of each donation and the position of the car in the red light. This is possible due the low number of donations per light. Once the performer has passed through all the vehicles, he/she will approach one of the hired (data-entry) observers to report the amount of money collected from each vehicle. All donations from any drivers including our hired drives were kept by the performer. We also guaranteed a payment to the performer of approximately 5 USD per hour, which is approximately four times the hourly amount of the Peruvian minimum wage.

We also hired three data entry assistants, who positioned themselves on both sides of the street with clip boards to register data of the characteristics of the vehicles as well as the people inside during the performance. In order to collect the donation data, we designed a special data entry paper form. This facilitates the information collection and minimizes the error rates. On a piece of A4 paper, we have boxes representing the place of each vehicle relative to the traffic light. Within each box there are relevant fields to be filled by the observers. After collecting donations, the performer approaches one of these data entry points and matches donation amounts into the vehicles depicted on the paper form. A copy of the data entry form can be seen in Appendix figure C.1.

The unit of our analysis is the vehicle since the number of individuals may vary in each vehicle and precise registration of passengers was not possible. The two outcome variables of interest are: donation decisions (binary) and donation amount (continuous).

The treatment variable is the binary identifying whether each vehicle has observed a donation from another vehicle. We do not measure the attention points of drivers, but we construct a proxy based on whether a donation happened within their range of sight.

For example, vehicle j is in range of sight of the vehicle i if j is positioned on the left, right, in front or diagonally in front of i. Figure 3.1 is a graphical representation of the definition of range of sight we use. In this Figure, vehicle i is indicated with a star, and the vehicles in range of sight are indicated with the red lines. If for vehicle i, any of the vehicles within these range of sight makes a donation, then the treatment variable takes the value of one. Otherwise, it is equal to zero.

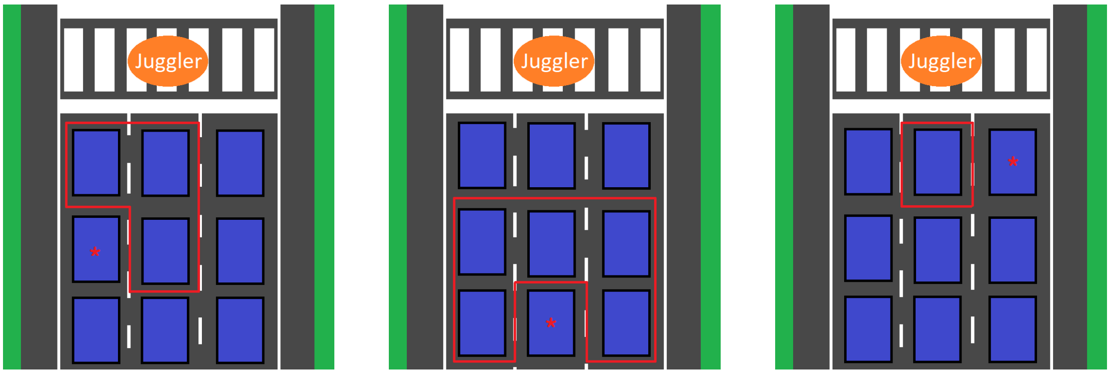


Figure 3.1: Range of Sight: Left, Middle, Right

Other covariates include demographic characteristics of the driver: age bracket (three categories), gender, number of passengers inside the vehicle, and the type of the vehicle (e.g. sedan, SUV, etc.)[4]

3.4 Results

We conducted experiments in three days in the month of October, 2019 and collected a total of 1,041 observations (excluding the drivers we hired). We used three stoplights in the city of Lima, Peru. The selected streets range from two to four lanes.

Table 3.1 presents: (a) the balance table and the summary statistics of drivers' and vehicles' characteristics and (b) raw mean estimates of the main outcome variables under treatment and control conditions.

As expected, we do not observe statistically significant differences in drivers' or vehicles' characteristics between the treatment and control groups. More specifically, on demographic features, only a small fraction of the drivers we observe are females

[4]Other types of vehicles include vans, pickups, trucks, and motorcycles.

(18.9% in control and 23.4% in treatment). About 64% of the drivers in our sessions are between 25 to 50 years old (estimated by our data-recorders). Regarding the vehicles, the average number of passengers per vehicle is around 1.5 for both groups; and most of the vehicles in the experiment are sedans (66.3% for control and 69.1% for treatment).

Panel (b) of Table 3.1 shows the summary statistics of donation probability and donation amount (in Peruvian Sol PEN) for both control and treated vehicles (treated vehicles observed at least one donation from another vehicle). Probability and average donation are both statistically higher for vehicles in the control condition (did not observe a peer donations). Probability of donation is 6.7% in the control group and 3.1% in the treatment group, while average donation is 0.066 PEN in the control and 0.021 PEN in the treatment.

Not reported in the table, conditional on making a donation, the average donation is 0.988 PEN in the control group. In the treatment group, the average donation is 0.667 PEN. In figure 3.2, we show the CDFs of the donations by treatment. After observing a peer donation, the distribution of drivers' donation decisions shift towards more zero donations (left panel). If we further zoom in at the positive donations only, we find that conditional on drivers donating a positive amount, the distribution of donations shifted towards a smaller donation magnitude after observing a peer donation.

3.4.1 Regression Analysis

To properly analyze the impact of peers' giving on people's donation behavior, we conducted a regression analysis. We use a Probit model to model probability of

Table 3.1: Balance Table of Characteristics and Main Outcome Variables by Treatment

Variable	Observed A Peer Donation=0 N/[Clusters]	Observed A Peer Donation=0 Mean/SE	Observed A Peer Donation=1 N/[Clusters]	Observed A Peer Donation=1 Mean/SE	T-test Difference (1)-(2)
Panel A: Characteristics					
Driver: Female	750 [139]	0.189 (0.014)	291 [88]	0.234 (0.022)	-0.044
Driver Age between 25-50	750 [139]	0.640 (0.018)	291 [88]	0.649 (0.027)	-0.009
Total Passengers	739 [139]	1.516 (0.052)	285 [88]	1.456 (0.047)	0.059
Vehicle Type: Sedan	750 [139]	0.663 (0.017)	291 [88]	0.691 (0.031)	-0.028
Panel B: Main Outcome Variables					
Donation Probability	750 [139]	0.067 (0.010)	291 [88]	0.031 (0.012)	0.036**
Donation	750 [139]	0.066 (0.012)	291 [88]	0.021 (0.008)	0.045***

Notes: The value displayed for t-tests are the differences in the means across the groups. Standard errors are clustered at each traffic light. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

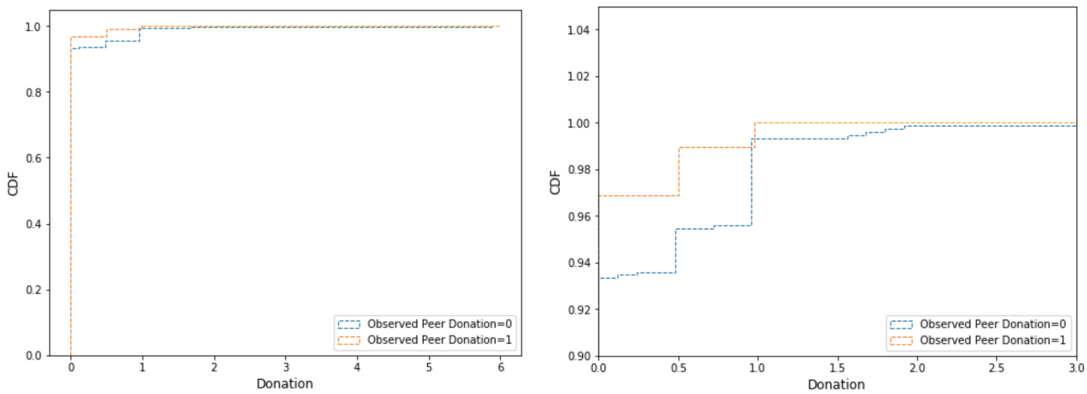


Figure 3.2: CDF Density of Donations by Treatment
Left Panel (all donation decisions), Right Panel (zoom in at positive donations only)

donation and a Tobit model to model the donation magnitude. We also include OLS regressions on donation magnitude in the Appendix.

We use the following specification in the analysis.

$$\begin{aligned} donation_i =& \beta_0 + \beta_1 ObservedAPeerDonation_i + \beta_2 LocationDummy_i \\ &+ \beta_3 ControlVars_i + \epsilon_i \end{aligned} \tag{3.1}$$

where $donation_i$ denotes an amount in Peruvian Sol (PEN). $ObservedAPeerDonation_i$ is 1 if another vehicle donated in the driver's range of sight and 0 otherwise. $LocationDummy_i$ indicate the location of the experiment, and $ControlVars_i$ include covariates such as the row the vehicle in the stoplight, the driver's gender, age, vehicle type, and number of passengers in the vehicle. Standard errors are clustered at each traffic light.

