ESSAYS ON REFERRAL PROGRAMS AND PREFERENCE ESTIMATION

Yupeng Chen

A DISSERTATION

in

Marketing

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2018

Supervisor of Dissertation

____

Raghuram Iyengar, Professor of Marketing

Graduate Group Chairperson

____

Catherine Schrand, Celia Z. Moh Professor, Professor of Accounting

Dissertation Committee

Eric T. Bradlow, The K.P. Chao Professor, Professor of Marketing, Professor of Economics, Professor of Education, Professor of Statistics

Christophe Van den Bulte, Gayfryd Steinberg Professor, Professor of Marketing

ESSAYS ON REFERRAL PROGRAMS AND PREFERENCE ESTIMATION

ACKNOWLEDGEMENT

My deepest gratitude goes to my advisor, Raghuram Iyengar, for his constant guidance and help. Raghu has been an extremely supportive mentor, who encourages me to explore research questions that I am passionate about and is always willing to share his insightful perspectives when I seek advice. I have learned from Raghu not only a great deal about how to conduct research, but also that as an academic I need to be always curious and passionate about research.

I am deeply indebted to my committee members, Eric Bradlow and Christophe Van den Bulte, for their invaluable insights and encouragement. I am also extremely grateful to my mentor, Jehoshua Eliashberg, for his continuous support and mentorship.

I am grateful to have met many good friends during the past years and their warm friendships is one huge reason why I have enjoyed so much of my time at Wharton. My most sincere appreciation goes to them.

Finally, I owe my deepest gratitude to my parents and my wife, Zhe, without whose love and support this dissertation would not have been possible.

ABSTRACT

ESSAYS ON REFERRAL PROGRAMS AND PREFERENCE ESTIMATION

Yupeng Chen

Raghuram Iyengar

In this dissertation, we study referral programs and preference estimation in two essays. In the first essay, we propose that a firm can enhance the effectiveness of its referral program by promoting better matching between referred customers and the firm. We develop three treatments aimed at promoting better matching, including (1) offering current customers a gift before inviting them to refer friends, (2) notifying current customers about the value that they have received from the firm before inviting them to refer friends, and (3) rewarding referring customers based on the value of their referred customers. We test these three treatments by conducting two field experiments in collaboration with a Chinese online financial services firm. We find that all three treatments substantially enhanced the effectiveness of the focal referral program, measured for each current customer as the total value of his referred customers. We also find that the enhancement was primarily driven by the acquisition of higher-value new customers rather than the acquisition of more new customers. In addition, we investigate customer heterogeneity in treatment effects and explore the mechanisms through which these treatments impacted customer referrals. In the second essay, we develop a new model for effective modeling of consumer heterogeneity in choice-based conjoint estimation. Assuming that most variations in consumers' partworth vectors are along a small number of orthogonal directions, we propose that shrinking the individual-level partworth vectors toward a low-dimensional affine subspace that is also inferred from data can be an effective approach to pooling information across consumers and modeling consumer heterogeneity. We develop a low-dimension learning model to implement this information pooling mechanism that builds on recent advances in rank minimization and machine learning. We evaluate the empirical performance of the low-dimension learning model using both simulation experiments and field choice-based conjoint data sets. We find that the low-dimension learning model overall outperforms multiple benchmark models in terms of both parameter recovery and predictive accuracy. While addressing two different marketing topics, both essays share a common theme - careful modeling of consumer heterogeneity plays a key role in understanding consumer behavior and developing effective marketing strategies.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1 : Introduction

Customer referral programs have been widely adopted by firms for new customer acquisition and have fueled the phenomenal growth of many companies, including Dropbox, Uber, and Airbnb. In the case of Dropbox, for instance, its referral program was largely responsible for the growth of its user base from 100,000 registered users in September 2008 to more than 4 million in January 2010. Referral programs provide firms an attractive method to acquire new customers, as they are cost-effective to run and provide access to prospective customers whom traditional marketing programs may not effectively reach (Berman, 2016), and are able to acquire higher-value new customers compared to other channels (Schmitt et al., 2011). Given the prominent role of referral programs in firms' customer acquisition practice, both researchers and managers are keen on answering the following question: How can firms enhance the effectiveness of their referral programs?

Conjoint analysis, and choice-based conjoint (CBC) in particular, has been the most popular method used by both researchers and practitioners to assess how consumers with heterogeneous preferences value different product or service attributes (Wittink and Cattin, 1989; Green and Srinivasan, 1990; Huber, 2004). The understanding of consumers' heterogeneous preferences plays a central role in a variety of marketing decisions, such as pricing, targeted promotions, differentiated product offerings, and market segmentation (Allenby and Rossi, 1998). One significant challenge facing conjoint researchers is that in most applications the amount of information elicited from each consumer is limited. To address the scarcity of information from each consumer and obtain accurate estimation of consumers' heterogeneous preferences, we need to answer the following question: How to pool information across consumers and model consumer heterogeneity?

