ABSTRACT

Title of Dissertation: THE INTERPLAY BETWEEN SOCIAL CONNECTIONS AND DIGITAL TECHNOLOGIES: THREE ESSAYS EXAMINING HEALTHY BEHAVIORS AND INCOME MOBILITY

Che-Wei Liu, Doctor of Philosophy, 2018

Dissertation directed by: Ritu Agarwal, Professor, Decision, Operations, and Information Technologies
Sunil Mithas, Professor, Decision, Operations, and Information Technologies

In the past few decades, digital technologies have profoundly altered virtually every aspect of human life. While the direct impact of digital technologies on individuals' economic welfare or personal behaviors has attracted considerable attention, the interplay of digital technologies with social connections remains underexplored. Indeed, regardless of whether formed offline or online, social connections in the form of personal ties and affiliations that have long been the bedrock of human society continue to shape human behaviors and outcomes. To the extent that digitization will only continue to grow in scale and scope, an understanding of such effects is important for scholars, practitioners, and policymakers. I address two overarching research questions in my dissertation: (1) Whether, and to what extent digital technologies affect individuals' economic welfare and habituated behavior, and (2) How social

connections such as personal ties and affiliations condition the impact of the digital technologies.

My studies are conducted in two distinct contexts: mobile interventions for health, and computer ownership for social and economic welfare. Drawing on diverse bodies of literature and using various econometric methods, I seek to answer questions related to how interventions orchestrated on mobile platforms help individuals form healthy behaviors, and how computer ownership affects long-term income mobility. In the first essay, I show that a social norms intervention on a mobile platform is effective in increasing individuals' physical activity. In the second study, I investigate how the motivational incentive of reciprocity can be leveraged to promote healthy behavior. Finally, in my third essay, I show that computer ownership generates both private and social returns (IT spillovers) on individuals' income mobility. All three papers then consider how individuals' social connections condition the direct effects of digital technologies. The first two studies explore how online social ties and social relationships moderate the impact of mobile interventions, and the third study examines how caste groups affect the positive spillover effects of computer ownership. Collectively, the three studies advance our understanding of the heterogeneous effects of digital technologies on individuals and provide implications for researchers and practitioners.

THE INTERPLAY BETWEEN SOCIAL CONNECTIONS AND DIGITAL TECHNOLOGIES: THREE ESSAYS EXAMINING HEALTHY BEHAVIORS AND INCOME MOBILITY

by

Che-Wei Liu

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2018

Advisory Committee:
Dr. Ritu Agarwal, Co-Chair
Dr. Sunil Mithas, Co-Chair
Dr. Guodong Gao
Dr. Peng Huang
Dr. Reeve Vanneman, Dean's Representative

Acknowledgements

First and foremost, I would like to express my deepest appreciation to my committee chair, Professor Ritu Agarwal, for her continuous encouragement and support. Without her persistent guidance, this dissertation would not have been possible. I also want to thank my co-chair, Professor Sunil Mithas, for his contribution to my Ph.D. study and research. I will always remember his guidance and late night phone conversations about our joint work.

I would like to thank Professors Guodong (Gordon) Gao, Peng Huang, and Reeve Vanneman for serving on my dissertation committee. I am truly indebted to Professor Gao, who always pushed me to think more deeply and provided invaluable insights to lay a great foundation for my dissertation. I want to thank Professor Huang; the empirical skills I learned from him will be invaluable for my career in the long-term. I am also grateful to Professor Vanneman. In particular, his comments and feedback on Essay 3 have greatly enhanced the quality of that chapter.

I would like to thank Professors Il-Horn Hann, Siva Viswanathan, Hank Lucas, and Kislaya Prasad for helping me explore research ideas at the initial stage of my Ph.D. journey. My senior, peer, and the junior Ph.D. students in the IS department: Yang Pan, Dongwon Lee, Tianshu Sun, Aishwarya Deep Shukla, Salman Aljazzaf, Lanfei Shi, Weiguang Wang, Raveesh Mayya, and Sabari Rajan, provided guidance and encouragement throughout the Ph.D. journey. I would like to express my gratitude also to Justina Blanco and Melissa Formby in the Ph.D. office for patiently answering all my questions, both substantive and trivial! Finally, I would like to thank my family

for their unconditional love and endless support. I am especially grateful to my wife, Qian, who supported me in every possible way during the completion of my dissertation.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and Overview

Digital technologies have become deeply embedded in our work and personal lives and have led to widespread changes in outcomes and behaviors. My research centers on two contexts where digital technologies play a critical role: the effects of the mobile interventions on physical activity and the effects of computer ownership on income mobility. Although the direct effect of digital technologies on individuals' economic welfare or healthy behavior has attracted considerable attention (Dolton and Makepeace 2004; Free et al. 2013), how digital technologies interplay with one's social connections remain underexplored. In many respects, humans are not solitary actors, but rather exist within interconnected groups. In the three studies of my dissertation, I explore the following two questions: (1) Whether, and to what extent digital technologies affect individuals' economic welfare and habituated behavior, and (2) How social connections such as personal ties and affiliations condition the impact of the digital technologies.

Collectively, the three studies draw on a large body of literature on social connections from a wide range of disciplines. The network approach originated in graph theory and has been widely used in the social sciences and psychology. A node in a social network could be individuals, groups, organizations, or even societies (Wasserman and Faust 1994) and a tie represents relationships among connected actors. Network theories argue that people's behaviors are not best predicted by individual attributes or demographic characteristics, but rather by the social networks in which they are embedded (Wellman 1997). They further suggest that the focus of social network analyses should be on the relationships (ties) among actors, rather than the

nodes' inner characteristics (Katz et al. 2004). My research does not address the phenomenon of how people create or maintain network ties, which have its roots in theories such as self-interest theory (Coleman 1988) or social exchange theory (Emerson 1972). Rather, I focus on what happens after a social network already exists – i.e., how social connections such as tie strength (Granovetter 1973) and embeddedness (Granovetter 1985) affect individuals' behaviors. With the proliferation of social network platforms, researchers now have an unprecedented opportunity to directly observe social relationships among actors and have started exploring the effects of digital connections in various contexts (Aral and Walker 2014; Bapna and Umyarov 2015).

In social networks, ties can assume a variety of forms, including proximity ties (e.g., same location), memberships ties (e.g., same clubs), kinship ties (e.g., siblings), affective ties (e.g., who likes whom), or cognitive ties (e.g., who knows whom) (Borgatti et al. 2009). Regarding operationalization of network ties, consistent with prevailing views in the literature, I use three types of social connections in my studies. First, the degree of social ties is probably one of the simplest measures of social connections and refers to the total number of friends one has. The volume of one's friends affects how an individual accesses directly connected resources or the extensity of being connected. Social ties have been linked to various outcomes such as video sharing (Yoganarasimhan 2012), musical service adoption (Bapna and Umyarov 2015), and successful fundings in P2P lending (Lin et al. 2013). Second, the level of closeness between a pair of users has been associated with social influence. Studies have examined how the embeddedness (number of common friends) moderates social

influence in product adoption (Aral and Walker 2014) and affects the level of trust in an investment game (Bapna et al. 2017). Third, social affiliations such as ethnicity or caste affect one's social capital, and in turn, affect outcomes. Individuals within the same group usually have a strong bond and relatively frictionless communications, and the ethnic group has shown to generate higher social capital that affects the survival rate of an entrepreneur (Kalnins and Chung 2006). Each of my studies discusses one or multiple types of social connections, and examines how social connections deepen our understanding of mobile interventions and computer ownership.

For my first essay, I study the conditional influence of social ties to explore heterogeneous treatment effects of a mobile intervention that aims at motivating individuals to exercise more. My intervention is comprised of a social norms message that seeks to motivate users' goal setting and goal attainment behaviors related to physical activity on a mobile platform. The purpose of my randomized field experiment is to test whether social norms information about what others in the community are doing induces more people to set a monthly goal of running a self-determined distance. To construct the social ties, I use the degree centrality as the number of social connections that have been extensively used in previous studies (Grewal et al. 2006; Hinz et al. 2015). In my examination of the effect of social norms on users' goal setting and goal attainment behaviors, I find a significant moderating role for social connectivity: individuals with higher levels of social connectivity are more susceptible to a social norms message. Further analysis reveals that individuals who have many followers (i.e., high in-degree) but do not follow many others (low out-degree) are the most susceptible to the social norms treatment. Strikingly, I find that social norms also

lead to a substantially lower rate of goal attainment as compared with the control message; this adverse effect is also heterogeneously experienced, conditional on the number of social ties. These findings have important implications for the safe design of interventions based on social norms.

My second essay investigates how the motivational incentive of reciprocity can be leveraged to promote healthy behavior. The maturity of wearable devices and smartphones has afforded us the ability to precisely track individuals' physical activity at relatively low cost. The increased observability, coupled with digital social networks that enable connections at an unprecedented scale, offer a unique opportunity to explore the power of reciprocity. I find that reciprocity-based incentive outperforms a self-interest based incentive to motivate individuals to exercise more. Compared to existing self-interest based incentives, this finding opens a new avenue for mechanism design in promoting healthy behavior. Moreover, I find that the magnitude of the reciprocity effect is contingent on the closeness between the senders and the receivers. Surprisingly, the closeness has an inverted U-shaped influence on the reciprocity effect. The effect is the strongest when the closeness is moderate and wanes when the closeness is either too strong or too weak. These findings suggest that reciprocity can be a potentially powerful mechanism for improving the effectiveness of mHealth interventions.

The final essay in my dissertation examines the effects of computer ownership and IT spillover effects on income mobility, which is an important metric to assess the extent of equality of opportunities for individuals to progress along economic and social ladders. I use two waves of nationally representative surveys from India between

2005 and 2011 (Desai and Vanneman 2016; Desai et al. 2016), and show that computer ownership is associated with a household's income mobility and that computer ownership in the same district generates positive spillover effects. Every 1% increase of owning a computer in a district leads to an increase of 0.11% in upward income mobility. I then leverage the caste system in India to further assess the existence of the spillover effects of computer ownership. The caste system is a social stratification system in India that is deeply entrenched in the fabric of society, it is responsible for the existence of unique personal affiliations, and there is an increasing trend to view caste as a form of ethnicity in India (Desai and Dubey 2012). Thus, I argue that the IT spillover effect would exist when individuals closely interact with each other. I show that higher computer ownership in a region from the same caste group is associated with an increase in upward mobility, but not from higher computer ownership of distant caste groups. These findings underscore the need to consider social interactions for a complete understanding of returns on digitization.

To conclude, my research not only focuses on the direct effect of digital technologies but also tests the conditioning effects of different facets of social connections. I examine whether the effects of a mobile intervention for physical activity goal setting would be conditional on the social connectivity, and whether reciprocity can foster healthy behavior and whether it is conditional on the social closeness between a sender and a receiver. I further use a social stratification system to validate the spillover effects of computer ownership. Together, these three studies provide a nuanced understanding of heterogeneous effects of digital technologies on different

groups of people depending on their social ties, and make theoretical contributions to the literature on the effects of IT.

Chapter 2: Unraveling the "Social" in Social Norms: The Conditioning Effect of User Connectivity

Abstract

Abundant empirical evidence supports the overall efficacy of social norms as a strategy to induce behavior change. However, very few studies examine how the effects of social norms are differentially manifest across individuals, especially in today's socially connected digital world. We conjecture that the effects of social norms are conditional on an individual's digital social ties, and we provide new empirical evidence from a randomized field experiment that included more than 7,000 individuals on an online physical activity community observed for a two-month period. In our examination of the effect of social norms on users' goal setting and goal attainment behaviors we find a significant moderating role for social connectivity: individuals with higher levels of social connectivity are more susceptible to a social norms message. Further analysis reveals that individuals who have many followers (i.e., high in-degree) but do not follow many others (low out-degree) are the most susceptible to the social norms treatment. Strikingly, we find that social norms also lead to a substantially lower rate of goal attainment as compared with the control message; this adverse effect is also heterogeneously experienced, conditional on the number of social ties. Our findings have important implications for the design of interventions based on social norms.

Keywords*: Social norms; social connections; goal setting; goal attainment; goal setting theory; heterogeneous treatment effect*

2.1. Introduction

Researchers and practitioners have long sought to exploit the power of interventions to "nudge" behavior change in many domains that have important societal implications. One such popular intervention is based on disseminating social norms, i.e., conveying information about behaviors that are widely prevalent among others. These social norms interventions, operating on the basic principle of "telling people about what lots of other people do" (Burchell et al. 2013, p. 1), have demonstrated efficacy in inducing behavior change in diverse traditional settings including towel reuse programs (Goldstein et al. 2008), energy consumption (Schultz et al. 2007),

excessive drinking (Wechsler et al. 2003), healthy eating, and weight loss (Bittner and Kulesz 2015; Napolitano et al. 2013).

While most studies on the efficacy of social norms have been conducted in traditional settings, in recent years, digital technologies have profoundly altered how one connects to society. Individuals are now embedded in vast virtual networks of friends, family, co-workers, and acquaintances enabled by digital social platforms, where they may "friend" someone, follow them, and be followed. Indeed, the prevalence of connectivity is a defining feature of modern social media platforms. A robust stream of research has shown that ubiquitous connectivity has significant impacts on individual actions and behavior (Aral and Walker 2014; Bapna and Umyarov 2015; Dewan et al. 2017).

This unprecedented digital connectivity has given rise to novel and unexplored questions about the effects of social norms. Since social norms are rooted in an individual's perception of what others do and in psychological theories of conformity (Burchell et al. 2013), how one is connected to others may play a significant role in their impact. In traditional settings such as towel reuse, energy consumption, or alcohol consumption, the behavior of "others" reflected in the norms is often not directly observable. By contrast, digital platforms today make it virtually effortless for users to learn about each other's actions, such as liking a campaign, writing a review, or posting a picture. In addition, the degree of connectivity, as well as the types of ties (whether they are one-directional, such as in-degree or out-degree, or reciprocated) can vary significantly among users of digital platforms, potentially complicating a user's susceptibility to the social norms. To the extent that individuals in the social network

care about the impressions they create (Tong et al. 2008), observability by others is likely to affect behavior (Exley and Naecker 2016; Kast et al. 2012). To our knowledge, the question of whether social norms exert differential influence across individuals in a social network has not been investigated.

We address this knowledge gap by examining the moderating role of digital social connections on the effects of social norms. We further explore whether the differential impact of the social norms can be related to the types of connections[1] that exist within a digital social network. Finally, given the growing use of social norms in nudging individuals' behavior, it is imperative to ensure their safe use and protect against unintended outcomes. We therefore examine whether the high connectivity might lead to an overly potent social norms effect and yield undesirable consequences.

Our empirical setting is the use of a social norms-based behavioral intervention to motivate goal setting for physical activity on a digital platform. We conducted a randomized field experiment in collaboration with one of the largest online physical activity communities in Taiwan. Like most digital platforms, this platform provides connectivity as a standard feature: a user can befriend other users and follow their posts and status updates. In the experiment, we sent over 7,000 users messages that encouraged them to set a self-determined monthly running distance goal. We randomized users to receive one of two message types: a *Control* message with standard language that refers to the benefits of setting goals, or a *Treatment* message that augments the control message with social norms by indicating the number of users

[1] We thank the AE and the reviewers for this insight.

in this community who set a goal in the pretreatment month.[2] In addition to the action of goal setting, the platform allows us to capture the actual running distance, i.e., the extent of goal attainment, over a two-month period. This important feature enables us to assess the degree to which the social norms result in miscalibration by the user, i.e., if runners over-shoot by setting a goal that exceeds their capabilities.

Our experiment yields three main findings. First, we find that connectivity plays an important role in moderating the effectiveness of social norms. For users who are not highly socially connected, the social norms message leads to a modest increase in the goal setting rate (1%) as compared to the control message. However, users who are highly socially connected respond significantly more vigorously to the social norms message, with a 7.7% increase in the goal setting rate as compared to the control message. Second, our analysis reveals that individuals who have many followers (i.e., high in-degree) but do not follow many others (low out-degree) are the most susceptible to the social norms treatment; however, the same results do not apply to individuals who elect to establish unreciprocated ties to many others but do not have many followers themselves. This result is consistent with our theoretical arguments that individuals who care about impression management are the most susceptible because their visibility is high. Third, we find that social norms also lead to a substantially lower rate of goal attainment than the control message (-17.0%). This effect is also heterogeneous among individuals. When a user is not highly socially connected, the social norms and control messages produce similar effects on goal attainment rate

[2] The study is part of a broader set of experiments in which we examined other treatments. Users in the different studies do not overlap.

(37.3% and 40.2%, respectively). For a highly socially connected user, the goal attainment rates are sharply different (32.1% for the social norms condition and 69.5% for the control). Further analysis shows that the highly socially connected individuals who are motivated by social norms to set goals have a propensity to set goals beyond their capabilities. In other words, social norms exert "too strong" an effect on these users, leading them to set unrealistic goals that are difficult to attain.

Our study makes useful theoretical and policy contributions. First, our work bridges the literature between social norms and social networks, adding to both streams of research. Despite the fact that the efficacy of social norms for behavior change has been established in varied contexts, most of the existing studies focus on the overall effects. As a result, we know surprisingly little about differences in individual susceptibility to its influence[3]. Our work represents the first exploration of the heterogeneous treatment effects of social norm interventions as conditioned by a user's social network characteristics, thereby extending our understanding of how social norms work.

Second, by drawing attention to the salience of digital connectivity for the effects of social norms, we offer a subtle insight for research on social networks. Following the work of Freeman (1978), it has been well known that a highly connected "central" user can influence the behavior of others, and empirical studies have established that opinion leaders can influence the actions and behaviors of others connected to them (Hinz et al. 2011; Iyengar et al. 2011). In this study, rather than the

[3] Some prior work in traditional settings also points to the existence of variation in the effects of social norms (Allcott 2011; Ferraro and Miranda 2013). However, these studies focus on individual demographic factors such as wealth or past behaviors.

influence of a highly connected user on others, we explore whether high social connectivity might also passively shape the focal users' decisions. We enrich this stream of research by showing that a user's connectivity has implications beyond the direct influence exerted on their friends; since people care that others observe their actions and care about the impressions others have of them, connectivity is also correlated with individuals' own decisions.

Finally, from a practice and policy perspective, our study has important implications for the effective and safe use of social norms in nudging people's behavior. Our finding that social norms do not affect all users equally should help optimize interventions towards those with high susceptibility based on easily observable measures (in our case, a user's connectivity on digital platforms). More importantly, the specific behavior we study, increasing physical activity, is highly consequential for individual and public health. Our finding that users in certain subgroups tend to "overshoot" in their goal setting is especially important, as it may help prevent harm to patients in health programs (Kent and Hayward 2007).

2.2. Background and Prior Literature

The conceptual background for our study is informed by three streams of literature, which we discuss below. First, we provide an overview of prior research on social norms and their efficacy in motivating diverse behaviors. Second, we present goal setting as a mechanism that can potentially help overcome the persistent human failing of self-control, which is frequently implicated in the existence of unhealthy behavior, by allowing individuals to self-regulate. Here we also discuss the possible consequences of failure to achieve a goal. Finally, we turn to variation in the effects of

social norms, explaining why in digitally connected communities the effects of social norms on goal setting are likely to be amplified for individuals with more social ties, and how the nature of ties (in-degree vs. out-degree) plays an important role.

2.2.1 Social Norms

Social norms encapsulate rules of behavior considered acceptable within a group. Mackie et al. (2012) defined social norms as "what people in some group believe to be *normal* in the group, that is, believed to be a typical action, an appropriate action or both" (p. 7). Theoretically, it has been argued that people conform to a norm in an attempt to enhance affiliation with the referent social group and become "liked" (Higgs 2015). As noted, substantial evidence supports the effectiveness of social norms for inducing specific behaviors. To illustrate, Gerber and Rogers (2009) examined the effect of social norms messages on voter turnout and found that descriptive social norms affect voting intention among citizens who vote less frequently. Goldstein et al. (2007) conducted two field experiments to examine the effect of a descriptive norm (i.e., the majority of guests reuse their towels) on an environmental conservation program. Their study revealed that signage containing normative appeals outperforms traditional appeals. Schultz et al. (2007) conducted a field experiment to examine the effect of a descriptive norm (doing what others do) combined with an injunctive norm (doing what others think one should do) on energy consumption and reported that an injunctive message eliminated the boomerang effect, where outperforming individuals regress to the average behavior. Chen et al. (2010) showed that social norms are useful to increase the number of online reviews of movies provided by users, while Burtch et al.'s (2017) results show that a combination of financial incentives and social norms is

the most effective way to increase the number and length of online reviews in an online clothing retailer setting.

In the specific context of healthy behavior, the focus of our study, limited prior research has examined the effects of social norms on inducing behavior change. Ball et al. (2010) conducted a survey-based study in which participants self-reported their physical activity and eating behaviors and the factors influencing them. Results supported the significance of social norms in inducing those healthy behavior, even after controlling for social support from the referent group. Lally et al. (2011) found that individuals indulge in unhealthy food habits because of misperceptions about peers' dietary behaviors and concluded that interventions to correct such a misperception have the potential to promote healthy diets via the social norms effect. Given the prevalence and power of social norms in traditional settings, it is reasonable to expect that individuals on a social network platform would be likewise affected by information communicating a norm that is prevalent among other users on the same platform.

It is important to note that such norms are distinct from other social mechanisms such as peer effects or homophily that focus on the influence or correlation among peers (Aral et al. 2009). While peer influence refers to the direct influence of one friend on another, homophily represents the tendency of similar individuals to be connected and make correlated choices (Dewan et al. 2017; Manski 1993). Unlike these two contexts that involve interactions among individuals, in our setting a third party – the social network platform – sends a message to users. We then observe how users in different groups respond to the corresponding message.

2.2.2 Social Norms and Goal Setting

Physical activity is a behavior where there is significant innate heterogeneity in individual capabilities based on a variety of biological factors, such as age and gender, and psychological factors, such as self-efficacy and self-motivation (Biddle and Mutrie 2007). In this context, setting goals for a physical activity regimen is a widely recommended motivational mechanism (Shilts et al. 2004). A self-imposed goal is considered an ideal motivational device because it allows individuals to calibrate their goals based on self-knowledge about their own capabilities. Goal setting has been investigated in empirical research spanning more than three decades and has been shown to be effective for performance enhancement (Locke and Latham 2002). It is an important means to address the persistent human failing of self-control and has been extensively studied in both psychology and behavioral economics. A goal serves as a reference point for performance, and failure to attain goals creates a psychological loss (Koch and Nafziger 2011). A goal can be useful even when not explicitly accompanied by rewards or punishments – individuals may use non-binding goals to self-regulate (Hsiaw 2013). Given the importance of establishing a goal as the first step in behavior change, it is arguably appropriate to communicate a social norm related to goals.

Shilts et al. (2004) conducted a detailed literature review on the effect of goal setting on physical activity and concluded that "moderate evidence indicates that implementing goal setting as a dietary or physical activity behavior change strategy is effective with adults" (p. 92). The study that is closest to our work is Ariely and Wertenbroch (2002), which reports results from a series of experiments on procrastination. Ariely and Wertenbroch found that people would self-impose a meaningful or costly deadline to overcome procrastination, but the self-imposed

deadline was not as effective as externally imposed ones. In contrast to their setting, our research focuses on first, whether an external message motivates a user to set a goal, and second, on a comparison of the goal induced by a general message relative to one induced by a social norms message.

While setting a goal has generally been associated with improvements in performance, research on the after-effects of goal failure has yielded inconclusive results. Goals may be counterproductive because failure leads to subsequent performance deterioration (Soman and Cheema 2004) and can be demotivating (Jones et al. 2013). Alternatively, individuals with autonomous motives may not simply give up, but may make reengagement plans to fulfill a new goal (Ntoumanis et al. 2014). In other words, people have different coping strategies when facing goal failure (Bittner and Zondervan 2015).

2.2.3 Social Connections and the Effects of Social Norms

As discussed previously, our review of the existing literature indicated strong evidence to support the "average" effect of social norms on individuals' behavior modification. However, to the extent that individuals possess heterogeneous characteristics, the expectation that social norms would yield identical effects for every individual has been challenged (Dorresteijn et al. 2011). Only a limited number of experimental studies examine heterogeneous treatment effects in social norms campaigns. In a campaign for energy conservation, Allcott (2011) finds a heterogeneous treatment effect across households. A household's pre-treatment level of energy consumption affects the magnitude of the treatment effect: the response to the social norms was amplified for the heavy electricity users. Costa and Kahn (2013)

find that households' political and environmental ideologies affect the response to social norms campaigns. Liberals tend to reduce energy consumption more than conservatives when exposed to a social norm. Ferraro and Miranda's (2013) results in the context of a water conservation program indicate that households' characteristics such as wealth and past usage behaviors yield heterogeneous treatment effects.

As we argued previously, dense digital connectivity has raised new questions about the heterogeneous effects of social norms. The social norms effect represents how our behaviors are shaped by how we think others behave (Burchell et al. 2013). A defining characteristic of digital social networks is the high level of observability and visibility across connections. When users change behavior to conform to a social norm communicated to them, this behavior is transparent to others in their network. Thus, individuals with different levels of visibility might react differently to the social norms based on the size of their audience. Studies have indeed suggested that observability can cause behavior change (Exley and Naecker 2016). For instance, Kast et al. (2012) showed that individuals' precautionary savings can be effectively increased by simply announcing personal saving goals to peers. Thus, we propose that a key factor moderating the response to social norms is the level of an individual's social connectivity, especially when individuals are highly socially connected. To the best of our knowledge, none of the previous studies have investigated how social norms could be moderated by individuals' social connections, an important aspect that has been repeatedly shown to be a powerful factor in human behaviors (Aral and Walker 2014; Susarla et al. 2012; Venkatesh et al. 2016).

Why would individuals with high social connectivity be more susceptible to a social norms message? Research suggests that individuals on social network platforms frequently engage in impression management (Tong et al. 2008), and prior work notes that impression management is one of the most important motivations for individuals to actively participate in such platforms (Krämer and Winter 2008). A user would thus rationally engage in activities that can enhance a positive image and avoid behaviors that would be harmful to their impression. For example, a commonly used strategy to create an impression is updating profile elements (Lampe et al. 2007; Walther et al. 2008), which represents signaling to gain social status in the network. Social network participants have even been shown to engage in atypical behaviors such as using conspicuous consumption as a signal to foster social status (Hinz et al. 2015).

Setting a goal, an action that is transparent in our setting, represents a signal as it is not only a challenge to the self, but also an explicit, visible act of conformance to the norm. Individuals who receive a social norms message have a choice in regard to whether they follow or ignore the norm. To the extent that the behavior of setting a goal is a positive act, highly socially connected individuals stand to gain more for their image. Not conforming to an accepted norm might incur costs – a group member who does not follow a norm is likely to be labeled as a deviant and runs the risk of being isolated if s/he continues to violate the norms (Hackman 1992). The costs of damaging one's impression because of non-conformance to a norm are obviously higher for individuals with more social connections and, as a result, their response to the social norms is likely to be stronger compared to those with fewer social connections.

We further propose that the benefits of conforming to the norms are not identical across users. Since the directed link determines the flow of information, it represents an individual's influence on others. In a directional social network platform, an individual can follow (i.e., have outgoing social ties) or be followed (incoming social ties) by another individual. Using incoming and outgoing ties as a source of variation, research has identified three user categories on Twitter, a directional social network platform (Dann 2010; Java et al. 2007): users who engage in information sharing (high followers, low followees), those who are information seeking (low followers, high followees), and those in friendship relationships (roughly equivalent followers and followees).

Of these categories, which users are likely to be more responsive to a social norms message? There are potentially two competing arguments. On the one hand, information seekers (high out-degree only) represent "sociability" (Rice et al. 1999), implying that they tend to access others in the network. According to Liebowitz (2006), high out-degree individuals tend to seek others' guidance and advice. Thus, information seekers might have a greater tendency to learn from others and follow others' behaviors. Since social norms convey information about what most others do, the tendency of information seekers might make them more susceptible to the social norms treatment. On the other hand, because individuals' behaviors are more visible to the incoming ties than the outgoing ties, the gains from impression management are likely to be stronger for information sharers (high in-degree only). Individuals with a high number of followers as compared to followees have a sizeable audience and are likely to have celebrity status (Krishnamurthy et al. 2008) on the social network platform.

Such individuals who "share information" would clearly benefit more from conforming to the norms to avoid poor impression management. Thus, this type of user might be more susceptible to the social norms message. Both arguments are plausible, prompting us to examine this issue empirically.

To summarize the conceptual background for this study, robust evidence supports the efficacy of social norms in diverse settings, although the literature on the use of social norms for physical activity is relatively modest. We draw on prior research in goal setting to argue that a social norms message encapsulating information about how others have established a goal for physical activity may address the self-control challenge that prevents individuals from undertaking healthy behavior. Although prior research has alluded to heterogeneous treatment effects for social norms, empirical work in this area is limited, and the conditioning effect of today's pervasive digital social connections has not been examined. Theoretical arguments based on individuals' desires to engage in impression management indicate that highly socially connected individuals would benefit more from conforming to the norms. We further suggest that norm conformance is not likely to be identical across users and present competing arguments for why both information sharers (high followers, low followees) and information seekers (low followers, high followees) may be susceptible to the social norms treatment. In the following section, we describe our experimental design and empirical study to examine these effects.

2.3. Experiment

2.3.1 Research Context

To estimate the causal effect of social norms on individuals' goal setting and subsequent physical activity, we conducted a randomized field experiment following the methodology proposed by List and Rasul (2011). We observed individuals' running behaviors across a period of time in a controlled environment, and these individuals were unaware that they were under observation. This design enables us to eliminate the undesirable Hawthorne effect (Adair 1984), also referred to as the observer effect, in which individuals change their behavior in response to their awareness of being observed.

Our field experiment was effected in cooperation with an online running platform, henceforth referred to as *RunningPlatform* to maintain anonymity. The platform is currently a leading physical activity platform in Taiwan, with more than 250,000 registered users. The platform provides services for runners who seek to keep track of their running activities and connect with friends. Its social network features allow us to investigate the interaction between social norms and social connections. Users in this online community can upload their running activity either through the website or by using the mobile app provided by the company. Although users share a common interest in developing a physical activity habit, participation does not mean that they are veteran runners. Rather, most users on the platform still suffer from insufficient physical activity (e.g., only around 3.9% of registered users run more than 30km in the pre-treatment month), making them a suitable target population for inducing the action of setting a physical activity goal.

The majority of users rely on the platform's mobile app to track their physical activities. Combining modern GPS technology with Google Maps, the platform shows users their route in detail for each running episode. Furthermore, the platform allows users to monitor their running pace, enabling them to adjust their training plans accordingly. For each running episode, the platform records not only the time spent running, but also detailed personal information such as average heart rate and even environmental information such as temperature and humidity. In addition, the platform shares information about marathon races in Taiwan. Individuals can then indicate whether they would like to attend the races and share their race experience with friends.

This physical activity community offers standard social network-related functions. Similar to Twitter, individuals on this platform can follow or be followed by other runners. This means that an individual's friend can be one she/he follows (followee), one who follows her/him (follower), or the two users could be mutual followers. A follower can see the followee's updated physical activity if the followee makes their running records available to the public or his/her friends. Each individual has his/her own personal page, which is similar to the personal wall feature on Facebook. The personal wall contains several pieces of information, including individuals' profile, pictures, friends, timeline of physical activity, monthly accumulated running distances, and monthly goal. One can also easily follow strangers by clicking the "follow" button on their personal wall.

2.3.2 Experiment Design

Prior to the start of our experiment, a goal-setting feature was already available on the platform. A user can set a monthly running distance goal for a given month by

inputting a self-determined number (in kilometers) at any time in the month. Our study seeks to examine the effectiveness of social norms messaging on users' willingness to set a goal and on their subsequent behaviors after setting a goal. We randomly assign individuals to two groups who received distinct messages, as listed in Table 2.1.[4] The baseline message conveys information about the benefit of setting up a goal – a 17% increase in running frequency, i.e., the number of times run in a given month, which is calculated based on historical data from this platform. For the treatment group, we then add the social norms information to the baseline message (see Table 2.1). Consistent with Burtch et al. (2017), we use descriptive social norms of what other people on the platform do by displaying the absolute number of individuals setting a goal in the pre-treatment month (e.g., there were 5,223 individuals setting up a goal in January and 5,715 individuals in February).

We intentionally elected not to include a "pure" social norms message that merely conveys the number of individuals setting a goal without describing the benefit of setting a goal. This decision was motivated by the rationale that highlighting social norms without underscoring their benefits is not common or feasible in practice. For example, in the towel reuse program studied by Goldstein et al. (2008), messages not only stated that the majority of hotel guests reuse towels, but also underscored that reusing towels can save the environment: "… to help save the environment by reusing

[4] The randomization procedure between the social norms and control groups helps us alleviate any concerns related to peer influence processes in the social network platform. Since individuals are randomly assigned to the treatment and control groups, their social connections are virtually identical, which means that if a social influence effect exists, all participants be equally affected by friends who set a goal. Therefore, when calculating the average treatment effects and heterogeneous treatment effects between these two groups, the social influence is canceled out. The difference between the two groups is therefore attributable to the social norms effect.

your towels during your stay" (p. 474). Allcott (2011), in attempts to modify energy consumption behavior, not only mentioned users' average energy consumption to convey a social norm but also provided tips emphasizing the benefit of reducing energy consumption: "Things you can do right now … SAVE UP TO $40" (p. 1084). Therefore, we use a pure benefit message group as a control group, then add a social norms message on top of the benefit message for the treatment group, making the use of a social norms message more appropriate in practice.

We further note that our control message serves an appropriate control because it is identical to the message received by the treatment group in every respect except for the addition of the social norm information. Finally, we do not include a "no message" group as the baseline group because the reminder effect of a message has been extensively studied and is well-known (Fjeldsoe et al. 2009; Hurling et al. 2007).[5]

Our experiment occurred over a two-month period in February and March 2016. When individuals logged in to the online community for the first time in a treatment month, they received a pop-up message. In any treatment month, an individual received only one message regardless of the device she used (website or mobile). Both website and mobile app users were directed to a page displaying the message, shown in Table 2.1, and a box to input a monthly goal. We did not force users to set a goal, and a user could simply close the window or click a return button to leave this page without setting a goal.

[5] We conduct additional analysis to determine if our treatment outperforms a pseudo-baseline control group that does not receive a reminder message. We construct a control group from the same time period in the previous year to account for any seasonal effect; we then compare goal-setting rates from February and March 2015 to the experimental data. We find that the social norms message (treatment) has the strongest effect on goal setting, followed by the control group message, and both combined together are much stronger than the no-message pseudo-baseline group.

Starting at 00:00 a.m. on February 1, 2016, every individual who logged in to the system, either through the website or the mobile app, was randomly assigned to one of the two groups. The randomization process was based on the user's identification number (*user_id*). This user ID is generated randomly when a user registers with the platform and is independent of users' preferences and capabilities.

2.3.3 Data

Since we are interested in estimating the impact on goal setting, we need to exclude users who are already self-motivated to set a goal in the pre-treatment month. Further, we seek to understand whether setting a goal increases individuals' performance in terms of total monthly running distance. Thus, we retain only individuals who have running records in the pre-treatment month (to allow for a comparison of pre- and post-treatment performance). We sent messages to a total of 7,196 individuals. As detailed in Table 2.2, there were 3,682 users in the *Control* group and 3,514 users in the *Social Norms* groups. Each group accounts for roughly 50% of the total sample.

Figure 2.1 summarizes the number of individuals in each group who logged into the system daily. As noted, each individual received only one message per month, which was delivered at the time of her first login during the treatment month. To simplify the visual representation, we retain only 29 days from both February and March and merge these numbers into one graph (2016 was a leap year, with 29 days in February). We drop the last two days of March to keep the graph parsimonious because the number of individuals receiving messages in those two days is small enough to

neglect.[6] As can be seen in the top half of Figure 2.1, most users received their messages in the first few days of the months when they logged in to the system. The bottom half of Figure 2.1 presents the proportion of individuals in each group. We see that the proportions are quite stable across time, offering some initial evidence for the success of the randomization procedure.

We conduct a series of t-tests (Table 2.2) as a randomization check by comparing means between the two groups for various characteristics, including age, gender, distance in the pre-treatment month, registered by Facebook account, log of the number of friends, log of the number of comments, log of the number of races attended, log of tenure from registration, and whether the message was read on a weekend or not. Users provide their age and gender during the registration step when they first use this platform. The platform tracks users' running activities across time, enabling us to capture running distance before and during the treatment month. When a user registers for an account, they can choose to use "Facebook login" to simplify the registration procedure, and we calculate the proportion of users who did so. We calculate the number of friends by summing the number of a user's followers and the number of followees (all-degree). Similar to Facebook functionality, users can leave comments on each other's running posts. Therefore, we can calculate the number of comments. We further calculate each user's tenure by counting the number of days elapsed from their registration date. The platform has also tracked all the major marathon races in Taiwan since it was first launched in 2011. Individuals who attend a race can indicate their participation, and we calculate the number of races users attended. We take log forms

[6] We drop 41 observations, which accounts for only 0.57% of the total sample, to draw the graph.

of these variables to reduce concerns regarding skewness. Finally, we calculate the proportion of users who read messages during a weekend.[7]

T-test results in Table 2.2 for these characteristics provide strong evidence of comparability and balance across the *Control* and *Social Norms* groups, affirming that the randomization process is successful. Among all the characteristics, only "% Registered by Facebook account" is modestly higher in the *Social Norms* group than in the *Control* group, but the difference (2.5%) is not practically significant. When comparing so many characteristics between two groups, one minor difference is not unusual (see Bloom et al. 2014 for more discussion). Furthermore, a joint test on all characteristics shows no significant difference. The F-test for all variables in the randomization yields a p-value of 0.20.

Table 2.3 shows the correlation of all the variables. Log(# of friends) and Log(# of comments) are highly correlated because both variables capture some form of social connection. Therefore, to avoid the multicollinearity problem, we do not include both variables in the formal models. Instead, we use Log(# of friends) in the main model and use Log(# of comments) to replace Log(# of friends) in the additional analysis to further demonstrate the robustness of our findings.

2.4. Results

2.4.1 Model-free Evidence

In Figure 2.2, we first visually depict the model-free evidence of the average treatment effect (ATE) on Goal Setting Rate and Goal Attainment Rate. In the *Control* group, 7.6% of users set a goal, while in the *Social Norms* group, 9.4% of users set a

[7] We later explain in more detail how we use this variable to address the selection problem.

goal, indicating a 23.7% increase in goal setting (t-test, $p < 0.01$). This result is consistent with the literature supporting the efficacy of social norms campaigns: individuals conform to the behavior of the majority in their referent group to set a goal. For Goal Attainment Rate, it is interesting to note that we find just the opposite. While individuals in the *Control* group have a moderate attainment rate (43.5%), users in the *Social Norms* group have a substantially lower success rate in reaching their self-determined goal (36.1%). This constitutes a drop of 17% in goal attainment. However, the effect is only weakly supported by the statistical test (t-test, $p < 0.1$). One plausible explanation for the weak result, which we investigate further below, may be that individuals are not uniformly affected by the social norms message.