Table 3.2 shows the regression analysis results. The Probit model indicates that vehicles whose driver or passengers observe a donation significantly decrease their probability of donating by around 4%. As expected, we also find that if the vehicle is positioned in the back rows (less visibility of the performer and less likely the performer will reach these rows), drivers are significantly less likely to donate and donate less. If the vehicle is an SUV, it is more likely to donate, arguably due to socioeconomic status.

The results for the donation amount (Tobit model) are consistent: when the driver or passengers of a vehicle see a donation from another vehicle in their range of sight, they significantly decreases their donations by around 0.9 PEN. Again, row position as well as type of vehicle also explain significantly these donation amounts.

We also performed regressions where we separate our treatment variable 'Ob-

Table 3.2: Regression Analysis on Donation Magnitude and Probability

	Probit (Probability)	Probit with Control Vars	Tobit (Magnitude)	Tobit with Control Vars
Observed A Peer Donation	-0.0357**	-.0399**	-0.822*	-0.914**
	(.0164)	(.0160)	(0.430)	(0.427)
Location: 1		-.0110		-0.247
		(.0142)		(0.298)
Location: 2		.0417		0.691*
		(.0264)		(0.394)
Row Position: 2		-.034		-0.610*
		(.0214)		(0.350)
Row Position: 3		-.0407**		-0.613*
		(.0199)		(0.327)
Row Position: 4		-.0678***		-1.591
		(.0260)		(1.033)
Driver: Female		-.00382		-0.130
		(.0177)		(0.337)
Age: 25-50		.0104		-0.0564
		(.0225)		(0.501)
Age: 50+		.00730		-0.0556
		(.0265)		(0.564)
Total Passenger		.00314		0.0564
		(.00390)		(0.0680)
Vehicle: SUV		.0393*		0.737**
		(.0209)		(0.367)
Vehicle: Van		.0227		0.329
		(.0392)		(0.578)
Vehicle: Pickup				-10.33
				(0)
Vehicle: Truck		.0341		0.512
		(.0614)		(0.808)
Vehicle: Motorcycle				-8.817
				(0)
Constant	-1.501***	-1.554***	-3.176***	-2.956***
	(0.0739)	(0.246)	(0.543)	(0.515)
Observations	1,041	962	1,041	974

Probit regression coefficients are in marginal effects.
* p<.1, ** p<.05, ***p<.01 (robust standard errors clustered at each traffic light, displayed in parentheses)

served A Peer Donation' into 'Observed A Hired Driver Donation', 'Observed A Sedan Donation', 'Observed An SUV Donation', and 'Observed Other Donation'. We find negative regression coefficients on all treatment variables. However, due to loss of power (too few observations in each category after the split), we only observe significance in the 'Observed A Hired Driver Donation' treatment, which has the largest number of observations.

3.5 Conclusion

This paper provides field evidence of peer influence on people's giving behavior. Our experiment results show that when drivers see another vehicle donate to the street performer in their range of sight, they are significantly less likely to donate (reduced by 4%), and average donation amount is significantly reduced by around 0.9 PEN. The crowding-out in the amount of donations due to observing a peer donation is rather large and almost close to 100%.

As we have seen in the previous literature, there are theoretical reasons that predict both positive impacts of others' giving behavior (e.g., social norms), as well as negative impacts (e.g., substitution or crowding-out). The majority of field experiments find evidence of positive impacts. That is, due to norms or related motivations, people regard others' giving and their own as complements. To our knowledge, almost all evidence of substitution effects come from empirical or lab studies, and there has not been evidence of negative impacts on giving in the field.

Our paper fills a gap in the literature by studying peer influence in the field in a developing country. Also, to our knowledge, ours is the first study documenting crowding-out effects in the field. There are several reasons why donation behavior in our context could differ from other studies. First, street performers are common in the developing world, and donating to them is rather rare in a natural setting (social norm). Second, unlike charities in the developed world where others' donations could serve as signals of the impact of that cause, in our studied context, individuals can determine the quality of the recipient's performance and the degree of need more directly. Third, individuals or drivers sitting in the vehicles are mostly anonymous to other vehicles, so image motivation is low. These conditions provide a potential setting in the field that is different from most of the donation contexts studies in previous field experiments. Last, the time to offer money after the performance may be quite short (30 seconds). If a vehicle offers money, there is less time for another one to do so, and the other vehicle may decide not to donate to avoid a car accident. The caveat in the current time constraint can be resolved in a future study where we utilize a traffic light that has longer stop time.

In future extensions of this study, we also plan to collect additional survey data to study more in depth the underlying mechanisms behind the substitution effects. We also plan to study how performers' socioeconomic characteristics could affect donation behavior differently. We plan to vary the performers' gender, race, and income-background appearances experimentally.

Bibliography

J. M. Abowd, F. Kramarz, S. Perez-Duarte, and I. Schmutte. A Formal Test of Assortative Matching in the Labor Market. NBER Working Papers 15546, National Bureau of Economic Research, Inc, 2009. URL https://ideas.repec.org/p/nbr/nberwo/15546.html.

C. Ahlin. Matching Patterns When Group Size Exceeds Two. *American Economic Journal: Microeconomics*, 9(1):352–384, 2017. URL https://ideas.repec.org/a/aea/aejmic/v9y2017i1p352-84.html.

F. Alpizar, F. Carlsson, and O. Johansson-Stenman. Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in costa rica. *Journal of Public Economics*, 92(5):1047 – 1060, 2008. ISSN 0047-2727. doi: https://doi.org/10.1016/j.jpubeco.2007.11.004. URL http://www.sciencedirect.com/science/article/pii/S0047272707001909.

A. Alsen, J. Runge, A. Drachen, and D. Klapper". Play with me? understanding and measuring the social aspect of casual gaming. *Association for the Advancement of Artificial Intelligence*, 2016.

J. Altonji, F. Hayashi, and L. Kotlikoff. Parental altruism and inter vivos transfers: Theory and evidence. *Journal of Political Economy*, 105(6):1121–1166, 1997. doi: 10.1086/516388. URL https://doi.org/10.1086/516388.

F. Andersson, S. Burgess, and J. I. Lane. Cities, matching and the productivity gains of agglomeration. *Journal of Urban Economics*, 61(1):112 – 128, 2007. ISSN 0094-1190. doi: https://doi.org/10.1016/j.jue.2006.06.005. URL http://www.sciencedirect.com/science/article/pii/S0094119006000623.

J. Andreoni. Giving with impure altruism: Applications to charity and ricardian equivalence. *Journal of Political Economy*, 97(6):1447–1458, 1989. doi: 10.1086/261662. URL https://doi.org/10.1086/261662.

J. Andreoni. Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, 100(401):464–477, 1990. ISSN 00130133, 14680297. URL http://www.jstor.org/stable/2234133.

J. Andreoni and R. Petrie. Public goods experiments without confidentiality: a glimpse into fund-raising. *Journal of Public Economics*, 88(7):1605 – 1623, 2004. ISSN 0047-2727. doi: https://doi.org/10.1016/S0047-2727(03)00040-9. URL http://www.sciencedirect.com/science/article/pii/S0047272703000409.

M. Andrews, L. Gill, T. Schank, and R. Upward. High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias. *Economics Letters*, 117(3):824 – 827, 2012. ISSN 0165-1765. doi: https://doi.org/10.1016/j.

econlet.2012.07.033. URL http://www.sciencedirect.com/science/article/pii/S0165176512004272.

AppAnnie. The State of Mobile 2019. Technical report, App Annie, 2018.

D. Ariely, A. Bracha, and S. Meier. Doing good or doing well? image motivation and monetary incentives in behaving prosocially. *American Economic Review*, 99(1):544–55, March 2009. doi: 10.1257/aer.99.1.544. URL http://www.aeaweb.org/articles?id=10.1257/aer.99.1.544.

N. Ashraf, O. Bandiera, and B. K. Jack. No margin, no mission? a field experiment on incentives for public service delivery. *Journal of Public Economics*, 120:1 – 17, 2014. ISSN 0047-2727. doi: https://doi.org/10.1016/j.jpubeco.2014.06.014. URL http://www.sciencedirect.com/science/article/pii/S0047272714001546.

W. Bao, Y. Rao, J. Wang, and D. Houser. Social exposure and trustworthiness: Experimental evidence. *Economics Letters*, 162:73 – 75, 2018. ISSN 0165-1765. doi: https://doi.org/10.1016/j.econlet.2017.10.017. URL http://www.sciencedirect.com/science/article/pii/S0165176517304330.

R. Bapna and A. Umyarov. Do your online friends make you pay? a randomized field experiment on peer influence in online social networks. *Manage. Sci.*, 61(8):1902–1920, Aug. 2015. ISSN 0025-1909. doi: 10.1287/mnsc.2014.2081. URL http://dx.doi.org/10.1287/mnsc.2014.2081.

R. Bapna, J. Ramaprasad, and A. Umyarov. Monetizing freemium communities: Does paying for premium increase social engagement? *MIS Quarterly: Management Information Systems*, 42(3):719–735, 9 2018. ISSN 0276-7783. doi: 10.25300/MISQ/2018/13592.

M. Bateson, D. Nettle, and G. Roberts. Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, 2(3):412–414, 2006. ISSN 1744-9561. doi: 10.1098/rsbl.2006.0509. URL http://rsbl.royalsocietypublishing.org/content/2/3/412.