In this dissertation, we aim to answer these two questions with substantial marketing implications in two essays. In the first essay (Chapter 2), we propose that a firm can increase the total value of new customers acquired through its referral program by promoting better matching between these new customers and the firm. Better matching can be promoted by two different processes: Motivating active matching, in which the firm motivates current customers to exert greater effort to screen their friends and refer good matches, or facilitating passive matching, in which the firm induces a higher proportion of high-value current customers to refer their friends who, by homophily, are more likely to be good matches than friends of low-value current customers. We propose three treatments aimed at promoting better matching: (1) offering current customers a gift before inviting them to refer friends, (2) notifying current customers about the value that they have received from the firm before inviting them to refer friends, and (3) rewarding referring customers based on the value of their referred customers.

We empirically test the three proposed treatments by conducting two field experiments in collaboration with a Chinese online financial services firm. Using data from the experiments, we find that all three treatments substantially increased the total value of referred customers, which validates their capacity in enhancing the effectiveness of the referral program. We also find that the effects of the three treatments were primarily driven by the acquisition of higher-value referred customers rather than the acquisition of more referred customers, indicating that these treatments promoted better matching as we expected. Moreover, we conduct a series of analyses to investigate the workings of these treatments, and find evidence suggesting that the value-based reward facilitated passive matching and had a larger impact on current customers of higher value.

Effective information pooling across consumers is critical for accurate conjoint estimation. In the marketing literature, researchers have primarily investigated three information pooling mechanisms: (1) shrinking the individual-level partworths toward the population mean (Lenk et al., 1996; Rossi et al., 1996; Evgeniou et al., 2007), (2) recovering segments in the population and shrinking the individual-level partworths toward their respective segment means (Allenby et al., 1998; Chen et al., 2017), and (3) approximating the individual-level partworths using discrete points (Kamakura and Russell, 1989; Ansari and Mela, 2003; Kim et al., 2004). In the second essay (Chapter 3), we propose an information pooling mechanism distinct from the three well-established mechanisms and develop a low-dimension learning model to operationalize such a mechanism. In our approach, we assume that most variations in consumers' heterogeneous partworths, viewed as vectors in an Euclidean space, are along a small number of orthogonal directions; consequently, consumers' partworth vectors all reside near some low-dimensional affine subspace of the Euclidean space (James et al., 2017). In this case, a natural mechanism for pooling information across consumers is to shrink the individual-level partworths toward this low-dimensional affine subspace, which is also inferred from the data. Our model implements such a low-dimension information pooling mechanism using a convex optimization framework in which both the distance between each partworth vector and the affine subspace as well as the dimension of the affine subspace are penalized. We further enhance our model by incorporating the convex regularization of Evgeniou et al. (2007) that shrinks the individual-level partworths toward the population mean.

We compare our low-dimension learning model and a restricted version of the model in which only the low-dimension information pooling mechanism is implemented to multiple benchmark models using simulation experiments and two field data sets. We find that the low-dimension learning model and its restricted version overall outperform the benchmark models both in terms of parameter recovery and predictive accuracy, and they demonstrate strong performance irrespective of whether the underlying assumption - the true individual-level partworths have a good low-dimensional affine subspace approximation - seems to hold or not. Therefore, the effectiveness of our low-dimension learning model (and the low-dimension information pooling mechanism) in recovering consumers' heterogeneous preferences is empirically validated. We also find that the performance of the low-dimension learning model is very close to that of its restricted version, suggesting that the incremental value of shrinking the individual-level partworths toward the population mean can be very limited when we are already shrinking the individual-level partworths toward a low-dimensional affine subspace.

CHAPTER 2 : Enhancing Effectiveness of Referral Programs by Promoting Better Matching: Evidence from Field Experiments

2.1. Introduction

Customer referral programs, in which a firm's current customers are rewarded for bringing in new customers, have been widely used by firms for customer acquisition and have fueled the phenomenal growth of many companies, including Dropbox, Uber, and Airbnb. In the case of Dropbox, for instance, a current user receives 500MB of free storage space for each friend invited to sign up for and install Dropbox. According to Drew Houston, Co-Founder and CEO of Dropbox, this referral program was largely responsible for the growth of Dropbox's user base from 100,000 registered users in September 2008 to more than 4 million in January 2010. Referral programs provide firms an attractive method to acquire new customers, as they are cost-effective to run and provide access to prospective customers whom traditional marketing programs may not effectively reach (Berman, 2016), and customers acquired through referral programs (i.e., referred customers) can be more valuable than customers acquired through other methods (Schmitt et al., 2011).