Earlier, we argued that the effects of social norms are likely to be conditioned by social connections. To examine this heterogeneity in treatment effects (HTE), we use the log of "# of friends" as an indicator of social connections, representing the sum of followers and followees. To quantify HTE, we follow Green and Kern (2012)'s approach and estimate average treatment effects among subgroups along with a continuous covariate. Ten subgroups are constructed based on the continuous covariate of social connections. A higher number indicates individuals in that group have more friends. For example, after taking the exponential function of the log of number of friends, individuals in subgroup 1 only have 1.1 friends, while those in subgroup 10 have on average 312.1 friends. Figure 2.3 shows raw differences between the *Control* group and *Social Norms* group for each subgroup. Here we see that the treatment effects of both Goal Setting Rate and Goal Attainment Rate are strongest for subgroup 10,

providing initial support for the conjecture that social connections moderate the treatment effects, especially when individuals are highly connected.

2.4.2 Regression Analyses

2.4.2.1 Regression Analyses for the Average Treatment Effects

Since the outcome variables of goal setting and goal attainment are dichotomous (setting a goal or not; attaining a goal or not), we use a Probit model to examine the treatment effect (Wooldridge 2010). Specifically, we consider the following model in the latent variable form:

$$Goal\ Setting_i^* = \alpha_1 + \beta_{11} * T_i + \beta_{12} * HS_i + \sum_{j=3}^{k} \beta_{1j*} Z_i + \varepsilon_{1i} \quad (1)$$
$$Goal\ Setting_i = 1\ [Goal\ Setting_i^* > 0]$$
$$Goal\ Attainment_i^* = \alpha_2 + \beta_{21} * T_i + \beta_{22} * HS_i + \sum_{j=3}^{k} \beta_{2j*} Z_i + \varepsilon_{2i} \quad (2)$$
$$Goal\ Attainment_i = 1\ [Goal\ Attainment_i^* > 0]$$

where *T* represents the treatment dummy, which equals 1 when the group is the *Social Norms* group and equals 0 when the group is the Control group. We also construct the variable *high social connection* (*HS*) to capture individuals' social connections. We define high social connection as those individuals whose number of friends places them in the top 10%, which is motivated by the initial findings of the model-free evidence in Figure 2.3.[8] We also control for a wide range of individual characteristics (*Z*), including age, gender, average daily distance in pre-treatment month, whether the account was

[8] To provide robustness, we used the following two alternative variables to construct social connections. First, we use the number of comments to construct another facet of social connections. The results are broadly similar to the results of using the number of friends (all-degree). Second, we construct a new social connections variable that only counts the number of mutual followers, since mutual following represents close relationships. The results are also similar to the results of using the number of friends (all-degree).

registered by using Facebook ID, log of the number of races, log of tenure since registration, and monthly dummy.

Regression results are in Table 2.4. Consistent with the model-free evidence, we find that the social norms treatment has a strong effect on the Goal Setting Rate. Table 2.4 Column (1) shows a positive coefficient of 0.115 ($p<0.01$) of the social norms treatment. The marginal effect is a 17.5% difference between the treatment and control groups. Regression results for the second outcome, Goal Attainment Rate, are shown in Table 2.4 Column (3). As with Goal Setting Rate, we see a difference between the *Social Norms* group and the *Control* group: the coefficient of the treatment is -0.179 and is only significant at the $p<0.1$ level. The marginal analysis shows a 6.6% drop of Goal Attainment Rate in the *Social Norms* group as compared to the rate in the *Control* group.

Overall, our results support the existence of a strong effect of social norms in motivating goal-setting behavior. More importantly, these results highlight the need to observe and pay attention to a subsequent effect. In this situation, an individual might elect to conform to the behavior of others because of normative pressures to do so; however, such conformity may then be accompanied by an adverse side effect, reflected in the lower Goal Attainment Rate. We discuss this finding further in section 2.4.5.

2.4.2.2 Quantile Treatment Effects of Social Connections

To better estimate the HTE across the spectrum of social connections, we follow Allcott (2011) to examine the Quantile Treatment Effects (QTEs). Similar to the approach in the model-free analysis, we use the log of "number of friends" (all-degree)

as an indicator of social connections and construct ten subgroups by the level of social connections. The QTE then estimates the differences between the *Social Norms* and *Control* groups in each of the ten corresponding quantiles. Figure 2.4 and Table A1 show the average treatment effects on each subpopulation for both Goal Setting and Goal Attainment outcomes. The number on the x-axis indicates each subgroup, with higher numbers reflecting a higher log of number of friends. We regress in each subgroup to calculate the conditional average treatment effects (CATE) using equations (1) and (2) by excluding the variable high social connection because this variable is directly related to the log of number of friends. Therefore, the y-axis in Figure 2.4 indicates the conditional average treatment effect. As can be seen, the treatment effect of Goal Setting is especially strong for subgroup 10, which has the highest number of friends. The CATE of Goal Setting in subgroup 10 is 8.3%, which is almost twice as large as the maximum of the conditional treatment effect for other subgroups.

The CATE of Goal Attainment in subgroup 10 is also strikingly strong. On average, individuals in the treatment condition of subgroup 10 have a 33.7% lower Goal Attainment Rate than individuals in the control condition. These plots vividly illustrate a sharp rise in Goal Setting for subgroup 10 and a steep decline in Goal Attainment for the same subgroup. Together, these two findings indicate that individuals with strong social connections are more susceptible to the social norms message and perform worse after receiving the treatment message. To provide further robustness, we follow the procedure of Feller and Holmes (2009) to calculate CATE by using generalized additive models. Figure A1 in the Appendix shows a broadly

similar finding, which provides additional confidence for the presence of HTE caused by social connections.

We note that although subgroup 10 only represents 10% of the sample, the low Goal Attainment Rate for this group might represent an economically substantial impact in other social network platforms. We have argued that such highly socially connected individuals are more susceptible to social norms because they are more visible to others. Individuals in subgroup 10 have an average of 312 friends in our target platform, but other popular social network platforms report levels of connectivity that easily exceed that of subgroup 10. For instance, individuals on Facebook in 2014 had an average of 350 Facebook friends in the United States[9], and active Twitter users have 707 followers on average[10]. With the growing numbers of online friends and the high penetration rates of the major social network platforms, the impact of the social norms treatment might be comparatively modest in a specialized setting like the *RunningPlatform* and greater in other settings such as Facebook or Twitter. Furthermore, individuals in subgroup 10 represent the most socially active group on the platform. If a social norms treatment negatively affects individuals in subgroup 10 and prompts them to reduce their engagement with the platform, then the treatment might adversely affect individuals' activities in other subgroups as well.

2.4.2.3 Regression Analyses for the Heterogeneous Treatment Effects

---

[9] Statista: Average number of Facebook friends of users in the United States as of February 2014. https://www.statista.com/statistics/232499/americans-who-use-social-networking-sites-several-times-per-day/

[10] KickFactory: The Average Twitter User Now has 707 Followers. https://kickfactory.com/blog/average-twitter-followers-updated-2016/

As with ATE, we use regression analyses for HTE to formally test our conjecture that individuals who are more socially connected are more susceptible to the social norms message. Specifically, consider the following model in the latent variable form:

$$Goal\ Setting_i^* = \alpha_1 + \beta_{11} * T_i + \beta_{12} * HS_i + \beta_{13} * T_i * HS_i + \sum_{j=4}^{k} \beta_{1j*} Z_i + \varepsilon_{1i} \quad (3)$$

$$Goal\ Setting_i = 1\ [Goal\ Setting_i^* > 0]$$

$$Goal\ Attainment_i^* = \alpha_2 + \beta_{21} * T_i + \beta_{22} * HS_i + \beta_{23} * T_i * HS_i + \sum_{j=4}^{k} \beta_{2j*} Z_i + \varepsilon_{2i} \quad (4)$$

$$Goal\ Attainment_i = 1\ [Goal\ Attainment_i^* > 0]$$

where all the variables are exactly the same as shown in equations (1) and (2). We add the moderating effect of high social connection (*HS*) on social norms (*T*) by interacting both terms. Following Baron and Kenny (1986) and Dawson (2014), when testing the moderation, the interaction term is the core element. If the term is significant, we show that the variable (*HS*) is a statistically significant moderator between the independent variable (*T*) and the outcome variable.

As shown in Table 2.4 Column (2), the interaction term's coefficient is 0.381 and significant at $p<0.01$ level. The positive sign of the interaction term in Table 2.4 Column (2) indicates a complementarity effect between the *Social Norms* intervention and *high social connection*: individuals who have high social connectivity have a greater chance of setting a goal when receiving a *Social Norms* message. After calculating the predictive margins, we find that for individuals who are not highly socially connected, the probability of setting a goal between the *Control* and the *Social*

Norms groups is very close (8.7% (SN) - 7.7% (Control) = 1.0%), but for individuals who are highly socially connected, the probability of setting a goal between the *Control* and the *Social Norms* groups is quite different (14.3% (SN) - 6.6% (Control) = 7.7%). Therefore, the complementarity effect between the *Social Norms* intervention and *high social connection* leads to an overall increase of 7.7% - 1.0% = 6.7% in goal setting.

For the Goal Attainment Rate, Table 2.4 Column (4) shows a strong interaction effect between *high social connection* and the *Social Norms* message. The negative sign of the interaction term in Table 2.4 Column (4) indicates a substitution effect between *Social Norms* intervention and *high social connection*. This suggests that individuals who have high social connectivity have a lower chance of attaining a goal when receiving the *Social Norms* message. Similarly, after calculating the predictive margins, we find that for individuals who are not highly socially connected, the probability of attaining a goal between the *Control* and the *Social Norms* groups is very similar (37.7% (SN) - 40.2% (Control) = -2.6%). In contrast, for individuals who are highly socially connected, the probability of attaining a goal between the *Control* and the *Social Norms* groups is drastically different (32.1% (SN) - 69.5% (Control) = -37.4%). Therefore, the substitution effect between *Social Norms* intervention and *high social connection* leads to an overall decrease of 34.8% ((-37.4%) - (-2.6%)) in goal attainment.

2.4.3 Addressing Selection Concerns

Although random assignment of treatments is the ideal approach to identify the treatment effects, analyzing the goal attainment effect in equations (2) and (4) might suffer from potential endogeneity bias because the outcomes of both equations are

distorted by nonrandom sample attrition. Although our treatment was assigned randomly in the original sample, the randomization does not hold for goal attainment because the actual estimated sample suffers from a selection bias issue, which can result in attrition bias with unknown direction. To correct the selection issue, we estimate the treatment effects by using the Heckman selection model (Heckman 1976; Heckman 1979). Since the outcome variable (Goal Attainment) is binary, we use a probit selection model, also known as heckprobit model, to correct for the selection problem (Greene 2012; Van de Ven and Van Praag 1981). Specifically, we use the following equations[11]:

$$\begin{aligned}\boldsymbol{Goal\ Attainment_i^{outcome}} &= \boldsymbol{\alpha_1 + \beta_{11} * T_i + \beta_{12} * HS_i + \beta_{13} * T_i * HS_i} \\ &+ \sum_{j=4}^{k} \boldsymbol{\beta_{1j*} Z_i} \\ &+ \boldsymbol{\varepsilon_i}\end{aligned} \tag{5.1}$$

$$\begin{aligned}\boldsymbol{Goal\ Attainment_i^{s\square lect}} &= \boldsymbol{\alpha_2 + \beta_{21} * T_i + \beta_{22} * HS_i + \beta_{23} * T_i * HS_i} \\ &+ \sum_{j=4}^{k} \boldsymbol{\beta_{2j*} Z_i} \\ &+ \boldsymbol{\beta_{25} * I_i + \mu_i}\end{aligned} \tag{5.2}$$

$$\boldsymbol{corr(\varepsilon, \mu) = \rho}$$

The dependent variable is not always observed. Rather, goal attainment is observed only when individuals choose to set a goal. When $\rho \neq 0$, the standard probit model for the first equation yields biased results, and we use the heckprobit model to provide consistent and asymptotically efficient estimates for all the parameters. For the

[11] For simplicity, we only present the heckprobit model to correct Equation (4). Simply removing an interaction term in Equations (5) can be used to correct Equation (2). It is also worth noting that the *Goal Attainment*^{Select} in Equation (5.2) of the heckprobit model is the binary value determined by the Goal Setting variable. In other words, those runners who selected to set the goals made their goal attainment observable.

model to be well identified, we need to satisfy the exclusion restriction, which requires that at least one variable appears with a non-zero coefficient in the selection model but not in the outcome equation. We construct a variable, *Is weekend* (*I*), to indicate whether an individual reads the message during a weekend or not. Previous literature has documented that day of the week can affect mood (McFarlane et al. 1988; Ryan et al. 2010) and human behaviors such as diet (An 2016), sedentary behaviors (Marshall et al. 2015), or online behaviors (Leunga et al. 2016). The differences between a weekend and a weekday can even be observed in one's physiological condition, likely because of the chronic work stress experienced during weekdays (Schlotz et al. 2004). A weekend is typically characterized by a relaxed atmosphere that is associated with close relations and leisure (Rybczynski and Glancy 1992), and a weekday is often linked with routine jobs, fatigue, stress, and work pressure (Schlotz et al. 2004; van Hooff et al. 2007). These differences would affect goal-setting behaviors because establishing a goal involves the development of an action plan to motivate oneself toward a goal, which is very likely to be different under different environmental contexts. Therefore, we conjecture that reading the message on the weekend would affect the goal-setting behavior differently from reading the message on a weekday. Finally, because goal attainment is a relatively longer-term behavior, there is no legitimate reason to believe that the specific type of day when the message is read would affect goal attainment, making it a suitable variable to use in the selection model.

Table 2.5 shows the results of addressing the selection issue. While Columns (1) and (2) show the results of the main effects, Columns (3) and (4) display the moderating effect. As expected, individuals who receive messages on the weekend

have a different likelihood of setting a goal (Columns (1) and (3)). Columns (2) and (4) show that the main effect and the moderating effect are still significant, and with a very similar magnitude as in the main model, reducing the concern regarding selection attrition. In fact, the p-values (Prob > chi^2) at the bottom of Table 2.5 are insignificant. Since ρ denotes the correlation between the error terms of the selection and outcome equations, this means the two stage equations are not statistically correlated and the selection issue is less of a concern.

2.4.4 Types of High Social Connections

We had earlier proposed competing arguments for different types of social ties. Prior research (Dann 2010; Java et al. 2007) has shown that users who engage in information sharing (high followers, low followees) are very different from those who are information seeking (low followers, high followees). Based on our previous arguments, information sharers might benefit more from conforming to the norm in order to make a good impression because they have a sizeable audience, but information seekers might also follow the norms because of their innate tendency to obtain guidance and information from others. We thus use the nature of the ties to further examine these theoretical arguments.

We categorized highly socially connected users into the following three types: only high in-degree users (high followers), only high out-degree users (high followees), and both high in-degree and out-degree users. We use the same threshold of 90% to differentiate high and low in-degree and out-degree. We then combine the baseline type, the low socially connected users, to create a dummy variable representing these four types of users. We replicate the analyses in equations (3) and (5) by replacing the

newly constructed variable with the original high social connection variable. Table 2.6 shows the results. We find that individuals who have only high in-degree are more likely to set a goal and less likely to achieve the goal, but the same results cannot be observed for individuals with only high out-degree. The finding further buttresses the argument that individuals who care more about impression management are more susceptible to the norms. Information seekers' natural tendencies might not be strong enough to be affected by the social norms intervention. For individuals with both high in-degree and out-degree, as might be expected, the results are very similar to the main findings.

2.4.5 The Consequences of Susceptibility to Social Norms

Our results related to the heterogeneous effects of social norms for highly socially connected individuals reveal an interesting puzzle: why are individuals in the treatment group with high social connectivity less likely to attain goals? In this section, we explore plausible explanations for this puzzle. To gain deeper insights into the differences in user behavior between the treatment and control groups, we first construct two identical samples by applying a matching procedure.[12] We use a simple matching algorithm on observable variables, including age, gender, distance in the pre-treatment month, number of friends, registered by FB account, number of comments, and number of races attended, to match individuals in the *Control* group with individuals in the *Social Norms* group. For each pair, we calculate the Mahalanobis distance between individuals in the *Control* group and individuals in the *Social Norms* group. The Mahalanobis distance is one of the most commonly used distance measures

[12] Our matching procedure is very similar to the procedure described in Dewan et al. (2017).

(De Maesschalck et al. 2000) and has been extensively used for matching (Leuven and Sianesi 2015; Rubin 1980). We iterate the matrix to find the smallest set of the sum of Mahalanobis distances between individuals in the *Control* group and individuals in the *Social Norms* group. We employ a one-to-one match between these two groups and find the two sets of individuals having the smallest sum of the Mahalanobis distance.

All 278 individuals in the *Control* group who set a goal remain labeled as the *Control* group; the matching procedure helps us isolate two subgroups in the *Social Norms* group because there are more individuals in the *Social Norms* group setting a goal. The first subgroup in the *Social Norms* group is called the *Matched* subgroup. This subgroup, as a result of the matching procedure, has exactly the same number of individuals as the *Control* group (278) and shares very similar attributes.[13] The second subgroup in the *Social Norms* group is the *Social Norms motivated* subgroup. These 52 individuals represent the incremental set of users who responded to the social norms message.[14] We estimated probit models to examine the effect of different groups on various outcomes:

$$Y_i^* = \alpha + \beta_1 * G_i + \sum_{j=2}^{k} \beta_{j*} Z_i + \varepsilon_i \qquad (6)$$
$$Y_i = 1\,[Y_i^* > 0]$$

where Y represents different dependent variables, including Goal Attainment, Log (# of friends), Goal minus Distance in Previous Month, and Run Next Month.[15] "Goal

[13] We confirm the balance between the *Control* group and the *Matched* subgroup by using a series of t-tests to ensure the quality of our matching. There are no statistically significant differences for all the covariates we used for matching. Further, there are no statistical differences for Goal Attainment Rate, average daily goal, and average daily distance in the pre-treatment month.

[14] We replicated the matching using a non-parametric procedure, "Coarsened Extract Matching (CEM)" (Blackwell et al. 2009). The results of CEM are broadly similar to the main findings.

[15] For outcome variables that are continuous, we use an OLS model for estimation.

minus Distance in Previous Month" is used to represent the deviation of a goal from one's own capability, and "Run Next Month" indicates whether individuals come back to the platform to upload their running records. *G* indicates a group dummy to represent *Control* group, *Matched* subgroup, and *Social Norms motivated* subgroup, with *Social Norms motivated* subgroup as the baseline group. *Z* represents individuals' characteristics used in the previous models.

Before examining the formal model, we visually compare the *Social Norms motivated* subgroup with the other two groups in Figure 2.5. Compared to the other two groups, individuals in the *Social Norms motivated* subgroup very clearly have a low Goal Attainment Rate. Only 25.0% of individuals in this subgroup attain their goal, as compared with 43.5% in the *Control* group and 38.1% in the *Matched* subgroup. The most striking difference between the *Social Norms motivated* subgroup and the other two groups is the log (# of friends). On average, after calculating the exponential function of the log of number of friends, individuals in the *Social Norms motivated* subgroup have 75 friends, but individuals in the other two groups have approximately 25 friends. Table 2.7 Columns (1) and (2) reinforce what we have found in the previous section: that individuals who have high social connections are more susceptible to a social norms message and perform worse.

Why does the *Social Norms motivated* subgroup have such a low attainment rate? We explore this question further by examining the value of the goals set by these three groups as compared to their previous running performance (see Figure 2.5). We calculate the average daily difference between the goal and the pre-treatment distance. The higher the difference, the more difficult it is to complete the goal. We clearly see

that individuals in the *Social Norms motivated* subgroup have a higher difference than those in the other two groups. Table 2.7 Column (3) shows the results. To attain their goals, individuals in the *Social Norms motivated* subgroup have to run 0.63 km daily more than individuals in the *Control* group and 0.51 km daily more than individuals in the *Matched* subgroup. This indicates individuals in the *Social Norms motivated* subgroup, as compared to individuals in the *Control* group and the *Matched* subgroup, overshoot by setting a goal that is hard to attain.

We offer the following explanation for why individuals in the *Social Norms motivated* subgroup tend to set unrealistic goals. As previously noted, goals are transparent in the online platform, therefore individuals can observe others and be observed by them, i.e., setting a goal is important for impression management. Individuals in the *Social Norms motivated* subgroup, because of their susceptibility to the social norms message, probably care more about what others think and are using the goal as a mechanism for self-promotion in their social network (Buffardi and Campbell 2008). Thus, in an attempt to portray an "idealized" self to impress their peers and set a high goal (Manago et al. 2008; Peluchette and Karl 2009), they may overlook the fact that their goals may be too challenging to be feasible. The S.M.A.R.T. rule of goal setting theory (in which R stands for "***R**ealistic*") emphasizes the importance of setting a goal that can be attained realistically, given available resources and individual capabilities (Doran 1981). Therefore, the social norms message appears to be a bit too "potent" for these individuals and pushes them to violate the *Realistic* rule.

The negative effects of setting an unrealistic goal for physical activity can extend beyond the inability to attain the goal. Our detailed data from Figure 2.5 reveals

another undesirable consequence of using social norms to motivate highly socially connected individuals to set a goal. A high proportion of individuals in the *Control* group and *Matched* subgroups (92.8% and 93.2%, respectively) keep posting running records in the post-treatment month. However, only 80.8% of individuals in the *Social Norms motivated* subgroup do so. The formal model in Table 2.7 Column (4) shows a similar difference in magnitude of 12.3% and 13.5%. It may be the case that failing to attain the goal creates frustration, or that the unrealistic goal causes unexpected injuries that prevent users from running in the post-treatment month. This finding is consistent with the goal-setting literature, which has repeatedly shown that while setting a challenging goal is beneficial for performance (Locke and Latham 2002), the gains are achieved only up to a point (Mento et al. 1987).

2.4.6 Additional Analyses for Robustness

It could be argued that our definition of goal attainment is overly restrictive in that we require individuals to meet or exceed their goals. We conducted additional analyses by redefining goal attainment by the level of accomplishment.[16] Specifically, we use 90%, 75%, and 50% of the goal to redefine goal attainment; for example, if an individual runs more than 75% of his/her goal, we treat this as goal attainment.

Table 2.8 shows the results of each of the heckprobit models. We find that our results still hold when we redefine goal attainment as 90% or 75%, but not 50%. This finding is consistent with earlier arguments that highly socially connected users are more likely to set an unrealistic goal that is beyond their capabilities, resulting in low goal attainment. We find that highly socially connected users still perform worse than

[16] We thank the AE for this suggestion.

other users in terms of goal attainment rates even when we relax the constraint by 75%, meaning that highly socially connected users might, on average, set goals unrealistically by more than 133% (1/0.75) of their capabilities. However, we do not see the same results when we relax the constraint by 50%, probably because even highly socially connected users would not set a goal that is beyond 200% (1/0.5) of their capabilities. These results add further nuance to our understanding of highly socially connected users: that they are likely to set an unrealistic goal, but only up to a point.

2.5. Discussion

2.5.1 Main Findings

Social norms campaigns, predicated on the innate human tendency to conform with what most other people do, have been extensively used to induce behavior change in many contexts, such as environmental protection (Goldstein et al. 2008) and energy consumption (Schultz et al. 2007). In contrast to these traditional settings, a digital platform with social network features has greatly increased each user's visibility and the ease of observing others. Given the widespread use of such platforms, we sought to unpack the interaction between social norms and social connections. We conducted a randomized field experiment with more than 7,000 individuals in an online physical activity community to understand the mechanisms underlying the heterogeneous treatment effects of social norms campaigns. Our research bridges the literature on social norms and social networks and provides the first evidence on how the effects of social norms are conditioned by an individual's social connections.

We confirm that social norms do have a strong effect in motivating behavioral changes related to goal setting, but the effects are not uniformly distributed among users in our sample. Our analysis of conditional average treatment effects shows that individuals who are highly socially connected are more susceptible to the social norms message. In particular, individuals in the top 10th percentile of social connectivity appear to be disproportionately influenced by social norms. For users who are highly socially connected, the social norms treatment causes a lift of 7.7% in goal setting as compared to the control message. However, for users with low social connections, the social norms message only induces a 1% increase in Goal Setting Rate relative to the control message.

Moreover, we find that dissimilar types of high social connections moderate the effect of social norms differently because variation in the types of social ties reflects differences in levels of visibility to others. We argued that individuals who engage in information sharing (high followers, low followees) would behave differently from those who are information seeking (low followers, high followees) and found that individuals who only have high in-degree are more susceptible to social norms, but the same does not apply to individuals who only have high out-degree. Previous literature has shown that individuals' behaviors can be conditional on the level of observability (Exley and Naecker 2016). The tendency to conform to the norm to set a goal is amplified by the level of visibility, making information sharing users more susceptible to the social norms message.

Finally, we find that the group receiving the social norms message has a substantially lower Goal Attainment Rate than the control group (-17%). This effect is

also similarly heterogeneously experienced by users. Goal Attainment Rates in the *Control* and *Social Norms* groups are almost identical for low socially connected individuals, but the differences become strikingly large for highly socially connected individuals (32.1% (SN), and 69.5% (Control)). Further analysis reveals that the highly socially connected users in the *Social Norms* group set more aggressive goals that are beyond their capabilities, probably because of the desire to impress their peers, resulting in a significantly lower Goal Attainment Rate. Such goal failures may plague individuals over a period of time (Brunstein and Gollwitzer 1996), resulting in negative effects for subsequent behaviors, such as a low running rate in the post-treatment month.

2.5.2 Limitations and Implications

We acknowledge several limitations to our study that offer opportunities for future work. First, we address the issue of generalizability. Our research was conducted in the context of health behavior with observable social connections on a platform for physical activity. To the degree that the mechanism driving the moderating effect is a desire for impression management, which is one of the most important reasons for individuals to participate in social network platforms (Krämer and Winter 2008), such effects should be generalizable to other social network platforms. However, we acknowledge that the moderating effect might not always be applicable to other contexts where a goal is less likely to promote a good impression. For example, our goal of increasing physical activity is a positive signal of desirable behavior; a goal focused on reduction of negative behavior (e.g., excess alcohol consumption or drug abuse) may have different dynamics. The response of individuals with high social

connectivity in such settings is not immediately obvious and warrants more study. Our results, then, pave the way for future studies to examine the moderating effects of high social connectivity on other interventions. Given the growing popularity of social network features, more empirical evidence can be obtained relatively easily in other settings such as diabetes monitoring or weight control.

Second, our target users are individuals who choose to participate in this physical activity platform, meaning that they are at least aware of the need for physical activity. Therefore, the generalizability of our findings to users who lack the initial motivation to exercise needs more investigation. However, given the prevalence of health-related mobile applications and the penetration of mobile technologies, it is reasonable to assume that many individuals might have experience using health-related mobile applications. For instance, according to a nationwide survey in the United States (Krebs and Duncan 2015), almost 60% of mobile users had downloaded a health-related app. Consequently, initial adoption might not be a major obstacle for the general public; rather, the challenge is sustaining the behavior necessary to foster a healthy habit (Krebs and Duncan 2015), which is the focal point of our research.

Third, practitioners have extensively used push notifications as a strategy to motivate users for behavior change. Whether the effects of such interventions attenuate after several attempts is an open question. We share this limitation with most social norms studies and call for more careful investigations of the effects of repeated treatments. Fourth, although we have a large sample for a randomized experiment, it is not sufficiently robust to be able to observe and draw a strong conclusion about the long-term effects for goal failure, an area in which the literature is inconclusive (Jones

et al. 2013; Ntoumanis et al. 2014). How do individuals respond differently to a goal that they set organically versus a goal induced by a social norms message? How do individuals handle their failure to reach different types of goals? These are important questions for future research.

Our study yields useful implications for both research and practice. First, our findings indicate that users are essentially different in their response to social norms messages. In addition to the level of social connectivity, there might be other moderating factors that are relevant in social norms campaigns, such as individuals' positions in a social network or their psychological traits. In the medical literature, there is much interest in determining who are "responders" and "non-responders" to interventions (Michelle et al. 2018). We call for more research to explore additional moderating mechanisms to help decision makers refine social norms campaigns on digital platforms. Second, when social norms are coupled with goal setting as a strategy to address the self-control problem, one needs to pay special attention to the possible negative outcomes for some users. To mitigate the powerful and potentially harmful effect of social norms pushing individuals to set unrealistic goals, one might consider suggesting a guideline of a maximum running distance, or incorporate a feedback system to reinforce users' perception of the framed messages. A suitable guideline might restrain users from setting an unrealistic goal, and appropriate design of the feedback system might increase users' goal commitment. Future research should investigate what factors can help alleviate the unexpected negative effects of social norms for some individuals.

Finally, a digital nudge as a strategy to change user behavior is increasingly popular today, especially in the form of push notifications on online platforms. While practitioners have routinely used a simple A-B test approach to examine the average treatment effects, our findings signify the need to go a step further to open the black box and examine heterogeneous treatment effects on different individuals. Platforms should consider multiple indicators to evaluate the trade-offs among interventions in order to fully reap the benefits of push notifications.

Chapter 3: Is It Better to Give Than to Receive? Leveraging Digital Social Connections for Healthy Behavior

Abstract

Motivating individuals to engage in healthy behavior has long persisted as a major challenge in society. Although self-interest based financial incentives have been widely deployed, there is a pressing need to improve their effectiveness. We investigate a new motivational incentive that is based on reciprocity and can be leveraged in conjunction with financial incentives to promote desirable behavior. We conducted a large randomized field experiment with over 1,700 pairs of users on a mobile social network platform. Individuals in our experiment receive a gift from their friends, and are asked to return this favor by participating in a challenge related to physical activity. We find that on average, reciprocity outperforms self-interest in motivating individuals to exercise more. Importantly, our results reveal that the magnitude of the reciprocity effect is contingent on the social closeness between senders and receivers. Interestingly, social closeness has an inverted U-shaped influence on the reciprocity effect. The effect is strongest when closeness is moderate, and wanes when closeness is either too strong or too weak. Compared to the widely used self-interest based financial incentives, our findings offer a potentially more powerful avenue for mechanism design in promoting healthy behavior. This mechanism can be implemented cost-effectively with improved precision for better outcomes using today's ubiquitous digital social connections and wearable devices.

Keywords*: Reciprocity, self-interest, healthy behavior, financial incentive, social connections, social closeness*

3.1. Introduction

Unhealthy lifestyles such as a poor diet, physical inactivity, and alcohol and tobacco abuse, have caused substantial economic costs (Scarborough et al. 2011). To curb the rapid growth in healthcare expenditure, numerous stakeholders including the government, employers, and health plans have increasingly emphasized incentivizing healthy behavior (Baicker et al. 2010). Financial incentives to individuals that invoke self-interest motivations have been extensively studied in a variety of settings such as weight loss, smoking cessation, and physical activity (Charness and Gneezy 2009; Finkelstein et al. 2008; John et al. 2011). However, despite the popularity of such

incentive programs, empirical findings are equivocal (Gneezy et al. 2011): although several studies show encouraging results (Halpern et al. 2015; Volpp et al. 2009), others (e.g., Cawley and Price 2013) report dropout rates as high as 68%. Blumenthal et al. (2013) also note that the effectiveness of financial incentive programs provided by the Centers for Medicare and Medicaid services show mixed results, and call for more investigations on improving the design of such interventions.

In this study, we propose and test a new mechanism in financial incentive design to promote healthy behavior: a reciprocity-based incentive. Reciprocity is a form of a social rule that underscores the importance of repaying what other people have provided. It encapsulates a fundamental aspect of human behavior and has been suggested as a universal norm that most societies endorse (Gouldner 1960). We test whether being *indebted* to repay a gift motivates individuals to exercise more than pure self-gain. Our work aims to leverage individuals' social connections to improve the effectiveness of financial incentives for changing behaviors related to physical activity. Given that social connections can be quite heterogeneous among friends, we also aim to identify the conditions under which reciprocity works best for in order to target motivations more precisely.

Although several studies have demonstrated that reciprocity has a significant effect on human behavior (Alpizar et al. 2008; Falk 2007; Fehr and Gächter 2000), this mechanism has seldom been incorporated into health-promoting programs. This is likely due to at least two challenges. One, there are limited means to effectively enable user interactions, which are a necessary condition for reciprocity. Obtaining information about individuals' social connections can be cost-prohibitive in an offline

setting and indeed, this limiation is a major cause for why prior research on reciprocity utilizes a third party previously unknown to the subjects to trigger a reciprocal response (see a summary in Appendix Table A2). And two, it is challenging to study reciprocity in a public health context due to limited behavior observability. Traditional methods for capturing healthy behavior rely on self-reports, which are not conducive to precise tracking of physical activity. Furthermore, detailed tracking of social connections and physical activity simultaneously can easily become cost-prohibitive for a large sample. As a result, extant understanding of how to leverage reciprocity to promote healthy behavior is limited.

Recent advances in information technologies allow us to address the above two challenges. The maturity of wearable devices and smartphones has afforded us the ability to precisely track individuals' physical activity at relatively low cost. The widespread proliferation of online social networks and mHealth platforms has greatly facilitated interactions among individuals. The ease of capturing such interactions makes the reciprocity-based incentives more viable today. Additionally, the maturity of GPS techniques on smartphones and sports watches increases the observability of individuals' physical activities. When coupled with digital social networks that enable connections at an unprecedented scale, it is now feasible to leverage reciprocity for promoting healthy behavior.

Our work departs from prior literature that examines the effects of reciprocity in various contexts such as donations (Alpizar et al. 2008; Falk 2007), labor markets (Falk et al. 1999; Gneezy and List 2006; Kube et al. 2012), and marketing (Maréchal and Thöni 2007), in three significant ways. First, as a result of the challenges described

above, previous studies have not used reciprocity as a commitment device to motivate behavior change. Our study takes advantage of the capabilities offered by online social networks and digital connections to demonstrate the feasibility of using reciprocity as a mechanism for healthy behavior change.

Second, studies that explore reciprocity largely do so by manipulating the form of the gifts. For instance, Falk (2007) shows that the donations increase by 17% in a small gift treatment and by 75% in a larger gift treatment as compared to a no-gift control. Other work (Kube et al. 2012) suggests that a non-monetary gift is a more effective trigger for invoking reciprocity than a monetary gift of equivalent value. The subsequent behaviors are used to quantify the magnitude of reciprocity when comparing the gift treatment and the no-gift control. As a result of the treatments examined in these studies, little is known about the *relative* performance of reciprocity versus the widely used self-interest motivation.[17] In contrast, we both propose a new form of a commitment device on behavior change in health and quantify its effects relative to self-interest. Finally, to the degree that reciprocity is meaningful only in the context of a dyad, surprisingly, most previous literature does not examine the influence of the sender, who is typically an unrelated third party, on the receiver's reciprocal behaviors. In our study, by changing the identity of the gift giver from the third-party to a receiver's online friend, we are able to explore how social closeness between a dyad, i.e., the nature of the interpersonal relationship, moderates the effects of reciprocity.

[17] A notable exception is presented in Chung and Narayandas (2017) that compare the unconditional compensations with the conditional compensation on salespersons' performance.

We conducted a large randomized controlled experiment with over 1,700 pairs of users in an online physical activity community of runners in Taiwan to address two research questions related to the efficacy of reciprocity as a mechanism for healthy behavior change: (1) Do reciprocity-based incentives compare favorably with the widely used self-interest incentives in promoting healthy behavior? (2) Since reciprocity is a form of a social rule to repay favors, how does the social closeness influence the effectiveness of a reciprocity-based incentive?

Our field experiment builds on the framework of a gift exchange process and includes three groups. A standard gift exchange procedure involves to give, to receive, and to reciprocate (Mauss and Halls 1954/2000). Therefore, a pair of online friends are randomly chosen: one serves as a sender, and the other as a receiver. In the first group, the Reciprocity Treatment group, a receiver receives a gift from a friend. To reciprocate, the receiver needs to complete a physical activity challenge (e.g., run 30 km in two weeks) to return an equivalent gift to the sender. In the second group, the Friend Control group, a receiver also receives a gift from a friend and is asked to complete the same physical activity challenge, but earns an equivalent gift for themselves. Both the reciprocated gift in the Reciprocity Treatment group and earned gift in the Friend Control group are provided by the platform. The only difference is the beneficiary. By comparing these two groups, we are able to quantify the differences between reciprocity-based and self-interest based incentives while controlling for the friend's effect. However, both groups are distinct from a traditional incentive program, in which the sender is an unknown third-party. Therefore, in the third group, the Baseline Control group, a receiver also receives a gift, but this time the sender is the

platform.[18] A receiver then completes the same physical activity challenge to earn an equivalent gift for themselves. The process is similar to a conventional program that provides enrollment rewards to encourage participation and completion rewards to motivate behavior change (e.g., Acland and Levy 2015; Charness and Gneezy 2009). By comparing the Reciprocity Treatment and the Baseline Control groups, we can examine whether a reciprocity-based incentive outperforms a conventional incentive program.

Previewing our findings, the field experiment yields several novel results. Our findings show that the reciprocity-based incentive outperforms the self-interest based incentive in motivating individuals to exercise more. Compared to a traditional incentive program, which was initiated by the platform (i.e., Reciprocity Treatment vs. Baseline Control groups), a reciprocity-based incentive initiated by friends leads to a sizeable 32.0% increase in challenge completion rate. Even after controlling for the gift giver's effect, the reciprocity-based design still outperforms self-interest incentive by 20.4% in challenge completion rate (i.e., Reciprocity Treatment vs. Friend Control groups). Thus, our study offers new findings for mechanism design in promoting healthy behavior.

We further find that the magnitude of the reciprocity effect is contingent on the social closeness between the senders and the receivers. Interestingly, social closeness has an inverted U-shaped influence on the reciprocity effect. The effect is strongest when closeness is moderate and attenuates when closeness is either too strong or too weak. We interpret this finding to suggest that for distant friends, reciprocity may be

[18] See more discussion in the experimental design section to see how we maintain balance across groups.

less effective because there is no strong need for participants to maintain a long-term relationship. For close friends, the weaker effect is likely because individuals do not worry about jeopardizing the relationship even if they fail to complete the challenge or, more practically, because individuals can compensate their friends through offline interactions.

Our study makes important theoretical and practical contributions. First, we extend existing evidence on the effects of reciprocity-based incentives to a novel and consequential context of healthy behavior. It is not clear whether prior findings can be readily generalized due to idiosyncrasies unique to this domain. Healthy behavior requires substantial effort; it is costlier to run 40 km than it is to, say, share a coupon code. Additionally, most existing studies on reciprocity examine activities that are either one-shot or span a short period of time. Indeed, Gneezy and List (2006) point out that the duration of behavior is a gap in the literature and show in a field experiment that reciprocity works only in the first few hours on workers' efforts. Our results provide empirical evidence that reciprocity can be effective for healthy behavior that takes time and effort to complete. We further contribute to literature on incentives by juxtaposing reciprocity with the commonly used self-interest based incentive, and providing a quantitative estimate of the differences in efficacy. In doing so, we advance understanding of an emerging stream of literature by comparing a social or behavioral intervention with financial incentives (Ashraf et al. 2014; Chung and Narayandas 2017; Kast et al. 2012).