G. S. Becker. A theory of marriage: Part i. *Journal of Political Economy*, 81(4):813–846, 1973. ISSN 00223808, 1537534X. URL http://www.jstor.org/stable/1831130.

R. Benabou and J. Tirole. Intrinsic and Extrinsic Motivation. *The Review of Economic Studies*, 70(3):489–520, 07 2003. ISSN 0034-6527. doi: 10.1111/1467-937X.00253. URL https://dx.doi.org/10.1111/1467-937X.00253.

R. Benabou and J. Tirole. Incentives and prosocial behavior. *American Economic Review*, 96(5):1652–1678, December 2006. doi: 10.1257/aer.96.5.1652. URL http://www.aeaweb.org/articles?id=10.1257/aer.96.5.1652.

P. D. Berger and N. I. Nasr. Customer lifetime value: Marketing models and applications. *Journal of Interactive Marketing*, 12(1):17–30, 1998.

B. D. Bernheim. A Theory of Conformity. *Journal of Political Economy*, 102(5):841–877, October 1994. URL https://ideas.repec.org/a/ucp/jpolec/v102y1994i5p841-77.html.

T. Besley and M. Ghatak. Competition and incentives with motivated agents. *American Economic Review*, 95(3):616–636, June 2005. doi: 10.1257/0002828054201413. URL http://www.aeaweb.org/articles?id=10.1257/0002828054201413.

G. E. Bolton and E. Katok. An experimental test of the crowding out hypothesis: The nature of beneficent behavior. *Journal of Economic Behavior Organization*, 37(3):315 – 331, 1998. ISSN 0167-2681. doi: https://doi.org/10.1016/S0167-2681(98)00091-2. URL http://www.sciencedirect.com/science/article/pii/S0167268198000912.

A. Bracha and L. Vesterlund. Mixed signals: Charity reporting when donations signal generosity and income. *Games and Economic Behavior*, 104:24 – 42, 2017. ISSN 0899-8256. doi: https://doi.org/10.1016/j.geb.2017.03.001. URL http://www.sciencedirect.com/science/article/pii/S0899825617300428.

T. Broberg, T. Ellingsen, and M. Johannesson. Is generosity involuntary? *Economics Letters*, 94(1):32 – 37, 2007. ISSN 0165-1765. doi: https://doi.org/10.1016/j.econlet.2006.07.006. URL http://www.sciencedirect.com/science/article/pii/S0165176506002527.

J. Carpenter, A. Liati, and B. Vickery. They come to play: Supply effects in an economic experiment. *Rationality and Society*, 22(1):83–102, 2010. doi: 10.1177/1043463109358486. URL https://doi.org/10.1177/1043463109358486.

T. N. Cason and V.-L. Mui. Social Influence in the Sequential Dictator Game. Monash

Economics Working Papers archive-37, Monash University, Department of Economics, 1998. URL https://ideas.repec.org/p/mos/moswps/archive-37.html.

C. Chai, E. L. Francis, and T. Xiao. Supply chain dynamics with assortative matching. 2016.

B. P. Chamberlain, Â. Cardoso, C. H. B. Liu, R. Pagliari, and M. P. Deisenroth. Customer life time value prediction using embeddings. *in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017.

N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357, 2002.

Y. Chen, F. M. Harper, J. Konstan, and S. X. Li. Social comparisons and contributions to online communities: A field experiment on movielens. *American Economic Review*, 100(4):1358–98, September 2010. doi: 10.1257/aer.100.4.1358. URL http://www.aeaweb.org/articles?id=10.1257/aer.100.4.1358.

M. Chica and W. Rand. Building agent-based decision support systems for word-of-mouth programs. a freemium application. *SSRN Electronic Journal*, 08 2016. doi: 10.2139/ssrn.2821171.

B. Choi, R. E. Kraut, and M. Fichman. Matching people and groups: Recruitment and selection in online games. 2008.

A. Chouldechova and A. Roth. The frontiers of fairness in machine learning. *arXiv preprint arXiv:1810.08810*, 2018.

K. Coussement, S. Lessmann, and G. Verstraeten. A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telecommunication industry. *Decision Support Systems*, 95, 11 2016.

J. Dana, D. M. Cain, and R. M. Dawes. What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes*, 100(2):193 – 201, 2006. ISSN 0749-5978. doi: https://doi.org/10.1016/j.obhdp.2005.10.001. URL http://www.sciencedirect.com/science/article/pii/S0749597805001330.

D. D. Davis. Rebate Subsidies, Matching Subsidies and Isolation Effects. Technical Report 1, 2006.

D. D. Davis, E. L. Millner, and R. J. Reilly. Subsidy schemes and charitable contributions: A closer look. *Experimental Economics*, 8(2):85–106, Jun 2005. ISSN 1573-6938. doi: 10.1007/s10683-005-0867-y. URL https://doi.org/10.1007/s10683-005-0867-y.

S. DellaVigna, J. A. List, and U. Malmendier. Testing for altruism and social pressure in charitable giving *. *The Quarterly Journal of Economics*, 127(1):1–56, 2012. doi: 10.1093/qje/qjr050. URL +http://dx.doi.org/10.1093/qje/qjr050.

S. N. Durlauf and A. Seshadri. Is assortative matching efficient? *Economic Theory*, 21

(2):475–493, 2003. ISSN 1432-0479. doi: 10.1007/s00199-002-0269-8. URL https://doi.org/10.1007/s00199-002-0269-8.

C. Eckel and P. Grossman. Subsidizing charitable contributions: A field test comparing matching and rebate subsidies. Framed Field Experiments 00145, The Field Experiments Website, 2005b. URL https://ideas.repec.org/p/feb/framed/00145.html.

C. Eckel, P. Grossman, and R. Johnston. An experimental test of the crowding out hypothesis. *Journal of Public Economics*, 89:1543–1560, 08 2005. doi: 10.1016/j.jpubeco.2004.05.012.

C. C. Eckel and P. J. Grossman. Rebate versus matching: does how we subsidize charitable contributions matter? *Journal of Public Economics*, 87(3):681 – 701, 2003. ISSN 0047-2727. doi: https://doi.org/10.1016/S0047-2727(01)00094-9. URL http://www.sciencedirect.com/science/article/pii/S0047272701000949.

C. C. Eckel and P. J. Grossman. Subsidizing charitable giving with rebates or matching: Further laboratory evidence. *Southern Economic Journal*, 72(4):794–807, 2005a. ISSN 00384038. URL http://www.jstor.org/stable/20111853.

C. C. Eckel and P. J. Grossman. *Do Donors Care about Subsidy Type? An Experimental Study*, pages 157–175. 2006.

M. Ekstrom. Do watching eyes affect charitable giving? evidence from a field experiment. *Experimental Economics*, 15(3):530–546, Sep 2012. ISSN 1573-6938. doi: 10.1007/s10683-011-9312-6. URL https://doi.org/10.1007/s10683-011-9312-6.

T. Ellingsen and M. Johannesson. Anticipated verbal feedback induces altruistic behavior. SSE/EFI Working Paper Series in Economics and Finance 668, 2007. URL http://hdl.handle.net/10419/56375.

C. Exley. Incentives for prosocial behavior: The role of reputations. *Management Science*, 64(5):2460–2471, 2018. doi: 10.1287/mnsc.2016.2685. URL https://doi.org/10.1287/mnsc.2016.2685.

P. S. Fader and B. G. S. Hardie. A note on deriving the pareto/nbd model and related expressions. 2005.

P. S. Fader, B. G. S. Hardie, and K. L. Lee. Rfm and clv: Using iso-value curves for customer base analysis. *Journal of Marketing Research*, 42(4):415–430, 2005.

A. Falk, U. Fischbacher, and S. Gachter. Living in two neighborhoods—social interaction effects in the laboratory. *Economic Inquiry*, 51(1):563–578, 2013. doi: 10.1111/j.1465-7295.2010.00332.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1465-7295.2010.00332.x.

E. Fehr and K. M. Schmidt. A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3):817–868, 1999. ISSN 00335533, 15314650. URL http://www.jstor.org/stable/2586885.

E. Filiz-Ozbay and E. Y. Ozbay. Effect of an audience in public goods provision. *Experimental Economics*, 17(2):200–214, 2014. doi: 10.1007/s10683-013-9363-y. Electronic

Supplementary Material The online version of this article (doi: 10.1007/s10683-013-9363-y) contains supplementary material, which is available to authorized users.

U. Fischbacher and S. Gachter. Social preferences, beliefs, and the dynamics of free riding in public good experiments. *American Economic Review*, 2010.

U. Fischbacher, S. Gachter, and E. Fehr. Are people conditionally cooperative? evidence from a public goods experiment. *Economics Letters*, 71(3):397 – 404, 2001. ISSN 0165-1765. doi: https://doi.org/10.1016/S0165-1765(01)00394-9. URL http://www.sciencedirect.com/science/article/pii/S0165176501003949.

B. S. Frey and S. Meier. Social comparisons and pro-social behavior: Testing "conditional cooperation" in a field experiment. *The American Economic Review*, 94(5):1717–1722, 2004. ISSN 00028282. URL http://www.jstor.org/stable/3592843.

S. Gachter and E. Fehr. Collective action as a social exchange. *Journal of Economic Behavior & Organization*, 39(4):341 – 369, 1999. ISSN 0167-2681. doi: https://doi.org/10.1016/S0167-2681(99)00045-1. URL http://www.sciencedirect.com/science/article/pii/S0167268199000451.

S. Gachter, D. Nosenzo, and M. Sefton. The impact of social comparisons on reciprocity. *The Scandinavian Journal of Economics*, 114(4):1346–1367, 2012. doi: 10.1111/j.1467-9442.2012.01730.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9442.2012.01730.x.

S. Gachter, D. Nosenzo, and M. Sefton. Peer effects in pro-social behavior: Social norms or social preferences? *Journal of the European Economic Association*, 11(3):548–573, 2013. doi: 10.1111/jeea.12015. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/jeea.12015.

GameAnalytics. Global Market Trends: Mobile Gaming Benchmarks. Technical report, GameAnalytics, 2019.

A. Ghose and S. P. Han. Estimating demand for mobile applications in the new economy. *Management Science*, 60(6):1470–1488, 2014. doi: 10.1287/mnsc.2014.1945. URL https://doi.org/10.1287/mnsc.2014.1945.

L. Goette and A. Stutzer. Blood donations and incentives: Evidence from a field experiment. *SSRN Electronic Journal*, 06 2008. doi: 10.2139/ssrn.1153406.

D. Goux and E. Maurin. Persistence of interindustry wage differentials: A reexamination using matched worker-firm panel data. *Journal of Labor Economics*, 17(3):492–533, 1999. ISSN 0734306X, 15375307. URL http://www.jstor.org/stable/10.1086/209929.

M. Gruetter and R. Lalive. The importance of firms in wage determination. *Labour Economics*, 16(2):149–160, 2009. URL https://ideas.repec.org/a/eee/labeco/v16y2009i2p149-160.html.

A. Gunnthorsdottir, D. Houser, and K. McCabe. Disposition, history and contributions in public goods experiments. *Journal of Economic Behavior Organization*, 62(2):304

– 315, 2007. ISSN 0167-2681. doi: https://doi.org/10.1016/j.jebo.2005.03.008. URL http://www.sciencedirect.com/science/article/pii/S0167268105002532.

S. Gupta, D. R. Lehmann, and J. A. Stuart. Valuing customers. *Journal of Marketing Research*, 41(1):7–18, 2004.

K. J. Haley and D. M. Fessler. Original article. *Evolution and Human Behavior*, 26(3): 245–256, 2005. doi: 10.1016/j.evolhumbehav.2005.01.002.

W. T. Harbaugh. What do donations buy?: A model of philanthropy based on prestige and warm glow. *Journal of Public Economics*, 67(2):269 – 284, 1998. ISSN 0047-2727. doi: https://doi.org/10.1016/S0047-2727(97)00062-5. URL http://www.sciencedirect.com/science/article/pii/S0047272797000625.

IRS. Charitable contribution deductions. URL https://www.irs.gov/charities-non-profits/charitable-organizations/charitable-contribution-deductions#:~:text=You%20may%20deduct%20charitable%20contributions,limitations%20apply%20in%20some%20cases.

A. Jones and J. Posnett. Charitable donations by uk households: evidence from the family expenditure survey. *Applied Economics*, 23(2):343–351, 1991. doi: 10.1080/00036849100000143. URL https://doi.org/10.1080/00036849100000143.

D. Karlan and M. A. McConnell. Hey look at me: The effect of giving circles on giving. *Journal of Economic Behavior & Organization*, 106:402 – 412, 2014. ISSN 0167-2681.

doi: https://doi.org/10.1016/j.jebo.2014.06.013. URL http://www.sciencedirect.com/science/article/pii/S0167268114002017.

H. Kitchen. Determinants of charitable donations in canada: a comparison over time. *Applied Economics*, 24(7):709–713, 1992. doi: 10.1080/00036849200000039. URL https://doi.org/10.1080/00036849200000039.

M. Kosfeld and S. Neckermann. Getting more work for nothing? symbolic awards and worker performance. *American Economic Journal: Microeconomics*, 3(3):86–99, August 2011. doi: 10.1257/mic.3.3.86. URL http://www.aeaweb.org/articles?id=10.1257/mic.3.3.86.

M. Kremer. The o-ring theory of economic development. *The Quarterly Journal of Economics*, 108(3):551–575, 1993. ISSN 00335533, 15314650. URL http://www.jstor.org/stable/2118400.

E. Krupka and R. A. Weber. The focusing and informational effects of norms on pro-social behavior. *Journal of Economic Psychology*, 30(3):307 – 320, 2009. ISSN 0167-4870. doi: https://doi.org/10.1016/j.joep.2008.11.005. URL http://www.sciencedirect.com/science/article/pii/S0167487008001128.

E. L. Krupka and R. A. Weber. Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, 11(3):495–524, 2013. doi: 10.1111/jeea.12006. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/jeea.12006.

E. L. Krupka, S. Leider, and M. Jiang. A meeting of the minds: Informal agreements and social norms. *Management Science*, 63(6):1708–1729, 2017. doi: 10.1287/mnsc.2016.2429. URL https://doi.org/10.1287/mnsc.2016.2429.

N. Lacetera and M. Macis. Do all material incentives for pro-social activities backfire? the response to cash and non-cash incentives for blood donations. *Journal of Economic Psychology*, 31(4):738 – 748, 2010a. ISSN 0167-4870. doi: https://doi.org/10.1016/j.joep.2010.05.007. URL http://www.sciencedirect.com/science/article/pii/S0167487010000735.

N. Lacetera and M. Macis. Social image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme. *Journal of Economic Behavior & Organization*, 76(2):225 – 237, 2010b. ISSN 0167-2681. doi: https://doi.org/10.1016/j.jebo.2010.08.007. URL http://www.sciencedirect.com/science/article/pii/S0167268110001526.

N. Lacetera, M. Macis, and R. Slonim. Will there be blood? incentives and displacement effects in pro-social behavior. *American Economic Journal: Economic Policy*, 4(1):186–223, February 2012. doi: 10.1257/pol.4.1.186. URL http://www.aeaweb.org/articles?id=10.1257/pol.4.1.186.

P. Legros and A. F. Newman. Beauty is a beast, frog is a prince: Assortative matching with nontransferabilities. *Econometrica*, 75(4):1073–1102, 2007. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/4502021.

A. Lemmens and C. Croux. Bagging and boosting classification trees to predict churn. *Journal of Marketing Research*, 43(2):276–286, 2006.

E. F. P. Luttmer. Group loyalty and the taste for redistribution. *Journal of Political Economy*, 109(3):500–528, 2001. doi: 10.1086/321019. URL https://doi.org/10.1086/321019.

D. Masclet, C. Noussair, S. Tucker, and M. C. Villeval. Monetary and Non-Monetary Punishment in the Voluntary Contributions Mechanism. *The American Economic Review*, 93(2):366 – 380, 2003. URL https://halshs.archives-ouvertes.fr/halshs-00151423.

S. Meier and A. Stutzer. Is volunteering rewarding in itself? *Economica*, 75(297):39–59, 2008. URL https://EconPapers.repec.org/RePEc:bla:econom:v:75:y:2008:i:297:p:39-59.

C. Mellstrom and M. Johannesson. Crowding out in blood donation: Was titmuss right? *Journal of the European Economic Association*, 6(4):845–863, 2008. ISSN 1542-4774. doi: 10.1162/JEEA.2008.6.4.845. URL http://dx.doi.org/10.1162/JEEA.2008.6.4.845.

R. Mendes, G. J. van den Berg, and M. Lindeboom. An empirical assessment of assortative matching in the labor market. *Labour Economics*, 17(6):919 – 929, 2010. ISSN 0927-5371. doi: https://doi.org/10.1016/j.labeco.2010.05.001. URL http://www.sciencedirect.com/science/article/pii/S0927537110000618.

G. E. Newman and Y. J. Shen. The counterintuitive effects of thank-you gifts on charitable giving. *Journal of Economic Psychology*, 33(5):973 – 983, 2012. ISSN 0167-4870. doi: https://doi.org/10.1016/j.joep.2012.05.002. URL http://www.sciencedirect.com/science/article/pii/S0167487012000530.

D. B. Nieborg. Crushing candy: The free-to-play game in its connective commodity form. *Social Media + Society*, 1(2), 2015.

G. Oestreicher-Singer and L. Zalmanson. Content or community? a digital business strategy for content providers in the social age. *MIS Q.*, 37(2):591–616, June 2013. ISSN 0276-7783. doi: 10.25300/MISQ/2013/37.2.12. URL https://doi.org/10.25300/MISQ/2013/37.2.12.

B. A. Olken, J. Onishi, and S. Wong. Should aid reward performance? evidence from a field experiment on health and education in indonesia. *American Economic Journal: Applied Economics*, 6(4):1–34, October 2014. doi: 10.1257/app.6.4.1. URL https://www.aeaweb.org/articles?id=10.1257/app.6.4.1.