We extend the literature on commitment devices such as peer pressure and temptation bundling (Kast et al. 2012; Milkman et al. 2013), by adding reciprocity as a

new form of a commitment device. With the help of digital technologies and the prevalence of online social networks, our study demonstrates that reciprocity can greatly improve the effectiveness of a conventional incentive program without incurring extra costs. Our study responds the call for "making incentive dollars go further" (Blumenthal et al. 2013). This provides a foundation for future studies such as those comparing the effects of different commitment devices or examining the synthesized effects when combining multiple commitment devices (e.g., increasing observability from peers, or loss aversion) with a reciprocity-based design.[19]

Our analysis of the conditioning role played by the interpersonal relationship on the effects of reciprocity offers a deeper understanding of the mechanism. Since most studies in the previous literature use a third party previously unknown to the subjects to initiate reciprocity (Appendix Table A2), it is not feasible for them to test how interpersonal relationships alter the effects of reciprocity. Empirical evidence suggests that social closeness plays a key role in human behaviors (Aral and Walker 2014; Bapna et al. 2017; Nitzan and Libai 2011), although findings about the precise nature of interpersonal relationships on reciprocal behaviors are mixed (Falk et al. 1999; Maréchal and Thöni 2007). For example, Maréchal and Thöni (2007) show that the first time visit of a customer, a proxy to gauge whether sellers and customers know each other before the treatment, would moderate the effects of gifts (free samples). The effect of a gift to increase sales is conditional on sellers and buyers knowing each other. In contrast, Chen et al. (2009) show, in a laboratory experiment setting, that

[19] For instance, Mittone and Ploner (2011) show that individuals would reciprocate more when being observed by peers.

participants' positive reciprocity is not affected by the existing relationship (strangers vs. friends). Our research extends and complements these studies by treating interpersonal relationships as existing along a continuum; a more realistic assumption. Our finding of an inverted U-shaped relationship between social closeness and the effects of reciprocity add further nuance to our understanding of interpersonal relationships and allow practitioners to identify optimal pairs to fully exert the power of reciprocity to achieve precision motivation for each individual.

From the perspective of practice, because our study directly leverages individuals' existing social connections, it has the potential to create value for both the platform and for participants. For the platform, a reciprocity-based incentive increases the bonds between pairs of users, an essential element for a social-network platform to thrive (Ellison 2007). For individuals, a reciprocity-based incentive not only enhances interpersonal relationships between senders and receivers, but might also enable the social support necessary to sustain engagement in healthy behavior in the long run (Kiernan et al. 2012).

3.2. Prior Literature

Our research context is similar to an incentivized wellness program, which aims to reduce health costs and improve employee well-being and productivity by providing inducements. Most wellness programs focus on healthy behaviors, such as exercise, diet, weight loss, and smoking cessation (Baicker et al. 2010; Naydeck et al. 2008; Osilla et al. 2012). Two trends in incentivized wellness programs are noteworthy. One, in light of the robust evidence supporting the value of such programs (e.g., Baicker et al. (2010) assess the effects of 32 wellness programs and show that every dollar spent

reduces $3.27 in medical costs and $2.73 in absenteeism costs), their use is accelerating. And two, online platforms are increasingly being used to orchestrate and deliver wellness programs. For example, Herman et al. (2006) test whether a financial incentive ($150 cash rebate) integrated into an online physical activity program is associated with improvement in health status.

We briefly review several streams of work that provide the conceptual background for our study. First, to address the problem of self-control, especially when an individual is able to acknowledge the problem (such as a lack of physical activity or being overweight), commitment devices may be used to prevent derailment from an intended course of action (Rogers et al. 2014). We review different types of commitment devices and discuss the differences between a hard commitment device and a soft commitment device. Second, since our experimental design has its roots in the power of reciprocity, we discuss gift exchange theory, which offers insights into why individuals reciprocate. Finally, we turn to the conditioning effects of individuals' social connections and present arguments for a curvilinear conditioning role played by the strength of the social closeness on the effects of reciprocity.

3.2.1 Commitment Devices

The difficulty of sustaining healthy habits is a manifestation of the more general problem of self-control, a human fallibility that has garnered considerable attention in both psychology and economics. Typical examples include failure to quit smoking (Hughes et al. 2004), inability to sustain new year's resolutions (Wiseman 2007), and procrastination in submitting homework (Schiming 2012). A self-control problem arises when individuals disproportionately weigh immediate costs against benefits in

decision-making (Ameriks et al. 2007). In the health field, many people fail to sustain a healthy life because of the temptation of short-term benefit, leading individuals to deviate from ideal behavior.

To avoid undesirable future selves (such as becoming indigent, obese, or afflicted by disease), individuals employ commitment devices to prevent lapses in self-control (Bryan et al. 2010). Such devices can be in the form of a formal or informal contract with oneself or close friends, where failure to adhere to contract terms would cause loss and suffering. A commitment device could be as simple as setting a goal, or as harsh as donating a substantial amount of money (e.g., $1000) to a place that one dislikes. Bryan et al. (2010) label commitment devices with financial penalties for failure or rewards for success as hard commitments. In contrast, a soft commitment refers to the psychological consequences of a device. Typical examples of hard commitment devices include late penalties on coursework when one fails to meet deadlines (Ariely and Wertenbroch 2002) or loss of deposits when one fails to quit smoking or lose weight (Giné et al. 2010; Halpern et al. 2012).

Building on the gift exchange process, our paper combines a hard commitment device (enrollment and completion incentives) and a soft commitment device (reciprocity) to foster healthy behavior and directly compares reciprocity with self-interest incentives. Our paper studies a growing body of literature that seeks to directly compare social or behavioral interventions with financial incentives. In particular, Kast et al. (2012) examine how a self-help peer group acts as a commitment device for precautionary savings. They focus on the effect of peer pressure, reflected in being observed by others, on individuals' ability to deal with the self-control problem. Ashraf

et al. (2014) conducted a field experiment to directly compare a pro-social motivation with a financial incentive motivation to promote HIV prevention through condom sales. They show that agents who receive non-financial rewards, a pro-social motivation facilitated by social comparisons, exert more effort than agents receiving financial rewards at promoting sales. Departing from prior work, our research directly leverages individuals' social connections to form a psychological debt to repay favors as a soft commitment device. Our research context is similar to Milkman et al.'s (2013). While they use temptation bundling as a soft commitment device to increase gym attendance rates, we use reciprocity to increase individuals' physical activity.

3.2.2 Gift Exchange Theory and Reciprocity

The theory of gift exchange has its roots in literature from anthropology. Mauss's (1954/2000) famous essay investigates the gift exchange process in a primitive society that involves a series of actions: to give, to receive, and more importantly, to reciprocate. One may give a return gift for the purpose of maintaining a social balance or fairness, showing gratitude, or rewarding generosity (Kolm 2000), and sometimes might even overcompensate the cost of the initial gift. Gift exchange engenders reciprocity because it helps agents to signal social distance and distinguish friends from non-friends (Levine 1998).

Gift exchange has been widely investigated in labor markets. Akerlof's (1982) gift exchange model suggests that firms pay more than the minimum wage to receive higher productivity from employees. In an experimental setting, Fehr et al. (1993) show that individuals exhibit a form of social preference, perhaps reciprocity, to help others in exchange for higher wages. In a field experiment setting, Kube et al. (2012) show

the differences between non-monetary and monetary gifts on workers' reciprocating efforts. In their experiment, individuals in a non-monetary group exert higher work performance than the control group, while individuals in a cash gift group experience no significant effect.

In the information systems literature, researchers have viewed open source and online communities through the lens of a gift economy (Bergquist and Ljungberg 2001). Wasko and Faraj (2005) discuss why individuals help strangers in computer-mediated discussion forums even when they know that their help would not be reciprocated. Reciprocity between two agents is often difficult because of the anonymous nature of interaction in the online forum. However, when individuals offer advice in such forums, the contributor does not expect direct reciprocation but rather expects the receivers to contribute to someone else someday. This form of reciprocity is labeled "generalized exchange" (Ekeh 1974). The sharing in the online environment is different from traditional gift exchange (Smith and Kollock 1999), and knowledge as a form of a public good in the online forum reflects shared moral obligation for generalized reciprocity and prosocial behavior rather than self-interest (Wasko and Faraj 2000).

Reciprocity implies that people behave more nicely and are more cooperative in response to friendly actions (Fehr and Gächter 2000). Calls for contributions from charities are one of the most common contexts where the power of reciprocity has been leveraged. Falk (2007) conducted a randomized field experiment in an event to solicit donations. The small gift group elicited 17% more donations, and the large gift group elicited 75% more donations than the no-gift control group. Alpizar et al. (2008) found

that tourists contribute more when receiving a small gift at a national park in Costa Rica. Likewise, marketers who provide free samples to increase sales also leverage the power of reciprocity. For instance, Bawa and Shoemaker (2004) show that consumers feel obliged to purchase when provided with a free sample. Beltramini (1992) found that giving a business gift to customers increases positive perceptions of key product attributes. Beyond donations and marketing, researchers have also tested reciprocity in a hospital setting. Currie et al. (2013) show that when patients bring small gifts such as bookmarks to physicians, doctors reciprocate with better service and fewer unnecessary prescriptions of antibiotics.

Why do individuals reciprocate when there is often no overt, material gain from doing so? They do so for three primary reasons: conforming to a universal norm, expressing gratitude, and avoiding psychological distress. First, reciprocity is essential to form human relationships (Gudeman 2001). A classic explanation offered by Gouldner (1960) argues that one should help those who have helped them. Such a norm has long been imprinted in individuals and internalized in most societies. Yan (2003) suggests that failure to fulfill the obligations of reciprocity violates ethics and may be viewed as immoral. Second, people reciprocate to show gratitude, an emotional appreciation for received benefits. Komter (2004) describes gratitude as an imperative force that compels us to return the received benefits. Palmatier et al. (2009) show that marketers' relationship management leads to customers' gratitude in such a way that it increases customers' purchase intentions. Third, failure to reciprocate may be costly as it generates psychological pressure in the form of feelings of guilt (Becker 1990). For

instance, Dahl et al. (2005) show that customers incur a feeling of guilt for not purchasing from a salesperson after establishing temporary social connectedness.

3.2.3 The Conditioning Effects of Social Closeness

The importance of social connections between a pair of users in an online social network is underscored by substantial research showing that connections are related to behaviors such as adoption, content sharing, or trust. Yoganarasimhan (2012) finds that network structure is a critical driver of the popularity of a video on Youtube, and Nitzan and Libai (2011) show that individuals are more likely to defect if connected by a defecting neighbor, and such effects would be influenced by tie strength and homophily. Katona et al. (2011) find that both the degree effects, i.e., individuals who are connected with many adopters, and the clustering effect, i.e., the density of users who are the adopters, affect a focal user's adoption. Bapna et al. (2017) show that degree of interaction, embeddedness, and being tagged in the same photo would affect individuals' trust.

Despite the growing body of research supporting the effects of social connections on behaviors, the conditional effects of social closeness on the impact of reciprocity have not been empirically tested. This oversight is likely because it is costly and hard to leverage individuals' social connections without the help of modern technologies and social network platforms. In this study, address this gap and explore how social closeness conditions the effects of reciprocity. To the extent that individuals reciprocate to follow universal norms and to conform to the social rules, it is a reasonable conjecture that the effects of reciprocity would be stronger when receivers regard the sender to be a close rather than a distant friend. However, we argue that such

a positive relationship would hold only up to a point, and then decay. Specifically, we suggest that social closeness between senders and receivers will have an inverted U-shaped relationship on the effects of reciprocity.

Our expectation for such a curvilinear effect is predicated on the individual utility generated from reciprocating a gift, and supported by findings from prior research. According to Fehr and Gächter (2000), positive reciprocity has long been embedded in the society, and the received favors create "feelings of indebtedness obliging many people to repay the psychological debt." However, the psychological debt should not be uniformly distributed and is likely to be conditional on the relationships between senders and receivers prior to receiving the gifts. Studies have shown that people respond differently when in close rather than distant relationships (Fiske 1992). Grant and Gino (2010) find that receivers' perceptions of social worth can motivate different levels of prosocial behaviors. Therefore, a receiver's perception of the gift would be affected by the social closeness to the sender and, in turn, affect the reciprocal response.

When receiving gifts from distant friends, the desire to return the favor is plausibly weak. A strong relationship might exert a more powerful effect than a distant tie does. Ryu and Feick (2007) show that the strength of an ego's social ties (strong tie vs. weak tie) affect the effectiveness of a referral program and argue that for strong tie relationships, individuals are more concerned about the other's welfare. For example, individuals are more inclined to make a referral for a strong than a weak tie (Frenzen and Nakamoto 1993). Aral and Walker (2014) demonstrate that the embeddedness or common social contexts between a pair of users on Facebook positively affects

application adoption. Furthermore, one reason people reciprocate is for long-term outcomes and the desire for future interaction (Homans 1974). Since distant friends are merely online acquaintances, the desire to maintain a long-term relationship and future interaction is likely low, thereby dampening the receiver's desire to reciprocate.

Likewise, when the gift giver is a very close friend, the need for reciprocation may also be low. As noted previously, one reason people reciprocate is the desire for future interaction (Homans 1974). With very close friends, the obligation to reciprocate may be attenuated because receivers and senders already have an intimate bond. In other words, the receiver believes that the relationship will persist even in the absence of a return gift. According to Gouldner (1960), the foundation of reciprocity is the embedded obligation to repay. Receivers feel *indebted* toward the senders until they have repaid the favors. For a very close friend, the perception of indebtedness is likely to be low because receivers are not afraid of jeopardizing the relationship even if they fail to repay their friends, or because receivers could reciprocate via other offline means. Furthermore, Chen et al. (2009) argue that during interactions with distant friends, the dominant mindset is "transactional:" i.e., individuals will often calculate the economic outcomes of behaviors. In contrast, with close friends, individuals bring a social/relational frame of mind that takes the psychological considerations of the exchange into account. With close friends, receivers may take the goodwill for granted. Therefore, both the level of gratitude as well as the experience of guilt may not be potent enough to invoke a stronger effect of reciprocity.

To summarize, we propose that for receivers who receive gifts from a sender who is moderately close to them, the effect of reciprocity on behavior would be the

strongest. When viewed through the lens of a long-term relationship, moderate closeness would invoke a higher inclination to maintain compared to a distant one (no desire) or a very close one (no need to do so, because the relationship is already close). For psychological debt, the level of guilt or appreciation would be low for a distant one (do not care) or a very close one (there may be other means or occasions to compensate). Therefore, the utility gained by repaying a favor would be the strongest for moderately close friends.

3.3. Methodology

We conducted a randomized field experiment on a mobile physical activity platform for runners, henceforth referred to as *RunningPlatform* to maintain anonymity. The platform provides a Twitter-like capability that allows individuals to follow or be followed by others. Using this feature we are able to find pairs of friends, a necessary condition to test reciprocity. To ensure the quality of the randomization process, we assign users to groups based on their online identification numbers. Participants in the study are not aware of the existence of our experiment. This helps us avoid the well-known "Hawthorne effect" (Adair 1984) threat, where subjects who are aware of being observed alter their behavior.

3.3.1. Experimental Design

To test the effects of reciprocity, we design an experiment that follows a standard gift exchange process: to give, to receive, and then to reciprocate. We first describe how we identify user dyads, one of whom serves as a sender and the other as a receiver. We observe individuals' running performance two weeks before the

intervention date, and use a threshold of 40 km to identify potential senders.[20] When users log into the platform, the system examines their running performance to determine if they qualify as senders. In both the web and mobile platforms, a qualified user would be directed to a page that describes the information presented in Table 3.1. On this page, we ask whether they would like to help their friends by forwarding a physical activity related challenge invitation to them. If a sender agrees to forward the message, the system *randomly* selects a friend who ran less than 40 km in the pre-treatment period from the sender's friend list.[21] This process ensures that the receivers are relatively weak runners. The threshold for differentiating between senders and receivers ensures that the same individual is never selected as both a sender and a receiver, which could potentially create a confound. Furthermore, since weak runners are the ones most in need of motivation, we purposely select weak runners to receive the challenge. Indeed, it is a common practice to use capable users as a role model to motivate incapable ones because capable individuals show how to perform a skill and exhibit that a goal is attainable and desirable (Morgenroth et al. 2015).[22] The receivers are then *randomly* assigned to one of three groups. A sender's friend will receive a gift (a virtual gold coin) from the sender and has a chance to win a second virtual gold coin if he or she completes the challenge. Each gold coin represents a chance to enter a lottery drawing, which provides rewards such as $200 sports watches, $150 sports

[20] The threshold is based on *Physical Activity Guidelines for Americans* in 2008.

[21] A sender's friend could be the sender's follower, followee, or mutual follower.

[22] Although we select senders from the "strong" (run more than 40km in two weeks) runners, we confirm that strong senders' weak friends are not different from weak runners' weak friends. Both of them do not run more than 10 km in the pre-treatment two-week period; thus, reducing the concern of our sample being unrepresentative.

shoes, $50 T-shirts and $30 socks. Figure 3.1 shows the rewards. Each sender is only paired with one receiver, and a receiver receives a gold coin only from one sender. Therefore, each pair of users is unique in our setting. The goal of the challenge is to motivate the receivers to increase running distances.

We design a between-subjects experiment. The receivers do not overlap among groups. In the first group, the treatment group, the receiver receives a virtual gold coin from the sender (a gift) and is offered the chance to participate in a running challenge to earn a gold coin to give back to the sender. This is our *Reciprocity Treatment* group. In the second group, the *Friend Control* group, the receiver receives a gold coin from the sender and is given a chance to participate in the same challenge, but in this treatment, they get to keep the second gold coin for themselves. In both groups, the gold coin earned by completing the challenge is provided by the platform, thus the only difference between the Reciprocity Treatment and Friend Control groups is the beneficiary of the extra coin: the sender (reciprocity) or self (self-interest). Our goal is to examine which mechanism will be more powerful in motivating receivers to complete the challenge.

We add a third group, in which the receivers receive a virtual gold coin (gift) from the platform and have a chance to participate in the same challenge to earn a second coin for themselves. We call this group the *Baseline Control* group. We include this Baseline Control group to mimic a conventional incentive program. Although the Friend Control group is the ideal control group to account for the gift giver's effect and helps us isolate the effect of reciprocity, most conventional incentive programs do not use individuals' social connections. Instead, a traditional incentive program would

typically offer enrollment rewards from a third-party (similar to a platform gift in the Baseline Control group) and then provide incentives for participants once they complete a challenge. For instance, Acland and Levy (2015) provide a first-week attendance reward and follow-up incentives for gym attendance and Charness and Gneezy (2009) offer a show-up fee and follow-up incentives for gym attendance. The Baseline Control group resembles such a traditional incentive program. If, for some reason, receiving a gift from a friend dampened rather than amplified the motivation to exercise, then even if we show that the Reciprocity Treatment group outperforms the Friend Control group, the practical significance would be limited because their effects might not even as good as the one in the Baseline Control group. The Baseline Control group thus serves as a basis for comparison, helping us measure the overall effects of summing up the effect of a friend's gift and the effect of reciprocity when compared with the Reciprocity Treatment group and exclude the possibility that a friend's gift would generate a negative effect when compared with the Friend Control group.

One potential concern in designing these three groups is the comparability issue. To ensure adequate balance, receivers in all groups are chosen from senders' friends lists. Recall that once a qualified user agrees to become a sender, a receiver is randomly selected from the sender's friend list and randomly assigned into one of the three groups. Receivers in the Reciprocity Treatment and Friend Control groups receive the sender's information, but such information is not disclosed to receivers in the Baseline Control group. Instead, users in the Baseline Control group only know that they receive a gold coin (gift) from *RunningPlatform*. This design ensures that the subjects in the three treatments are equivalent, and our findings are driven by the

treatments rather than selection.

Participants in all three groups receive the same challenge. According to the Physical Activity Guidelines for Americans (2008), individuals gain health benefits if they exercise for at least 150 minutes per week. One can gain additional health benefits when increasing aerobic physical activity to 300 minutes. Based on these guidelines, we ask individuals who have run less than 25 km in the pre-treatment period to complete a 30 km challenge over a two-week period, and we ask individuals who have run between 25 km and 40 km in the pre-treatment period to complete a 45 km challenge over a two-week period. We avoid using a customized goal for each individual because our objective is to promote physical activity that meets recommended professional guidelines. Table 3.2 shows the complete messages received by all the groups.

To test for the moderating effect of the closeness between senders and receivers, our experimental design disallowed senders from selecting their receivers. Rather, a friend is *randomly* picked from the sender's friend list. This ensures that there is no selection on the senders' side to pick receivers, and the closeness of the dyad is orthogonal to the treatment. Such a design allows us to test the conditional role of social closeness on the effects of reciprocity.

3.3.2. Experimental Timeline

Figure 3.2 shows the timeline of the experiment, conducted over six weeks in July, August, and September 2017. In the pre-treatment period (7/24-8/6), we recorded all individuals' running distances to be used for identifying senders and receivers in the next stage. During the pairing period (8/7-8/20), we calculated the total running

distance for each user in the pre-treatment period to identify qualified senders, i.e., those whose pre-treatment running distance exceeded 40 km. For senders who agreed to forward this message, we sent a push notification to a randomly selected friend (the receivers) whose pre-treatment performance (total distance) had to be less than 40 km. Upon receiving the message, a receiver can decide to accept or decline this challenge. If a receiver does not make a decision right away, she/he can return to that page to decide to participate or not before the end of the pairing period. Once receivers agreed to accept the challenge, they undertake the challenge during the treatment period (8/21-9/3). Each period is exactly two weeks.

One day before the treatment period (at 9:00 am on 8/20), we sent a reminder message to each of the participants in all three groups about their upcoming challenge. The message communicated the time period for the challenge, the distance they were supposed to run, and the list of rewards. Each group gets unique information about the beneficiary after completing the challenge. Table 3.3 shows complete messages received by all the groups.

During the two-week challenge period, we sent push notifications to the challengers thrice a week at 9:00 am on Tuesday, Thursday, and Saturday. In contrast to daily push notifications, such a strategy has the advantage of not overly annoying users while maintaining a consistent reminder effect. All participants received messages about their accumulated distance. For individuals who had not completed the challenge, we encouraged them to keep running to win their friends a gold coin (Reciprocity Treatment group) or win themselves a second gold coin (Baseline Control and Friend Control groups). For individuals who completed the challenge, we notified

them that their friends would win a gold coin (Reciprocity Treatment group) or they would win a second gold coin (Baseline Control and Friend Control groups). Individuals who completed the challenge did not receive the reminder message again. Table 3.4 shows complete messages for all the groups.

At the end of the treatment period, we surveyed participants to elicit their opinions about this event and subjective measures of social closeness. To increase the response rate, we provided an incentive for users who complete the questionnaire; a chance to be entered into another lottery drawing for a $200 sports watch. The questionnaires were customized for different individuals based on their experimental group.

3.3.3. Sample Description

In total, 3,502 users qualified as senders, of which 2,744 agreed to forward the message to their friends (receivers) during the pairing period. Interestingly, a large majority (78.4%) were willing to serve as a sender to encourage their friends to participate in our challenge. Once a sender agreed to participate, we randomly assigned the receiver into one of the three groups based on the receiver's online identification number (*user_id*). Among qualified users, 1,738 receivers logged in to the platform and read the message during the pairing period. There were 571 receivers in the Baseline Control group, 590 receivers in the Friend Control group, and 577 receivers in the Reciprocity Treatment group, respectively. Each group accounts for 32.85%, 33.95%, and 33.20% of total receivers. The 1,738 pairs constitute our final sample.

The most crucial aspect of this randomized controlled trial is to ensure balance among these three groups. First, since we have individuals' GPS information, we know

their location during the pre-treatment period. Figure 3.3 provides a geographical visualization of most of the participants[23] who provide GPS information in each county or city. We use their latest running records to draw this graph. Individuals in all the groups are approximately evenly concentrated in cities such as Taipei, New Taipei (north), Taichung (center), and Kaohsiung (south). This provides us with preliminary confidence in our randomization process. We conduct a series of ANOVA analyses to formally test the balance among these three groups (see Table 3.5). Since we pair a sender and a receiver to create dyads, extra caution is needed to ensure the balance of sender-receiver dyads across the three groups. Therefore, we show both senders' and receiver's characteristics in Table 3.5, including pre-treatment distance, age, tenure from registration date, the number of races they participated in, the proportion of female users, and the number of online friends. The F-statistic for each variable provides further confidence in our randomization process.[24] Finally, since the challenge type is based on users' capabilities, individuals might be assigned the 30 km or the 45 km challenge. The results, again, show support for appropriate balance among groups.

3.4. Results

3.4.1. Experimental Results

Upon receiving the message, receivers can choose to accept or decline the challenge. The acceptance rates are 86.3% (Baseline Control group), 82.7% (Friend Control group), and 83.9% (Reciprocity Treatment group), respectively. Pairwise t-

[23] We draw this graph by using 93.1% of participants who have GPS information.

[24] Although there is no statistical difference of the pre-treatment distances in the 30 km challenge types or the 45 km challenge types, respectively, we do find statistical differences when pooling these two types together. We address this problem by matching the groups and by controlling the inverse-probability weighting (IPW) in the model.

tests show no statistical differences between acceptance rates. Major individual characteristics such as age, gender, tenure, and the number of friends are also statistically identical across groups.[25] Nevertheless, as a further caution against a selection problem, we include the declined users in the analyses. Our dependent variable is *Complete_Challenge*, which is equal to 1 when an individual's total running distance during the treatment period is greater than the total distance of their predefined challenge (either 30 km or 45 km); 0 otherwise. Since declining the challenge indicates no chance to win the second gold coin, we treat this case as a failure to complete the challenge.[26] This allows us to avoid self-selection problem when calculating the average treatment effects (ATE).

We find that 30.6% of individuals in the Baseline Control group, 34.4% of individuals in the Friend Control group, and 40.4% of individuals in the Reciprocity Treatment group complete the challenge. Further analysis using t-tests show that the completion rate of the Reciprocity Treatment group is statistically higher than that of the Baseline Control group ($p < 0.01$) and that of the Friend Control group ($p < 0.05$). This indicates that the motivation to earn a gold coin for a friend (Reciprocity Treatment group) is stronger than the traditional incentive program (Baseline Control group) and also greater than the motivation to earn a coin for oneself while controlling the effects of a friend's gift (Friend Control group).

We formally test the average treatment effects with a Probit model:

---

[25] The pre-treatment distance difference is addressed later.

[26] In the additional analysis, we treat those declined individuals as normal participants and calculate whether they complete the challenge or not based on their predefined challenges. The results are similar to the main models.

$$Y_i^* = \beta_0 + \beta_1 * Treatment_i + \Sigma\beta_r Z_i + \varepsilon_i \quad (1)$$

$$Y_i = 1\,[Y_i^* > 0]$$

where Y_i^* is the key dependent variable, *Complete_Challenge*, indicating whether an individual completes the challenge or not. $Treatment_i$ is the Reciprocity Treatment group as compared to the Friend Control or Baseline Control groups, respectively.[27] Z_i indicates both senders' and receivers' characteristics, including pre-treatment distance, the number of friends, gender (female), age, tenure, and the number of races participated, and the type of challenge (30 km or 45 km) that the participant was assigned to. ε_i is the error term for individual observation i. Results are reported in Table 3.6. In the odd columns we show the results of comparing the Reciprocity Treatment and Baseline Control groups, and in the even columns the results of comparing the Reciprocity Treatment and Friend Control groups. In Column (1), the Baseline Control group is the default group. The results show that the Reciprocity Treatment group's completion rate is positive and significant at $p < 0.01$ level as compared to the Baseline Control group. The marginal effect of reciprocity is 9.8%. Since the completion rate of the Baseline Control group is 30.6%, reciprocity leads to a sizable 32.0% increase in the completion rate. The regression shows that changing the identity of the gift giver from the platform (a third party) to a friend and leveraging the effect of reciprocity can dramatically improve the performance of a traditional incentive program without incurring extra costs. In Column (2), the Friend Control group is the default group. The results show that the Reciprocity Treatment group's

[27] The results are broadly similar if we include three groups in the same models. We present them separately to better illustrate the effects of reciprocity.

completion rate is positive and significant at $p < 0.05$ level. The marginal effect of reciprocity is 7.0%. Since the completion rate of the Friend Control group is 34.3%, reciprocity leads to a 20.4% increase in the completion rate. The regression shows that after controlling for the effect of the gift giver, the pure effect of reciprocity still outperforms an incentive scheme that is based on a self-interest mechanism.

Recall that we treat individuals who do not participate in our treatment as a failure in Table 3.6 Columns (1) and (2). One potential concern may be that individuals who decline our invitation were also affected by the message in regard to their physical activity and treating them as failures might not be appropriate. To address this concern, we release this constraint by treating them in the same way as individuals who participate. As long as their running distance in the treatment period is greater than the challenge distance, we treat them as having completed the challenge. The completion rates after removing the constraint are 33.6%, 36.6%, and 42.4% in the Baseline Control, Friend Control, and Reciprocity Treatment groups, respectively. Columns (3) and (4) show the results. Similar to the main results, the Reciprocity Treatment group outperforms both Baseline Control ($p < 0.01$) and Friend Control groups ($p < 0.05$). The marginal effects of the Reciprocity Treatment group are 8.83% in Column (3) and 7.02% in Column (4), respectively. Thus, our results are robust to alternate measures of *Complete_Challenge*.

In Columns (1) to (4), we included participants who declined in the analyses. Although the overarching goal of our treatments is to evaluate the average impacts of a reciprocity-based intervention on healthy behavior, it is particularly important for a policymaker (the platform) to explicitly evaluate the effects of the intervention on those

who choose to participate. Therefore, we are interested in quantifying the average treatment effects in the treated (ATT). We conduct the analysis by excluding individuals who decline our invitations. Table 3.6 Columns (5) and (6) show the results. Although the average completion rates are higher when we calculate these rates on the subset of treated participants (35.5% in the Baseline Control group, 41.6% in the Friend Control group and 48.1% in the Reciprocity Treatment group), the marginal effects for the Reciprocity Treatment group as compared to the mean completion rates in the Baseline Control and Friend Control groups are similar to the main models (37.0% increase as compared to the Baseline Control group and 18.6% increase as compared to the Friend Control group). Since more than 84% of participants accepted the challenge, it is not surprising to find a similar magnitude of ATE and ATT.

Figure 3.4 shows the cumulative challenge completion rate across the 14-day treatment period. The Reciprocity Treatment group outperforms the other two groups throughout. Over the course of the treatment period, there are a higher proportion of individuals in the Reciprocity Treatment group who complete the challenge than individuals in the other two groups. This shows that the effect of reciprocity is consistent across time and that the reciprocity-based incentive does not generate a short-term booster effect such as running in the last few days to complete the challenge. Our results are therefore informative for the missing link of the role of duration in reciprocity that Gneezy and List (2006) highlight. This is especially essential to help individuals to foster a long-term healthy habit.

3.4.2. The Curvilinear Relationship of Closeness and Reciprocity

Because reciprocity is fundamentally a response to social rules, we expected that its effects would not be uniformly distributed across receivers in the Reciprocity Treatment group; rather they would be conditional on the social closeness between senders and receivers. To examine this conjecture, we construct a closeness index (*ClosenessIndex*), drawing on insights from prior work, that comprises the interaction between senders and receivers to capture social contexts (Aral and Walker 2014). We create this index as a summative measure of the following variables (each of which is binary): whether a sender and a receiver are in a common online running club, whether a sender and a receiver have attended the same marathon race before, whether a sender and a receiver mutually follow each other, whether a receiver has pressed *like* on a sender's running record in the past half year[28]. The common online running club and common races reflect the shared experience between a sender and a receiver (Aral and Walker 2014). Mutually following each other is a variable constructed from the network structure and has been used to represent the strength of a tie (e.g., Shi et al. 2014). The *likes* behaviors represent receiver's positive association with the sender (Kosinski et al. 2013), representing receivers' fondness toward senders.[29] We avoid using the continuous form of the number of *likes* because it would be affected by the length of the relationship. The range of the closeness index is from zero to four, with zero indicating low closeness and four representing a close relationship.

---

[28] The *like* feature in the platform is similar to Facebook's feature. Results are robust to different time periods (3-months and 1-year) for constructing this measure.

[29] It could be argued that senders' *like* on receivers also represents a positive signal of closeness between pairs. Adding an extra variable, whether a sender has pressed *like* on a receiver's running record in the past half year, yields similar results.

We test the curvilinear relationship in the Reciprocity Treatment group by using the following model suggested by Haans et al. (2016):

$$Y_i^* = \beta_0 + \beta_1 * ClosenessIndex_i + \beta_2 * ClosenessIndex_i^2 + \Sigma\beta_r Z_i + \varepsilon_i \qquad (3)$$

$$Y_i = 1\,[Y_i^* > 0]$$

where the dependent variable and the control variables are the same as in the main models. Note that we disallowed senders to pick friends during the pairing period, and receivers are *randomly* selected from a sender's friends list. Therefore, the social closeness (*ClosenessIndex*) between senders and receivers is exogenous in our setting. To confirm this, we conducted ANOVA analyses to examine both senders' and receivers' pre-treatment distances across different levels of closeness. The F-statistics show that both senders' ($p = 0.57$) and receivers' pre-treatment distances ($p = 0.45$) are balanced among different levels of closeness. Table 3.7 Column (1) shows the results of equation (3). The significant and negative sign of the squared term of the *ClosenessIndex* indicates an inverted U-shaped relationship ($p < 0.01$).[30] Although it is necessary for the squared term to be significant in order to establish the presence of a U-shaped relationship, it alone is not sufficient (Haans et al. 2016). We conducted a formal test for an inverted U-shaped relationship as described by Lind and Mehlum (2010) and confirmed that there is an inverted U-shaped effect of *ClosenessIndex* on completion rate.

Figure 3.5 visually illustrates this relationship. We can clearly discern that when senders and receivers are very distant and very close, the challenge completion rates

[30] To alleviate the concern of skewness, we also test the model by removing or replacing extreme values and obtain consistent results.

are roughly 30%, but when senders and receivers are moderately close, the challenge completion rate reaches above 45%. As noted previously, we interpret this finding as follows: for distant friends, reciprocity may be less effective because there is no strong need for participants to maintain a long-term relationship or seek future interaction. For close friends, the weaker effect is likely because individuals do not worry about jeopardizing the relationship even if they fail to complete the challenge or because they can easily pay back the favor through other offline interactions. We also examine the curvilinear relationship for the other two groups. As we would expect, such an inverted U-shaped relationship does not exist in the Baseline Control and Friend Control groups because there is no concern about maintaining a long-term relationship between senders and receivers when failing to complete the challenge.

3.4.3. Addressing the Diminishing Marginal Returns Concern

One might legitimately conjecture that the Reciprocity Treatment group outperforms the Friend Control group because of the law of diminishing marginal returns. That is, individuals in the Friend Control group might value the second virtual gold coin less because they already own one and, as a result, exert less effort to complete the challenge. If that were the case, the effect of reciprocity could be overestimated. To examine this possibility, we tested another treatment during the pilot test conducted in May 2017 prior to the full experiment. Individuals in this group do not receive a gold coin ("a gift"); rather, they only receive a message from the platform inviting them to complete the challenge and win a gold coin. Thus, the receiver does not have a gold coin to start with (i.e., no endowment), and should value the challenge

more. This helps us assess the effectiveness of the first gold coin in the self-interest scheme. We call this group the *No Gift* group.

In the pilot test, we only included the Friend Control and Reciprocity Treatment groups, in addition to the No Gift group. A similar concern in designing these three groups is the comparability issue. As in the full experiment, a receiver was randomly selected from the sender's friends list and randomly assigned into one of these three groups. Receivers in the Reciprocity Treatment and Friend Control groups receive the sender's information, but such information is not disclosed to receivers in the No Gift group. Users in the No Gift group only see a challenge invitation from the platform.

Individuals in the No Gift or Friend Control group are facing the same challenge, and winning the challenge would give them the same reward (a virtual gold coin). If diminishing marginal returns exist, then the second virtual gold coin should be valued less than the first virtual gold coin sent from a friend. Therefore, the motivation of the Friend Control group users to win the second gold coin should be weaker than that of the No Gift group users. However, contrary to this conjecture, we find that the Friend Control group (36.3%) outperforms the No Gift group (30.9%), although the result was not statistically significant; likely as a result of the smaller sample size. This helps alleviate the concern about potential diminishing marginal returns.

3.4.4. Addressing Imbalance in the Pre-treatment Distances

3.4.4.1. Matching

Although we use individuals' online user identification that should be orthogonal to individuals' characteristics to do the randomization, we find that individuals in the Friend Control group run more than the Reciprocity Treatment group

in the pre-treatment period.[31] One might be concerned that since runners' performance is an important factor in challenge completion, the imbalance between the Friend Control and Reciprocity Treatment groups might bias outcomes. We conduct additional analyses to address this issue. However, before showing the additional analyses, we note that since individuals in the Friend Control group have higher performance than individuals in the Reciprocity Treatment group, individuals' completion challenge rate in the Friend Control group should be higher than individuals in the Reciprocity Treatment group because it is easier to complete the challenge.[32] Therefore, our results do not overestimate the effects of reciprocity, rather, if pre-treatment performance affects our results, we underestimate the effects of reciprocity.

We perform two additional sets of analyses to further alleviate this concern: propensity score matching (PSM) and inverse probability weighting (IPW). We use PSM to match the Friend Control and Reciprocity Treatment groups using individuals' characteristics to calculate the propensity score in the first stage Probit model, including the pre-treatment distance, the number of friends, gender (female), age, tenure, the number of races participated, and the types of challenge (30 km or 45 km) that the participant was assigned to. Before matching, individuals in the Friend Control group run 20.13 km in the pre-treatment period, which is higher than the 18.59 km of individuals in the Reciprocity Treatment group in the pre-treatment period ($p < 0.05$).

---

[31] Since Friend Control is the ideal control group that controls the effects of friend's gifts, we only discuss Friend Control and Reciprocity Treatment in this section.

[32] To calculate the level of difficulty in completing a challenge, we calculate the difference between individuals' pre-treatment distances and the challenge goal. The higher this number, the harder the challenge. On average, individuals in the Friend Control group only need to run 14.88 km to complete the challenge and individuals in the Reciprocity Treatment group need to run 15.67 km to do so. The differences are statistically significant, indicating that individuals in the Friend Control group are likely to find it relatively easier to complete their challenges than those in the Reciprocity Treatment group.

After matching, individuals in the Friend Control and Reciprocity Treatment groups run 18.81 km and 19.29 km in the pre-treatment period, respectively. The difference is not statistically significant.[33] We confirm that all the other characteristics are not statistically different in Table 3.8. Table 3.9 shows the regression results. Column (1) shows that individuals' completion rate in the Reciprocity Treatment group is higher than that in the Friend Control group ($p < 0.05$). The marginal effects show an 8.9% difference. Considering the average completion rate in the matched Friend Control group is 32.4%, the effect of reciprocity results in a 27.5% increase. The increase is higher than what we estimated in the main model (20.4%), which further support the arguments that our main results are indeed underestimating the effects of reciprocity.

An alternative method to address the problem of imbalance is to employ inverse-probability weighting (IPW). We use IPW to reduce the bias of unweighted estimators. We deployed a two-step approach to calculate IPW estimators: a Probit model to estimate inverse-probability weights by using individuals' characteristics; followed by computing the average treatments effects using the estimated inverse-probability weights. After including IPW in the model, we find that the average treatment effect between the Friend Control and Reciprocity Treatment groups is 7.0% ($p < 0.05$), which is roughly equivalent in magnitude to the results from the main model.

3.4.4.2. Performance Comparisons

The purpose of our experiment is to examine if individuals increase their performance in terms of total running distances under different manipulations. Since

[33] In addition, the difference between individuals' pre-treatment distances and the goal in the two groups is not statistically different. This means the level of difficulty in completing the challenge, on average, is the same when using the matched sample.

there is an imbalance in pre-treatment distances among groups, directly comparing individuals' performance during the treatment period between groups might be misleading. As before, we first employ the PSM to find a matched sample. Before presenting results, we visually examine the cumulative probability of individuals' total running distances in the treatment period using the matched samples. Figure 3.6 shows that the cumulative probability of the Reciprocity Treatment group is shifted to the right as compared to that of the Friend Control group[34], showing better overall performance than the Friend Control group. The horizontal line indicates a cumulative probability at 50%. The median distances are 20.0 km and 26.3 km in the Friend Control and Reciprocity Treatment groups, respectively.

Results in Table 3.9 Column (2) show that individuals' performance[35](total running distance in the treatment period) in the Reciprocity Treatment group is superior to that in the Friend Control group when using the matched sample. On average, individuals in the Reciprocity Treatment group can run 3.06 km more than those in the Friend Control group ($p < 0.05$). Considering the average distance that subjects in the Friend Control group ran in the treatment period (26.55 km), this represents an 11.50% increase. Likewise, we employ IPW in the model and find that the average treatment effect between the Reciprocity Treatment and Friend Control groups is 2.26 km.

3.4.5. Robustness Checks

In the results reported thus far, our findings remain remarkably consistent across a variety of models (e.g., addressing imbalance through the use of PSM

---

34 We exclude top 5% of users in both groups to better present the graph.

35 We also analyzed other measures performance, viz., improvement from the pre-treatment period and deviation from the challenge. The results hold for these measures as well.

matching and IPW analysis, using performance as an alternative measure to examine the treatment effect through the OLS model) and alternative measures (such as those for the closeness index). We further conduct an additional series of robustness checks to eliminate alternative explanations.

3.4.5.1. Mitigating the Potential of Cheating Behavior

While an important aspect of the novelty of this study is our ability to track individuals' running records using digital devices such as smartphones or wearable technologies, there may be a concern related to the potential for cheating. Despite the fact that the platform supports multiple devices, some users may develop a habit of using the traditional method, which is to manually input the distance they run through the web or mobile interface. To the degree these are self-reported, one might question the veracity of the data. To alleviate this concern, we retain individuals who only use the digital devices to automatically record running distances and examine their outcomes. There are 9.3%, 9.2%, and 9.0% of individuals in the Baseline Control, Friend Control, and Reciprocity Treatment groups, respectively, who have some manual running records during the treatment period. The proportions are not statistically different. After excluding those individuals, we find that the challenge completion rates are 28.6%, 32.3%, and 38.7% in the Baseline Control, Friend Control, and Reciprocity Treatment groups, respectively. Table 3.10 Columns (1) and (2) show the results. The results are qualitatively similar to the main results. As we have shown previously, the Reciprocity Treatment group is statistically better than the Baseline Control group ($p < 0.01$) and Friend Control group ($p < 0.05$).

Besides manually manipulating running records, there are other possible cheating methods such as driving a car while turning on the mobile application. To alleviate this concern, we only retain individuals who keep all their records publicly visible. The platform allows individuals to keep their records private or reveal their running records to the public and to the user's friends. We assume that the visibility of running records creates a moral obligation to tell the truth. Studies have suggested that the observability can change individuals' behaviors (Exley and Naecker 2016), and publicly violating prevailing social norms can result in shame or guilt (Bicchieri 2005). We keep individuals who always reveal their records to the public or their friends and run a subsample analysis. Only 6.7%, 5.6%, and 7.3% of individuals in the Baseline Control, Friend Control, and Reciprocity Treatment groups, respectively, posted some private running records during the treatment period. The proportions are not statistically different. After excluding those individuals, we find that the challenge completion rates are 30.2%, 34.3%, and 39.8% in the Baseline Control, Friend Control, and Reciprocity Treatment groups, respectively. Table 3.10 Columns (3) and (4) show the results. Consistent with the main findings, the Reciprocity Treatment group significantly outperforms the Baseline Control group ($p < 0.01$) and the Friend Control group ($p < 0.05$) in terms of challenge completion rate.

3.4.5.2. Retaining Persistent Weak Runners

A potential concern with our treatment is that we only observe individuals' performance for two weeks and use this pre-treatment performance as a criterion (below 40 km) to find the weak runners as receivers. Since it is possible that individuals' running behaviors might fluctuate across time, we might inadvertently

choose a receiver who is actually a capable runner but taking a short break during the pre-treatment period. To address this concern, we trace individuals' running behaviors back to eight weeks (roughly two months) and only retain individuals who registered at least eight weeks before the treatment period. We apply the same criterion and include only those individuals whose average two-week running distance over eight weeks is smaller than 40 km in our subsample. The results of the subsample analyses are shown in Table 3.10 Columns (5) and (6) and are statistically significant. The marginal effects show an 8.0% and a 6.6% difference, respectively. Considering the average completion rates are 29.4% in Baseline Control group and 31.40% in Friend Control group, this is a 27.2% and 21.0% increase. The results indicate that the treatment effect of reciprocity on persistent weak runners is as strong as the effect on the whole sample.

3.4.5.3. Alternative Measures of Closeness

Our theoretical arguments proposed a key conditioning role for the social closeness between the senders and the receivers. We used a multi-dimensional index to operationalize closeness, and tested the robustness of this relationship using alternative operationalization of the index. To further ensure the robustness of this index, we re-estimate the relationship using alternative measures suggested in prior literature. We measure the number of days senders and receivers know each other (*DaysKnowEachOther*[36]) to gauge the length of their friendship (Gilbert and Karahalios 2009) as a proxy for social distance. Chen et al. (2009) also use the length of time participants know the other person to measure the level of closeness of

[36] We take a natural logarithm transformation of this variable to address skewness.

relationships between friends. We calculate the number of days since a sender and a receiver first built a social connection on the platform (either followed or be followed) until the last date of pre-treatment period. Table 3.7 Column (2) shows the results. Consistent with *ClosenessIndex*, *DaysKnowEachOther* shows an inverted U-shaped relationship of reciprocity on completion rate ($p < 0.05$).

A second alternative to conceptualize social closeness is to consider the network structure. The number of common friends between senders and receivers (*CommonFriends*[37]) has long been used to measure the level of trust and cooperation in relationships (see more discussion in Aral and Walker 2014). To measure *CommonFriends*, we compare all the friends[38] of a pair comprised of a receiver and a sender and calculate the *number of common friends* shared by receivers and senders. Table 3.7 Column (3) shows the results. The number of common friends shows a similar inverted U-Shaped relationship ($p < 0.05$).

The Pearson correlation coefficient between *ClosenessIndex* and *DaysKnowEachOther* is 0.21 ($p < 0.01$), and that between *ClosenessIndex* and *CommonFriends* is 0.39 ($p < 0.01$), indicating small to medium strength correlations between *ClosenessIndex* and the other two variables (Cohen 1988). We see that *ClosenessIndex* correlates with these two measures, but not to a large extent. The results suggest that although, as might be expected, the shared social contexts between senders and receivers (*ClosenessIndex*) are positively correlated with other types of interpersonal relationship, they are not indistinguishable from each other. This

[37] We take a natural logarithm transformation of this variable to address skewness.

[38] A friend could be a follower, followee, or mutual follower.

alleviates the potential concern that all three variables measure the same facet of the social closeness between senders and receivers.

3.4.5.4. Falsification Test

Independent of the magnitude, individuals in each of our groups increase their performance, indicating the success of our manipulation. One potential concern might be that the observed improvement is not a result of individuals' intrinsic motivation, but rather caused by unobserved external environmental factors such as good weather. We address this concern as follows. We find one group of users that can suitably serve as a control for such external factors. Because each receiver is paired with a sender; we are able to use the sender as a control group to see whether they also improve their performance during this period. Since the senders' role is to simply forward a message to their friends, it is unlikely that they would change their behaviors because of our intervention. Figure 3.7 shows that the senders do not improve. In fact, senders show a modest decline in performance. The results further support the efficacy of the intervention.

We further confirm if the weather conditions in the pre-treatment and treatment periods are similar. The average temperatures in both periods are 28.9°C (83.9°F) and 28.6°C (83.5°F). The average precipitation hour is 2.8 hours and 2.6 hours for a total of 14 days, and the humidity level is 79.4% and 78.2%. Since the weather conditions are similar between these two periods, the increased distance can be attributed to our treatments with greater confidence.

3.4.6. Post-treatment Surveys

Our main analyses have focused on directly comparing the effects of the treatments through observable, objective measures. While observability increases the rigor of our analysis, it does not provide deep insights into study participants' perceptions about the treatment, such as their motivations to participate in the challenge, the subjective perception of closeness, and their emotions during the treatment period. To gain such insights that may further enhance interpretation of findings, we conducted a post-treatment survey that was distributed electronically to all individuals participating in the challenge. A total of 526 participants completed the survey for an overall response rate of 35.9%; and 33.9%, 39.5%, and 34.3% in Baseline Control, Friend Control, and Reciprocity Treatment groups, respectively.[39] Not surprisingly, individuals who complete the challenges were more likely to respond: the challenge completion rate is 73.0% among all the respondents.

Using the survey data, we first explore what motivates individuals to participate in the challenge. We asked respondents to select all possible reasons from a pick-list of four, five, and seven potential motivations in the Baseline Control, Friend Control, and Reciprocity Treatment groups. On top of the four general motivations such as interesting and healthy, one additional motivation related to social comparison was added in the Friend Control group, and two additional reciprocity related motivations were added in the Reciprocity Treatment group. In the Baseline Control and Friend Control groups, "this challenge is beneficial for my health" ranks as the top reason to participate (62.9% in the Baseline Control and 62.7% in the Friend Control groups),

[39] We provided an incentive for completing the survey (a lottery to win a $200 sports watch) to increase the response rate.

followed by "this challenge is interesting" (50.9% in the Baseline Control and 56.5% in the Friend Control groups). In the Reciprocity Treatment group, these two reasons receive an affirmative response from 53.0% (Healthy) and 49.4% (Interesting), respectively. Besides general reasons to attend the challenge, we added two items that related to reciprocity in the Reciprocity Treatment group: showing gratitude (Palmatier et al. 2009) or avoiding guilt (Dahl et al. 2005). There were 74 votes (44.6%) for "I feel guilty toward my friend who sends me a gold coin if I did not win a gold coin" and 107 votes (64.5%) for "I would like to express my gratitude to my friend who sends me a gold coin" among 166 respondents in Reciprocity Treatment group. Although these motivations are not mutually exclusive, the statistics indicate that the majority of participants in the Reciprocity Treatment group undertake the challenge as a way to express gratitude, which is especially beneficial in our setting because it helps strengthen the social connections between senders and receivers, an essential element for long-term health of participants and for the sustainability of the platform.

To the extent that our study is motivated by the need to induce healthy behavior using interventions that are effective, it is important to understand if there are potential negative side-effects associated with the treatments. One possible side-effect of using the reciprocity to motivate individuals is to incur negative feeling such as stress, frustration, or guilt if receivers cannot complete the challenge. We ask respondents to indicate how they felt during the challenge, including both negative and positive emotions such as finding the challenge stressful, experiencing self-doubt, feeling frustrated, experiencing guilt, and feeling joyful, energetic, and content (Shaver et al. 1987). Independent of their success in completing the challenge, there are no statistical

differences between the Friend Control and Reciprocity Treatment groups for both negative and positive emotions. These results alleviate the problem of possible side effects associated with the use of a reciprocity-based incentive.

Finally, in the main analysis and robustness checks, we used three objective measures to gauge social closeness. In the survey, we complement these measures with the "Inclusion of Other in the Self" (IOS) scale, which provides a subjective assessment of the perceived closeness of a relationship (Aron et al. 1992). We ask respondents (who are the receivers) to select the nature of the relationship between them and their sender "friend" in the experiment from seven graphs (see Appendix Figure A2). In each of the graphs, one circle refers to the participant and the other to the friend who sent the participant a gold coin. Different degrees of overlap correspond to different levels of perceived closeness in the relationship, quantified on a scale of 1-7. We use this measure to test the inverted U-shaped relationship the social closeness on reciprocity. Table 3.7 Column (4) shows the results. The results are consistent; as before, we find an inverted U-shaped relationship ($p < 0.05$). Our finding of the key conditioning effect of closeness and the specific form of this effect is consistent across multiple observable and subjective measures, further enhancing confidence in the findings.

3.5. Discussion

Finding effective interventions that can motivate positive behavior change for health-related behaviors has the potential to enhance economic and societal welfare. Researchers and policymakers are increasingly asking for more evidence in support of what incentives work and under what conditions (Blumenthal et al. 2013). We proposed a relatively cost-effective mechanism, reciprocity, that can be bundled with a traditional

financial incentive based intervention and tested its effects in a large-scale randomized controlled trial with over 1,700 pairs of users. We find that reciprocity outperforms self-interest in terms of challenge completion rates and running distances. We show that a friend's gift and a reciprocity-based incentive (the Reciprocity Treatment group) can boost performance relative to a conventional incentive program (the Baseline Control group) by 32.0%. We further find that, after controlling for the effects of a friend's gift, a reciprocity-based incentive can still improve the challenge completion rate by 20.4% as compared to a self-interest based incentive. Our results provide strong evidence to support the effectiveness of a reciprocity-based incentive, a new form of commitment device for healthy behavior change.

A key result from our study is that the effects of reciprocity and heterogeneously manifest, conditional on the nature of the interpersonal relationship between the gift exchange dyad. We find an inverted U-shaped relationship of the social closeness between senders and receivers on challenge completion rates. We interpret these results through the lens of relationship maintenance. For distant friends, the aspiration to build a long-term relationship is low and for close friends, a strong relationship is already established with or without completing this challenge to maintain it. In this study, we *randomly* pick friends as receivers for senders to test the conditional effects of social closeness. If we did not manipulate the randomization process, a reasonable conjecture is that a sender would choose a friend she/he is close to. However, our results show that such close pairs might not yield the best results because the effects are inferior to moderately close pairs. Through proper training, a machine learning algorithm will help

practitioners to identify optimal pairs to precisely motivate each individual and maximize the effects of reciprocity in real-world settings.

We acknowledge some limitations of our field experiment and discuss promising opportunities for future work. First, the treatment period for our challenge only spans two weeks. This duration might not be long enough to foster a long-term healthy habit, although we note that this observation period is longer and the observed behavior more effortful than the majority of prior studies on reciprocity. While we believe that the need to reciprocate to friends is essential to maintain the interpersonal relationship, whether the same value of gifts and the obligation to return can be generalized to a longer-period challenge is still an empirical question. An interesting question for future research would be to explore if multiple reciprocity-based interventions are unnecessary in order for long-term habits to develop.

Second, while most gift-exchange literature uses a monetary incentive to induce behavior changes, studies exploring the effects of non-monetary incentives have started to attract attention (Kube et al. 2012). Our study used a non-monetary incentive, a virtual gold coin for a lottery drawing to win sports accessories. A lottery represents an uncertain gift with a probability of winning gifts, in contrast to a guaranteed gift such as water bottle. It is not clear whether the effect of reciprocity would be conditional on the types of gifts or what would be an optimal gift to increase individuals' performance. Our study can be extended by changing the value of the gifts (small vs. large gifts) and the types of gifts (monetary vs. non-monetary gifts; certain vs. non-certain gifts).

Third, our experiment was conducted in the setting of an online platform. Although the platform provides detailed information about individuals' digital social

connections and running records, we are not able to observe activities that happened offline. For instance, we are not able to tell whether individuals who fail to reciprocate their friends' gift would compensate their friends offline, such as buying senders a meal. The unobservability of offline behaviors might lead to an underestimation of our experimental results, which is a typical constraint for a less controllable but more realistic field experiment as compared to a lab experiment.

Fourth, from the game design perspective, there are other conditions worthy of study. For instance, our current research design cannot fully differentiate between reciprocity and altruism. Although we argued that the treatment effect is caused by the desire to return the gift out a sense of obligation, we cannot completely rule out the possibility that the behavior is a result of altruism. A future direction to separate these two effects might be to include another group, in which individuals can earn a virtual gold coin and have a right to decide whether they want to keep this earned virtual gold coin, reciprocate back to the senders, or forward to other friends. Relatedly, our design does not allow us to determine whether a sender-receiver (pair) design is better than a non-pair design. For example, one can design another control group to randomly pick individuals to participate in the challenge and win two virtual gold coins (double incentives). The overall spending would be exactly the same as the treatment group's spending, but this design does not use dyads. Future studies could build on the evidence presented here to conduct additional experiments that are able to isolate such effects.

Fifth, our experiment was conducted in physical activity context. While we believe that the underlying psychological and economic mechanisms of reciprocity would hold in other health contexts, we acknowledge that human behaviors could be

strongly affected by many other factors. More empirical evidence is needed to examine whether reciprocity outperforms self-interest in fields such as glucose monitoring or hypertension control.

Finally, since the pairs of users in our platform are already online friends, our test of the curvilinear relationship of closeness is shifted to the right because we are not able to observe a pair of complete strangers. Although we cannot find a pair of complete strangers, the limitation might not be a major concern. In this study, we have shown that a moderate closeness between senders and receivers would be stronger than a distant one. There is no valid reason to believe that the obligation to return favors to total strangers would be stronger than to moderately close friends. Furthermore, to leverage the power of reciprocity, we need to rely on individuals' social connections. A pair comprised of complete strangers would not be suitable to apply a reciprocity-based incentive.

In conclusion, this study harnesses the power of modern wearable and mobile technologies, which are increasingly being used in wellness programs (Claxton et al. 2016; Handel and Kolstad 2017). These devices, technologies, and social networks allow us to precisely track users' behaviors and pair senders and receivers. Leveraging these features, we designed a new incentive based on reciprocity and conducted a randomized controlled experiment to test its effectiveness. We provide evidence that a reciprocity-based design can substantially improve the cost-effectiveness of a conventional incentive program and if identified appropriately, judicious choice of pairs of senders and receivers would further strengthen such effects. From the perspective of both effective resource utilization and improved outcomes, it is

increasingly important to gain a better understanding of how interventions can be more accurately customized for individuals. Our field experiment provides useful evidence of such precision targeting.

Chapter 4: Does Computer Ownership Cause Income Mobility? Spillover Effects and the Role of Caste in India

Abstract

Income mobility is an important metric to assess the equality of opportunity to move along the economic and social ladders, and has significant implications for both public policy and individual well-being. This study uses a newly available and high-quality panel data to study income mobility in India between 2005 and 2011, and the role of computer ownership and caste to explain observed patterns in income mobility of households. We discuss determinants of income mobility and assess the extent to which computer ownership influenced income mobility in India at a time when it experienced significant adoption and diffusion of computers. Our main findings show that households experiencing a positive change in computer ownership have a higher probability of moving upward in terms of their income quintiles, indicating a possible role of technology to overcome the poverty trap, among other reasons. We also find that beyond private returns to focal households, computer ownership also creates social returns to other households in the same neighborhood. Our analyses of information technology (IT) externalities and spillovers suggest that computerization from the same caste group in a neighborhood has a significantly positive spillover effect on upward income mobility. These findings underscore the need to consider social interactions and social embeddedness for a more complete understanding of returns on digitization. We discuss the implications of our findings for research and policies related to digitization, human capital, and income mobility.

Keywords: *Computer ownership, India, Income mobility, Technology literacy, IT spillover, Caste*

4.1. Introduction

Reducing poverty and income inequality have long been major concerns for policy makers. The traditional measures for gauging the distribution of income such as Gini coefficient or Theil index for cross-sectional data (Deininger and Squire 1996; Theil 1967) provide a useful but static view of income inequality. Therefore, increasingly there is a growing attention on income mobility (Chetty et al. 2014; Lukiyanova and Oshchepkov 2012), which is considered a more dynamic way to gauge inequality in income distribution over time. An interesting way to conceptualize income mobility is through Joseph Schumpeter's metaphor of rooms in a hotel (cited

in Sawhill and Condon 1992; Schumpeter 1955, p. 126). Imagine that at time t1, the rooms in the hotel are occupied, but the occupants receive unequal accommodations regarding the quality of rooms. Suppose we revisit occupants who still live in the hotel at a later time t2, mobility will inform us the extent to which occupants had opportunities to move between rooms over time. Fairness means that occupants living in shabby rooms had a chance to move to the ordinary or luxurious rooms, and people living in the luxurious rooms did not have an entitlement to occupy them all the time.

The main constraint to study a dynamic view of income distribution is the paucity of relevant panel data from the same entity at two or more points of time to analyze the change in economic well-being such as income or consumption. It is therefore not a surprise that despite significant progress in understanding the value of IT at multiple levels, we know very little whether computers contribute to income mobility, a critical metric from a policy and individual viewpoint. To the extent many emerging economies are deploying computers and trying to digitize their economy, it is important to know whether computer ownership can help with upward income mobility. This question is particularly important in the Indian context where the government is pushing toward further digitization as part of its "Digital India" initiative (Economic Survey of India 2015; Ghosh 2015; Liu and Mithas 2016; Srivastava and Roche 2015).[40]

In this study, we utilize two recent waves of nationally representative and very high quality Indian Human Development Survey (IHDS) data from India between 2005 and 2011 to construct a dynamic view of income mobility in India (Desai and

[40] Also see http://www.digitalindia.gov.in/content/programme-pillars.

Vanneman 2016; Desai et al. 2016). Our work focuses on micro-level income mobility, which examines income change between two time points for the same households. We focus on the issue of income mobility as it relates to information technology (IT), extending and complementing prior literature which tends to study income mobility considering human capital, social economic status, or labor market (Buchinsky et al. 2003; Cuesta et al. 2011; Dartanto and Nurkholis 2013; Woolard and Klasen 2005).

Our study makes at least two contributions. First, this study is perhaps among the first in the information systems literature to investigate the impact of computer ownership on income mobility after controlling for determinants that have been extensively studied such as human capital and economic status. Prior literature documents private returns regarding wage premiums due to computer literacy mainly from a static viewpoint (Dolton and Makepeace 2004; Krueger 1993); our research extends that to provide evidence from the dynamic perspective. Second, in addition to examining private returns to computer ownership, we view the relationship between computer ownership and income mobility through the lens of the social interaction process.

We pose the following question: Does digitization generate any positive externalities or social returns when it comes to income mobility? In other words, does the proportions of computer ownership in the same community (e.g., district in India) influence focal family's economic mobility? To preview our key results, we find that the computer ownership brings not only private returns to households but also social returns to other households. The direct effect of owning a computer is associated with a 13.4% increase in the probability of moving upward between 2005 and 2011. In

addition, every 1% increase in computer ownership in a district is associated with an increase of 0.11% in upward income mobility, providing support for the notion of positive externalities in computer ownership. For further insights, we decompose the spillover effects in two parts: the effect from the same caste or religious group and the effect from other caste or religious groups. For these analyses, we leverage the presence of caste system in India, a system of social stratification that similar to other social stratification systems limits interactions among different social groups. We argue that IT spillover should be stronger when interacting members belong to the same caste or religious group than when members belong to other caste or religious groups. Our empirical results show support for the argument that an increase in the proportion of households from same caste or religious groups owning computers exerts a positive and strong effect on other households' income mobility. Every 1% increase in owning a computer in a district from the same caste group is associated with an increase of 0.10% in upward mobility. However, the effect of computer ownerships from the other caste groups is not statistically different from zero. A Wald test show a significant difference of computer ownership in a district between the same caste group and different caste groups. Taken together, these results indicate that social returns of computer ownership exist when there is an effective social interaction process.

4.2. Background and Theoretical Framework

4.2.1 Prior Literature

Three streams of literature frame the background of our strudy. First, because of its centrality and policy importance, income mobility has received significant scholarly attention. Researchers have conducted studies in many countries around the

world such as China (Ding and Wang 2008; Shi et al. 2010), Russia (Lukiyanova and Oshchepkov 2012), Britain (Jarvis and Jenkins 1995), France (Buchinsky et al. 2003), and Latin America countries (Cuesta et al. 2011).

Among studies that focus on India, Fields (2007) conducted a comprehensive literature review of income mobility. His review shows that most studies investigated mobility issues in India using three principal data sources: National Council for Applied Economic Research (NCAER) data between 1968 and 1970, the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) data in 1970s and 1980s, and data from Palanpur (a village in Moradabad district of Uttar Pradesh) between 1950s to 1990s. For example, Gaiha (1988) uses NCAER panel data between 1968 to 1970 to show that access to cultivable land and modern agricultural inputs influence upward mobility for cultivating poor. Bhide and Mehta (2004) use NCAER data set in the early 1970s and 1980s to investigate determinants of income mobility such as tribal status, demographic composition, and literacy. Walker and Ryan (1990) use panel data in the 1970s and 1980s from ICRISAT to examine how the education of father, caste, inherited land, quality of inherited land affect respondents' income mobility. Lanjouw and Stern (2003) use data between 1950s to 1990s from Palanpur, to investigate the role of growth of the rural non-farm economy in determining income-poverty and improvement in living standards of the poor. More recently, Gautam et al. (2012) use Additional Rural Income Survey (ARIS) and Rural Economic and Demographics (REDS) surveys in 1999 and 2007 to examine income, consumption, and asset mobility. The mobility in their research was associated with land classes, agro-climate zones, agricultural profit classes, assets classes, caste groups, and gender groups. For

instance, landless or marginal farmers were found to be the most vulnerable groups regarding income or assets mobility.[41]

Second, previous literature discusses how access to information technologies helps to alleviate various societal problems such as digital inequality, mortality, and well-being of the general public. Government initiatives play critical roles to reduce the problem of digital inequality for socioeconomically disadvantaged groups in both developing and developed countries (Hsieh et al. 2011; Venkatesh and Sykes 2013). Hsieh et al. (2011) argue that social capital influences adoption of information technologies, especially for socioeconomically disadvantaged groups. Researchers argue that access to digital resources might reduce costs and improve health outcomes (Agarwal et al. 2010). For instance, Venkatesh and Sykes (2013) show that IT intervention reduce infant mortality rate and argue that IT could complement other investments such as medical facilities. Recent work suggests that IT empowers individuals living in the rural areas in the developing countries to improve their living standards (Jha et al. 2016; Leong et al. 2016). This is because IT increases individuals' income by creating opportunities to access new market (e.g. e-commence, Leong et al. 2016), or a variety of financial services (e.g. mobile money ecosystem, Kendall et al. 2011).

Third, increasingly spillover effects or social returns to IT are beginning to receive attention in recent years. Previous literature investigates knowledge or technology spillover effects in terms of productivity or innovation at several levels such

[41] Table A3 in Appendix provides a brief review of these studies and how our study makes unique contributions.

as firms, industries, or countries. Job hopping, foreign investment, or technology diffusion are possible mechanisms to transfer knowledge from one entity to another. For instance, Mun and Nadiri (2002) investigate information technology externalities between 1984 and 2000 in the United States and find that computerization of one industry's customer and suppliers industries lead to a reduction in both labor and material costs of the focal industry. Han et al. (2011) show that IT investments from suppliers industries lead to an increase in the productivity of downstream industries, and find that IT intensity and competitiveness of downstream industries moderate such spillover effects. Chang and Gurbaxani (2012) analyze the spillover effect from an outsourcing perspective and show that firms would benefit from IT service firms because of their accumulated knowledge. The underlying reason is that IT-related knowledge makes the external vendors achieve higher levels of efficiency in delivering IT services. Cheng and Nault (2007) investigate the effects of information technology investments made upstream on the effects to downstream productivity by arguing that the competition on the supply side will generate extra benefits that can be passed on downstream. Their empirical results show that a 10.5% increase in suppliers' IT capital lead to an increase of suppliers' output by 0.63% to 0.70%. Tambe and Hitt (2013) argue that the mobility of IT workers generates significantly positive spillover effects on firm productivity. In short, previous literature discusses IT spillover effects mostly from a productivity perspective whereby firms increase productivity through an accumulation of technological know-how, and IT spillover increases the output productivity through improved products or services.

Our review of the prior literature suggests that few prior studies have examined income mobility from an IT perspective, even though there is a vigorous debate in the academic literature (Acemoglu 2003) and practitioner press (Long 2015; Rotman 2014) on the role of IT in influencing income inequality and living standards. In addition, most studies used data before the year 2000 in India and mainly focus on rural areas, and hence they have little to say whether recent adoption and diffusion of computers in the Indian economy (Nilekani 2008; Parker et al. 2016) may have contributed to income mobility, something that we address in this study.

4.2.2 Hypotheses

Wider adoption of personal computers and related technologies in the last three decades has improved the way people communicate, learn, and earn a living. There are more than 2 billion personal computers in use globally in 2015.[42] Information technologies connect populations beyond geographical and time constraints and bring radical transformation to the society.

Why will computer ownership affect income mobility? First, computer ownership can lead to an increase in computer literacy, which may increase wage premiums in the long run. Chua et al. (1999) show that computer experience can be accumulated through owning a computer at home. Venkatesh and Vitalari (1987) show that continuous use of computer through adequate training can reduce deficiency of computer knowledge, and find that early adopters of computers at home extensively use word processing and business applications for work-related tasks, i.e. individuals

[42] Global installed base of PC from 2005-2015. Retrieved from http://www.statista.com/statistics/203624/global-installed-base-of-pcs/

gain digital literacy from computer ownerships and gain skills to handle work-related tasks and business applications. Such accumulation of computer literacy can result in salary premium at work. For instance, Krueger (1993) used Current Population Survey (CPS) to evaluate the returns to computer use, and show that workers who know how to use computers earn 10% to 15% more than workers have no computer literacy. The direct benefit of computer on wages has been documented in mostly cross-sectional data in various contexts such as West Germany, Canada, the Netherlands, and the United Kingdom (Arabsheibani and Marin 2006; Oosterbeek 1997; Reilly 1995).

Second, computers ownerships at home can reduce computer anxiety and provide flexibility in work-related tasks. Ogunkola (2008) argues that computer ownership can build confidence that mitigates fear and anxiety about computers. Computer anxiety is an aversion of interacting with computers or thinking about computers (Poynton 2005), which is especially common in middle-aged adults and elderly (Dyck and Smither 1994). Fear of computer leads to an inferior status for individuals in a digital world. For instance, Link and Marz (2006) show that students who lack basic computer skills were skeptical about e-learning and refused to use the existing e-learning offerings at one university, and in turn lost the opportunity to reap benefits from the e-learning systems. Furthermore, information technologies make it feasible to supplement one's income by working at home. Venkatesh and Vitalari (1992) show that supplemental work by using computers at home has a positive effect on flexibility and household income and is negatively related to commuting time.

The differences between computer owners and non-owners might persist over time because of the accumulated computer literacy that reinforces individuals'

computer skills. Computer expertise gained from computer use at home may affect computer use at future workplace (Facer et al. 2001). Therefore, we hypothesize that a household with a computer would enjoy a higher income than a household without a computer. Such effect would lead to a disparity of income between computer owners and non-owners, and may affect their socio-economic outcomes over time. Therefore, we hypothesize:

H1: A positive change in computer ownership leads to upward income mobility.

We argue that computerization in a region leads to positive social returns for the following reasons. First, computer diffusion is an engine of growth and a source of labor market changes. When technology evolves, it spreads throughout the economy and brings generalized productivity gains (Bresnahan and Trajtenberg 1995). David et al. (1997) show that the spread of computer technology explains the increase of growth in the relative demand for more-skilled workers. Therefore, regional computer adoption rate leads to productivity gains and a possible change in local labor market, bringing long-term economic benefit to the regions. Furthermore, information technology fosters human capital. Studies of human capital show that an accumulation of human capital of a person not only affects the earnings for that individual but also the earnings of other individuals (Manda et al. 2002). Moretti (2004) suggests that increasing the aggregated stock of human capability in one region would lead to a positive effect on society that was not fully reflected in the private returns of individuals.

Second, computerization generates knowledge spillover, which is key to generate social returns in a region. Knowledge diffusion plays a critical role in increasing growth. For example, R&D investments and the production of knowledge

create not only private returns for the investors but also create a long-term social return by growing the economy (Jones and Williams 1998). Leon-Ledesma (2005) shows how international trade drives R&D diffusion and productivity growth. Acs et al. (2009) discuss the positive relationship between knowledge and the degree of entrepreneurship, and demonstrate that knowledge generated endogenously results in spillover effects that allow other individuals to identify and exploit opportunities. Likewise, computer ownership can increase owners' digital literacy, which flows among individuals in a region raising the level of local computer knowledge. Since digital literacy has been linked to an increase of individuals' wages (Krueger 1993), we argue that computer knowledge spillover effects derived from the growth of computer ownership would also increase income mobility in the neighborhood due to spillover effects.

Third, learning from earlier adopters reduces costs for the late adopters. Heckman (2000) argues that learning begets learning and earlier technology adopters make the late adopters easier to learn a new technology. Such learning processes can be found in various contexts. For instance, Saxenian (1996) shows that individuals learn from each other when moving between firms and industries in Silicon Valley. An increase in computer ownerships in a region makes it easy for late adopters to imitate and shortens the time to diffuse computer knowledge. Furthermore, the spillovers of technology not only benefit the late adopters but also the early adopters. For example, interoperability of computer software brings benefits to both existing adopters and future adopters (Katz and Shapiro 1986). An increase of computer ownership in a region reduces the costs of learning and facilitates knowledge transfer, leading to an

acceleration of digital literacy. Knowledge flows from highly endowed individuals to less endowed individuals create a positive spillover effect from technology owners to non-technology owners. Therefore, we hypothesize:

H2: There is a positive spillover effect of computers ownership in a neighborhood on upward income mobility.

One key reason for the spillover effects is the process of social interactions. Han et al. (2011) argue that positive IT spillover effects in the form of IT-enabled innovations that be transferred to downstream industries are enabled by interactions among industry players. It is well-known that one's social connections affect one's behaviors in various setting. Individuals' social networks would affect ones' access to various resources, such as finance (Aldrich et al. 1987), information, or labor resources (Aidis et al. 2008). For instance, Bandiera et al. (2010) demonstrate how workers' behaviors are affected by their socially connected friends. Mas and Moretti (2009) study the peer effect in a supermarket and find that there is a positive spillovers effect of introducing a highly productive staff into the chain.

Since social connections affect one's decisions and outcomes, we expect caste in India to play an important role in how IT spillover operates when it comes to income mobility. The word caste derives from Portuguese casta, meaning breed or race, and caste has long existed in India. Caste is a form of a social stratification system that has gone through significant changes since India's independence but still plays a significant role in modern Indian society. Desai and Dubey (2012) discuss caste and social distance and depict the relationship among castes as follows: "there is an increasing tendency to view caste as a form of ethnicity in which castes compete with each other for power

and proudly brandish their own narratives of origin (p. 7)." The caste stratification in India has also influenced other religious or ethnic groups such as Muslims, Christians, and Sikhs (Desai and Kulkarni 2008).

Caste affects not only dietary habits, educational opportunities, economic advantages, but also interactions with members of other castes as well. Gille (2011) uses social caste in India to construct weighting matrix to examine education spillover effects on farm productivity and finds that caste is a valid indicator for social interactions because households are influenced more by the educational level of neighbors from the same caste than from other castes. This argument builds on the assumption that a close neighbor has a stronger influence than distant ones (Goux and Maurin 2007). People are influenced differently by their close caste groups. The spillovers effect would happen when individuals interact with each other, and the caste system can potentially impede interactions and communications among members of different castes while facilitating those among members of the same caste or religious groups. Thus, households are more likely to be influenced by neighbors in the same caste group through daily interaction. Therefore, we hypothesize:

> *H3: Compared to computer ownership from other groups in a neighborhood, there is a positive spillover effect of owning computers from the same group in a neighborhood on upward income mobility.*

4.3. Method

4.3.1 Data

We use two waves of a nationally representative survey from India: Indian Human Development Survey (IHDS) 2005 and IHDS 2011. These two surveys were

conducted by National Council of Applied Economic Research (NCAER), New Delhi in collaboration with the University of Maryland in 2004–2005 (Desai et al. 2016) and 2011-2012 (Desai and Vanneman 2016). The data necessary for our study became available in 2015 when IHDS released data from the second wave of the survey, and we use the latest versions of data released in 2016. Respondents in this survey are typically the household head or a knowledgeable person who knows about the household economic situation, and the response rate of IHDS is as high as 92% (Sonalde et al. 2010). Both surveys contain a rich amount of information, including household employment, total income, consumption expenditure, social network, residence, caste and religion group, household size, etc. The IHDS 2005 survey covers 26,734 households (143,374 individuals) and 14,820 households (72,380 individuals) in rural and urban areas respectively and covers 1,503 villages and 971 urban neighborhoods across India. About 83% of these households were resurveyed in 2011-2012. The IHDS 2011 survey consists of 14,573 households (135,118 individuals) and 27,579 households (69,450 individuals) in rural and urban areas respectively.

The high quality of IHDS data make it comparable to other datasets as Census and NFHS, and IHDS data have been used in several studies (Adams 2008; Desai and Adams 2006; Dutta 2015; Frijters and Sharma 2012; Sen and Noon 2007; Vanneman et al. 2008) to investigate poverty rate, education disparity, and learning outcomes. To create a balanced panel between 2005 and 2011, we take steps to cleanse data and dropped observations with missing data. The final sample we constructed is a balanced one by matching the household ID from two waves of surveys. The sample covers 64,428 households across two years (32,214 in each year) for both rural and urban areas

in 33 states.