M. Orne. On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. *American Psychologist*, 17 (11), 1962.

T. Page, L. Putterman, and B. Unel. Voluntary association in public goods experiments: Reciprocity, mimicry and efficiency*. *The Economic Journal*, 115(506):1032–1053,

2005. doi: 10.1111/j.1468-0297.2005.01031.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0297.2005.01031.x.

A. Payne. Does the government crowd-out private donations? new evidence from a sample of non-profit firms. *Journal of Public Economics*, 69(3):323 – 345, 1998. ISSN 0047-2727. doi: https://doi.org/10.1016/S0047-2727(98)00005-X. URL http://www.sciencedirect.com/science/article/pii/S004727279800005X.

A. Payne. Measuring the Effect of Federal Research Funding on Private Donations at Research Universities: Is Federal Research Funding More than a Substitute for Private Donations? *International Tax and Public Finance*, 8(5):731–751, November 2001. doi: 10.1023/A:1012843227003. URL https://ideas.repec.org/a/kap/itaxpf/v8y2001i5p731-751.html.

P. Pei Chen, A. Guitart, A. Fernández del Río, and Á. Periáñez. Customer Lifetime Value in Video Games Using Deep Learning and Parametric Models. *arXiv e-prints*, 2018.

M. Rege and K. Telle. The impact of social approval and framing on cooperation in public good situations. *Journal of Public Economics*, 88(7):1625 – 1644, 2004. ISSN 0047-2727. doi: https://doi.org/10.1016/S0047-2727(03)00021-5. URL http://www.sciencedirect.com/science/article/pii/S0047272703000215.

W. Reinartz and V. Kumar. The impact of customer relationship characteristics on profitable lifetime duration. *Journal of Marketing*, 67:77–99, 2003.

E. Reuben and A. Riedl. Enforcement of contribution norms in public good games with heterogeneous populations. *Games and Economic Behavior*, 77(1):122 – 137, 2013. ISSN 0899-8256. doi: https://doi.org/10.1016/j.geb.2012.10.001. URL http://www.sciencedirect.com/science/article/pii/S0899825612001492.

R. D. Roberts. A positive model of private charity and public transfers. *Journal of Political Economy*, 92(1):136–48, 1984. URL https://EconPapers.repec.org/RePEc:ucp:jpolec:v:92:y:1984:i:1:p:136-48.

O. A. Rud and J. P. Rabanal. Efficient investment via assortative matching. 2015.

J. Runge, S. Wagner, D. Klapper, and J. Claussen. Freemium pricing: Evidence from a large-scale field experiment. *Academy of Management Proceedings*, 2017:11533, 01 2017.

A. Rustichini and U. Gneezy. Pay Enough or Don't Pay at All*. *The Quarterly Journal of Economics*, 115(3):791–810, 08 2000. ISSN 0033-5533. doi: 10.1162/003355300554917. URL https://dx.doi.org/10.1162/003355300554917.

A. SavikhinSamek and R. M. Sheremeta. Recognizing contributors: an experiment on public goods. *Experimental Economics*, 17(4):673–690, Dec 2014. ISSN 1573-6938. doi: 10.1007/s10683-013-9389-1. URL https://doi.org/10.1007/s10683-013-9389-1.

J. Shang and R. Croson. A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods. *The Economic Journal*, 119(540):1422–1439, 2009. doi: 10.1111/j.1468-0297.2009.02267.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0297.2009.02267.x.

R. Shimer and L. Smith. Assortative matching and search. *Econometrica*, 68(2):343–369, 2000. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/2999430.

R. Sifa, F. Hadiji, J. Runge, A. Drachen, K. Kersting, and C. Bauckhage. Predicting purchase decisions in mobile free-to-play games. 2015.

R. Sifa, J. Runge, C. Bauckhage, , and D. Klapper. Customer lifetime value prediction in non-contractual freemium settings: Chasing high-value users using deep neural networks and SMOTE. *In Proceedings of the 51st Hawaii International Conference on System Sciences*, pages 923–932, 2018.

S. Smith, F. Windmeijer, and E. Wright. Peer effects in charitable giving: Evidence from the (running) field. *The Economic Journal*, 125(585):1053–1071, 2015. doi: 10.1111/ecoj.12114. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ecoj.12114.

A. R. Soetevent. Anonymity in giving in a natural context—a field experiment in 30 churches. *Journal of Public Economics*, 89(11):2301 – 2323, 2005. ISSN 0047-2727. doi: https://doi.org/10.1016/j.jpubeco.2004.11.002. URL http://www.sciencedirect.com/science/article/pii/S0047272704001756.

R. Steinberg. Does government spending crowd out donations? interpreting the evidence.

Annals of Public and Cooperative Economics, 62:591 – 612, 03 1991. doi: 10.1111/j.1467-8292.1991.tb01369.x.

S. Torres, P. Portugal, J. T. Addison, and P. Guimaraes. The sources of wage variation and the direction of assortative matching: Evidence from a three-way high-dimensional fixed effects regression model. *Labour Economics*, 54:47 – 60, 2018. ISSN 0927-5371. doi: https://doi.org/10.1016/j.labeco.2018.06.004. URL http://www.sciencedirect.com/science/article/pii/S0927537117300489.

A. Vanderveld, A. Pandey, A. Han, and R. Parekh. An engagement-based customer lifetime value system for e-commerce. *in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 293–302, 08 2016.

P. Verhoef, P. Spring, J. Hoekstra, and P. Leeflang. The commercial use of segmentation and predictive modeling techniques for database marketing in the netherlands. *Decision Support Systems*, 34(4):471–481, 2003.

P. Warr. Pareto optimal redistribution and private charity. *Journal of Public Economics*, 19(1):131–138, 1982. URL https://EconPapers.repec.org/RePEc:eee:pubeco:v:19:y:1982:i:1:p:131-138.

G. M. Weiss. Mining with rarity: A unifying framework. *SIGKDD Explor. Newsl.*, 6(1):7–19, 2004. ISSN 1931-0145.

S. D. Woodcock. Match Effects. Discussion Papers dp07-13, Department of Economics, Simon Fraser University, Aug. 2007. URL https://ideas.repec.org/p/sfu/sfudps/dp07-13.html.

S. T. Yen. An econometric analysis of household donations in the usa. *Applied Economics Letters*, 9(13):837–841, 2002. doi: 10.1080/13504850210148189. URL https://doi.org/10.1080/13504850210148189.

Appendix A

Supplement to Chapter One

In this section, we report the impact of the assortative matching system on all the outcome variables we test. The results are consistent with the ones we report in the main text.

On the user-level, tables A.7 and A.8 report the IV regression results on game-specific metrics and productivity metrics using all outcome variables, while tables A.9 and A.10 report the net effects analysis using the same outcome variables from our OLS regressions.

Table A.7: User-Level IV Regression on Game Metrics (Full Table)

	(a) Predicted Payers						
	Revenue	Gem Spend	Message	Gift	Weekly Events	Retention	Satisfaction
Intercept	-2.2**	8.5***	-1.3	-2.1	1.6	-2.3***	0.11
	(1.1)	(3.1)	(1.2)	(1.8)	(1.8)	(0.62)	(0.66)
$TActivity_{pred}$	0.062	0.17	0.36***	0.15*	0.26***	0.078***	0.033
	(0.046)	(0.14)	(0.053)	(0.078)	(0.08)	(0.027)	(0.03)
R-squared	0.04	0.02	0.02	0.02	0.04	0.02	0.04
Observations	17965	17965	17965	17965	17965	17972	5580

	(b) Predicted Non-Payers						
	Revenue	Gem Spend	Message	Gift	Weekly Events	Retention	Satisfaction
Intercept	-2.1***	4.6***	-2.1***	-0.97	-1.1*	-2.1***	0.98
	(0.24)	(1.4)	(0.42)	(0.66)	(0.65)	(0.23)	(0.61)
$TActivity_{pred}$	-0.011	-0.22	0.12	-0.16	0.24	0.037	-0.26
	(0.088)	(0.51)	(0.15)	(0.24)	(0.24)	(0.085)	(0.18)
R-squared	0.02	0.02	0.02	0.01	0.04	0.01	0.06
Observations	13645	13645	13645	13645	13645	13646	2465

* $p<.1$, ** $p<.05$, *** $p<.01$ (standard errors in parentheses)
User-level IV regressions include user device group controls, acquisition channel controls, weekly event controls, and day of week controls

Table A.8: User-Level IV Regression on Productivity Metrics (Full Table)

	(a) Predicted Payers			
	Time Spent	Campaign Starts	Campaign Wins	Success Rate
$Intercept$	6.3***	4.7***	4.7***	1***
	(0.77)	(0.66)	(0.65)	(0.064)
$TActivity_{pred}$	0.081**	0.093**	0.073**	0.0026
	(0.038)	(0.043)	(0.037)	(0.0052)
$Campaign$				-0.00083***
$Starts_{day14}$				(2.1e-05)
R-squared	0.015	0.026	0.027	0.42
Observations	17965	17965	17965	17965