4.3.2 Variables

Our unit of analysis is a household. The *Household income* measure is derived from more than fifty questions, including agricultural, non-agricultural, pensions, and government benefits. We deflate household income in 2011 back to 2005 equivalents by using Consumer Price Index (CPI).[43] Merely using the household income without considering the household compositions can be misleading. Therefore, we follow previous literature in poverty and wealth literature and account for household size and composition (Lukiyanova and Oshchepkov 2012; Roberts 2001; Woolard and Klasen 2005) in our models. We use the well-known adjusted household income, Adult Equivalent Income (AEI), to construct our key dependent variable, calculated as follows:

$$\text{Adult Equivalent Income (AEI)} = \frac{Household\ income}{(Adults + 0.5\ X\ Children)^{0.9}} \qquad (1)$$

Where *Adults* means the number of adults in the household older than 18 years of age, *Children* is the number of children whose age is below 18. We follow Woolard and Klasen (2005) to set a scaling parameter as 0.9 to reflect modest economies of scale. We take a log of the Adult Equivalent Income (AEI) variable.

We used transition matrices to examine the movement of income over two-time points. We divided households' incomes into m groups (deciles or quintiles) in both periods. Let P be a matrix of m X m transitions, P_{ij} means the percentage of units whose income class fall in class i at first period and j at second period. Therefore, in the

[43] Inflation, consumer prices (annual %) http://data.worldbank.org/indicator/FP.CPI.TOTL.ZG

transition matrices, the rows represent initial income classes at the first period and the columns indicate the final income classes at the second period. P_{ij} $(i = j)$ represents the percentage of households who still remain in the original income class, which are known as 'no mobility (immobile)', while P_{ij} $(i \neq j)$ represents households move from original income class to a new income class, which are called 'mobiles'. Among those mobiles units, households move from i to j ($i<j$) are regarded as moving upward and households move from i to j ($i>j$) are regarded as moving downward.

$$P = \begin{bmatrix} P_{11} & \cdots & P_{1m} \\ \vdots & \ddots & \vdots \\ P_{m1} & \cdots & P_{mm} \end{bmatrix} \tag{2}$$

We categorized household into a quintile (m=5) as suggested by Woolard and Klasen (2005). Therefore, each household in a given year is sorted from top to bottom, and categorized into one out of five categories. Based on the transition matrix, we constructed our main dependent variable, Income Mobility (*Mobility*). *Mobility*, which equals to 0, represents households becoming worse-off (P_{ij} $(i > j)$); *Mobility*, which equals to 1, represents households staying the same in terms of their stratum (P_{ij} $(i = j)$); and *Mobility*, which equals to 2, represents households becoming better-off (P_{ij} $(i < j)$).

Table 4.1 provides a summary of variables used in this study. The variables can be categorized into two broad types: one for endowment effect in the base year (2005), and one for change, which was calculated by subtracting value in 2005 from the value in 2011.

The average *Mobility* approximately equals to 1 because we observe the same households across years (the percentage of households moving upwards equal to the

percentage of households moving downwards). Because we deflated household income by CPI, the positive change in AEI indicates an upward movement or economic progress.

In our dataset, only a very small percentage of households (0.9%) owned a computer in 2005, but we observed a surge of owning a computer (6.0%) from 2005 to 2011. We also calculated the proportion of households who own computers at the district level (the district level in India is similar to a county in the United States). We further calculated the proportion of technology ownership at the district level based on the households' castes and religions. We computed two sets of figures for each household: the proportion of households owning computers in the **same** caste or religious group, and the proportion of households owning computers in **other** caste or religious groups. According to IHDS's survey, a household could belong to one of the following caste or religion: Brahmin; Forward caste; Other Backward Classes (OBC); Dalit; Adivasi; Muslim; Christian, Sikh, and Jain.

Table 4.1 provides other important indicators such as human capital and economic status variables. There were roughly 25.4% of persons in a household have jobs in 2005. On average, there were approximately 5.9 persons in a household. The percentages of male adults and female adults in a household fell between 32.7% and 32.5%. Household heads' age was approximately 48.2 years old with only 9.1% of households' heads were female. Our sample households experienced a minor increase of the share of individuals in a household having a job (4.2%). The increase of share of male and female adults was quite minor, ranging from 1.1% to 4.3%. There were roughly 8.4% of households experiencing a change from male to female household

head. We controlled households' economic status. Households, on average, had 7.6 durable goods. There were 50.2% of households who own land, and 27.5% of households lived in urban in the base year. We also examined the extent of social harmony in a region. IHDS survey asked the respondents whether people generally get along with each other well or have a conflict with others. On average, 54.9% of households indicated that their villages or neighborhoods got along well.

Tables 4.2 provide a more comprehensive picture of a household owning computers in 2005 and 2011. The computer ownerships saw a significant increase in the period. Among those who did not have a computer in 2005, there were 6.0% of households owning a computer in 2011. Among those who did own a computer in 2005, a modest proportion of household report not having computers in 2011 (33.4%). Table 4.3 shows detail information of computer ownership and income mobility among castes and religious groups. All groups experienced a sharp increase in computer ownership from 2005 to 2011. Among them, OBC, Dalit, and Christian, Sikh, Jain have higher income mobility than the other groups.

4.3.3 Modeling Determinants of Mobility Change

Following Dartanto and Nurkholis (2013), we use ordered logit model to investigate determinants of mobility change given that the dependent variable (*Mobility*) is inherently ordered. The dependent variable represents the change of mobility for a household based on the transition matrix (moving upward, no mobility, or moving downward).[44] The initial value related variables represent households'

[44] We construct a more granular mobility index by subtracting Quintile position in 2005 from Quintile position in 2011 to create a dependent variable with 9 levels of mobility. For simplicity, we present results only 3 levels of mobility. Both dependent variables yield similar results.

conditions in the base year and are useful to examine whether such endowment effects in the base year persist across time. The change-related variables represent changes in households' assets as well as demographic composition and employment, which might be reasonably be seen as exogenous to the mobility (Woolard and Klasen 2005).[45]

To test our hypotheses, we use the following models:

$$Y_i^* = \alpha + \beta_1 * COC_i + \beta_2 * COE_i + \Sigma\beta_3 * X_i + \varepsilon_i \quad (3)$$

$$Y_i^* = \alpha + \beta_1 * COC_i + \beta_2 * COE_i + \beta_3 * DISTCOC_i + \beta_4 * DISTCOE_i + \Sigma\beta_5 * X_i + \varepsilon_i \quad (4)$$

$$Y_i^* = \alpha + \beta_1 * COC_i + \beta_2 * COE_i + \beta_3 * DISTSGCOC_i + \beta_4 * DISTSGCOE_i + \beta_5 * DISTCOITC_i + \beta_6 * DISTOGCOE_i + \Sigma\beta_7 * X_i + \varepsilon_i \quad (5)$$

Where

Y_i^* is the unobserved dependent variable such that:

$Y_i = 0$ if $Y_i^* \leq \mu_0$ (0 represents households' welfare become worse (P_{ij} $(i > j)$)),
$Y_i = 1$ if $\mu_0 \leq Y_i^* \leq \mu_1$ (1 represents households' welfare remain the same (P_{ij} $(i = j)$)),
$Y_i = 2$ if $\mu_1 \leq Y_i^* \leq \mu_2$ (2 represents households' welfare become better (P_{ij} $(i < j)$))

Each major construct could be decomposed into two parts: the change between 2005 and 2011, and the endowment in 2005. While the change-related variables are variables of interest, the endowment related variables are used to control initial status. COC_{i} is a positive change in computer ownership, which means a household has no computer in 2005 and owns it in 2011. COE_{i} indicates computer ownership in 2005. Similarly, $DISTCOC_i$ and $DISTCOE_i$ represent the change and endowment of proportion of households owning a computer in a district. $DIST\boldsymbol{SG}COC_i$ and $DIST\boldsymbol{SG}COE_i$ represent the change and the endowment of proportion of households in the ***same caste group*** owning a computer in a district. $DIST\boldsymbol{OG}COC_i$ and $DIST\boldsymbol{OG}COE_i$

[45] We do not include changes in physical assets in the model because it is likely to be endogenous. We deal with the selection issue of owning a computer in the later section.

represent the change and the endowment of proportion of households in ***other caste groups*** owning a computer in a district. X_i is a vector of both the change and endowment of all the human capital, and economic variables, including Adult Equivalent Income (AEI), household size, share of persons has jobs in a household, share of male adults, share of female adults, household head's age, household head's gender, number of durables owned by a household, land owner or not, and living in the rural or urban area. In the ordered logit model, ε_i has a standard logistic distribution.

4.3.4 Selection of Computer Ownership

The independent variable, computer ownership, can potentially suffer from a self-selection problem because households made decisions based on their needs and preferences. Therefore, the self-selection problem can lead to potentially biased statistical inferences. When the endogenous variable is binary, it is common to use the Heckman two-stage model (Heckman 1976) to address the endogeneity associated with computer ownership. More specifically, we use an endogenous treatment regression to solve this issue. The endogenous treatment regression will first have a selection equation using a Probit specification and then a second-stage for outcome equation. It allows a correlation structure between the unobservables that affect the treatment and the unobservables that affect the potential outcomes:

$$Y_i^* = \alpha + \beta_1 * COC_i + \beta_2 * COE_i + \Sigma\beta_3 * X_i + \varepsilon_i$$

$$COC_i = \begin{cases} 1, & if\ \boldsymbol{\gamma Z_{it}} + u_i > 0 \\ 0, & otherwise \end{cases} \quad (6)$$

where $\boldsymbol{Z_i}$ include one variable that affects households' decision on computer ownership but does not directly influence households' income mobility and other control variables that are used in the second-stage. The error terms ϵ_{it} and u_{it} follow a bivariate normal

distribution. The exogenous variable is the distance in kilometers to the nearest government or private Higher Secondary/Inter College. Individual who live close to a high school have a higher probability to access school resources. Pal et al. (2007) indicates that "computer-aided learning (CAL) projects are active in over 20,000 public primary schools in India. … include a computer center with 3-5 machines set up per primary school." These computer centers together with school libraries complement regional and national library systems in India (Mahajan 2010), and may provide access to some digital resources for public use. This increase in digital literacy can lead to an increase in computer ownership, which satisfies the requirement of a valid exogenous variable. Distance to school has been extensively used as an exogenous variation in the returns to education literature (Currie and Moretti 2003; Kane and Rouse 1995). There is no reason to believe the distance to school would be directly correlated with income mobility. Carneiro et al. (2016) argue the distance variable was not endogenously determined and showed that "distance to the nearest secondary school is uncorrelated with direct determinants of wages" (p.4).

4.4. Results

4.4.1 Income Mobility

Table 4.4 presents a two-stage view of income mobility in India between 2005 and 2011. The quintile mobility matrix is derived from sorting AEI in the samples in 2005 and 2011. Quintile 1 means the lowest tier (bottom 20%), and quintile 5 represents the highest income class (top 20%). We find that 35.5% of households in the lowest quintile in 2005 remain the poorest in 2011, and 51.8% of households stay in the richest quintile. The lowest 35.5% of households are trapped in poverty, but another 26.7% of

the lowest quintile moves one quintile and another 37.8% moves to the middle or top quintile in 2011. Given that the top and bottom quintile households can only move down or up in this transition matrix, it is not surprising that the top and bottom quintile in the trace is higher than the middle ones. Table 4.5 shows the distribution of the dependent variable: households' income mobility. We find that approximately 33.2% of households stay in the same quintile from 2005 to 2011.

4.4.2 Determinants of Income Mobility

Hypothesis 1, which argues that positive change in computer ownership is associated with upward mobility, is supported. Column 1 of Table 4.6 shows the result of the ordered logit model, where we examine the direct effect of positive change of owning a computer, and control for the endowment effect of computer ownership in the base year. The positive sign of computer ownership in the base year indicates a possible technology impediment for poverty trap. It shows household with an endowment of technology takes advantage of technological knowledge to advance their income level. Change in computer ownership leads to upward mobility. The positive sign indicates households gaining computer ownerships enjoy upward mobility. We calculate the marginal effects to interpret the results. Households with a positive change in computer ownership experience an increase of 13.4% in upward mobility. As we argued in the previous section, owning a computer might increase computer literacy and help with upward income mobility if household members leverage the benefits from computer use.

We find support for Hypothesis 2, which posits that computer ownership has positive externalities on income mobility. We consider the spillover effect of computer

ownership at the district level. Column 2 of Table 4.6 shows the spillover effect and shows that an increase of computer ownerships in a district is associated with an increase in income mobility. To better interpret the results, we calculate the marginal effect of computer ownership at the district level. Every 1% increase of owning a computer in a district leads to an increase of 0.11% in upward income mobility.

Hypotheses 3 argues that households gain positive spillover effects of computer ownership in a district from the same caste group than other caste groups, and we find support for the hypothesis. Column 3 of Table 4.6 shows that computer ownership from the same group has a significantly positive effect, yet computer ownership from the distant groups has an insignificant effect. A Wald test show a significant difference in computer ownership in a district between the same caste and different caste groups (probability > chi^2 is smaller than 0.01). Every 1% increase of owning a computer in a district from the same caste group leads to an increase of 0.10% in upward mobility.

Other variables in our models show relationships along expected direction, consistent with prior literature. From Table 4.6, we observe a negative relationship between initial income and subsequent income change. Such phenomenon of moving toward the mean has been extensively documented in previous research (Shi et al. 2010; Woolard and Klasen 2005). Households having more number of household members would have a negative effect on subsequent income change, suggesting that larger households might have a harder time to improve their economic position. Similar findings are reported by Haddad and Ahmed (2003) and Woolard and Klasen (2005).

4.4.3 Alternative Models Using Changes in Adult Equivalent Income

As suggested by Woolard and Klasen (2005), an alternative model to measure

household mobility can be "derived directly from the standard household utility maximization model with adult equivalent household income as a money metric measure of utility (p. 878)." Following Woolard and Klasen (2005), we use the change of the adult equivalent income (AEI) between 2005 and 2011 as an alternative dependent variable to run parsimonious OLS models, which means we reproduced the models from equation (3) to equation (5) by using ΔY_i as dependent variable, where ΔY_i = log(AEI_{iy}) – log (AEI_{ix}), and x = 2005 and y = 2011.

Table 4.7 show the result of the models. Similar to previous ordered logit models, households experiencing a change in owning a computer show a positive effect on change of AEI. Furthermore, we find support for IT spillover effects at the district level, and the effects mainly come from the increase of computer ownership in a district in the same caste group.

4.4.4 Correcting the Self-selection of Computer Ownership

To correct the self-selection problems associated with computer ownership, we employ two strategies. First, we test the endogenous treatment regression model as specified in equation (6). Table 4.8 shows the results. We test two dependent variables: Mobility and $\Delta lnAEI$. Column 1 and Column 3 show the coefficient of the selection equation, while Column 2 and Column 4 present the coefficients of the outcome equation. As expected, distance to high school negatively correlated with the adoptions of computers ($p < 0.01$). The longer distance to high school, the less likely a household positively changes to own a computer. Our main results hold after accounting for self-selection problems, and change in owning a computer leads to an increase in Mobility and $\Delta lnAEI$ ($p<0.01$). We use a Wald test of ρ, which rejects the null hypothesis of no

correlation between the errors in the selection equation and the errors in the outcome equation. This shows that computer ownership does suffers from self-selection issues and therefore needs to be corrected.

Second, the self-selection problem results in a non-randomized sample. We use a propensity score matching (PSM) approach to address this problem (for further details on propensity score approach, see Mithas et al. 2006; Mithas and Krishnan 2009; Rosenbaum and Rubin 1983). We use PSM to match households that choose to own a computer and households that chose not to do so. We compare households in our sample to ensure that treated and untreated subjects have similar propensity scores in terms of their observed characteristics. To estimate the propensity score, we employ a probit model in the first stage to calculate the probability of a household owing computer with observable variables such as wealth index (number of durables), the number of individuals in a household, and years of education. We perform a one-to-one nearest-neighbor matching without replacement and identify control and treatment observations on a common support region. After matching two samples, there are 1,912 observations included from treated and untreated groups respectively. Table 4.9 shows that the average treatment effect on the treated (ATT) of household mobility is approximately 0.120 with a t-statistics of 5.37. This indicates that households with a positive change in computer ownership experience upward mobility compared to households who did not experience a positive change in computer ownership. Table 4.10 provides a broadly similar result as indicated by Table 4.9.

4.4.5 Robustness Checks

We conduct a series of additional robustness checks. First, we use a reduced

sample to test our model to rule out the possibility of outliers generated by small villages in rural areas or small neighborhoods in urban areas. Therefore, we keep villages or neighborhoods that have more than 57 households (excludes 0.5 standard deviations less than the mean[46]) to do a robustness check. The reduced sample has 27,772 households in each year. Table 4.11 presents our findings. Similar to what we discussed in the previous section, the results in Table 4.11 fully support H1 to H3.

Second, one might argue that the positive network externalities of computer ownership are merely a representation of economic growth. It is the economic growth that leads to an upward income mobility, and the increase of computer ownership in a region is only a byproduct. To alleviate this concern, we create a variable, regional economic growth, and control for this variable. We calculate the average increase of household income in a district level and then take a log form of this variable. Table 4.12 presents the results. As expected, the regional economic growth leads to a strong upward income mobility. All the hypotheses are still supported even after controlling for regional economic growth in each district.[47]

Third, we argue that owning a computer at home leads to an increase in computer literacy, and in turns benefit individuals at work. Since home computers are considered a shared good because they are often shared by multiple users[48], the benefits brought by computer ownership might be larger when there are more individuals with

---

46 Using other thresholds such as excluding one standard deviation yield similar results.

47 Besides calculating regional economic growth using our sample from the same questionnaire, we also collected data on gross state domestic product (GSDP); our findings remain unchanged even when we control for state-level economic growth.

48 For instance, one survey showed that 30% of the respondents share their computers with their children (Furnell et al. 2007).

jobs in a household. Observing family members interacting with computers may lead to a reduction in fear to use computers (Beckers and Schmidt 2001). With a higher proportion of individuals having a job, family members who face computer related challenge could help each other and create an environment that facilitates digital literacy among family members.

Table 4.13 shows the results to test these arguments. The interaction term involving computer ownership and share of persons with jobs is positive and statistically significant ($p<0.01$), implying that a household with higher shares of persons with jobs experiences a higher level of upward mobility when it experiences a positive change in computer ownership. Figure 4.1 shows the moderating effects visually. Households might have three outcomes: moving upward (green area), no mobility (gray area), and moving downward (red area). As the share of persons with jobs in a household increases, the probability of a household's moving upward also increases. However, the slopes of moving upwards are different between households having negative or no change of computer and households having a positive change in computer ownership. For example, households having negative or no change of computers have 27.4% of chance to move upward when only having 20% of individuals having jobs in a household, the chance increase to 35.8% when having 80% of individuals having jobs in a household. This indicates a (35.8%-27.4%) / 27.4% = 40.5% of increase. On the contrary, households having a positive change of computers experience an increase from 45.5% to 71.6% when moving from 20% to 80% of individuals having jobs in a household. This shows a (71.6%-45.5%) / 45.5% = 57.4% of increase. The results echo the positive sign of moderating effect between computer

ownership and share of persons with jobs in the previous table.

Finally, because IT spillover effects depend on extensive social interactions among individuals, such effects will be stronger in an environment where people get along well with each other than in an environment with conflicts. Columns 1, 2, and 3 in Table 4.14 provide evidence that supports this argument. In Column 1, the interaction term between the change of computer ownership in a district and get along well is positive and statistically significant at 1% level. This means the marginal effects of the regional IT spillover are conditional on the level of regional harmony. While the marginal effect of every 1% increase of IT spillover leads to a 0.02% increase in upward mobility in a region where harmony level ranks at a 25 percentile, the marginal effect of IT spillover could be as high as 0.32% in a region where harmony level ranks at 75 percentile. In Column 2, the interaction term between the change of computer ownership in a district from the same group and get along well is not significant, probably because individuals have a strong bond for the same group regardless of the conflict level in a district. Therefore, the IT spillover effect is not conditional on the regional harmony. In contrast, in Column 3, the positive significance of the interaction term indicates IT spillover effect exists even from distant caste groups when individuals get along well with each other. A similar calculation of the marginal effects shows that the IT spillover effect from distant groups is -0.11% in a region where harmony level ranks at a 25 percentile but is 0.12% in a region where harmony level ranks at a 75 percentile supporting the argument that the IT spillover effects would be amplified under harmonious environment.

4.5. Discussion

4.5.1 Main Findings

Our goal in this study was to assess the extent to which computer ownership leads to upward income mobility. We examined the role of IT in income mobility, accounting for other well-documented factors such as initial income, and socio-economic status. Our models suggest that households experiencing a positive change in owning a computer experience upward mobility between 2005 and 2011. We also find that households with a positive change in computer ownership enjoy an even higher upward mobility when they have a higher share of persons with jobs, probably because job market rewards individuals with computer skills and households' members benefit from an increase in computer literacy.

Furthermore, our findings suggest that computer ownership generates not only private returns to households but also social returns to their neighborhoods. Our focus here was on the indirect effect of computer ownership or spillover effects, which are externalities that are experienced by unrelated third parties. The positive externalities would flow from high knowledge endowed users to the low ones through social interaction. Our results show support for the positive spillover effects of computer ownership. We use caste as an indicator of social distance given the role that caste plays in India as a distinct social stratification system, and show that the computer ownership in the same caste group generates a stronger positive spillover effect on income mobility than that from other caste groups in a district. We also show that a friendly community facilitates the spillover effect, and such an effect applies even from distant caste groups. Since the knowledge transfer would be smoother under harmonious environments, the IT spillover effects are amplified when people get along well with

each other. We provide a series of robustness checks, which include alternative models, endogenous treatment effects, propensity score matching, a reduced sample that excludes outliers, and the control of economic growth. Together these models provide confidence that owning a computer leads to upward income mobility.

In summary, we document a positive direct effect of digitization on income mobility, suggesting that information accessibility plays an important role in household's economic status and mobility. Besides the private return of digitization, we also see a social return of digitization in the form of generating positive spillover effect. We further decompose the spillover effect into proximate and distant ones based on the social stratification system in India. The results suggest that digitization generates a positive spillover effect, plausibly through intensive social interaction. As we mentioned in the previous section that caste system, a socially stratified system, in India could form communication impediment among individuals from different groups, which in turn yields a positive spillover effect from the same group but the insignificant effect from other caste groups.

Our research contributes to a broad stream of literature that relates to computer literacy and digital divide (Hsieh et al. 2008; Srivastava and Shainesh 2015; Venkatesh and Sykes 2013). We extend from a static view of computer ownership to a dynamic view across time, and examine divergence of income mobility between computer owners and computer non-owners. Owning a computer at home provides benefits at work (Facer et al. 2001) because it increases computer literacy and reduces computer anxiety (Chua et al. 1999). We find that computer endowment might be related to income mobility, creating a poverty trap for families with less digital resources. We

also find positive change in computer ownership leads to upward income mobility. Extending from a one-period snapshot of computer impacts on income, our work incorporates a time dimension to provide a dynamic view of such effect. The results indicate such effect might persist over time, increasing the income disparity between technology owners and non-owners, creating digital-divide.

4.5.2 Implication for Research

Our findings suggest implications for further research. First, the benefit of IT on improving the productivities of organizations has been well established in the IS field but studies that discuss the linkage between IT and income mobility are scant. To the best of our knowledge, this research is among the first to analyze the direct effect of computer ownership on income mobility and the spillover effects in the form of social impact. Although our research focuses only on computer ownership, further research should concentrate on other technologies. For example, Smartphone usage and Internet access could drastically improve the richness of information retrieval and sharply reduces the time of receiving such information. One would expect even a stronger spillover effect among households in the same district for smartphone usage or Internet access.

Second, this research examines the effect of a positive externality of computer ownership and speculates that the spillover effect might come from human interaction through their social network. A follow-up study could take a closer look at such interactions directly to provide a granular policy intervention. Beyond ones' caste, future research could focus on other factors such as ethnicity or individuals' social ties. Furthermore, individuals' offline social ties and online social ties might have different

effects on the process of IT spillover.

Third, India in the past two decades has gone through rapid economic growth and experienced significant IT deployment, while becoming the leading IT and business process outsourcing provider in the world. Whether the positive returns of computer ownership on income mobility would hold in other contexts and emerging economies is an empirical question. We call for more research to investigate the direct and spillover effects of digitization in other countries for generalizability.

4.5.3 Implication for Practice

The findings of this study provide implications for understanding the social impacts of public policies related to information technology. The positive spillover effect of computer ownership provides some justification for a government subsidy for computer use in public education. Our study is particularly relevant for emerging economies such as India, China, and Brazil because these countries are currently experiencing a technology revolution along with their economic development. A better understanding of the externalities of technology use and adoption may be informative for policymaking for government and think tanks.

Besides the direct effect and spillover effects of technology use, this study also finds a digital divide among different socio-economic and caste groups. We observe that users in some location such as living in an urban area, or users from specific caste group own a higher endowment of technologies. A critical issue for the policy may be to reduce the disparity among various socio-economic groups. Spillovers of computer ownership provide a new angle for technology policies, especially when there is a trade off with other policies. Our results suggest that technology policies should not evaluate

private returns but also social returns. Therefore, to fully reap the social welfare of disseminating positive knowledge spillover, policymakers should pay attention to reducing the cost of information exchange and removing the obstacles to interactions among different ethnic groups and cultural backgrounds. Our research suggests that providing education in computer courses at school and digital training programs for job seekers can reduce disparity due to digital literacy between technology owners and non-owners to assist with upward mobility.

To conclude, this study suggests that the benefits of digitization should not be measured only in terms of tangible measures such as productivity, but there is a need to also look at intangible benefits such as spillover effect on income mobility through social interactions. Our results suggest that households benefit from an increase in the extent of digitization in the same district in terms of upward income mobility, especially from the socially close group.

Tables

Table 2. 1 Message Design of each Group

Group	Message
Control	Goals help you form a habit of running; the statistics show runners who set up a goal can increase the number of times they run in a month by 17%. Please set up a goal.
Social Norms	Goals help you form a habit of running; the statistics show runners who set up a goal can increase the number of times they run in a month by 17%. Last month, [Number] runners set up a goal. Please set up a goal.

Note: [Number] is 5,223 for the treatment in February and is 5,715 for the treatment in March.

Table 2. 2 Randomization Check

	Control				Social Norms				t-stat
# of participants	3682				3514				
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	P-values
Age	38.99	7.92	16.00	67.00	38.87	7.95	16.00	69.00	0.52
Gender (percent of female)	0.19	0.39	0.00	1.00	0.19	0.39	0.00	1.00	0.90
Daily distance in pre-treatment month	2.20	2.56	0.00	22.21	2.21	2.50	0.00	23.43	0.87
% Registered by Facebook account	0.47	0.50	0.00	1.00	0.50	0.50	0.00	1.00	0.05
Log (# of friends)	2.98	1.61	0.00	7.84	3.04	1.63	0.00	7.93	0.12
High social connection	0.10	0.29	0.00	1.00	0.10	0.31	0.00	1.00	0.17
Log (# of races)	1.12	1.30	0.00	4.96	1.10	1.27	0.00	4.70	0.48
Log (# of comments)	1.87	2.22	0.00	10.05	1.88	2.24	0.00	10.30	0.85
Log (tenure)	6.24	0.82	3.43	7.48	6.20	0.82	3.71	7.48	0.08
Is weekend	0.19	0.39	0.00	1.00	0.19	0.39	0.00	1.00	0.96

Table 2. 3 Correlation Table

	1	2	3	4	5	6	7	8	9
1. Age	1.00								
2. Gender (percent of female)	-0.08*	1.00							
3. Distance in pre-treatment month	0.12*	-0.11*	1.00						
4. % Registered by Facebook account	-0.12*	0.07*	-0.15*	1.00					
5. Log (# of friends)	-0.03*	-0.04*	0.36*	0.09*	1.00				
6. Log (# of races)	-0.00	-0.09*	0.34*	-0.22*	0.50*	1.00			
7. Log (# of comments)	0.03*	-0.02	0.42*	-0.21*	0.70*	0.59*	1.00		
8. Log (tenure)	0.02	-0.12*	0.24*	-0.28*	0.49*	0.51*	0.48*	1.00	
9. Is Weekend	-0.04*	0.00	-0.14*	0.00	-0.11*	-0.08*	-0.13*	-0.02*	1.00
N	7196								

* $p < .05$

Table 2. 4 Treatment Effects on Goal Setting and Goal Attainment

	(1) Goal Setting	(2) Goal Setting	(3) Goal Attainment	(4) Goal Attainment
Social Norms	0.115***	0.070	-0.179*	-0.070
	(0.043)	(0.046)	(0.106)	(0.113)
High social connection	0.137*	-0.084	0.148	0.786***
	(0.072)	(0.111)	(0.163)	(0.278)
Social Norms X		0.381***		-0.944***
High social connection		(0.136)		(0.332)
Age	-0.001	-0.001	0.009	0.009
	(0.003)	(0.003)	(0.007)	(0.007)
Gender (female)	-0.092	-0.094	-0.146	-0.135
	(0.057)	(0.057)	(0.142)	(0.143)
Pre-treatment running distance	-0.003	-0.002	0.108***	0.111***
	(0.009)	(0.009)	(0.024)	(0.025)
Registered by FB account	0.042	0.043	-0.009	0.004
	(0.045)	(0.045)	(0.111)	(0.112)
Log (# of races)	0.146***	0.148***	-0.093*	-0.100*
	(0.021)	(0.021)	(0.052)	(0.052)
Log (tenure)	-0.157***	-0.158***	-0.022	-0.009
	(0.032)	(0.032)	(0.075)	(0.075)
_cons	-0.714***	-0.679***	-0.392	-0.555
	(0.231)	(0.230)	(0.533)	(0.538)
Monthly dummies	Yes	Yes	Yes	Yes
N	7196	7196	608	608
Prob > chi^2	0.00	0.00	0.00	0.00
Log pseudolikelihood	-2038.56	-2034.49	-389.86	-385.98

Note: Baseline group is *Control* group
Robust standard errors in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$

Table 2. 5 Addressing the Selection Issue

	(1) Selection Eqn.	(2) Outcome Eqn.	(3) Selection Eqn.	(4) Outcome Eqn.
	Goal Setting	Goal Attainment	Goal Setting	Goal Attainment
Is weekend	-0.236***		-0.234***	
	(0.061)		(0.061)	
Social Norms	0.115***	-0.199*	0.070	-0.079
	(0.043)	(0.113)	(0.046)	(0.121)
High social connection	0.130*	0.116	-1.212**	0.787***
	(0.072)	(0.193)	(0.497)	(0.278)
Social Norms X			0.375***	-0.978***
High social connection			(0.136)	(0.358)
Age	-0.001	0.009	-0.001	0.009
	(0.003)	(0.007)	(0.003)	(0.007)
Gender (female)	-0.095*	-0.122	-0.097*	-0.121
	(0.058)	(0.162)	(0.058)	(0.163)
Pre-treatment running distance	-0.007	0.105***	-0.006	0.110***
	(0.009)	(0.028)	(0.009)	(0.026)
Registered by FB account	0.039	-0.017	0.039	-0.002
	(0.045)	(0.112)	(0.045)	(0.114)
Log (# of races)	0.143***	-0.121	0.145***	-0.118
	(0.021)	(0.092)	(0.021)	(0.104)
Log (tenure)	-0.154***	0.012	-0.155***	0.013
	(0.032)	(0.129)	(0.032)	(0.135)
_cons	-0.665***	-0.075	-0.630***	-0.150
	(0.231)	(1.101)	(0.231)	(0.143)
Monthly dummies	Yes	Yes	Yes	Yes
N	7196		7196	
Prob > chi^2	0.75		0.85	
Log pseudolikelihood	-2420.43		-2412.68	

Note: Baseline group is *Control* group
Robust standard errors in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$

Table 2. 6 Different Types of High Social Connections

	(1) Goal Setting	(2) Goal Attainment
Social Norms	0.070	-0.082
	(0.046)	(0.134)
High social connection (Only High Out-degree)	-0.011	0.429
	(0.309)	(0.709)
High social connection (Only High In-degree)	-0.481	5.475***
	(0.449)	(0.563)
High social connection (Both High)	-0.068	0.756**
	(0.119)	(0.329)
Social Norms X High social connection (Only High Out-degree)	0.615	-0.462
	(0.379)	(0.778)
Social Norms X High social connection (Only High In-degree)	0.960*	-4.908***
	(0.535)	(0.750)
Social Norms X High social connection (Both High)	0.296**	-1.059***
	(0.149)	(0.377)
Age	-0.001	0.009
	(0.003)	(0.007)
Gender (female)	-0.093	-0.152
	(0.058)	(0.161)
Pre-treatment running distance	-0.001	0.112***
	(0.009)	(0.026)
Registered by FB account	0.045	0.008
	(0.045)	(0.115)
Log (# of races)	0.148***	-0.113
	(0.021)	(0.100)
Log (tenure)	-0.158***	0.007
	(0.032)	(0.129)
_cons	-0.682***	-0.395
	(0.231)	(1.135)
Monthly dummies	Yes	Yes
N	7196	7196
Prob > chi^2	0.00	0.87
Log pseudolikelihood	-2032.52	-2413.88

Note: Baseline group of the treatment variable is *Control* group; baseline group of social connections variable is *low social connected individuals*; we suppressed the results of the selection model for Column (2) of heckprobit model to maintain simplicity.
Robust standard errors in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$

Table 2. 7 Models for Subgroups (Matching)

	(1) Probit	(2) OLS	(3) OLS	(4) Probit
	Goal Attainment	Log (# of friends)	Goal minus Distance in Previous Month	Run Next Month
Control group	0.543**	-0.482**	-0.626**	0.678***
	(0.220)	(0.190)	(0.282)	(0.258)
Matched subgroup	0.417*	-0.526***	-0.509*	0.776***
	(0.220)	(0.189)	(0.268)	(0.257)
Age	0.009	0.005	-0.019	0.012
	(0.007)	(0.007)	(0.013)	(0.011)
Gender (female)	-0.126	0.272**	-0.310	0.093
	(0.143)	(0.125)	(0.209)	(0.222)
Registered by FB account	0.018	0.826***	-0.015	0.004
	(0.113)	(0.101)	(0.180)	(0.174)
Log (# of races)	-0.088*	0.349***	-0.057	-0.094
	(0.052)	(0.050)	(0.091)	(0.092)
Log (tenure)	-0.016	0.783***	-0.003	0.195
	(0.075)	(0.071)	(0.124)	(0.122)
_cons	-0.998*	-2.506***	2.457***	-1.302
	(0.572)	(0.501)	(0.936)	(0.820)
High social connection	Yes	No	Yes	Yes
Pre-treatment running distance	Yes	Yes	No	Yes
Monthly dummies	Yes	Yes	Yes	Yes
N	608	608	608	608
Prob > chi^2	0.00	0.00	0.24	0.00
Log pseudolikelihood	-387.95	-963.26	-1285.34	-143.48

Note: Baseline group is *Social Norms motivated* subgroup
Robust standard errors in parentheses * p < .1, ** p < .05, *** p < .01

Table 2. 8 Redefining Goal Attainment

	(1) Goal Attainment (90%)	(2) Goal Attainment (75%)	(3) Goal Attainment (50%)
Social Norms	-0.036	-0.070	-0.106
	(0.136)	(0.134)	(0.077)
High social connection	0.693**	0.631*	0.431
	(0.312)	(0.348)	(0.358)
Social Norms X	-0.962***	-0.735**	-0.535
High social connection	(0.352)	(0.371)	(0.374)
Age	0.004	0.004	-0.000
	(0.007)	(0.007)	(0.005)
Gender (female)	-0.111	-0.008	0.088
	(0.153)	(0.154)	(0.105)
Pre-treatment running distance	0.158***	0.182***	0.143***
	(0.026)	(0.032)	(0.034)
Registered by FB account	-0.017	-0.095	-0.121
	(0.114)	(0.114)	(0.085)
Log (# of races)	-0.099	-0.116	-0.126***
	(0.108)	(0.102)	(0.039)
Log (tenure)	-0.052	0.007	0.068
	(0.129)	(0.132)	(0.063)
_cons	-0.081	0.184	1.738***
	(1.091)	(1.116)	(0.404)
Monthly dummies	Yes	Yes	Yes
N	7196	7196	7196
Prob > chi2	0.99	0.87	0.13
Log pseudolikelihood	-2423.92	-2411.02	-2323.33

Note: Baseline group is *Control* group; we suppressed the results of the selection model for all columns of heckprobit model to maintain simplicity.
Robust standard errors in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$

Table 3. 1 Message to Qualified Senders

Group	Message
Sender	Congratulations on being selected to participate in the "You Exercise, We Send a Gift" event: Since you have run more than 40 km in the past two weeks, your friend will get a chance to win a gold coin. We find some of your friends' exercise frequency is not as good as it should be. We want to raise a challenge for your friend to help them to form an exercise habit. To encourage them to participate in our challenge, we would like to provide him/her a chance to win gold coins. Each gold coin represents a chance to enter into the lottery drawing. The following prizes are being offered: [Rewards Figure[49]] It is not the gift that counts, but the thought behind it. Your challenge would be much more meaningful for your friend than a challenge raised by *RunningPlatform*. If you agree to help us, we will randomly find a qualified friend and present him/her a challenge (We have a limited quota, please take this opportunity now). Agree, press here to pick your lucky friend / Disagree

Table 3. 2 Message to Receivers

Group	Message
Baseline Control Group	*RunningPlatform* is now holding a "You Exercise, We Send a Gift" event. *RunningPlatform* is committed to providing better service to every runner, so we now sent you a gold coin. We would like you to accept our challenge. If you win the challenge, you will earn a second gold coin. Challenge: From 2017/8/21 to 2017/9/3 (a two-week period), you can win this challenge if your accumulated distance exceeds *W* km. Each gold coin represents a chance to enter into the drawing. You have already received a gold coin from *RunningPlatform*. The second gold coin will double your chance of winning the lottery. The following prizes are being offered: [Rewards Figure] We will notify you at the end of this event. The winner can leave shipping information or go to the store to pick up the prize. Accept challenge / Not accept
Friend Control Group	*RunningPlatform* is now holding a "You Exercise, We Send a Gift" event. Your friend (*XXX*) ran 40 km in the past two weeks to earn you a gold coin, and has now sent you this gold coin. We would like you to accept our challenge. If you win the challenge, you will earn a second gold coin. Challenge: From 2017/8/21 to 2017/9/3 (a two-week period), you can win this challenge if your accumulated distance exceeds *W* km. Each gold coin represents a chance to enter into the drawing. You have already received a gold coin from your friend (*XXX*). The second gold coin will double your chance of winning the lottery. The following prizes are being offered: [Rewards Figure] We will notify you at the end of this event. The winner can leave shipping information or go to the store to pick up the prize. Accept challenge / Not accept

[49] Figure 3.1

Reciprocity Treatment Group	*RunningPlatform* is now holding a "You Exercise, We Send a Gift" event. Your friend (*XXX*) ran 40 km in the past two weeks to earn you a gold coin, and has now sent you this gold coin. We would like you to accept our challenge. If you win the challenge, you will return your friend a gold coin. Challenge: From 2017/8/21 to 2017/9/3 (a two-week period), you can win this challenge if your accumulated distance exceeds *W* km. Each gold coin represents a chance to enter into the lottery drawing. You have already received a gold coin from your friend (*XXX*). Winning this challenge can help your friend (*XXX*) win a gold coin to enter into the drawing. The following prizes are being offered: [Rewards Figure] We will notify you at the end of this event. The winner can leave shipping information or go to the store to pick up the prize. Accept challenge / Not accept

Note: *W* could be either 30 km or 45 km depending on the pre-treatment distance a runner ran. *XXX* represents a sender's online name. *RunningPlatform* is a pseudo name of the platform to maintain anonymity.

Table 3. 3 Reminder Messages of the Upcoming Challenge

Group	Message
Baseline Control Group	Your challenge is going to start. From 8/21 to 9/3 (a two-week period), you can win this challenge if your accumulated distance exceeds *W* km. Each gold coin represents a chance to enter into the drawing. You have already received a gold coin from *RunningPlatform*. The second gold coin will double your chance of winning the lottery. The following prizes are being offered: [Rewards Figure]
Friend Control Group	Your challenge is going to start. From 8/21 to 9/3 (a two-week period), you can win this challenge if your accumulated distance exceeds *W* km. Each gold coin represents a chance to enter into the drawing. You have already received a gold coin from your friend (*XXX*). The second gold coin will double your chance of winning the lottery. The following prizes are being offered: [Rewards Figure]
Reciprocity Treatment Group	Your challenge is going to start. From 8/21 to 9/3 (a two-week period), you can win this challenge if your accumulated distance exceeds *W* km. Each gold coin represents a chance to enter into the drawing. You have already received a gold coin from your friend (*XXX*). Winning this challenge can help your friend (xxx) win a gold coin to enter into the drawing. The following prizes are being offered: [Rewards Figure]

Note: *W* could be either 30 km or 45 km depending on the pre-treatment distance a runner ran. *XXX* represents sender's online name. *RunningPlatform* is a pseudo name of the platform to maintain anonymity.

Table 3. 4 Reminder Messages of the Challenge Progress

Group	Complete Challenge	Message
Baseline Control Group	Yes	Your accumulated distance: *X* km. You have completed this challenge. You will win a second gold coin. Each gold coin represents a chance to enter into the drawing. The following gifts are being offered, and we will draw lucky winners after this event. [Rewards Figure]

	No	Your accumulated distance: *X* km. You have not completed this challenge. Please keep on going to win yourself a second gold coin. Each gold coin represents a chance to enter into the drawing. The following gifts are being offered, and we will draw lucky winners after this event. [Rewards Figure]
Friend Control Group	Yes	Your accumulated distance: *X* km. You have completed this challenge. You will win a second gold coin. Each gold coin represents a chance to enter into the drawing. The following gifts are being offered, and we will draw lucky winners after this event. [Rewards Figure]
	No	Your accumulated distance: *X* km. You have not completed this challenge. Please keep on going to win yourself a second gold coin. Each gold coin represents a chance to enter into the drawing. The following gifts are being offered, and we will draw lucky winners after this event. [Rewards Figure]
Reciprocity Treatment Group	Yes	Your accumulated distance: *X* km. You have completed this challenge. Your friend will win a gold coin. Each gold coin represents a chance to enter into the drawing. The following gifts are being offered, and we will draw lucky winners after this event. [Rewards Figure]
	No	Your accumulated distance: *X* km. You have not completed this challenge. Please keep on going to win your friend a gold coin. Each gold coin represents a chance to enter into the drawing. The following gifts are being offered, and we will draw lucky winners after this event. [Rewards Figure]

Note: *X* indicates the total running distance.

Table 3. 5 Randomization Check

	Baseline Control	Friend Control	Reciprocity Treatment	ANOVA (p-value)
N	571	590	577	N/A
(Sender) Pre-treatment Distance (km)	82.49	79.41	77.59	0.09
(Sender) Age	41.68	41.34	41.61	0.72
(Sender) Tenure (days)	872.80	853.92	820.29	0.28
(Sender) # of Race	13.82	12.45	11.63	0.15
(Sender) % Female	17.7%	15.3%	16.6%	0.53
(Sender) # of Friends	129.34	132.39	126.25	0.87
(Receiver) Pre-treatment Distance (km)*	18.74	20.13	18.59	0.02
(Receiver) Pre-treatment Distance for 30 km Group (km)	13.07	13.91	13.27	0.17
(Receiver) Pre-treatment Distance for 45 km Group (km)	32.34	32.54	31.99	0.47
(Receiver) Age	40.06	40.23	40.55	0.51
(Receiver) Tenure (days)	789.58	805.75	771.14	0.57
(Receiver) # of Race	7.99	7.13	7.30	0.50
(Receiver) % Female	31.7%	26.4%	29.6%	0.14
(Receiver) # of Friends	88.40	92.05	88.78	0.92
(Receiver) Proportion of 30 km Challenge	70.6%	66.6%	71.6%	0.15

* We address the imbalance problem by using propensity score matching and by controlling inverse-probability weighting.

Table 3. 6 Regression Results of Completion Rate

	(1)	(2)	(3)	(4)	(5)	(6)
	Main Models		Redefine Completion of Challenge for the Declined Participants		Average Treatment Effect on the Treated (ATT)	
	Complete Challenge	Complete Challenge	Complete Challenge	Complete Challenge	Complete Challenge	Complete Challenge
Baseline group	Baseline Control	Friend Control	Baseline Control	Friend Control	Baseline Control	Friend Control
Reciprocity Treatment	0.28***	0.20**	0.25***	0.19**	0.35***	0.21**
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Pre-treatment Distance	0.03***	0.04***	0.04***	0.04***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
# of Friends	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female	-0.04	-0.14	-0.07	-0.17**	-0.04	-0.18*
	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Age	0.02***	0.01**	0.02***	0.02***	0.03***	0.02***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Tenure	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
# of Race	0.01***	0.01***	0.01***	0.01***	0.01*	0.01**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
_cons	-1.89***	-1.56***	-1.93***	-1.60***	-1.97***	-1.59***
	(0.31)	(0.30)	(0.31)	(0.30)	(0.33)	(0.33)
Challenge Type Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Sender's Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
N	1148	1167	1148	1167	977	972
Log pseudolikelihood	-708.30	-732.22	-720.30	-738.07	-624.28	-633.23
Chi^2	74.53	71.06	80.91	81.22	73.89	62.38

Robust standard errors in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$

Note: Baseline group is ***Baseline Control Group*** in (1), (3), (5); Baseline group is ***Friend Control Group*** in (2), (4), (6); Sender's characteristics include pre-treatment running distance, the number of friends, gender (female), age, tenure, and the number of races participated.

Table 3. 7 Curvilinear Relationship of Closeness

	(1) Complete Challenge	(2) Complete Challenge	(3) Complete Challenge	(4) Complete Challenge
ClosenessIndex	0.43***			
	(0.15)			
ClosenessIndex squared	-0.11***			
	(0.04)			
DaysKnowEachOther		0.66*		
		(0.39)		
DaysKnowEachOther squared		-0.08**		
		(0.04)		
CommonFriends			0.31**	
			(0.15)	
CommonFriends squared			-0.08**	
			(0.04)	
IOS				0.47*
				(0.26)
IOS squared				-0.06**
				(0.03)
Pre-treatment Distance	0.04***	0.04***	0.04***	0.07***
	(0.01)	(0.01)	(0.01)	(0.02)
# of Friends	0.00	0.00	0.00	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)
Female	-0.09	-0.11	-0.12	-0.28
	(0.12)	(0.12)	(0.12)	(0.28)
Age	0.01*	0.02**	0.01*	0.04**
	(0.01)	(0.01)	(0.01)	(0.02)
Tenure	-0.00	0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
# of Race	0.01*	0.01*	0.01*	0.00
	(0.00)	(0.00)	(0.00)	(0.01)
_cons	-1.75***	-2.99***	-1.58***	-3.14***
	(0.44)	(1.09)	(0.43)	(1.04)
Challenge Type Dummies	Yes	Yes	Yes	Yes
Sender's Characteristics	Yes	Yes	Yes	Yes
N	577	577	577	166
Log pseudolikelihood	-367.29	-368.11	-369.40	-71.41
Chi^2	43.32	39.46	38.33	31.35

Robust standard errors in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$

Note: Sender's characteristics include pre-treatment running distance, the number of friends, gender (female), age, tenure, and the number of races participated.

Table 3. 8 Randomization Check for the Matched Sample

	Friend Control	Reciprocity Treatment	T-test (p-value)
N	531	531	N/A
(Sender) Pre-treatment Distance (km)	79.12	77.51	0.48
(Sender) Age	41.38	41.68	0.51
(Sender) Tenure (days)	856.82	828.85	0.42
(Sender) # of Race	12.59	11.74	0.45
(Sender) % Female	15.1%	16.4%	0.56
(Sender) # of Friends	130.87	125.11	0.64
(Receiver) Pre-treatment Distance (km)	18.81	19.29	0.44
(Receiver) Pre-treatment Distance for 30 km group (km)	13.73	13.72	0.97
(Receiver) Pre-treatment Distance for 45 km group (km)	31.70	32.10	0.40
(Receiver) Age	40.51	40.23	0.53
(Receiver) Tenure (days)	784.74	800.12	0.65
(Receiver) # of Race	7.15	7.23	0.91
(Receiver) % Female	28.4%	27.9%	0.84
(Receiver) # of Friends	87.08	91.52	0.69
(Receiver) Proportion of 30 km Challenge	71.8%	69.7%	0.46

Table 3. 9 Regression Results for the Matched Sample

	(1)	(2)
	Complete Challenge	Running Distance (Treatment Period)
	Probit Model	OLS Model
Baseline group	Friend Control	Friend Control
Reciprocity Treatment	0.25***	3.06**
	(0.08)	(1.37)
Pre-treatment Distance	0.03***	0.71***
	(0.01)	(0.11)
# of Friends	0.00	0.00
	(0.00)	(0.00)
Female	-0.11	-3.13**
	(0.09)	(1.51)
Age	0.02***	0.38***
	(0.01)	(0.11)
Tenure	-0.00	-0.00*
	(0.00)	(0.00)
# of Race	0.01***	0.38***
	(0.00)	(0.08)
_cons	-1.52***	0.22
	(0.32)	(5.65)
Challenge Type Dummy	Yes	Yes
Sender's Characteristics	Yes	Yes
N	1062	1062
R^2		0.185
Log pseudolikelihood	-665.47	
Chi^2	61.64	

Robust standard errors in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$

Note: Baseline group is ***Friend Control Group***; Sender's characteristics include pre-treatment running distance, the number of friends, gender (female), age, tenure, and the number of races participated.

Table 3. 10 Robustness Checks

	(1) Complete Challenge	(2) Complete Challenge	(3) Complete Challenge	(4) Complete Challenge	(5) Complete Challenge	(6) Complete Challenge
Baseline group	Baseline Control	Friend Control	Baseline Control	Friend Control	Baseline Control	Friend Control
Reciprocity Treatment	0.29***	0.21**	0.27***	0.19**	0.23***	0.19**
	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)
Pre-treatment Distance	0.04***	0.04***	0.04***	0.04***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
# of Friends	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female	-0.05	-0.11	-0.04	-0.15	-0.02	-0.06
	(0.09)	(0.09)	(0.09)	(0.09)	(0.10)	(0.10)
Age	0.02***	0.01**	0.02***	0.02***	0.02***	0.01**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Tenure	-0.00	0.00	-0.00	-0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
# of Race	0.01**	0.01***	0.01**	0.01***	0.01**	0.01**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
_cons	-1.96***	-1.55***	-2.02***	-1.73***	-2.18***	-1.59***
	(0.33)	(0.32)	(0.32)	(0.32)	(0.35)	(0.35)
Challenge Type Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Sender's Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
N	1043	1061	1068	1092	942	956
Log pseudolikelihood	-627.56	-654.90	-652.72	-678.70	-566.19	-588.49
Chi^2	74.31	63.75	74.23	73.95	63.93	45.25

Robust standard errors in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$

Note: Baseline group is ***Baseline Control Group*** in (1), (3), (5); Baseline group is ***Friend Control Group*** in (2), (4), (6); Sender's characteristics include pre-treatment running distance, the number of friends, gender (female), age, tenure, and the number of races participated; The sample at each column is as follows. Columns (1) and (2): exclude users who upload manual running records; Columns (3) and (4): exclude users who upload private running records; Columns (5) and (6): keep persistent weak runners.

Table 4. 1 Descriptive Statistics

	Mean	S.D.
Income variables		
Mobility	0.978	0.817
Δlog (Adult Equivalent Income)	0.423	1.048
Log (Adult Equivalent Income) in base year	9.075	0.962
IT variables		
Household owns a computer in base year	0.009	0.096
Positive change in computer ownership (Not owning a computer in base year, and owning a computer in the second period.)	0.060	0.237
% of households in a district own a computer in base year	0.009	0.019
Change in % own a computer in a district	0.054	0.064
% of (same group) households in a district own a computer in base year	0.009	0.035
Change in % (same group) own a computer in a district	0.055	0.099
% of (other groups) households in a district own a computer in base year	0.011	0.033
Change in % (other groups) own a computer in a district	0.058	0.083
Human capital variables		
Share of persons with jobs in base year	0.254	0.239
Household size in base year	5.872	3.038
Share of male adults in base year	0.327	0.165
Share of female adults in base year	0.325	0.147
Age of household head in base year	48.205	13.338
Whether a household head is a female in base year	0.091	0.288
Economic and social status variables		
Number of durables owned by household in base year	7.576	3.711
Land owner in base year	0.502	0.500
Urban in base year	0.275	0.446
% Get along well in a district	0.549	0.289
Variation variables		
Change in share of persons with jobs	0.042	0.268
Change in household size	-0.997	3.036
Change in share of male adults	0.011	0.200
Change in share of female adults	0.043	0.168
Change from male to female household head	0.084	0.277
Change in % get along well in a district	0.032	0.413

Note: N = 32,214

Table 4. 2 Change in Households Owning Computer across Years

	Owns Computer in 2011		
Owns Computer in 2005	No	Yes	Total
No	29,993	1,919	31,912
	93.99%	6.01%	100%
Yes	101	201	302
	33.44%	66.56%	100%
Total	30,094	2,120	32,214
	93.42%	6.58%	100%

Table 4. 3 Statistics by Castes and Religious Groups

	Frequency		Owns a computer		Mobility
Groups	2005	2011	2005	2011	
Brahmin	1744	1714	2.4%	15.6%	0.908
Forward caste	4923	5045	2.1%	12.7%	0.929
Other Backward Classes (OBC)	11213	11045	0.6%	4.6%	0.992
Dalit	6837	6885	0.3%	2.9%	1.020
Adivasi	2899	2985	0.5%	4.2%	0.951
Muslim	3592	3619	0.8%	5.1%	0.948
Christian, Sikh, Jain	1006	921	3.2%	21.0%	1.079

Table 4. 4 Quintile Mobility Matrix by Adult Equivalent Income (AEI)

	2011 Quintile					
2005 Quintile	1 (bottom 20%)	2	3	4	5 (top 20%)	Total
1 (bottom 20%)	35.46%	26.70%	18.88%	12.89%	6.07%	100%
2	29.29%	26.75%	22.56%	14.84%	6.57%	100%
3	20.26%	23.33%	24.85%	21.21%	10.35%	100%
4	12.97%	16.61%	21.63%	27.11%	21.68%	100%
5 (top 20%)	6.13%	7.65%	12.75%	21.67%	51.81%	100%

Table 4. 5 Household Mobility Variation Based on Quintile Mobility Matrix

	Mobility
0 (Moving downward)	11,118
	34.51%
1 (No Mobility)	10,691
	33.19%
2 (Moving upward)	10,405
	32.30%
Total	32,214
	100%

Table 4. 6 Determinants of Households' Change in Mobility (OLogit, DV: Income Mobility)

	(1) Mobility	(2) Mobility	(3) Mobility
Change in own a computer (H1)	0.769***	0.761***	0.757***
	(0.043)	(0.044)	(0.044)
Household owns a computer in base year	0.715***	0.743***	0.713***
	(0.084)	(0.085)	(0.085)
Change in % own a computer in a district (H2)		0.638**	
		(0.273)	
% of households in a district own a computer in base year		-1.393*	
		(0.747)	
Change in % (same group) own a computer in a district (H3)			0.565***
			(0.136)
Change in % (other groups) own a computer in a district (H3)			-0.167
			(0.205)
% of (same group) households in a district own a computer in base year			0.088
			(0.348)
% of (other groups) households in a district own a computer in base year			-0.709*
			(0.379)
Share of persons with jobs in base year	0.706***	0.706***	0.712***
	(0.066)	(0.066)	(0.066)
Adult equivalent income in base year	-1.367***	-1.368***	-1.369***
	(0.021)	(0.021)	(0.021)
Household size in base year	-0.027***	-0.027***	-0.027***
	(0.007)	(0.007)	(0.007)
Share of male adults in base year	0.505***	0.505***	0.505***
	(0.100)	(0.100)	(0.100)
Share of female adults in base year	-0.180*	-0.181*	-0.180*
	(0.104)	(0.104)	(0.104)
Head age in base year	-0.004***	-0.004***	-0.004***
	(0.001)	(0.001)	(0.001)
Number of durables owned by household in base year	0.112***	0.112***	0.111***
	(0.005)	(0.005)	(0.005)
Household Head is female in base year	0.042	0.042	0.039
	(0.044)	(0.044)	(0.044)
Land owner in base year	-0.035	-0.032	-0.033
	(0.028)	(0.028)	(0.028)
Urban in base year	0.197***	0.191***	0.194***
	(0.032)	(0.034)	(0.033)
Change in household size	-0.031***	-0.031***	-0.031***
	(0.006)	(0.006)	(0.006)
Change in share of male adults in a household	0.704***	0.702***	0.699***
	(0.076)	(0.076)	(0.076)
Change in share of female adults in a household	0.002	0.001	-0.001
	(0.081)	(0.081)	(0.081)
Change from male to female household head	0.128***	0.127***	0.125***
	(0.045)	(0.045)	(0.045)
Change in share of persons with jobs	1.225***	1.223***	1.228***
	(0.053)	(0.053)	(0.053)
State dummies	Yes	Yes	Yes
Group dummies	Yes	Yes	Yes
Observations	32214	32214	32214
Log pseudolikelihood	-30676	-30672	-30666

Robust standard errors in parentheses
$^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$

Table 4. 7 Determinants of Households' Change in ln(AEI) (OLS, DV: $\Delta lnAEI$)

	(1) $\Delta lnAEI$	(2) $\Delta lnAEI$	(3) $\Delta lnAEI$
Change in own a computer (H1)	0.473***	0.470***	0.467***
	(0.022)	(0.022)	(0.022)
Household owns a computer in base year	0.329***	0.341***	0.332***
	(0.052)	(0.053)	(0.053)
Change in % own a computer in a district (H2)		0.285**	
		(0.116)	
% of households in a district own a computer in base year		-0.608*	
		(0.340)	
Change in % (same group) own a computer in a district (H3)			0.271***
			(0.058)
Change in % (other groups) own a computer in a district (H3)			-0.109
			(0.085)
% of (same group) households in a district own a computer in base year			-0.136
			(0.150)
% of (other groups) households in a district own a computer in base year			-0.311**
			(0.144)
Share of persons with jobs in base year	0.409***	0.409***	0.413***
	(0.028)	(0.028)	(0.028)
Adult equivalent income in base year	-0.800***	-0.800***	-0.801***
	(0.007)	(0.007)	(0.007)
Household size in base year	-0.016***	-0.016***	-0.016***
	(0.003)	(0.003)	(0.003)
Share of male adults in base year	0.221***	0.221***	0.221***
	(0.042)	(0.042)	(0.042)
Share of female adults in base year	-0.089**	-0.089**	-0.090**
	(0.043)	(0.043)	(0.043)
Head age in base year	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)
Number of durables owned by household in base year	0.073***	0.072***	0.072***
	(0.002)	(0.002)	(0.002)
Household Head is female in base year	0.049***	0.049***	0.048***
	(0.018)	(0.017)	(0.018)
Land owner in base year	-0.007	-0.006	-0.007
	(0.011)	(0.011)	(0.011)
Urban in base year	0.141***	0.138***	0.142***
	(0.012)	(0.013)	(0.013)
Change in household size	-0.015***	-0.015***	-0.015***
	(0.002)	(0.002)	(0.002)
Change in share of male adults in a household	0.222***	0.222***	0.220***
	(0.032)	(0.032)	(0.032)
Change in share of female adults in a household	-0.052	-0.053	-0.054
	(0.034)	(0.034)	(0.034)
Change from male to female household head	0.064***	0.064***	0.064***
	(0.018)	(0.018)	(0.018)
Change in share of persons with jobs	0.628***	0.627***	0.629***
	(0.022)	(0.022)	(0.022)
Constant	7.566***	7.552***	7.569***
	(0.079)	(0.080)	(0.081)
State dummies	Yes	Yes	Yes
Group dummies	Yes	Yes	Yes
Observations	32214	32214	32214
R^2	0.410	0.410	0.411

Robust standard errors in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$

Table 4. 8 Endogenous Treatment Regression Models for Computer Ownership

	(1) Computer Ownership	(2) Mobility	(3) Computer Ownership	(4) $\Delta lnAEI$
	Selection Equation	Outcome Equation	Selection Equation	Outcome Equation
Change in own a computer		1.083***		0.567***
		(0.392)		(0.065)
Distance to high school in base year	-0.013***		-0.014***	
	(0.004)		(0.005)	
Household owns a computer in base year	-5.895***	0.326***	-6.411***	0.326***
	(1.043)	(0.113)	(0.126)	(0.117)
Share of persons in a household with jobs in base year	-0.552***	0.155***	-0.755***	0.295***
	(0.210)	(0.029)	(0.169)	(0.035)
Adult equivalent income in base year	0.267***	-0.479***	0.216***	-0.813***
	(0.053)	(0.010)	(0.034)	(0.008)
Household size in base year	0.034**	-0.003	0.048***	-0.011***
	(0.016)	(0.003)	(0.011)	(0.003)
Share of male adults in a household in base year	0.383*	0.157***	0.480**	0.192***
	(0.233)	(0.045)	(0.203)	(0.054)
Share of female adults in a household in base year	0.180	-0.036	-0.010	-0.111**
	(0.208)	(0.046)	(0.226)	(0.054)
Head age in base year	0.002	-0.002***	0.000	-0.002***
	(0.002)	(0.000)	(0.002)	(0.001)
Number of durables owned by household in base year	0.151***	0.018***	0.169***	0.070***
	(0.028)	(0.005)	(0.010)	(0.003)
Household Head is female in base year	0.212***	-0.003	0.234***	0.068***
	(0.081)	(0.021)	(0.083)	(0.023)
Land owner in base year	-0.076	-0.027**	-0.148***	0.002
	(0.069)	(0.012)	(0.055)	(0.013)
Urban in base year	-0.024	-0.065	0.049	0.005
	(0.168)	(0.061)	(0.169)	(0.060)
Change in household size	0.073***	-0.010***	0.085***	-0.011***
	(0.016)	(0.003)	(0.009)	(0.003)
Change in share of male adults in a household	0.984***	0.226***	1.312***	0.239***
	(0.349)	(0.041)	(0.150)	(0.040)
Change in share of female adults in a household	0.606***	0.006	0.641***	-0.014
	(0.182)	(0.038)	(0.165)	(0.042)
Change from male to female household head	0.030	0.054***	0.139	0.098***
	(0.114)	(0.021)	(0.086)	(0.023)
Change in share of persons in a household with jobs	-0.516***	0.384***	-0.608***	0.554***
	(0.123)	(0.023)	(0.127)	(0.027)
Constant	-5.541***	5.496***	-5.022***	7.788***
	(0.449)	(0.111)	(0.408)	(0.106)
State dummies	Yes	Yes	Yes	Yes
Group dummies	Yes	Yes	Yes	Yes
Observations	21448	21448	21448	21448
Wald test of ρ (prob > chi^2)	0.08		0.10	

Robust standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 4. 9 Average Treatment Effect on the Treated using Propensity Score Matching Approach (Mobility)

		(1)	(2)	(3)
		Mobility	Mobility	Difference in Mobility
		Change of Owning Computer	Others	
Before matching	Mean	0.994	0.977	0.017
	Freq.	1,919	30,295	(0.91)
After Matching	Mean	0.995	0.875	0.120***
	Freq.	1,912	1,912	(5.37)

t-statistics in parentheses *** p < .01

Table 4. 10 Average Treatment Effect on the Treated using Propensity Score Matching Approach ($\Delta lnAEI$)

		(1)	(2)	(3)
		$\Delta lnAEI$	$\Delta lnAEI$	Difference in $\Delta lnAEI$
		Change of Owning Computer	Others	
Before matching	Mean	0.540	0.416	0.124***
	Freq.	1,919	30,295	(5.04)
After Matching	Mean	0.543	0.304	0.239***
	Freq.	1,912	1,912	(7.01)

t-statistics in parentheses *** p < .01

Table 4. 11 Robustness Check: Remove Outliers

	(1) Mobility	(2) Mobility	(3) Mobility
Change in own a computer (H1)	0.797***	0.788***	0.782***
	(0.047)	(0.047)	(0.047)
Household owns a computer in base year	0.750***	0.761***	0.745***
	(0.094)	(0.095)	(0.096)
Change in % own a computer in a district (H2)		0.820**	
		(0.339)	
% of households in a district own a computer in base year		-0.221	
		(1.001)	
Change in % (same group) own a computer in a district (H3)			0.673***
			(0.163)
Change in % (other groups) own a computer in a district (H3)			-0.037
			(0.253)
% of (same group) households in a district own a computer in base year			0.371
			(0.393)
% of (other groups) households in a district own a computer in base year			-0.610
			(0.410)
Share of persons with jobs in base year	0.729***	0.731***	0.735***
	(0.072)	(0.072)	(0.072)
Adult equivalent income in base year	-1.371***	-1.373***	-1.374***
	(0.023)	(0.023)	(0.023)
Household size in base year	-0.024***	-0.024***	-0.023***
	(0.007)	(0.007)	(0.007)
Share of male adults in base year	0.490***	0.485***	0.485***
	(0.108)	(0.108)	(0.108)
Share of female adults in base year	-0.212*	-0.213*	-0.213*
	(0.112)	(0.112)	(0.112)
Head age in base year	-0.004***	-0.004***	-0.004***
	(0.001)	(0.001)	(0.001)
Number of durables owned by household in base year	0.112***	0.110***	0.110***
	(0.006)	(0.006)	(0.006)
Household Head is female in base year	0.056	0.056	0.052
	(0.047)	(0.047)	(0.047)
Land owner in base year	-0.035	-0.030	-0.032
	(0.029)	(0.029)	(0.029)
Urban in base year	0.151***	0.131***	0.138***
	(0.036)	(0.038)	(0.038)
Change in household size	-0.026***	-0.026***	-0.026***
	(0.006)	(0.006)	(0.006)
Change in share of male adults in a household	0.703***	0.700***	0.696***
	(0.082)	(0.082)	(0.082)
Change in share of female adults in a household	0.002	-0.000	-0.003
	(0.087)	(0.087)	(0.087)
Change from male to female household head	0.168***	0.167***	0.165***
	(0.048)	(0.048)	(0.048)
Change in share of persons with jobs	1.225***	1.225***	1.229***
	(0.057)	(0.057)	(0.057)
State dummies	Yes	Yes	Yes
Group dummies	Yes	Yes	Yes
Observations	27772	27772	27772
Log pseudolikelihood	-26440	-26437	-26431

Robust standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 4. 12 Robustness Check: Control Regional Economic Growth

	(1) Mobility	(2) Mobility	(3) Mobility
Change in own a computer (H1)	0.767***	0.757***	0.753***
	(0.044)	(0.044)	(0.044)
Household owns a computer in base year	0.739***	0.742***	0.724***
	(0.084)	(0.085)	(0.085)
Change in % own a computer in a district (H2)		0.689**	
		(0.277)	
% of households in a district own a computer in base year		0.033	
		(0.749)	
Change in % (same group) own a computer in a district (H3)			0.614***
			(0.138)
Change in % (other groups) own a computer in a district (H3)			-0.064
			(0.207)
% of (same group) households in a district own a computer in base year			0.213
			(0.350)
% of (other groups) households in a district own a computer in base year			0.048
			(0.381)
Share of persons with jobs in base year	0.735***	0.735***	0.740***
	(0.067)	(0.067)	(0.067)
Adult equivalent income in base year	-1.339***	-1.340***	-1.341***
	(0.021)	(0.021)	(0.021)
Economic growth in a district	1.362***	1.364***	1.366***
	(0.048)	(0.048)	(0.048)
Household size in base year	-0.032***	-0.032***	-0.032***
	(0.007)	(0.007)	(0.007)
Share of male adults in base year	0.467***	0.463***	0.462***
	(0.100)	(0.100)	(0.100)
Share of female adults in base year	-0.225**	-0.227**	-0.229**
	(0.105)	(0.105)	(0.105)
Head age in base year	-0.003***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)
Number of durables owned by household in base year	0.109***	0.108***	0.107***
	(0.005)	(0.005)	(0.005)
Household Head is female in base year	0.048	0.048	0.046
	(0.044)	(0.044)	(0.044)
Land owner in base year	-0.024	-0.018	-0.019
	(0.028)	(0.028)	(0.028)
Urban in base year	0.259***	0.238***	0.239***
	(0.032)	(0.034)	(0.034)
Change in household size	-0.040***	-0.040***	-0.040***
	(0.006)	(0.006)	(0.006)
Change in share of male adults in a household	0.667***	0.665***	0.661***
	(0.076)	(0.076)	(0.076)
Change in share of female adults in a household	-0.041	-0.042	-0.045
	(0.082)	(0.082)	(0.082)
Change from male to female household head	0.128***	0.128***	0.126***
	(0.045)	(0.045)	(0.045)
Change in share of persons with jobs	1.252***	1.250***	1.255***
	(0.053)	(0.053)	(0.054)
State dummies	Yes	Yes	Yes
Group dummies	Yes	Yes	Yes
Observations	32214	32214	32214
Log pseudolikelihood	-30259	-30256	-30249

Robust standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 4. 13 Moderating Effect of Share of Persons with Jobs in a Household

	(2) Mobility	(2) $\Delta lnAEI$
Change in own a computer	0.553***	0.394***
	(0.060)	(0.031)
Household owns a computer in base year	0.722***	0.331***
	(0.084)	(0.052)
Change in own a computer X Share of persons with jobs in base year	1.195***	0.456***
	(0.208)	(0.108)
Share of persons with jobs in base year	0.654***	0.392***
	(0.068)	(0.028)
Adult equivalent income in base year	-1.369***	-0.801***
	(0.021)	(0.007)
Household size in base year	-0.027***	-0.016***
	(0.007)	(0.003)
Share of male adults in base year	0.509***	0.223***
	(0.100)	(0.042)
Share of female adults in base year	-0.172*	-0.087**
	(0.104)	(0.043)
Head age in base year	-0.004***	-0.002***
	(0.001)	(0.000)
Number of durables owned by household in base year	0.112***	0.072***
	(0.005)	(0.002)
Household Head is female in base year	0.043	0.049***
	(0.044)	(0.018)
Land owner in base year	-0.035	-0.007
	(0.028)	(0.011)
Urban in base year	0.191***	0.139***
	(0.032)	(0.012)
Change in household size	-0.031***	-0.015***
	(0.006)	(0.002)
Change in share of male adults in a household	0.713***	0.225***
	(0.076)	(0.032)
Change in share of female adults in a household	0.006	-0.051
	(0.081)	(0.034)
Change from male to female household head	0.130***	0.065***
	(0.045)	(0.018)
Change in share of persons with jobs	1.218***	0.625***
	(0.053)	(0.022)
State dummies	Yes	Yes
Group dummies	Yes	Yes
Observations	32214	32214
Log pseudolikelihood / R^2	-30665	0.410

Robust standard errors in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$

Table 4. 14 Spillover Effects Moderated by Regional Harmony

	(1) Mobility	(2) Mobility	(3) Mobility
Change in own a computer	0.762***	0.761***	0.760***
	(0.044)	(0.044)	(0.044)
Household owns a computer in base year	0.743***	0.720***	0.708***
	(0.085)	(0.085)	(0.085)
Change in % own a computer in a district	-1.380***		
	(0.494)		
% of households in a district own a computer in base year	-1.921**		
	(0.770)		
Change in % (same group) own a computer in a district		0.246	0.600***
		(0.294)	(0.136)
Change in % (other groups) own a computer in a district		-0.143	-1.886***
		(0.205)	(0.414)
% of (same group) households in a district own a computer in base year		0.141	0.009
		(0.348)	(0.352)
% of (other groups) households in a district own a computer in base year		-0.659*	-0.690*
		(0.383)	(0.382)
Share of persons with jobs in base year	0.713***	0.711***	0.717***
	(0.067)	(0.067)	(0.067)
Adult equivalent income in base year	-1.369***	-1.368***	-1.369***
	(0.021)	(0.021)	(0.021)
% get along well in a district in base year	0.150*	0.296***	0.169**
	(0.081)	(0.075)	(0.079)
Change in % get along well in a district	0.258***	0.260***	0.260***
	(0.052)	(0.052)	(0.052)
Change in % own a computer in a district X % get along well in a district in base year	3.979***		
	(0.799)		
Change in % (same group) own a computer in a district X % get along well in a district in base year		0.578	
		(0.448)	
Change in % (other groups) own a computer in a district X % get along well in a district in base year			3.235***
			(0.658)
Household size in base year	-0.027***	-0.026***	-0.027***
	(0.007)	(0.007)	(0.007)
Share of male adults in base year	0.510***	0.508***	0.510***
	(0.100)	(0.100)	(0.100)
Share of female adults in base year	-0.185*	-0.181*	-0.184*
	(0.104)	(0.104)	(0.104)
Head age in base year	-0.004***	-0.004***	-0.004***
	(0.001)	(0.001)	(0.001)
Number of durables owned by household in base year	0.112***	0.110***	0.111***
	(0.005)	(0.005)	(0.005)
Household Head is female in base year	0.044	0.042	0.040
	(0.044)	(0.044)	(0.044)
Land owner in base year	-0.030	-0.030	-0.034
	(0.028)	(0.028)	(0.028)
Urban in base year	0.196***	0.193***	0.193***
	(0.034)	(0.034)	(0.034)
Change in household size	-0.031***	-0.031***	-0.031***
	(0.006)	(0.006)	(0.006)
Change in share of male adults in a household	0.702***	0.697***	0.699***
	(0.076)	(0.076)	(0.076)
Change in share of female adults in a household	0.000	-0.002	-0.000
	(0.081)	(0.081)	(0.081)
Change from male to female household head	0.124***	0.125***	0.124***
	(0.045)	(0.045)	(0.045)
Change in share of persons with jobs	1.231***	1.231***	1.235***
	(0.053)	(0.053)	(0.053)
State dummies	Yes	Yes	Yes
Group dummies	Yes	Yes	Yes
Observations	32214	32214	32214
Log pseudolikelihood	-30647	-30652	-30641

Standard errors in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$

Figures

Figure 2. 1 Number of individuals in *Control* group and *Social Norms* group per day