	(b) Predicted Non-Payers			
	Time Spent	Campaign Starts	Campaign Wins	Success Rate
$Intercept$	4.8***	3.2***	3.2***	0.98***
	(0.44)	(0.43)	(0.41)	(0.05)
$TActivity_{pred}$	-0.073	-0.067	-0.089	-0.003
	(0.13)	(0.14)	(0.13)	(0.019)
$Campaign$				-0.00088***
$Starts_{day14}$				(2.5e-05)
R-squared	0.027	0.038	0.042	0.28
Observations	13645	13645	13645	13645

* p<.1, ** p<.05, ***p<.01 (standard errors in parentheses)
User-level IV regressions include user device group controls, acquisition channel controls, weekly event controls, and day of week controls

Table A.9: User-Level Net Effects on Game-Specific Metrics (Full Table)

	Revenue	Gem Spend	Message	Gift	Weekly Events	Retention	Satisfaction
$Intercept$	-1.9***	5.3***	-1.7***	-0.87	-0.024	-2***	0.84**
	(0.33)	(1.2)	(0.42)	(0.63)	(0.64)	(0.22)	(0.41)
$SystemON(1)$	0.025	0.067	0.25***	0.086**	0.18***	0.046***	0.0059
	(0.023)	(0.082)	(0.029)	(0.043)	(0.044)	(0.015)	(0.021)
$NonPayer(1)$	-0.24***	-0.62***	-0.2***	-0.38***	-0.35***	-0.056***	-0.13***
	(0.023)	(0.08)	(0.028)	(0.043)	(0.043)	(0.015)	(0.027)
$ON*$	0.0045	-0.032	-0.24***	-0.091*	-0.14***	-0.025	-0.0061
$NonPayer$	(0.028)	(0.1)	(0.035)	(0.053)	(0.054)	(0.019)	(0.035)
R-squared	0.046	0.023	0.028	0.033	0.055	0.022	0.038
Observations	31610	31610	31610	31610	31610	31618	8045
Net Effects T-Test	0.019	0.038	0.1***	0.033	0.086***	0.025***	0.0023
	(0.014)	(0.049)	(0.017)	(0.026)	(0.026)	(0.0091)	(0.013)

* p<.1, ** p<.05, ***p<.01 (standard errors in parentheses)
User-level net effect regressions include user device group controls, acquisition channel controls, weekly event controls, and day of week controls
T-test hypothesis: $0.7 * SystemOn + 0.3 * (ON * NonPayer) = 0$

Table A.10: User-Level Net Effects on Productivity Metrics (Full Table)

	Time Spent	Campaign Starts	Campaign Wins	Success Rate
Intercept	5.2***	3.6***	3.5***	0.98***
	(0.38)	(0.37)	(0.36)	(0.041)
SystemON	0.044*	0.057**	0.041*	-0.0005
	(0.023)	(0.025)	(0.022)	(0.0031)
NonPayer	-0.16***	-0.19***	-0.18***	-0.012***
	(0.021)	(0.023)	(0.021)	(0.0028)
$ON *$ $NonPayer$	-0.038	-0.058**	-0.045*	0.0024
	(0.027)	(0.029)	(0.026)	(0.0038)
$Campaign$ $Starts_{day14}$				-0.00084***
				(1.7e-05)
R-squared	0.024	0.045	0.048	0.37
Observations	31610	31610	31610	31610
Net Effects T-Test	0.019	0.022	0.015***	0.0004
	(0.013)	(0.015)	(0.013)	(0.839)

* p<.1, ** p<.05, ***p<.01 (standard errors in parentheses)
User-level net effect regressions include user device group controls, acquisition channel controls, weekly event controls, and day of week controls
T-test hypothesis: $0.7 * SystemOn + 0.3 * (ON * NonPayer) = 0$

On the team-level, tables A.11 and A.12 report the impact of the assortative matching system on the team average engagement and productivity of all team members, while tables A.13 and A.14 record those effects using existing team members only. Finally, tables A.15 and A.16 contain the overall net effects of the matching system on the team-level.

Table A.11: Team-Level Analysis on Game-Specific Metrics (Average of All Team Members: Full Table)

	Revenue	Weekly Event	Message	Gift
Intercept	0.77***	4.1***	0.93***	1.3***
	(0.16)	(0.16)	(0.12)	(0.13)
$T_{LowActive}$	-1.7***	-0.59***	-2***	-0.9***
	(0.12)	(0.092)	(0.096)	(0.082)
$ln(NPayer)$	0.19***	0.1***	0.088***	0.073***
	(0.034)	(0.029)	(0.03)	(0.025)
$ln(NPayer)*$	-0.089**	-0.15***	-0.071**	-0.038
$T_{LowActive}$	(0.039)	(0.037)	(0.032)	(0.031)
$NNewJoins$	-0.033***	-0.018***	-0.013***	-0.013***
	(0.003)	(0.0027)	(0.0023)	(0.0025)
Team Controls	✓	✓	✓	✓
R-squared	0.52	0.31	0.68	0.4
Observations	4635	4635	4635	4635

* p<.1, ** p<.05, ***p<.01 (standard errors in parentheses)
Team control variables include pre-experiment 14-day cumulative campaign starts, messaging, gifts, revenue, and age and size the day before the experiment

Table A.12: Team-Level Analysis on Productivity Metrics (Average of All Team Members: Full Table)

	Time Spent	Campaign Starts	Campaign Wins	Success Rate
Intercept	5.5***	4.4***	4.3***	0.78***
	(0.16)	(0.15)	(0.15)	(0.018)
$T_{LowActive}$	-0.64***	-0.63***	-0.63***	-0.002
	(0.087)	(0.085)	(0.084)	(0.0094)
$ln(NPayer)$	0.013	0.044*	0.04	-0.0012
	(0.027)	(0.026)	(0.026)	(0.0027)
$ln(NPayer)*$	0.0096	-0.045	-0.046	0.0054
$T_{LowActive}$	(0.038)	(0.035)	(0.035)	(0.0039)
$NNewJoins$	-0.0028	-0.01***	-0.01***	0.00017
	(0.0024)	(0.0023)	(0.0023)	(0.00026)
Team Controls	✓	✓	✓	✓
R-squared	0.23	0.29	0.29	0.044
Observations	4635	4635	4635	4635

* p<.1, ** p<.05, ***p<.01 (standard errors in parentheses)
Team control variables include pre-experiment 14-day cumulative campaign starts, messaging, gifts, revenue, and age and size the day before the experiment

Table A.13: Team-Level Analysis on Game-Specific Metrics (Average of Existing Team Members: Full Table)

	Revenue	Weekly Event	Message	Gift
$Intercept$	0.67***	4.1***	0.92***	1.3***
	(0.16)	(0.18)	(0.13)	(0.15)
$T_{LowActive}$	-1.7***	-0.56***	-2***	-0.83***
	(0.12)	(0.098)	(0.1)	(0.091)
$ln(NPayer)$	0.24***	0.12***	0.15***	0.15***
	(0.037)	(0.032)	(0.032)	(0.03)
$ln(NPayer)*$	-0.18***	-0.15***	-0.13***	-0.13***
$T_{LowActive}$	(0.041)	(0.04)	(0.034)	(0.035)
$NNewJoins$	-0.03***	-0.01***	-0.011***	-0.014***
	(0.0033)	(0.0032)	(0.0026)	(0.0031)
$TeamControls$	✓	✓	✓	✓
R-squared	0.51	0.27	0.66	0.37
Observations	4635	4635	4635	4635

* p<.1, ** p<.05, ***p<.01 (standard errors in parentheses)
Team control variables include pre-experiment 14-day cumulative campaign starts, messaging, gifts, revenue, and age and size the day before the experiment

Table A.14: Team-Level Analysis on Productivity Metrics (Average of Existing Team Members: Full Table)

	Time Spent	Campaign Starts	Campaign Wins	Success Rate
Intercept	5.4***	4.3***	4.2***	0.77***
	(0.17)	(0.17)	(0.17)	(0.02)
$T_{LowActive}$	-0.6***	-0.58***	-0.58***	-0.0019
	(0.091)	(0.092)	(0.091)	(0.0098)
$ln(NPayer)$	0.058**	0.082***	0.082***	0.0029
	(0.03)	(0.029)	(0.029)	(0.0031)
$ln(NPayer)*$	-0.041	-0.087**	-0.091**	0.0017
$T_{LowActive}$	(0.039)	(0.038)	(0.038)	(0.0042)
$NNewJoins$	-0.001	-0.0059**	-0.0063**	-0.00012
	(0.0028)	(0.0028)	(0.0028)	(0.00031)
Team Controls	✓	✓	✓	✓
R-squared	0.23	0.26	0.27	0.046
Observations	4635	4635	4635	4635

* p<.1, ** p<.05, ***p<.01 (standard errors in parentheses)
Team control variables include pre-experiment 14-day cumulative campaign starts, messaging, gifts, revenue, and age and size the day before the experiment

Table A.15: Team-Level Net Effects Analysis on Game-Specific Metrics (All Team Members: Full Table)