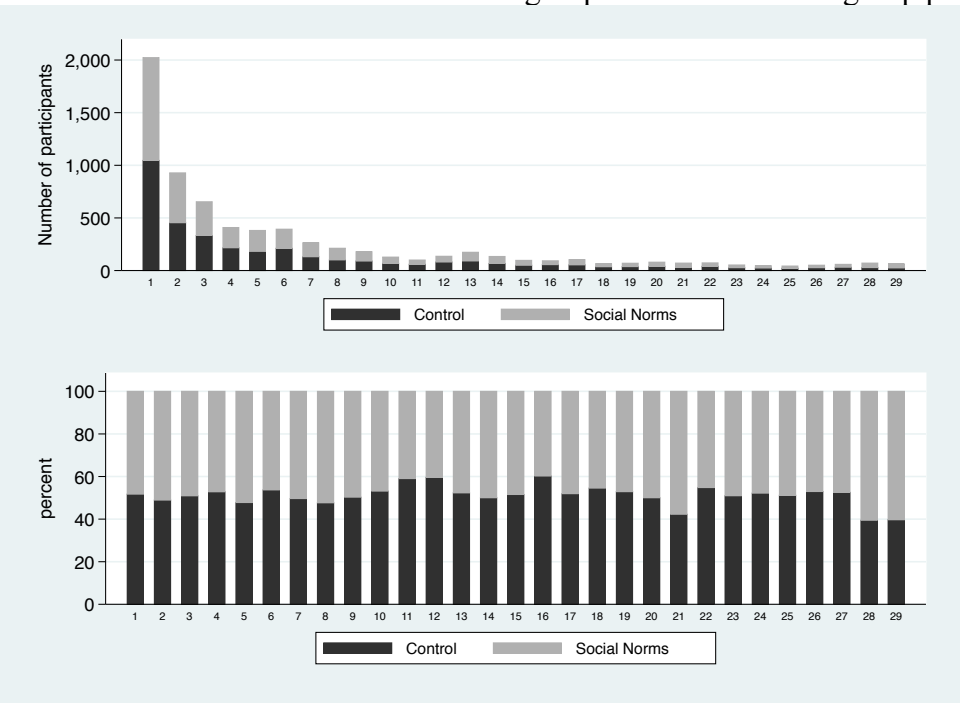


Figure 2. 2 Goal Setting Rate and Goal Attainment Rate

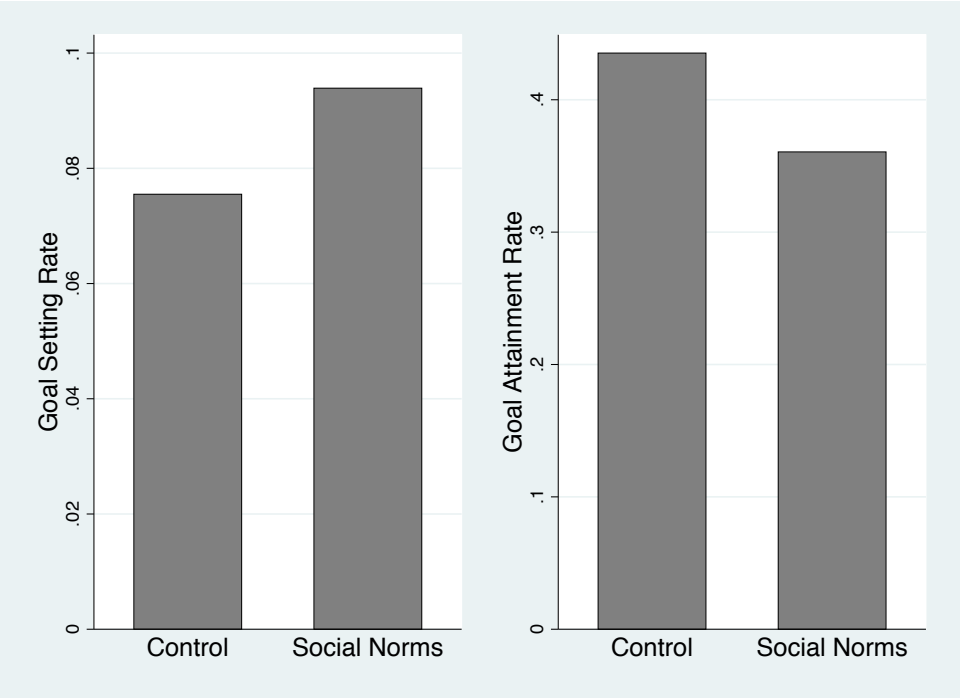


Figure 2. 3 Average Treatment Effects by subgroup

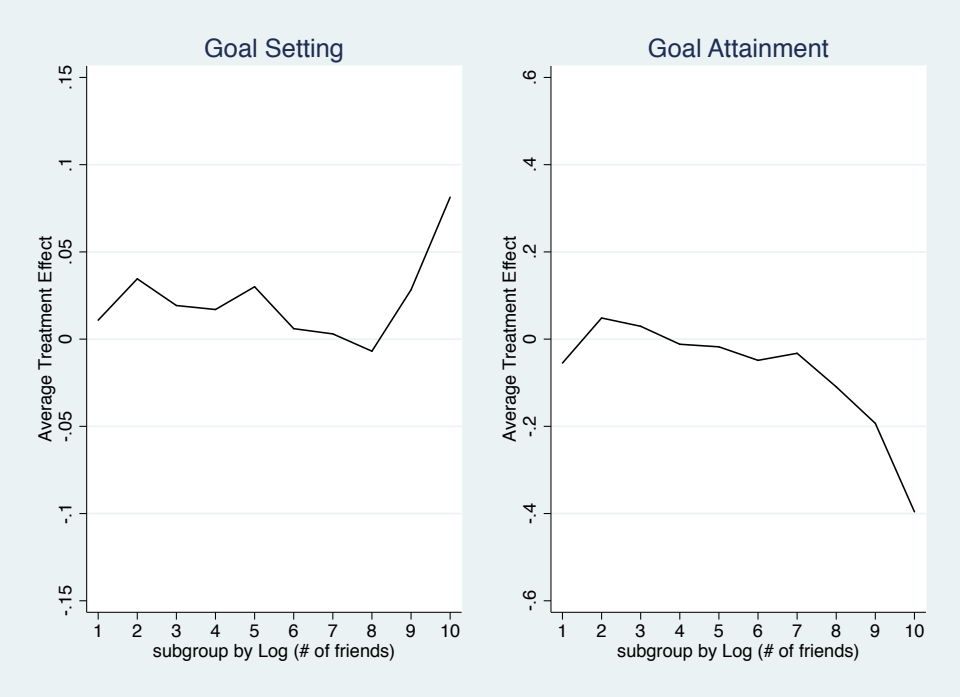


Figure 2. 4 Conditional Average Treatment Effects

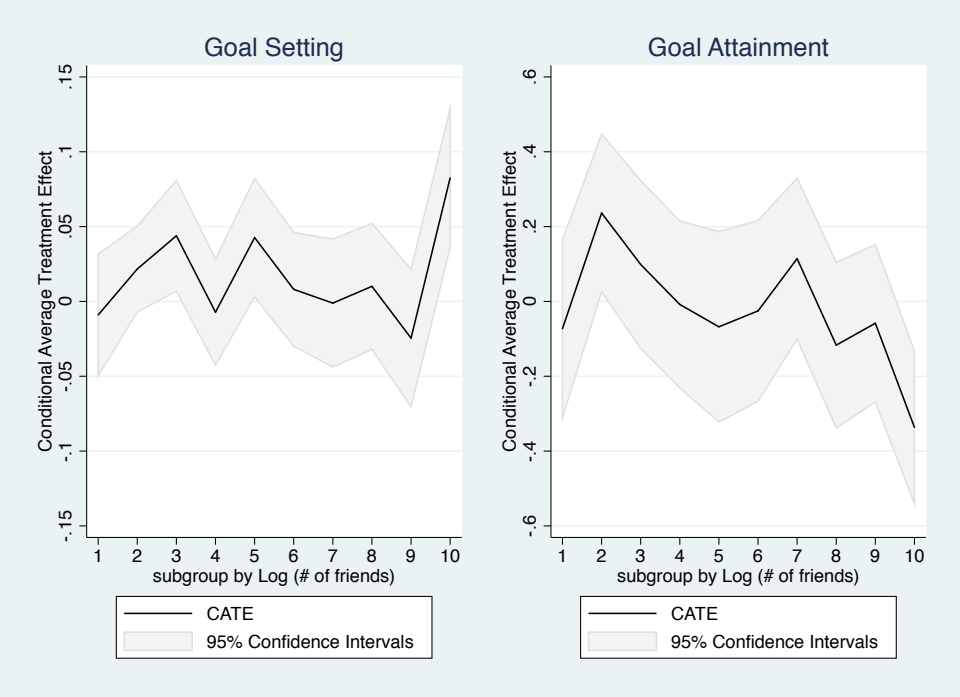


Figure 2. 5 Differences among Groups

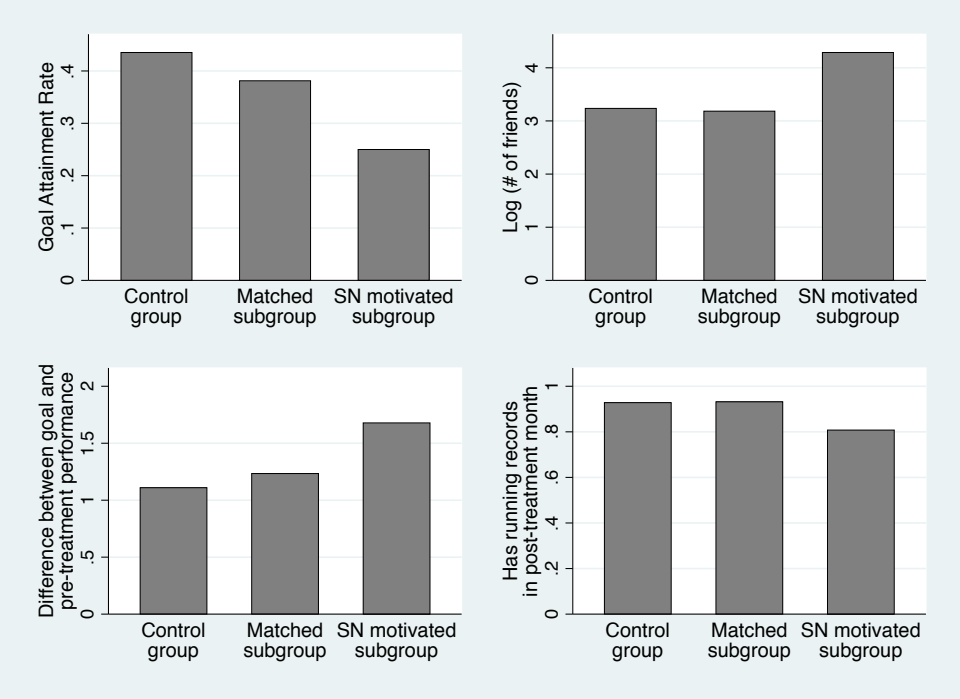


Figure 3. 1 Rewards offered in Lottery Drawing

Figure 3. 2 Timeline of the Experiment

Figure 3. 3 Geographic Distribution of Individuals by Groups

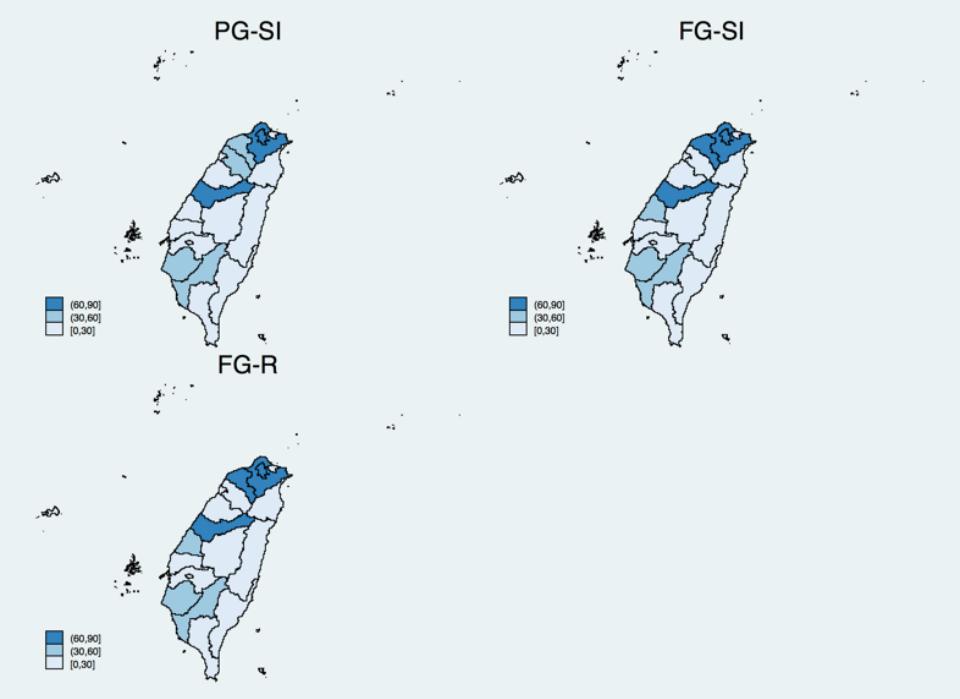


Figure 3. 4 Cumulative Challenge Completion Rates

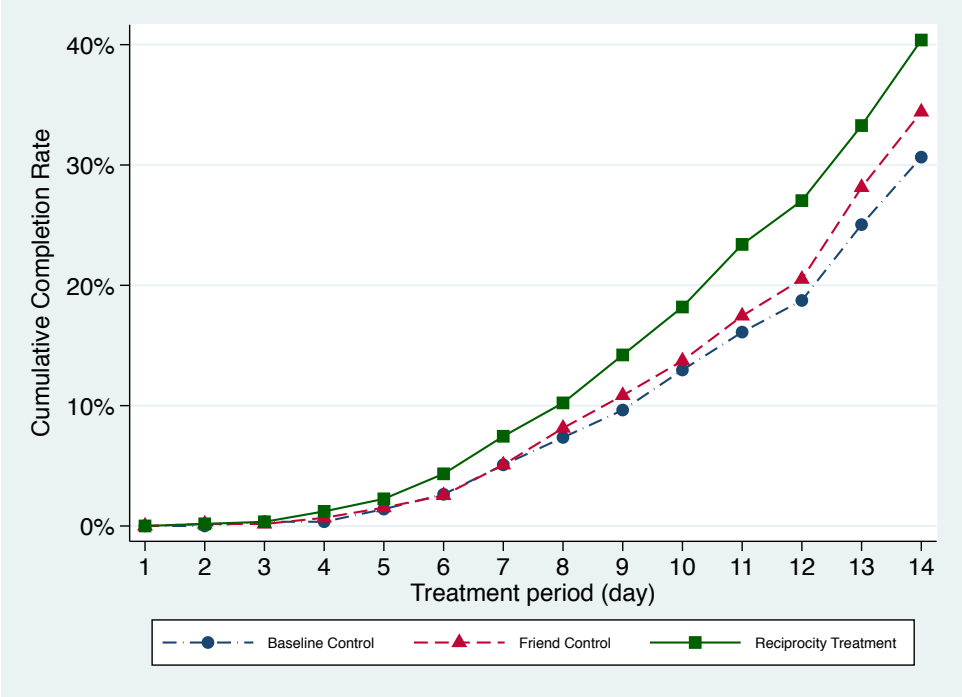


Figure 3. 5 Curvilinear Relationship of Social Closeness

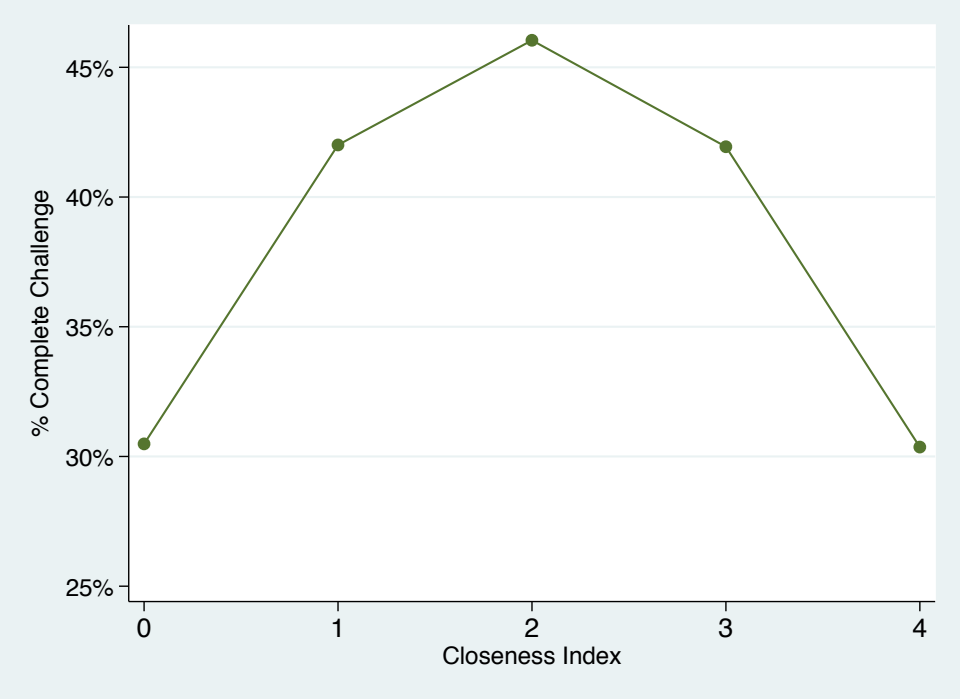


Figure 3. 6 Cumulative Probability of Total Distance in the Treatment Period

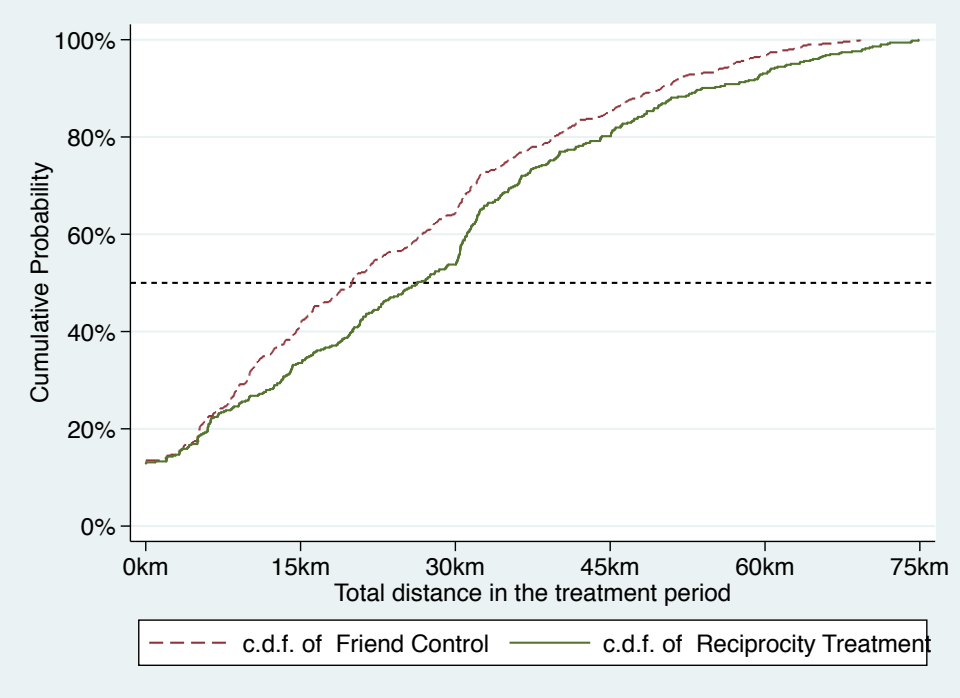


Figure 3. 7 Change of Performance across Time

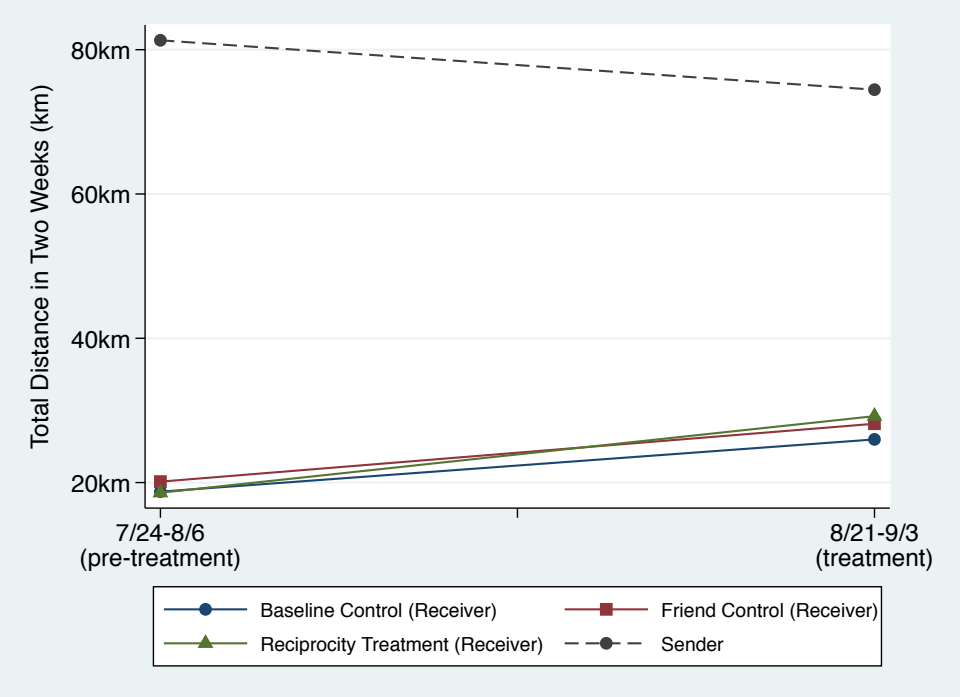
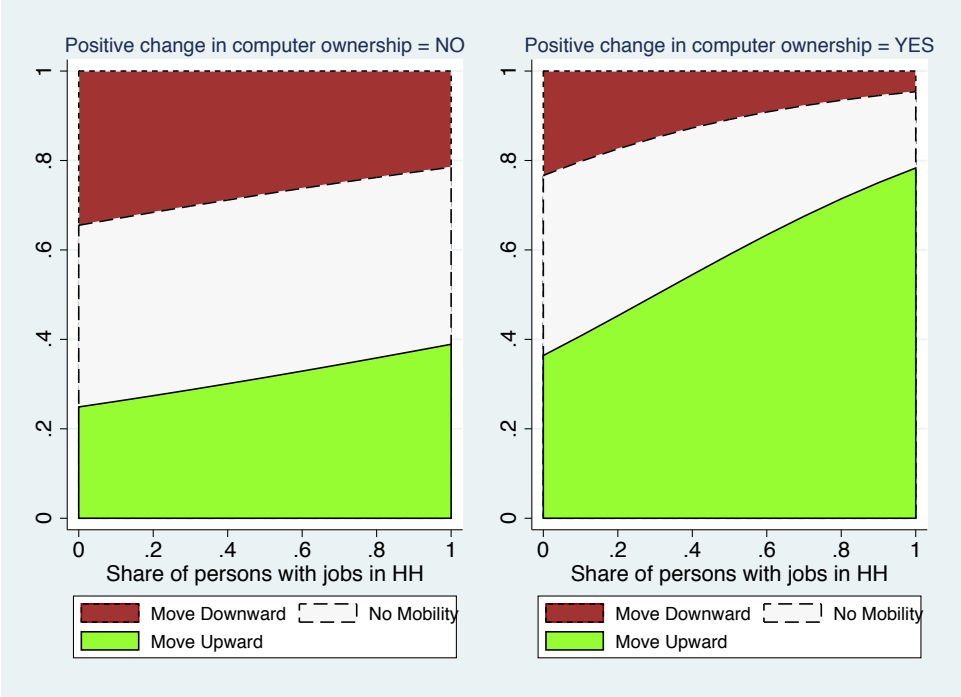


Figure 4. 1 How Share of Persons with Jobs in a Household Moderates the Relationship between Positive Change in Computer Ownership and Income Mobility



Appendices

Table A1. Summary statistics of each subgroup

Subgroup	1	2	3	4	5	6	7	8	9	10
N	720	720	720	720	720	720	720	720	720	716
Conditional Average Treatment Effects (Goal Setting)	-0.9%	2.2%	4.4%	-0.7%	4.3%	0.8%	-0.1%	1.0%	-2.5%	8.3%
Conditional Average Treatment Effects (Goal Attainment)	-7.3%	23.6%	9.9%	-0.8%	-6.8%	-2.5%	11.4%	-11.7%	-5.8%	-33.7%
Age	39.72	40.05	39.19	38.34	38.13	38.06	38.51	39.21	39.23	38.86
Gender (percent of female)	0.23	0.20	0.22	0.21	0.21	0.19	0.15	0.15	0.16	0.20
Daily distance in pre-treatment month	1.16	1.62	1.47	1.44	1.70	1.76	2.27	2.65	3.57	4.40
% Registered by Facebook account	0.24	0.33	0.55	0.57	0.62	0.61	0.55	0.52	0.44	0.42
Log (# of friends)	0.08	1.16	1.94	2.46	2.88	3.27	3.67	4.14	4.72	5.74
Log (# of races)	0.17	0.44	0.61	0.66	0.92	1.07	1.38	1.58	1.93	2.33
Log (tenure)	5.46	5.86	5.92	6.01	6.21	6.32	6.45	6.53	6.67	6.76

Table A2. Review of Gift Exchange Literature (Empirical Studies)

Authors	**Relationship between senders and receivers**	**Gifts**	**Induced Behaviors**	**Related Findings**
Alpizar et al. (2008)	Senders are unknown to the receivers	Non-monetary gifts (maps)	Donation	Giving gifts increases the likelihood of donations, but is far from profitable considering the costs of gifts.
Currie et al. (2013)		Non-monetary gifts (bookmarks)	Physicians' prescription	Physicians demonstrate reciprocity with better service when receiving small gifts from patients.
Chung and Narayandas (2017)		Monetary gifts	Workers' performance	The unconditional compensations increase sales force performance, but the magnitude is less than half that of the conditional compensations.
Falk (2007)		Non-monetary gifts (postcards)	Donation	Gifts treatments as compared to non-gifts treatment can elicit more donations in a sizable manner. Besides that, large gifts elicit more donations than small gifts.
Gneezy and List (2006)		Monetary gifts	Workers' performance	Gift treatment performs better than non-gift treatment only in the first few hours.
Kube et al. (2012)		Monetary gifts vs. Non-monetary gifts (bottle)	Workers' performance	The non-monetary gifts elicit stronger reciprocity than the monetary gifts.
Chen et al. (2009)	Senders are either strangers or friends	Monetary gifts	Monetary returns	Participants' positive reciprocity was not affected by the existing relationship (strangers vs. friends) between senders and receivers.
Falk et al. (1999)	Senders are unknown to the receivers, but social relationship is manipulated.	Monetary gifts	Workers' performance	Both social approval (participants seated Face-to-Face) and social pressure (participants allowed to discuss after the last round of the experiment) treatments do not statistically different from the baseline group to induce reciprocal behaviors.

Maréchal and Thöni (2007)	Senders and receiver might know each other before the treatment.	Non-monetary gifts (free sample)	Buyers' purchases	Free sample increases reciprocated behaviors, and the results are conditional on whether buyers and sellers know each other before.
This study	Senders are receivers' online friends	Non-monetary gifts (sports accessories)	Participants' physical activity	Reciprocity-based incentive outperforms self-interest based incentive on runners' physical activity. The social closeness between senders and receivers has an inverted U-shaped relationship on the effects of reciprocity.

Table A3. Review of Research on Income Mobility in India

Study	Source & Period	Sample	Main findings
Gaiha (1988)	National Council for Applied Economic Research (NCAER) data between 1968 and 1970	4,118 rural households	Access to cultivatable land and modern agricultural inputs play important roles to improve upward mobility for rural poor.
Bhide and Mehta (2004)	NCAER data set in the early 1970s and 1980s	3,319 rural households	This study shows that caste status is not an important factor of income mobility, but tribal status and demographic composition of households are. Furthermore, literacy, ownership of a house, increase in cultivated area and income from livestock increase the chance to escape from poverty.
Walker and Ryan (1990)	International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) data in 1970s and 1980s	219 rural households	Caste, inherited land, the number of bullocks inherited are important determinants for the income mobility.
Lanjouw and Stern (2003)	Data from Palanpur (a village in Moradabad district of Uttar Pradesh) between 1950s to 1990s	The village had 1,133 people, divided into 193 households in 1993 (last survey).	Village populations, agricultural practices, and occupational diversification are three forces of change for the village economy. They suggest that the non-farm economy plays a critical role in reducing poverty.
Gautam et al. (2012)	Additional Rural Income Survey (ARIS) and Rural Economic and Demographics (REDS) surveys in 1999 and 2007	5,885 rural households	The mobility is associated with land classes, agro-climate zones, agricultural profit classes, assets classes, caste groups, and gender groups. For instance, landless or marginal farmers are the most vulnerable groups regarding income or assets mobility.
Our research	Indian Human Development Survey (IHDS) 2005 and 2011.	23,358 rural households and 8,856 urban households	Computer ownership generates private returns for upward income mobility and creates social returns (spillover effects) for households in the same district, especially for the same caste group.

Figure A1. Conditional Average Treatment Effect by Generalized Additive Model

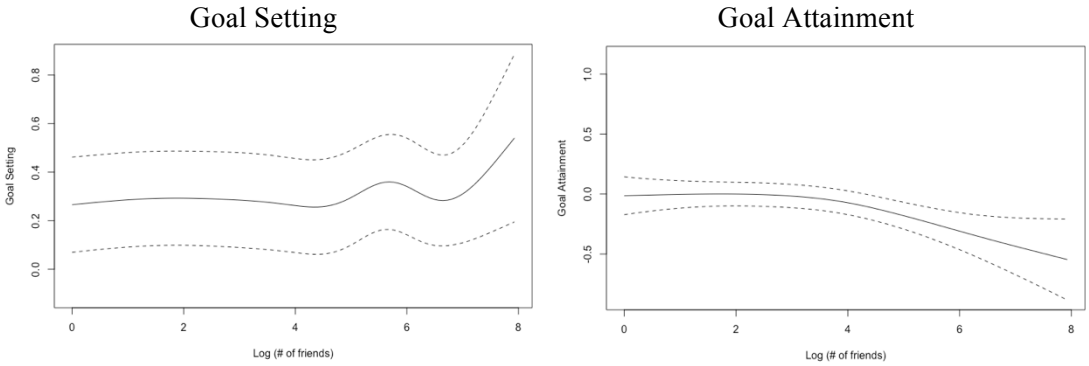
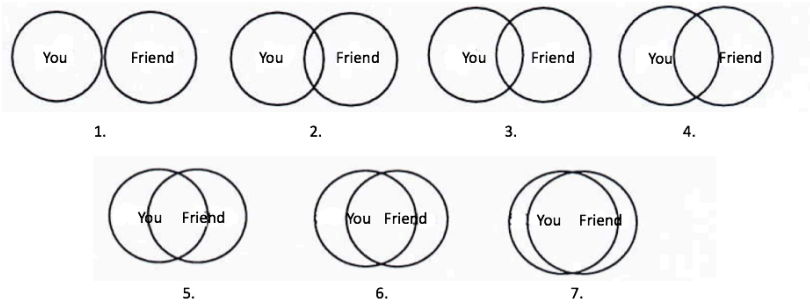


Figure A2. Inclusion of Other in the Self-scale (IOS)



Bibliography

Acemoglu, D. 2003. "Technology and Inequality," in: *NBER Reporter*.

Acland, D., and Levy, M. R. 2015. "Naiveté, Projection Bias, and Habit Formation in Gym Attendance," *Management Science* (61:1), pp. 146-160.

Acs, Z. J., Braunerhjelm, P., Audretsch, D. B., and Carlsson, B. 2009. "The Knowledge Spillover Theory of Entrepreneurship," *Small Business Economics* (32:1), pp. 15-30.

Adair, J. G. 1984. "The Hawthorne Effect: A Reconsideration of the Methodological Artifact," *Journal of Applied Psychology* (69:2), pp. 334-345.

Adams, C. 2008. "Educational Inequality in Institutionalized Social Stratification Systems: The Case of Caste-Based Educational Inequality in India," *Annual Meetings of the Population Association of America*, New Orleans, LA.

Agarwal, R., Gao, G., DesRoches, C., and Jha, A. K. 2010. "Research Commentary—the Digital Transformation of Healthcare: Current Status and the Road Ahead," *Information Systems Research* (21:4), pp. 796-809.

Aidis, R., Estrin, S., and Mickiewicz, T. 2008. "Institutions and Entrepreneurship Development in Russia: A Comparative Perspective," *Journal of Business Venturing* (23:6), pp. 656-672.

Akerlof, G. A. 1982. "Labor Contracts as Partial Gift Exchange," *The Quarterly Journal of Economics* (97:4), pp. 543-569.

Aldrich, H., Rosen, B., and Woodward, W. 1987. "The Impact of Social Networks on Business Foundings and Profit: A Longitudinal Study," *Frontiers of Entrepreneurship Research* (7:154), p. 68.

Allcott, H. 2011. "Social Norms and Energy Conservation," *Journal of Public Economics* (95:9), pp. 1082-1095.

Alpizar, F., Carlsson, F., and Johansson-Stenman, O. 2008. "Anonymity, Reciprocity, and Conformity: Evidence from Voluntary Contributions to a National Park in Costa Rica," *Journal of Public Economics* (92:5), pp. 1047-1060.

Ameriks, J., Caplin, A., Leahy, J., and Tyler, T. 2007. "Measuring Self-Control Problems," *American Economic Review* (97:3), pp. 966-972.

An, R. 2016. "Weekend-Weekday Differences in Diet among US Adults, 2003–2012," *Annals of Epidemiology* (26:1), pp. 57-65.

Arabsheibani, G. R., and Marin, A. 2006. "If Not Computers Then What? Returns to Computer Use in the UK Revisited," *Applied Economics* (38:21), pp. 2461-2467.

Aral, S., Muchnik, L., and Sundararajan, A. 2009. "Distinguishing Influence-Based Contagion from Homophily-Driven Diffusion in Dynamic Networks," *Proceedings of the National Academy of Sciences* (106:51), pp. 21544-21549.

Aral, S., and Walker, D. 2014. "Tie Strength, Embeddedness, and Social Influence: A Large-Scale Networked Experiment," *Management Science* (60:6), pp. 1352-1370.

Ariely, D., and Wertenbroch, K. 2002. "Procrastination, Deadlines, and Performance: Self-Control by Precommitment," *Psychological Science* (13:3), pp. 219-224.

Aron, A., Aron, E. N., and Smollan, D. 1992. "Inclusion of Other in the Self Scale and the Structure of Interpersonal Closeness," *Journal of Personality and Social Psychology* (63:4), p. 596.

Ashraf, N., Bandiera, O., and Jack, B. K. 2014. "No Margin, No Mission? A Field Experiment on Incentives for Public Service Delivery," *Journal of Public Economics* (120), pp. 1-17.
Baicker, K., Cutler, D., and Song, Z. 2010. "Workplace Wellness Programs Can Generate Savings," *Health Affairs* (29:2), pp. 304-311.
Ball, K., Jeffery, R. W., Abbott, G., McNaughton, S. A., and Crawford, D. 2010. "Is Healthy Behavior Contagious: Associations of Social Norms with Physical Activity and Healthy Eating," *Int J Behav Nutr Phys Act* (7:1), pp. 86-95.
Bandiera, O., Barankay, I., and Rasul, I. 2010. "Social Incentives in the Workplace," *The Review of Economic Studies* (77:2), pp. 417-458.
Bapna, R., Gupta, A., Rice, S., and Sundararajan, A. 2017. "Trust and the Strength of Ties in Online Social Networks: An Exploratory Field Experiment," *MIS Quarterly* (41:1).
Bapna, R., and Umyarov, A. 2015. "Do Your Online Friends Make You Pay? A Randomized Field Experiment on Peer Influence in Online Social Networks," *Management Science* (61:8), pp. 1920-1920.
Baron, R. M., and Kenny, D. A. 1986. "The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology* (51:6), p. 1173.
Bawa, K., and Shoemaker, R. 2004. "The Effects of Free Sample Promotions on Incremental Brand Sales," *Marketing Science* (23:3), pp. 345-363.
Becker, L. C. 1990. *Reciprocity* Chicago: University of Chicago Press.
Beckers, J. J., and Schmidt, H. G. 2001. "The Structure of Computer Anxiety: A Six-Factor Model," *Computers in Human Behavior* (17:1), pp. 35-49.
Beltramini, R. F. 1992. "Exploring the Effectiveness of Business Gifts: A Controlled Field Experiment," *Journal of the Academy of Marketing Science* (20:1), pp. 87-91.
Bergquist, M., and Ljungberg, J. 2001. "The Power of Gifts: Organizing Social Relationships in Open Source Communities," *Information Systems Journal* (11:4), pp. 305-320.
Bhide, S., and Mehta, A. K. 2004. "Chronic Poverty in Rural India: Issues and Findings from Panel Data," *Journal of Human Development* (5:2), pp. 195-209.
Bicchieri, C. 2005. *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press.
Biddle, S. J., and Mutrie, N. 2007. *Psychology of Physical Activity: Determinants, Well-Being and Interventions*, (2 ed.). Routledge.
Bittner, J. V., and Kulesz, M. M. 2015. "Health Promotion Messages: The Role of Social Presence for Food Choices," *Appetite* (87), pp. 336-343.
Bittner, J. V., and Zondervan, R. 2015. "Substitution Behaviours after an Interruption: Goal Failure and Goal Orientations," *Academy of Management Proceedings*: Academy of Management, p. 13508.
Blackwell, M., Iacus, S., King, G., and Porro, G. 2009. "CEM: Coarsened Exact Matching in Stata," *Stata Journal* (9:4), p. 524.

Bloom, N., Liang, J., Roberts, J., and Ying, Z. J. 2014. "Does Working from Home Work? Evidence from a Chinese Experiment," *The Quarterly Journal of Economics* (130:1), pp. 165-218.

Blumenthal, K. J., Saulsgiver, K. A., Norton, L., Troxel, A. B., Anarella, J. P., Gesten, F. C., Chernew, M. E., and Volpp, K. G. 2013. "Medicaid Incentive Programs to Encourage Healthy Behavior Show Mixed Results to Date and Should Be Studied and Improved," *Health Affairs* (32:3), pp. 497-507.

Borgatti, S. P., Mehra, A., Brass, D. J., and Labianca, G. 2009. "Network Analysis in the Social Sciences," *Science* (323:5916), pp. 892-895.

Bresnahan, T. F., and Trajtenberg, M. 1995. "General Purpose Technologies 'Engines of Growth'?," *Journal of Econometrics* (65:1), pp. 83-108.

Brunstein, J. C., and Gollwitzer, P. M. 1996. "Effects of Failure on Subsequent Performance: The Importance of Self-Defining Goals," *Journal of Personality and Social Psychology* (70:2), p. 395.

Bryan, G., Karlan, D., and Nelson, S. 2010. "Commitment Devices," *Annu. Rev. Econ.* (2:1), pp. 671-698.

Buchinsky, M., Fields, G., Fougere, D., and Kramarz, F. 2003. "Francs or Ranks? Earnings Mobility in France, 1967-1999."

Buffardi, L. E., and Campbell, W. K. 2008. "Narcissism and Social Networking Web Sites," *Personality and Social Psychology Bulletin* (34:10), pp. 1303-1314.

Burchell, K., Rettie, R., and Patel, K. 2013. "Marketing Social Norms: Social Marketing and the 'Social Norm Approach'," *Journal of Consumer Behaviour* (12:1), pp. 1-9.

Burtch, G., Hong, Y., Bapna, R., and Griskevicius, V. 2017. "Stimulating Online Reviews by Combining Financial Incentives and Social Norms," *Management Science* (Forthcoming).

Carneiro, P., Lokshin, M., and Umapathi, N. 2016. "Average and Marginal Returns to Upper Secondary Schooling in Indonesia," *Journal of Applied Econometrics* (32:1), pp. 16-36.

Cawley, J., and Price, J. A. 2013. "A Case Study of a Workplace Wellness Program That Offers Financial Incentives for Weight Loss," *Journal of Health Economics* (32:5), pp. 794-803.

Chang, Y. B., and Gurbaxani, V. 2012. "Information Technology Outsourcing, Knowledge Transfer, and Firm Productivity: An Empirical Analysis," *MIS Quarterly* (36:4), pp. 1043-1053.

Charness, G., and Gneezy, U. 2009. "Incentives to Exercise," *Econometrica* (77:3), pp. 909-931.

Chen, Y., Harper, F. M., Konstan, J., and Xin Li, S. 2010. "Social Comparisons and Contributions to Online Communities: A Field Experiment on Movielens," *American Economic Review* (100:4), pp. 1358-1398.

Chen, Y.-R., Chen, X.-P., and Portnoy, R. 2009. "To Whom Do Positive Norm and Negative Norm of Reciprocity Apply? Effects of Inequitable Offer, Relationship, and Relational-Self Orientation," *Journal of Experimental Social Psychology* (45:1), pp. 24-34.

Cheng, Z., and Nault, B. R. 2007. "Industry Level Supplier-Driven IT Spillovers," *Management Science* (53:8), pp. 1199-1216.

Chetty, R., Hendren, N., Kline, P., and Saez, E. 2014. "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States," *The Quarterly Journal of Economics* (129:4), pp. 1553-1623.

Chua, S. L., Chen, D.-T., and Wong, A. F. 1999. "Computer Anxiety and Its Correlates: A Meta-Analysis,," *Computers in Human Behavior* (15:5), pp. 609-623.

Chung, D. J., and Narayandas, D. 2017. "Incentives Versus Reciprocity: Insights from a Field Experiment," *Journal of Marketing Research* (54:4), pp. 511-524.

Claxton, G., Rae, M., Long, M., Damico, A., Whitmore, H., and Foster, G. 2016. "Health Benefits in 2016: Family Premiums Rose Modestly, and Offer Rates Remained Stable," *Health Affairs* (35:10), pp. 1908-1917.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*, (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Coleman, J. S. 1988. "Social Capital in the Creation of Human Capital," *American Journal of Sociology* (94), pp. S95-S120.

Costa, D. L., and Kahn, M. E. 2013. "Energy Conservation "Nudges" and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment," *Journal of the European Economic Association* (11:3), pp. 680-702.

Cuesta, J., Ñopo, H., and Pizzolitto, G. 2011. "Using Pseudo-Panels to Measure Income Mobility in Latin America," *Review of Income and Wealth* (57:2), pp. 224-246.

Currie, J., Lin, W., and Meng, J. 2013. "Social Networks and Externalities from Gift Exchange: Evidence from a Field Experiment," *Journal of Public Economics* (107), pp. 19-30.

Currie, J., and Moretti, E. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings," *The Quarterly Journal of Economics* (118:4), pp. 1495-1532.

Dahl, D. W., Honea, H., and Manchanda, R. V. 2005. "Three Rs of Interpersonal Consumer Guilt: Relationship, Reciprocity, Reparation," *Journal of Consumer Psychology* (15:4), pp. 307-315.

Dann, S. 2010. "Twitter Content Classification," *First Monday* (15:12).

Dartanto, T., and Nurkholis. 2013. "The Determinants of Poverty Dynamics in Indonesia: Evidence from Panel Data," *Bulletin of Indonesian Economic Studies* (49:1), pp. 61-84.

David, H., Katz, L. F., and Krueger, A. B. 1997. "Computing Inequality: Have Computers Changed the Labor Market?," National Bureau of Economic Research.

Dawson, J. F. 2014. "Moderation in Management Research: What, Why, When, and How," *Journal of Business and Psychology* (29:1), pp. 1-19.

De Maesschalck, R., Jouan-Rimbaud, D., and Massart, D. L. 2000. "The Mahalanobis Distance," *Chemometrics and Intelligent Laboratory Systems* (50:1), pp. 1-18.

Deininger, K., and Squire, L. 1996. "A New Data Set Measuring Income Inequality," *The World Bank Economic Review* (10:3), pp. 565-591.

Desai, S., and Adams, C. 2006. "In the Margins: Social Inequality in Children's Educational Outcomes in India," *Meetings of the Population Association of America*, Los Angeles, CA.

Desai, S., and Dubey, A. 2012. "Caste in 21st Century India: Competing Narratives," *Economic and Political Weekly* (46:11), p. 40.

Desai, S., and Kulkarni, V. 2008. "Changing Educational Inequalities in India in the Context of Affirmative Action," *Demography* (45:2), pp. 245-270.

Desai, S., and Vanneman, R. 2016. "India Human Development Survey-II (IHDS-II), 2011-12," in: *ICPSR36151-v5. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]*.

Desai, S., Vanneman, R., and National Council of Applied Economic Research, New Delhi,. 2016. "India Human Development Survey (IHDS), 2005," in: *ICPSR22626-v11. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]*.

Dewan, S., Ho, Y.-J., and Ramaprasad, J. 2017. "Popularity or Proximity: Characterizing the Nature of Social Influence in an Online Music Community," *Information Systems Research* (28:1), pp. 117-136.

Ding, N., and Wang, Y. 2008. "Household Income Mobility in China and Its Decomposition," *China Economic Review* (19:3), pp. 373-380.

Dolton, P., and Makepeace, G. 2004. "Computer Use and Earnings in Britain," *The Economic Journal* (114:494), pp. C117-C129.

Doran, G. T. 1981. "There's a Smart Way to Write Management's Goals and Objectives," *Management Review* (70:11), pp. 35-36.

Dorresteijn, J. A., Visseren, F. L., Ridker, P. M., Wassink, A. M., Paynter, N. P., Steyerberg, E. W., van der Graaf, Y., and Cook, N. R. 2011. "Estimating Treatment Effects for Individual Patients Based on the Results of Randomised Clinical Trials," *BMJ* (343), p. d5888.

Dutta, S. 2015. "Identifying Single or Multiple Poverty Trap: An Application to Indian Household Panel Data," *Social Indicators Research* (120:1), pp. 157-179.

Dyck, J. L., and Smither, J. A.-A. 1994. "Age Differences in Computer Anxiety: The Role of Computer Experience, Gender and Education," *Journal of Educational Computing Research* (10:3), pp. 239-248.

Economic Survey of India. 2015. "Economic Survey of India 2015, Available at http://www.indiabudget.nic.in/Survey.Asp," Finance Ministry, Government of India, New Delhi.

Ekeh, P. P. 1974. *Social Exchange Theory: The Two Traditions*. Cambridge, MA: Harvard University Press.

Ellison, N. B. 2007. "Social Network Sites: Definition, History, and Scholarship," *Journal of Computer-Mediated Communication* (13:1), pp. 210-230.

Emerson, R. M. 1972. "Exchange Theory, Part I: A Psychological Basis for Social Exchange," *Sociological Theories in Progress* (2), pp. 38-57.

Exley, C. L., and Naecker, J. K. 2016. "Observability Increases the Demand for Commitment Devices," *Management Science* (Articles in Advance).

Facer, K., Sutherland, R., Furlong, R., and Furlong, J. 2001. "What's the Point of Using Computers? The Development of Young People's Computer Expertise in the Home," *New Media & Society* (3:2), pp. 199-219.
Falk, A. 2007. "Gift Exchange in the Field," *Econometrica* (75:5), pp. 1501-1511.
Falk, A., Gächter, S., and Kovács, J. 1999. "Intrinsic Motivation and Extrinsic Incentives in a Repeated Game with Incomplete Contracts," *Journal of Economic Psychology* (20:3), pp. 251-284.
Fehr, E., and Gächter, S. 2000. "Fairness and Retaliation: The Economics of Reciprocity," *The Journal of Economic Perspectives* (14:3), pp. 159-181.
Fehr, E., Kirchsteiger, G., and Riedl, A. 1993. "Does Fairness Prevent Market Clearing? An Experimental Investigation," *The Quarterly Journal of Economics* (108:2), pp. 437-459.
Feller, A., and Holmes, C. C. 2009. "Beyond Toplines: Heterogeneous Treatment Effects in Randomized Experiments." Working Paper, Oxford University.
Ferraro, P. J., and Miranda, J. J. 2013. "Heterogeneous Treatment Effects and Mechanisms in Information-Based Environmental Policies: Evidence from a Large-Scale Field Experiment," *Resource and Energy Economics* (35:3), pp. 356-379.
Fields, G. S. 2007. "Income Mobility in India."
Finkelstein, E. A., Brown, D. S., Brown, D. R., and Buchner, D. M. 2008. "A Randomized Study of Financial Incentives to Increase Physical Activity among Sedentary Older Adults," *Preventive Medicine* (47:2), pp. 182-187.
Fiske, A. P. 1992. "The Four Elementary Forms of Sociality: Framework for a Unified Theory of Social Relations," *Psychological Review* (99:4), p. 689.
Fjeldsoe, B. S., Marshall, A. L., and Miller, Y. D. 2009. "Behavior Change Interventions Delivered by Mobile Telephone Short-Message Service," *American Journal of Preventive Medicine* (36:2), pp. 165-173.
Free, C., Phillips, G., Galli, L., Watson, L., Felix, L., Edwards, P., Patel, V., and Haines, A. 2013. "The Effectiveness of Mobile-Health Technology-Based Health Behaviour Change or Disease Management Interventions for Health Care Consumers: A Systematic Review," *PLoS Med* (10:1), p. e1001362.
Freeman, L. C. 1978. "Centrality in Social Networks Conceptual Clarification," *Social Networks* (1:3), pp. 215-239.
Frenzen, J., and Nakamoto, K. 1993. "Structure, Cooperation, and the Flow of Market Information," *Journal of Consumer Research* (20:3), pp. 360-375.
Frijters, P., and Sharma, A. 2012. "Dowry-Stature Link: Theory and Evidence," *41st Australian Conference of Economists*, Melbourne, Australia: University of Victoria.
Furnell, S., Bryant, P., and Phippen, A. D. 2007. "Assessing the Security Perceptions of Personal Internet Users," *Computers & Security* (26:5), pp. 410-417.
Gaiha, R. 1988. "Income Mobility in Rural India," *Economic Development and Cultural Change* (36:2), pp. 279-302.
Gautam, M., Nagarajan, H. K., and Pradhan, K. C. 2012. *Income, Consumption and Asset Mobility in Indian Rural Households : Evidence from Aris/ Reds Surveys*. National Council of Applied Economic Research (NCAER).

Gerber, A. S., and Rogers, T. 2009. "Descriptive Social Norms and Motivation to Vote: Everybody's Voting and So Should You," *The Journal of Politics* (71:01), pp. 178-191.

Ghosh, S. 2015. "Narendra Modi to Launch Digital India Week on 1 July, http://www.livemint.com/Politics/Zhmiym6byip3fd3ojrb1jk/Narendra-Modi-to-Launch-Digital-India-Week-on-1-July.Html," in: *Livemint (27 June 2015)*.

Gilbert, E., and Karahalios, K. 2009. "Predicting Tie Strength with Social Media," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: ACM, pp. 211-220.

Gille, V. 2011. "Education Spillovers in Farm Productivity: Empirical Evidence in Rural India," *Proceedings of the German Development Economics Conference*, Berlin: Verein für Socialpolitik, Research Committee Development Economics.

Giné, X., Karlan, D., and Zinman, J. 2010. "Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation," *American Economic Journal: Applied Economics* (2:4), pp. 213-235.

Gneezy, U., and List, J. A. 2006. "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments," *Econometrica* (74:5), pp. 1365-1384.

Gneezy, U., Meier, S., and Rey-Biel, P. 2011. "When and Why Incentives (Don't) Work to Modify Behavior," *The Journal of Economic Perspectives* (25:4), pp. 191-209.

Goldstein, N. J., Cialdini, R. B., and Griskevicius, V. 2008. "A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels," *Journal of Consumer Research* (35:3), pp. 472-482.

Goldstein, N. J., Griskevicius, V., and Cialdini, R. B. 2007. "Invoking Social Norms a Social Psychology Perspective on Improving Hotels' Linen-Reuse Programs," *Cornell Hotel and Restaurant Administration Quarterly* (48:2), pp. 145-150.

Gouldner, A. W. 1960. "The Norm of Reciprocity: A Preliminary Statement," *American Sociological Review* (25:2), pp. 161-178.

Goux, D., and Maurin, E. 2007. "Close Neighbours Matter: Neighbourhood Effects on Early Performance at School," *The Economic Journal* (117:523), pp. 1193-1215.

Granovetter, M. 1985. "Economic Action and Social Structure: The Problem of Embeddedness," *American Journal of Sociology* (91:3), pp. 481-510.

Granovetter, M. S. 1973. "The Strength of Weak Ties," *American Journal of Sociology* (78:6), pp. 1360-1380.

Grant, A. M., and Gino, F. 2010. "A Little Thanks Goes a Long Way: Explaining Why Gratitude Expressions Motivate Prosocial Behavior," *Journal of Personality and Social Psychology* (98:6), p. 946.

Green, D. P., and Kern, H. L. 2012. "Modeling Heterogeneous Treatment Effects in Survey Experiments with Bayesian Additive Regression Trees," *Public Opinion Quarterly* (76:3), pp. 491-511.

Greene, W. H. 2012. *Econometric Analysis*, (7th ed.). Upper Saddle River, NJ: Prentice Hall:

Grewal, R., Lilien, G. L., and Mallapragada, G. 2006. "Location, Location, Location: How Network Embeddedness Affects Project Success in Open Source Systems," *Management Science* (52:7), pp. 1043-1056.

Gudeman, S. 2001. *The Anthropology of Economy: Community, Market, and Culture Malden*. Oxford: Blackwell.

Haans, R. F., Pieters, C., and He, Z. L. 2016. "Thinking About U: Theorizing and Testing U-and Inverted U-Shaped Relationships in Strategy Research," *Strategic Management Journal* (37:7), pp. 1177-1195.

Hackman, J. R. 1992. *Group Influences on Individuals in Organizations*, (2 ed.). Consulting Psychologists Press.

Haddad, L., and Ahmed, A. 2003. "Chronic and Transitory Poverty: Evidence from Egypt, 1997–99," *World Development* (31:1), pp. 71-85.

Halpern, S. D., Asch, D. A., and Volpp, K. G. 2012. "Commitment Contracts as a Way to Health," *BMJ* (344), p. e522.

Halpern, S. D., French, B., Small, D. S., Saulsgiver, K., Harhay, M. O., Audrain-McGovern, J., Loewenstein, G., Brennan, T. A., Asch, D. A., and Volpp, K. G. 2015. "Randomized Trial of Four Financial-Incentive Programs for Smoking Cessation," *New England Journal of Medicine* (372:22), pp. 2108-2117.

Han, K., Chang, Y. B., and Hahn, J. 2011. "Information Technology Spillover and Productivity: The Role of Information Technology Intensity and Competition," *Journal of Management Information Systems* (28:1), pp. 115-146.

Handel, B., and Kolstad, J. 2017. "Wearable Technologies and Health Behaviors: New Data and New Methods to Understand Population Health," *American Economic Review* (107:5), pp. 481-485.

Heckman, J. J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models," *Annals of Economic and Social Measurement,* (5:4), pp. 475-492.

Heckman, J. J. 1979. "Sample Selection Bias as a Specification Error," (47:1), pp. 153-161.

Heckman, J. J. 2000. "Policies to Foster Human Capital," *Research in Economics* (54:1), pp. 3-56.

Herman, C. W., Musich, S., Lu, C., Sill, S., Young, J. M., and Edington, D. W. 2006. "Effectiveness of an Incentive-Based Online Physical Activity Intervention on Employee Health Status," *Journal of Occupational and Environmental Medicine* (48:9), pp. 889-895.

Higgs, S. 2015. "Social Norms and Their Influence on Eating Behaviours," *Appetite* (86), pp. 38-44.

Hinz, O., Skiera, B., Barrot, C., and Becker, J. U. 2011. "Seeding Strategies for Viral Marketing: An Empirical Comparison," *Journal of Marketing* (75:6), pp. 55-71.

Hinz, O., Spann, M., and Hann, I.-H. 2015. "Research Note—Can't Buy Me Love… or Can I? Social Capital Attainment through Conspicuous Consumption in Virtual Environments," *Information Systems Research* (26:4), pp. 859-870.

Homans, G. C. 1974. *Social Behavior: Its Elementary Forms*. New York: Harcourt Brace Jovanovich.

Hsiaw, A. 2013. "Goal-Setting and Self-Control," *Journal of Economic Theory* (148:2), pp. 601-626.

Hsieh, J. P.-A., Rai, A., and Keil, M. 2008. "Understanding Digital Inequality: Comparing Continued Use Behavioral Models of the Socio-Economically Advantaged and Disadvantaged," *MIS Quarterly* (32:1), pp. 97-126.

Hsieh, J. P.-A., Rai, A., and Keil, M. 2011. "Addressing Digital Inequality for the Socioeconomically Disadvantaged through Government Initiatives: Forms of Capital That Affect Ict Utilization," *Information Systems Research* (22:2), pp. 233-253.

Hughes, J. R., Keely, J., and Naud, S. 2004. "Shape of the Relapse Curve and Long-Term Abstinence among Untreated Smokers," *Addiction* (99:1), pp. 29-38.

Hurling, R., Catt, M., De Boni, M., Fairley, B. W., Hurst, T., Murray, P., Richardson, A., and Sodhi, J. S. 2007. "Using Internet and Mobile Phone Technology to Deliver an Automated Physical Activity Program: Randomized Controlled Trial," *Journal of Medical Internet Research* (9:2).

Iyengar, R., Van den Bulte, C., and Valente, T. W. 2011. "Opinion Leadership and Social Contagion in New Product Diffusion," *Marketing Science* (30:2), pp. 195-212.

Jarvis, S., and Jenkins, S. P. 1995. "Do the Poor Stay Poor? New Evidence About Income Dynamics from the British Household Panel Survey," *Occasional Paper*: ESRC Research Centre on Micro-social Change.

Java, A., Song, X., Finin, T., and Tseng, B. 2007. "Why We Twitter: Understanding Microblogging Usage and Communities," *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*: ACM, pp. 56-65.

Jha, S. K., Pinsonneault, A., and Dube, L. 2016. "The Evolution of an ICT Platform-Enabled Ecosystem for Poverty Alleviation: The Case of Ekutir," *MIS Quarterly* (40:2), pp. 431-445.

John, L. K., Loewenstein, G., Troxel, A. B., Norton, L., Fassbender, J. E., and Volpp, K. G. 2011. "Financial Incentives for Extended Weight Loss: A Randomized, Controlled Trial," *Journal of General Internal Medicine* (26:6), pp. 621-626.

Jones, C. I., and Williams, J. C. 1998. "Measuring the Social Return to R&D," *The Quarterly Journal of Economics* (113:4), pp. 1119-1135.

Jones, N. P., Papadakis, A. A., Orr, C. A., and Strauman, T. J. 2013. "Cognitive Processes in Response to Goal Failure: A Study of Ruminative Thought and Its Affective Consequences," *Journal of Social and Clinical Psychology* (32:5), pp. 482-503.

Kalnins, A., and Chung, W. 2006. "Social Capital, Geography, and Survival: Gujarati Immigrant Entrepreneurs in the U.S. Lodging Industry," *Management Science* (52:2), pp. 233-247.

Kane, T. J., and Rouse, C. E. 1995. "Labor-Market Returns to Two-and Four-Year College," *American Economic Review* (85:3), pp. 600-614.

Kast, F., Meier, S., and Pomeranz, D. 2012. "Under-Savers Anonymous: Evidence on Self-Help Groups and Peer Pressure as a Savings Commitment Device," National Bureau of Economic Research.

Katona, Z., Zubcsek, P. P., and Sarvary, M. 2011. "Network Effects and Personal Influences: The Diffusion of an Online Social Network," *Journal of Marketing Research* (48:3), pp. 425-443.

Katz, M. L., and Shapiro, C. 1986. "Technology Adoption in the Presence of Network Externalities," *The Journal of Political Economy* (4), pp. 822-841.

Katz, N., Lazer, D., Arrow, H., and Contractor, N. 2004. "Network Theory and Small Groups," *Small Group Research* (35:3), pp. 307-332.

Kendall, J., Maurer, B., Machoka, P., and Veniard, C. 2011. "An Emerging Platform: From Money Transfer System to Mobile Money Ecosystem," *Innovations* (6:4), pp. 49-64.

Kent, D. M., and Hayward, R. A. 2007. "Limitations of Applying Summary Results of Clinical Trials to Individual Patients: The Need for Risk Stratification," *JAMA* (298:10), pp. 1209-1212.

Kiernan, M., Moore, S. D., Schoffman, D. E., Lee, K., King, A. C., Taylor, C. B., Kiernan, N. E., and Perri, M. G. 2012. "Social Support for Healthy Behaviors: Scale Psychometrics and Prediction of Weight Loss among Women in a Behavioral Program," *Obesity* (20:4), pp. 756-764.

Koch, A. K., and Nafziger, J. 2011. "Self-Regulation through Goal Setting*," *The Scandinavian Journal of Economics* (113:1), pp. 212-227.

Kolm, S. 2000. "Introduction to the Economics of Altruism, Giving, and Reciprocity." Norwegian School of Economics and Business Administration.

Komter, A. E. 2004. *Gratitude and Gift Exchange*. New York: Oxford University Press, pp. 195–213.

Kosinski, M., Stillwell, D., and Graepel, T. 2013. "Private Traits and Attributes Are Predictable from Digital Records of Human Behavior," *Proceedings of the National Academy of Sciences* (110:15), pp. 5802-5805.

Krämer, N. C., and Winter, S. 2008. "Impression Management 2.0: The Relationship of Self-Esteem, Extraversion, Self-Efficacy, and Self-Presentation within Social Networking Sites," *Journal of Media Psychology: Theories, Methods, and Applications* (20:3), p. 106.

Krebs, P., and Duncan, D. T. 2015. "Health App Use among US Mobile Phone Owners: A National Survey," *JMIR mHealth and uHealth* (3:4), p. e101.

Krishnamurthy, B., Gill, P., and Arlitt, M. 2008. "A Few Chirps About Twitter," *Proceedings of the First Workshop on Online Social Networks*: ACM, pp. 19-24.

Krueger, A. B. 1993. "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989," *The Quarterly Journal of Economics* (108:1), pp. 33-60.

Kube, S., Maréchal, M. A., and Puppea, C. 2012. "The Currency of Reciprocity: Gift Exchange in the Workplace," *American Economic Review* (102:4), pp. 1644-1662.

Lally, P., Bartle, N., and Wardle, J. 2011. "Social Norms and Diet in Adolescents," *Appetite* (57:3), pp. 623-627.

Lampe, C. A., Ellison, N., and Steinfield, C. 2007. "A Familiar Face (Book): Profile Elements as Signals in an Online Social Network," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: ACM, pp. 435-444.

Lanjouw, P., and Stern, N. 2003. *Opportunities Off the Farm as a Springboard out of Rural Poverty: Five Decades of Development in an Indian Village*. Springer.

Leon-Ledesma, M. A. 2005. "Exports, Product Differentiation and Knowledge Spillovers," *Open Economies Review* (16:4), pp. 363-379.

Leong, C. M. L., Pan, S.-L., Newell, S., and Cui, L. 2016. "The Emergence of Self-Organizing E-Commerce Ecosystems in Remote Villages of China: A Tale of Digital Empowerment for Rural Development," *Mis Quarterly* (40:2), pp. 475-484.

Leunga, R., Lawb, R., Masieroc, L., and Hatterd, R. 2016. "Behavior of Online Visitors to Hotel Icon: A Weekday-Weekend Analysis," *e-Review of Tourism Research (eRTR)* (7).

Leuven, E., and Sianesi, B. 2015. "Psmatch2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing," in: *Statistical Software Components*.

Levine, D. K. 1998. "Modeling Altruism and Spitefulness in Experiments," *Review of Economic Dynamics* (1:3), pp. 593-622.

Liebowitz, J. 2006. "Keynote Paper: Developing Knowledge and Learning Strategies in Mobile Organisations," *International Journal of Mobile Learning and Organisation* (1:1), pp. 5-14.

Lin, M., Prabhala, N. R., and Viswanathan, S. 2013. "Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending," *Management Science* (59:1), pp. 17-35.

Lind, J. T., and Mehlum, H. 2010. "With or without U? The Appropriate Test for a U-Shaped Relationship," *Oxford Bulletin of Economics and Statistics* (72:1), pp. 109-118.

Link, T. M., and Marz, R. 2006. "Computer Literacy and Attitudes Towards E-Learning among First Year Medical Students," *BMC Medical Education* (6:1), p. 34.

List, J. A., and Rasul, I. 2011. "Field Experiments in Labor Economics," *Handbook of Labor Economics* (4), pp. 103-228.

Liu, C.-W., and Mithas, S. 2016. "The Returns to Digital Skills: Evidence from India, 2005-2011," in: *Proceedings of the 35th International Conference on Information Systems (Dec 11-14),* P. Agerfalk, N. Levina and S.S. Kien (eds.). Dublin, Ireland: Association for Information Systems.

Locke, E. A., and Latham, G. P. 2002. "Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey," *American Psychologist* (57:9), p. 705.

Long, H. 2015. "Stephen Hawking: Technology Is Making Inequality Worse." from http://money.cnn.com/2015/10/12/news/economy/stephen-hawking-technology-inequality/

Lukiyanova, A., and Oshchepkov, A. 2012. "Income Mobility in Russia (2000–2005)," *Economic Systems* (36:1), pp. 46-64.

Mackie, G., Moneti, F., Denny, E., and Shakya, H. 2012. "What Are Social Norms? How Are They Measured," in: *University of California at San Diego-UNICEF Working Paper, San Diego*.

Mahajan, P. 2010. "School Libraries in India: Present-Day Scenario," in: *Library Philosophy and Practice (e-journal)*. p. 371.

Manago, A. M., Graham, M. B., Greenfield, P. M., and Salimkhan, G. 2008. "Self-Presentation and Gender on Myspace," *Journal of Applied Developmental Psychology* (29:6), pp. 446-458.

Manda, D. K., Mwabu, G. M., and Kimenyi, M. S. 2002. *Human Capital Externalities and Returns to Education in Kenya*. Social Sector Division, Kenya Institute for Public Policy Research and Analysis.

Manski, C. F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies* (60:3), pp. 531-542.

Maréchal, M. A., and Thöni, C. 2007. "Do Managers Reciprocate? Field Experimental Evidence from a Competitive Market." Retrieved from SSRN at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=982079 on 09/11/2017.

Marshall, S., Kerr, J., Carlson, J., Cadmus-Bertram, L., Patterson, R., Wasilenko, K., Crist, K., Rosenberg, D., and Natarajan, L. 2015. "Patterns of Weekday and Weekend Sedentary Behavior among Older Adults," *Journal of Aging and Physical Activity* (23:4), pp. 534-541.

Mas, A., and Moretti, E. 2009. "Peers at Work," *American Economic Review* (99:1), pp. 112-145.

Mauss, M., and Halls, W. D. 1954/2000. *The Gift: The Form and Reason for Exchange in Archaic Societies*. WW Norton & Company (Original work published 1954).

McFarlane, J., Martin, C. L., and Williams, T. M. 1988. "Mood Fluctuations," *Psychology of Women Quarterly* (12:2), pp. 201-223.

Mento, A. J., Steel, R. P., and Karren, R. J. 1987. "A Meta-Analytic Study of the Effects of Goal Setting on Task Performance: 1966–1984," *Organizational Behavior and Human Decision Processes* (39:1), pp. 52-83.

Michelle, D., Kenyon, C., Guodong, G., Timothy, X., Ritu, A., Arie, K., and Nanette, S. 2018. "Individual Differences in Regulatory Mode Moderate the Effectiveness of a Pilot Mhealth Trial for Diabetes Management among Older Veterans.," *PLoS ONE* (13:3).

Milkman, K. L., Minson, J. A., and Volpp, K. G. 2013. "Holding the Hunger Games Hostage at the Gym: An Evaluation of Temptation Bundling," *Management Science* (60:2), pp. 283-299.

Mithas, S., Almirall, D., and Krishnan, M. S. 2006. "Do CRM Systems Cause One-to-One Marketing Effectiveness?," *Statistical Science* (21:2), pp. 223-233.

Mithas, S., and Krishnan, M. S. 2009. "From Association to Causation Via a Potential Outcomes Approach," *Information Systems Research* (20:2), pp. 295-313.

Mittone, L., and Ploner, M. 2011. "Peer Pressure, Social Spillovers, and Reciprocity: An Experimental Analysis," *Experimental Economics* (14:2), pp. 203-222.

Moretti, E. 2004. "Human Capital Externalities in Cities," *Handbook of Regional and Urban Economics* (4), pp. 2243-2291.

Morgenroth, T., Ryan, M. K., and Peters, K. 2015. "The Motivational Theory of Role Modeling: How Role Models Influence Role Aspirants' Goals," *Review of General Psychology* (19:4), p. 465.

Mun, S.-B., and Nadiri, M. I. 2002. "Information Technology Externalities: Empirical Evidence from 42 US Industries," National bureau of economic research.

Napolitano, M. A., Hayes, S., Bennett, G. G., Ives, A. K., and Foster, G. D. 2013. "Using Facebook and Text Messaging to Deliver a Weight Loss Program to College Students," *Obesity* (21:1), pp. 25-31.

Naydeck, B. L., Pearson, J. A., Ozminkowski, R. J., Day, B. T., and Goetzel, R. Z. 2008. "The Impact of the Highmark Employee Wellness Programs on 4-Year Healthcare Costs," *Journal of Occupational and Environmental Medicine* (50:2), pp. 146-156.

Nilekani, N. 2008. *Imagining India*. New Delhi: Penguin (Allen Lane).

Nitzan, I., and Libai, B. 2011. "Social Effects on Customer Retention," *Journal of Marketing* (75:6), pp. 24-38.

Ntoumanis, N., Healy, L. C., Sedikides, C., Smith, A. L., and Duda, J. L. 2014. "Self-Regulatory Responses to Unattainable Goals: The Role of Goal Motives," *Self and Identity* (13:5), pp. 594-612.

Ogunkola, B. J. 2008. "Computer Attitude, Ownership and Use as Predictors of Computer Literacy of Science Teachers in Nigeria," *International Journal of Environmental and Science Education* (3:2), pp. 53-57.

Oosterbeek, H. 1997. "Returns from Computer Use: A Simple Test on the Productivity Interpretation," *Economics Letters* (55:2), pp. 273-277.

Osilla, K. C., Van Busum, K., Schnyer, C., Larkin, J. W., Eibner, C., and Mattke, S. 2012. "Systematic Review of the Impact of Worksite Wellness Programs," *American Journal of Managed Care* (18:2), pp. e68-81.

Pal, J., Lakshmanan, M., and Toyama, K. 2007. "“My Child Will Be Respected”: Parental Perspectives on Computers in Rural India," *Proc. of 2nd IEEE/ACM International Conference on Information and Communication Technologies and Development*: IEEE, pp. 1-9.

Palmatier, R. W., Jarvis, C. B., Bechkoff, J. R., and Kardes, F. R. 2009. "The Role of Customer Gratitude in Relationship Marketing," *Journal of Marketing* (73:5), pp. 1-18.

Parker, C., Ramdas, K., and Savva, N. 2016. "Is IT Enough? Evidence from a Natural Experiment in India's Agriculture Markets," *Management Science* (62:9), pp. 2481-2503.

Peluchette, J., and Karl, K. 2009. "Examining Students' Intended Image on Facebook:“What Were They Thinking?!”," *Journal of Education for Business* (85:1), pp. 30-37.

Poynton, T. A. 2005. "Computer Literacy across the Lifespan: A Review with Implications for Educators," *Computers in Human Behavior* (21:6), pp. 861-872.

Reilly, K. T. 1995. "Human Capital and Information: The Employer Size-Wage Effect," *Journal of Human Resources* (30), pp. 1-18.

Rice, R. E., Collins-Jarvis, L., and Zydney-Walker, S. 1999. "Individual and Structural Influences on Information Technology Helping Relationships," *Journal of Applied Communication Research* (27:4), pp. 285-309.
Roberts, B. J. 2001. "Chronic and Transitory Poverty in Post-Apartheid South Africa: Evidence from Kwazulu-Natal," *Journal of Poverty* (5:4), pp. 1-28.
Rogers, T., Milkman, K. L., and Volpp, K. G. 2014. "Commitment Devices: Using Initiatives to Change Behavior," *JAMA* (311:20), pp. 2065-2066.
Rosenbaum, P. R., and Rubin, D. B. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika* (70:1), pp. 41-55.
Rotman, D. 2014. "Technology and Inequality." from https://www.technologyreview.com/s/531726/technology-and-inequality/
Rubin, D. B. 1980. "Bias Reduction Using Mahalanobis-Metric Matching," *Biometrics* (36:2), pp. 293-298.
Ryan, R. M., Bernstein, J. H., and Brown, K. W. 2010. "Weekends, Work, and Well-Being: Psychological Need Satisfactions and Day of the Week Effects on Mood, Vitality, and Physical Symptoms," *Journal of Social and Clinical Psychology* (29:1), pp. 95-122.
Rybczynski, W., and Glancy, M. 1992. "Waiting for the Weekend," *Journal of Leisure Research* (24:3), p. 295.
Ryu, G., and Feick, L. 2007. "A Penny for Your Thoughts: Referral Reward Programs and Referral Likelihood," *Journal of Marketing* (71:1), pp. 84-94.
Sawhill, I. V., and Condon, M. 1992. "Is US Income Inequality Really Growing? Sorting out the Fairness Question," *Policy Bites* (13).
Saxenian, A. 1996. *Beyond Boundaries: Open Labor Markets and Learning in Silicon Valley*. New York: Oxford University Press. pp. 23-39.
Scarborough, P., Bhatnagar, P., Wickramasinghe, K. K., Allender, S., Foster, C., and Rayner, M. 2011. "The Economic Burden of Ill Health Due to Diet, Physical Inactivity, Smoking, Alcohol and Obesity in the Uk: An Update to 2006–07 Nhs Costs," *Journal of Public Health* (33:4), pp. 527-535.
Schiming, R. C. 2012. "Patterns of Homework Initiation for Web-Based Activities in Economics: A Study of Academic Procrastination," *Journal for Economic Educators* (12:1), pp. 13-25.
Schlotz, W., Hellhammer, J., Schulz, P., and Stone, A. A. 2004. "Perceived Work Overload and Chronic Worrying Predict Weekend–Weekday Differences in the Cortisol Awakening Response," *Psychosomatic Medicine* (66:2), pp. 207-214.
Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., and Griskevicius, V. 2007. "The Constructive, Destructive, and Reconstructive Power of Social Norms," *Psychological Science* (18:5), pp. 429-434.
Schumpeter, J. A. 1955. *Imperialism and Social Classes: Two Essays*. Ludwig von Mises Institute.
Sen, M., and Noon, J. 2007. "Living Arrangement: How Does IT Relate to the Health of the Elderly in India?," *Annual Meeting of Population Association of America, New York, NY*.

Shaver, P., Schwartz, J., Kirson, D., and O'connor, C. 1987. "Emotion Knowledge: Further Exploration of a Prototype Approach," *Journal of Personality and Social Psychology* (52:6), p. 1061.
Shi, X., Liu, X., Nuetah, A., and Xin, X. 2010. "Determinants of Household Income Mobility in Rural China," *China & World Economy* (18:2), pp. 41-59.
Shi, Z., Rui, H., and Whinston, A. B. 2014. "Content Sharing in a Social Broadcasting Environment: Evidence from Twitter," *MIS Quarterly* (38:1), pp. 123-142.
Shilts, M. K., Horowitz, M., and Townsend, M. S. 2004. "Goal Setting as a Strategy for Dietary and Physical Activity Behavior Change: A Review of the Literature," *American Journal of Health Promotion* (19:2), pp. 81-93.
Smith, M. A., and Kollock, P. 1999. *Communities in Cyberspace*. Psychology Press.
Soman, D., and Cheema, A. 2004. "When Goals Are Counterproductive: The Effects of Violation of a Behavioral Goal on Subsequent Performance," *Journal of Consumer Research* (31:1), pp. 52-62.
Sonalde, D., Amaresh, D., Joshi, B. L., Mitali, S., Abusaleh, S., and Reeve, V. 2010. "Human Development in India: Challenges for a Society in Transition." New Delhi: Oxford University Press.
Srivastava, M., and Roche, E. 2015. "Top Executives of Tech Giants Pledge Support for Digital India Initiative." http://www.livemint.com/Industry/tzzEYXpYHC33kBZD9m7wFI/Top-executives-of-tech-giants-pledge-support-for-Digital-Ind.html," in: Livemint (28 Sep 2015).
Srivastava, S. C., and Shainesh, G. 2015. "Bridging the Service Divide through Digitally Enabled Service Innovations: Evidence from Indian Healthcare Service Providers," *MIS Quarterly* (39:1), pp. 245-267.
Susarla, A., Oh, J.-H., and Tan, Y. 2012. "Social Networks and the Diffusion of User-Generated Content: Evidence from Youtube," *Information Systems Research* (23:1), pp. 23-41.
Tambe, P., and Hitt, L. M. 2013. "Job Hopping, Information Technology Spillovers, and Productivity Growth," *Management Science* (60:2), pp. 338-355.
Theil, H. 1967. "Economics and Information Theory," Amsterdam: North-Holland.
Tong, S. T., Van Der Heide, B., Langwell, L., and Walther, J. B. 2008. "Too Much of a Good Thing? The Relationship between Number of Friends and Interpersonal Impressions on Facebook," *Journal of Computer-Mediated Communication* (13:3), pp. 531-549.
Van de Ven, W. P., and Van Praag, B. M. 1981. "The Demand for Deductibles in Private Health Insurance: A Probit Model with Sample Selection," *Journal of Econometrics* (17:2), pp. 229-252.
van Hooff, M. L., Geurts, S. A., Kompier, M. A., and Taris, T. W. 2007. "Workdays, in-between Workdays and the Weekend: A Diary Study on Effort and Recovery," *International Archives of Occupational and Environmental Health* (80:7), pp. 599-613.
Vanneman, R., Balakrishna, K., Barnes, D., Naumoff, K., Smith, K., and Uma, R. 2008. "Measuring Indoor Air Pollution and Lung Functioning in Indian Field

Settings," *Annual Meetings of the Population Association of America*, New York, NY.
Venkatesh, A., and Vitalari, N. 1987. "A Post-Adoption Analysis of Computing in the Home," *Journal of Economic Psychology* (8:2), pp. 161-180.
Venkatesh, A., and Vitalari, N. P. 1992. "An Emerging Distributed Work Arrangement: An Investigation of Computer-Based Supplemental Work at Home," *Management Science* (38:12), pp. 1687-1706.
Venkatesh, V., Rai, A., Sykes, T. A., and Aljafari, R. 2016. "Combating Infant Mortality in Rural India: Evidence from a Field Study of Ehealth Kiosk Imlementations," *MIS Quarterly* (40:2), pp. 353-380.
Venkatesh, V., and Sykes, T. A. 2013. "Digital Divide Initiative Success in Developing Countries: A Longitudinal Field Study in a Village in India," *Information Systems Research* (24:2), pp. 239-260.
Volpp, K. G., Troxel, A. B., Pauly, M. V., Glick, H. A., Puig, A., Asch, D. A., Galvin, R., Zhu, J., Wan, F., and DeGuzman, J. 2009. "Randomized, Controlled Trial of Financial Incentives for Smoking Cessation," *New England Journal of Medicine* (360:7), pp. 699-709.
Walker, T. S., and Ryan, J. G. 1990. *Village and Household Economics in India's Semi-Arid Tropics*. Johns Hopkins University Press.
Walther, J. B., Van Der Heide, B., Kim, S. Y., Westerman, D., and Tong, S. T. 2008. "The Role of Friends' Appearance and Behavior on Evaluations of Individuals on Facebook: Are We Known by the Company We Keep?," *Human Communication Research* (34:1), pp. 28-49.
Wasko, M. M., and Faraj, S. 2000. "'"It Is What One Does": Why People Participate and Help Others in Electronic Communities of Practice," *Journal of Strategic Information Systems* (9:2), pp. 155-173.
Wasko, M. M., and Faraj, S. 2005. "Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice," *MIS Quarterly* (29:1), pp. 35-57.
Wasserman, S., and Faust, K. 1994. *Social Network Analysis: Methods and Applications*. Cambridge University Press.
Wechsler, H., Nelson, T. E., Lee, J. E., Seibring, M., Lewis, C., and Keeling, R. P. 2003. "Perception and Reality: A National Evaluation of Social Norms Marketing Interventions to Reduce College Students' Heavy Alcohol Use," *Journal of Studies on Alcohol* (64:4), pp. 484-494.
Wellman, B. 1997. "Structural Analysis: From Method and Metaphor to Theory and Substance," *Contemporary Studies in Sociology* (15), pp. 19-61.
Wiseman, R. 2007. "New Year's Resolution Experiment." from http://www.richardwiseman.com/quirkology/new/USA/Experiment_resolution.shtml
Woolard, I., and Klasen, S. 2005. "Determinants of Income Mobility and Household Poverty Dynamics in South Africa," *Journal of Development Studies* (41:5), pp. 865-897.
Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT press.

Yan, Y. 2003. *Private Life under Socialism: Love, Intimacy, and Family Change in a Chinese Village, 1949-1999*. Stanford University Press.

Yoganarasimhan, H. 2012. "Impact of Social Network Structure on Content Propagation: A Study Using Youtube Data," *Quantitative Marketing and Economics* (10:1), pp. 111-150.