	Revenue	Weekly Event	Message	Gift
$Intercept$	2.3***	5.8***	2.4***	2.9***
	(0.2)	(0.19)	(0.16)	(0.17)
$T_{LowActive}$	-3.1***	-1.5***	-3.2***	-1.9***
	(0.15)	(0.11)	(0.12)	(0.11)
$NPayer*$	0.022***	0.01**	0.018***	0.015***
$T_{HighActive}$	(0.0052)	(0.0041)	(0.0044)	(0.0039)
$NPayer*$	0.079***	0.0045	0.03***	0.055***
$T_{LowActive}$	(0.016)	(0.013)	(0.012)	(0.012)
$NNonPayer*$	-0.095***	-0.028***	-0.04***	-0.046***
$T_{HighActive}$	(0.017)	(0.01)	(0.012)	(0.01)
$NNonPayer*$	0.026	0.022*	0.031**	0.041***
$T_{LowActive}$	(0.018)	(0.012)	(0.013)	(0.012)
$TeamControls$	✓	✓	✓	✓
R-squared	0.58	0.45	0.68	0.52
Observations	4635	4635	4635	4635
Net Effect T-Test	0.0081**	0.0059*	0.0076**	0.0045
	(0.004)	(0.003)	(0.003)	(0.003)

* p<.1, ** p<.05, ***p<.01 (standard errors in parentheses)

Team control variables include pre-experiment 14-day cumulative campaign starts, messaging, gifts, revenue, and age and size the day before the experiment

T-test hypothesis: $0.7 * 0.7 * (0.8 * NPayer * T_{HighActive} - 0.2 * NPayer * T_{LowActive}) + *0.3 * 0.3 * (0.2 * NNonPayer * T_{LowActive} - 0.8 * NNonPayer * T_{HighActive})$

Table A.16: Team-Level Net Effects Analysis on Productivity Metrics (All Team Members: Full Table)

	Time Spent	Campaign Starts	Campaign Wins	Success Rate
$Intercept$	7.2***	6.1***	6***	0.78***
	(0.18)	(0.17)	(0.17)	(0.017)
$T_{LowActive}$	-1.6***	-1.5***	-1.5***	-0.011
	(0.11)	(0.11)	(0.1)	(0.0086)
$NPayer*$	0.011***	0.01**	0.0096**	1.2e-05
$T_{HighActive}$	(0.0041)	(0.0039)	(0.0039)	(0.00025)
$NPayer*$	0.043***	0.029**	0.028**	0.0018*
$T_{LowActive}$	(0.013)	(0.012)	(0.012)	(0.001)
$NNonPayer*$	-0.013	-0.023**	-0.023**	-0.0016 ***
$T_{HighActive}$	(0.0094)	(0.0096)	(0.0096)	(0.00056)
$NNonPayer*$	0.028***	0.029***	0.029***	0.002***
$T_{LowActive}$	(0.011)	(0.011)	(0.011)	(0.0007)
$Campaign$				4.3e-06***
$Starts_{day14}$				(8.6e-07)
$TeamControls$	✓	✓	✓	✓
R-squared	0.42	0.45	0.45	0.045
Observations	4635	4635	4635	4635
Net Effect T-Test	0.0016	0.0033	0.0032	-2.437e-05
	(0.003)	(0.003)	(0.003)	(0.000)

* p<.1, ** p<.05, ***p<.01 (standard errors in parentheses)

Team control variables include pre-experiment 14-day cumulative campaign starts, messaging, gifts, revenue, and age and size the day before the experiment

T-test hypothesis: $0.7 * 0.7 * (0.8 * NPayer * T_{HighActive} - 0.2 * NPayer * T_{LowActive}) + *0.3 * 0.3 * (0.2 * NNonPayer * T_{LowActive} - 0.8 * NNonPayer * T_{HighActive})$

Appendix B

Supplement to Chapter Two

Table B.1 presents the probit regression of correlations between demographic characteristics and subjects' perception of public rebate.

B.1 Experiment Instructions

You are now taking part in an economics experiment, funded by the UCSC LEEPS Lab. This is a short experiment, please take your time to carefully read the instructions before making your decisions. By participating in this session, you will earn a $4 participation fee. During the experiment, depending on your decisions, you might receive additional payoffs. Immediately after the experiment, your total earnings will be paid to you via Venmo.

During the experiment, communication between participants will not be allowed. If you have questions, please send the experimenter a private Zoom message.

Table B.1: Correlations of demographic characteristics to subject types

	0 (Negative Image); 1 (Non-Negative Image)
Gender: Male	-0.334***
	(0.0844)
Gender: Other	-0.287
	(0.233)
Age	-0.0299
	(0.0194)
Charity Passion	0.0819***
	(0.0199)
Donate in Last Year	-0.0305
	(0.0794)
Work: Part-time	0.341***
	(0.0874)
Work: Full-time	0.209
	(0.150)
FB use: once a week	-0.528***
	(0.105)
FB use: once a month	-0.130
	(0.184)
FB use: rarely	-0.0978
	(0.0969)
Constant	0.188
	(0.477)
Observations	1,120

* p<.1, ** p<.05, ***p<.01 (robust standard errors in parentheses)

B.1.1 Specific Instructions

1. You will participate in a square counting task to earn your income.

2. You will choose a charity of your preference from a list of six charities.

3. You will make eight (8) consecutive donation decisions.

- In each decision, you will have the opportunity to donate part of your income to the charity of your choice. That is, your task is to decide how to allocate your income between the charity and your personal account.
- Each of the eight decisions will occur under different rules. Before each decision, you will be informed of the rules that apply to that particular decision.
- Some donations will be public (your name and donation will be posted on the LEEPS Facebook Page and will be visible to everyone). Some donations will be private (your name and donation will not be posted/shared anywhere, not even the charity).
- Some donations will receive a rebate (you get money back) from the LEEPS Lab.
- Some donations will not receive any rebate. The rebate information may be public or private, depending on your donation visibility.
- At the beginning of each decision, you will always have your original balance of participation fee + task income in your cash account. All eight decisions are independent, so your money and donation from one decision will NOT impact or transfer to the next.

After all your 8 decisions, the computer will randomly choose one decision and implement it. That is, the donation amount you indicated in the chosen decision will be placed, and you will receive the remaining amount.

Donations are real. The LEEPS Lab will add up all participants' donations and make one donation to each charity. We will not mention individual names to charities when making the donations, but we will email you a receipt afterwards.

Please note that the rebate offered by LEEPS will NOT apply to donations outside of this experiment.

B.1.2 List of Charities with Detailed Descriptions

Figures B.1 and B.2 show the list of charities presented to subjects during the experiment.

Figure B.1: List of Charities1

Choose Your Charity

Here is a list of charities and funds. You have **10** minutes to to read and learn more about their missions and activities. Given the special COVID-19 circumstances, we've selected charities and organizations that are on the front line fighting against the virus. We are also presenting trusted local charities in the Santa Cruz area. After you are done reading, please select **one** charity. If you choose to donate to it in any of the upcoming 8 decisions, we will send your donation to the selected charity.

UCSF COVID-19 Response Fund

University of California, San Francisco is among the top health sciences universities in the world. We are coordinating with colleagues across the Bay Area and Northern California to care for patients. As the COVID-19 outbreak expands, teams throughout UCSF's hospitals, clinics, and research labs are actively monitoring and responding to the evolving situation.

Your gift will be used to address the needs of patients and caregivers impacted by COVID-19, and to support emerging areas of greatest need, such as:

- Expanding diagnostic and testing capacity.
- Securing our ability to cover extended staff time and benefits for health care providers working directly on the crisis.
- Utilizing next-generation sequencing and gene-editing technologies to develop novel methods for diagnosing the disease.
- Using cell mapping and computational tools to uncover new or existing therapeutics for treating COVID-19.

Santa Cruz County Animal Shelter

SANTA CRUZ COUNTY ANIMAL SHELTER
OPEN **DOOR.** OPEN **HEART.**

The Santa Cruz County Animal Shelter is a non-profit Joint Powers Authority that provides 24-hour animal rescue and is Santa Cruz County's only full service, open-admission animal shelter. We rescue and assure temporary shelter, veterinary and humane care for approximately 5,000 stray, unwanted, abandoned, mistreated, neglected and injured animals every year.

The support of our community is crucial to sustaining and expanding our services. Effective October, 2019, we are initiating a three-stage expansion at our 7th Avenue/Rodriguez Street site.

- Phase I - 2020: Expand and Improve Shelter's Main Building to double the veterinary procedure area, create additional meeting & work space, relocate and improve cat and rabbit housing areas, open a community cat room, upgrade public reception areas and improve access to canine play yards.
- Phase II - 2021: New Cat Cafe – a cat adoption center with coffee shop service.
- Phase III - 2022: New Training and Education Center for owners, pets and professional animal handlers will offer counseling for adopters, a behavioral hotline, and a range of workshops for the public.

WHO COVID-19 Response Fund

The world is facing an unprecedented challenge with communities and economies everywhere affected by the growing COVID-19 pandemic. The world is coming together to combat the COVID-19 pandemic bringing governments, organizations from across industries and sectors and individuals together to help respond to this global outbreak.

The World Health Organization (WHO) is leading and coordinating the global effort with a range of partners, supporting countries to prevent, detect, and respond to the pandemic.

Our Goals:

- Putting in place activities to Track and understand the spread of the virus;
- Buying and ship essential supplies such as masks, gloves and protective wear for frontline workers;
- Producing guidance for the general public and for particular groups on measures to take to prevent the spread of the disease;
- Accelerating efforts to develop vaccines, tests and treatments.

UCSC COVID-19 Slug Support Campaign

These unprecedented times are challenging for all of us—but for some, the situation is dire. This is why our campus has created the COVID-19 Slug Support Campaign. Together, our Banana Slug community can help support students who are experiencing financial or personal crisis because of the infection and the drastic measures our state and country have had to take to curb it.

Figure B.2: List of Charities2

6/2/2020 Choose Your Charity

Your gift today will help our slug community to:

- Provide food stability through digital gift cards, Instacart purchases, and meal swipes to campus dining halls that remain open.
- Fund students in need of emergency housing and provide rent support for students who may have lost wages due to reduced work hours.
- Provide distance learning resources through a new laptop lending program called Slug Tech.

Santa Cruz Homeless Garden Project

The Homeless Garden Project provides job training, transitional employment and support services to people who are experiencing homelessness. HGP's vibrant education and volunteer program for the broad community blends formal, experiential and service-learning. The programs take place in our 3-acre organic farm and related enterprises.

In May of 1990, the Citizens Committee for the Homeless, a Santa Cruz County non-profit, began a new project by opening the gates of an organic garden on Pelton Avenue. The Homeless Garden Project would provide job-training and meaningful work in a therapeutic environment. The Project began as a place to provide sanctuary, refuge and meaningful work within the healing environment of the organic farm.

Our Values

- The capacity of every individual for growth and renewal
- The joy that comes from growing and sharing healthy food
- The well-being created by vibrant social and natural ecosystems

Año Nuevo Research at UCSC

Año Nuevo Reserve is one of the University of California's 39 Natural Reserves. It is located a short drive north of Santa Cruz and is an important breeding site for an incredible diversity of marine mammals and birds. Nearly 100,000 tourists visit the Reserve every year to see the northern elephant seals on docent-led tours.

The close proximity of the Reserve to UC Santa Cruz and several other universities make it a hotspot for undergraduate experiential learning. Students obtain hands-on research experience while working on world-class scientific research projects, resulting in dozens of peer-reviewed publlications annually.

Recent research topics include:

- Contaminant bioaccumulation in top predators
- The impact of acoustic disturbance on marine mammals
- Using animals as ocean sensors to study both animal behavior and the ocean environment

Charity:

Next

https://ruizhi-zhang-otree-experiment.herokuapp.com/p/phj07p1j/Donations/Charity/4/ 2/2

Appendix C

Supplement to Chapter Three

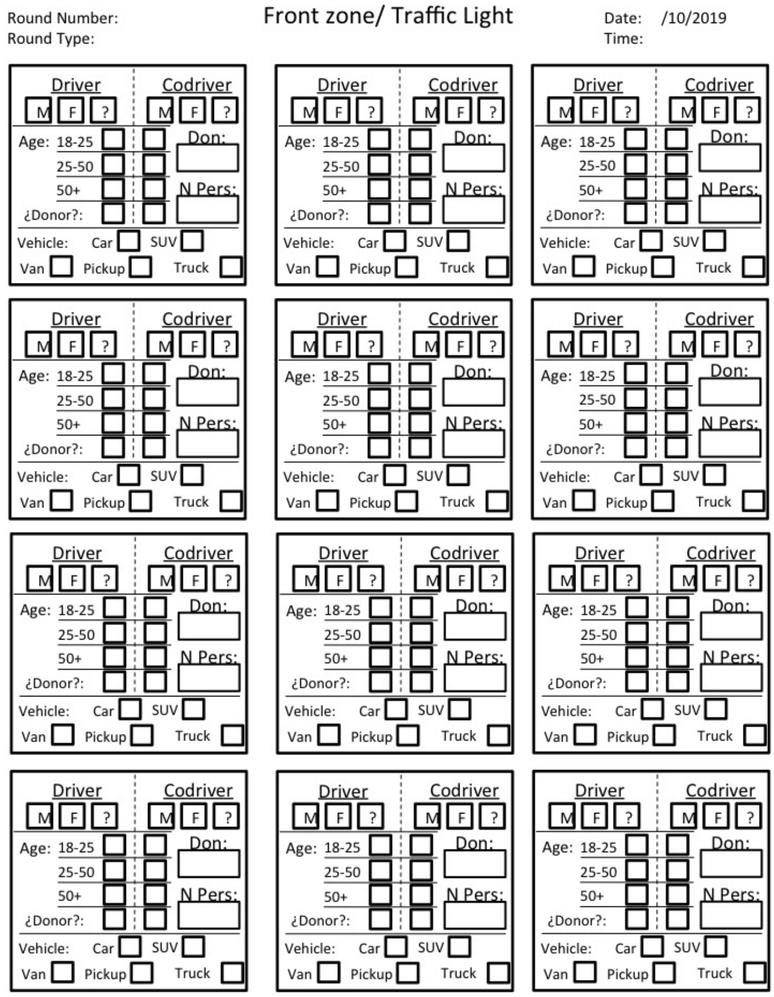
Table C.1 presents OLS regression results of the model specified in equation 3.1.

Figure C.1 displays the data-recording booklet we used during the experiment for a three-lane street. Each vehicle is represented by a box, where the data-recorder fills in the demographic characteristics as well as donations received according to the position of each vehicle.

Table C.1: OLS Regression on Donation Magnitude

	OLS (Donation Magnitude)	OLS with Control Vars
Observed A Peer Donation	-0.0452***	-0.0413***
	(0.0156)	(0.0147)
Location: 1		-0.0165
		(0.0152)
Location: 2		0.0579
		(0.0434)
Row Position: 2		-0.0453**
		(0.0179)
Row Position: 3		-0.0255
		(0.0265)
Row Position: 4		-0.104*
		(0.0534)
Driver: Female		-0.0189
		(0.0231)
Age: 25-50		-0.0488
		(0.0627)
Age: 50+		-0.0408
		(0.0657)
Total Passenger		0.00508
		(0.00679)
Vehicle: SUV		0.0644*
		(0.0355)
Vehicle: Van		0.00954
		(0.0326)
Vehicle: Pickup		-0.0721***
		(0.0231)
Vehicle: Truck		0.0248
		(0.0456)
Vehicle: Motorcycle		-0.0111
		(0.0203)
Constant	0.0659***	0.106**
	(0.0121)	(0.0528)
Observations	1,041	974
R-squared	0.005	0.028

* p<.1, ** p<.05, ***p<.01 (robust standard errors clustered at each traffic light, displayed in parentheses)


Round Number: Front zone/ Traffic Light Date: /10/2019
Round Type: Time:

Driver Codriver
M F ? M F ?
Age: 18-25 Don:
25-50
50+ N Pers:
¿Donor?:
Vehicle: Car SUV
Van Pickup Truck

Driver Codriver
M F ? M F ?
Age: 18-25 Don:
25-50
50+ N Pers:
¿Donor?:
Vehicle: Car SUV
Van Pickup Truck

Driver Codriver
M F ? M F ?
Age: 18-25 Don:
25-50
50+ N Pers:
¿Donor?:
Vehicle: Car SUV
Van Pickup Truck

Driver Codriver
M F ? M F ?
Age: 18-25 Don:
25-50
50+ N Pers:
¿Donor?:
Vehicle: Car SUV
Van Pickup Truck

Driver Codriver
M F ? M F ?
Age: 18-25 Don:
25-50
50+ N Pers:
¿Donor?:
Vehicle: Car SUV
Van Pickup Truck

Driver Codriver
M F ? M F ?
Age: 18-25 Don:
25-50
50+ N Pers:
¿Donor?:
Vehicle: Car SUV
Van Pickup Truck

Driver Codriver
M F ? M F ?
Age: 18-25 Don:
25-50
50+ N Pers:
¿Donor?:
Vehicle: Car SUV
Van Pickup Truck

Driver Codriver
M F ? M F ?
Age: 18-25 Don:
25-50
50+ N Pers:
¿Donor?:
Vehicle: Car SUV
Van Pickup Truck

Driver Codriver
M F ? M F ?
Age: 18-25 Don:
25-50
50+ N Pers:
¿Donor?:
Vehicle: Car SUV
Van Pickup Truck

Driver Codriver
M F ? M F ?
Age: 18-25 Don:
25-50
50+ N Pers:
¿Donor?:
Vehicle: Car SUV
Van Pickup Truck

Driver Codriver
M F ? M F ?
Age: 18-25 Don:
25-50
50+ N Pers:
¿Donor?:
Vehicle: Car SUV
Van Pickup Truck

Driver Codriver
M F ? M F ?
Age: 18-25 Don:
25-50
50+ N Pers:
¿Donor?:
Vehicle: Car SUV
Van Pickup Truck

Figure C.1: The data recording booklet we used in our experiment