Given the prominent role of referral programs in firms' customer acquisition practice, both researchers and managers are keen on understanding factors driving customer referrals and ways to leverage such factors to enhance the effectiveness of referral programs. In general, customers make referrals because (1) they want to earn an economic reward, (2) they want to help their friends make better choices and/or manage their friends' impressions of them, and (3) they like the firm and want to help its business.[1] In the marketing and information systems literatures, researchers have empirically investigated the roles of the first two factors in customer referrals. For instance, Ryu and Feick (2007) find that the impact of the referral reward on a lab participant's stated referral likelihood is moderated by the relationship between the participant and the targeted prospective customer as well

[1] https://rewardstream.com/blog/social-referrals-psychology-why-do-people-share/

as brand strength. Bapna et al. (2016) vary the allocation of the referral reward between the referring customer and the referred customer in a field experiment, and find that both equally splitting the reward and allocating all the reward to the referred customer lead to more successful referrals than allocating all the reward to the referring customer. Jung et al. (2017) compare different framings of a call-to-action message encouraging customer referrals in a field experiment, and find that the altruistic framing leads to a higher referral likelihood, more referrals, and more purchases from referred customers compared to the egoistic and equitable framings. This stream of research has two limitations. First, it has not examined the role of customers' intention to help the firm in making referrals, and the referral rewards being studied are all contingent only on the acquisition of the referred customer - the possibility of rewarding the referring customer based on the value of his referred customers has not been explored. Second, when assessing the impact of the design and communication of referral programs on their effectiveness, this stream of research has focused on measures corresponding to the number of referred customers and has not taken into account the value of each referred customer. While the number of referred customers is an important aspect of the effectiveness of referral programs, firms also care about the value of referred customers (Schmitt et al., 2011) and information about the value of referred customers is critical to our theoretical understanding of customer referrals (Van den Bulte et al., 2018).

In this essay, we operationalize the effectiveness of referral programs using the total value of all referred customers, which takes into account both the number and value of referred customers, and explore how firms can increase the total value of referred customers.[2] Specifically, we aim to answer the following question: Given a referral program that rewards a referring customer contingent on the acquisition of the referred customer (as is the case for most existing referral programs), can we develop treatments that can be applied to the re-

[2]Conceptually, we assess the value of a referred customer using her customer lifetime value (CLV). Empirically, however, since we do not have access to the margin information in our application setting (a Chinese online financial services firm), we will use a referred customer's investment amount as a proxy for her value.

ferral program to increase the total value of referred customers? Clearly, an important lever for increasing the total value of referred customers is making the referral program acquire higher-value new customers.[3] To shed light on how this lever can be used, we draw on the recent work in marketing on referral programs and customer value, which has proposed and identified better matching as a critical mechanism through which referral programs are able to acquire high-value new customers (Schmitt et al., 2011; Van den Bulte et al., 2018).[4] Better matching refers to the phenomenon that referred customers match with the firm better than customers acquired through other methods do, and there are two distinct matching processes that can be at work: active matching and passive matching. Active matching involves current customers' deliberate screening of their friends and diligent matching of those friends who they think may be a good fit to the firm. Specifically, active matching relies on that current customers, by knowing both the needs of their friends and the offerings of the firm, are likely to be better informed to assess the match between their friends and the firm than the two parties themselves. When current customers are properly motivated, they will screen their friends based on their match with the firm and refer good matches to the firm; moreover, when referring those friends, they may also exert effort to facilitate the matching by educating their friends about the offerings of the firm. Eventually, the firm is able to acquire high-value new customers from the social networks of its current customers. Passive matching, on the other hand, operates through homophily: As current customers have an above-average chance of being a good match and they tend to refer people who are similar to themselves, referred customers are more likely to match well with the firm compared to nonreferred customers. For our research question, in order to make the referral program acquire higher-value new customers, the firm may consider designing and implementing treatments aimed at promoting better matching between referred customers and the firm. Promoting better matching can be achieved by motivating active matching,

[3]Another lever is making the referral program acquire more new customers, which has been investigated in the literature (Ryu and Feick, 2007; Bapna et al., 2016; Jung et al., 2017).

[4]In the economics and sociology literatures, better matching has also been established as a key mechanism through which firms derive value from employee referral programs, a recruiting device that rewards current employees for referring job candidates for hire (Montgomery, 1991; Fernandez et al., 2000; Castilla, 2005; Beaman and Magruder, 2012; Brown et al., 2016; Pallais and Sands, 2016).

in which the firm motivates current customers to exert greater effort to screen their friends and refer good matches, or by facilitating passive matching, in which the firm induces a higher proportion of high-value current customers to refer their friends who, by homophily, are more likely to be good matches than friends of low-value current customers.

We propose three treatments aimed at promoting better matching. The first two treatments, which we term as the gift treatment and the notification treatment, are designed to induce reciprocity from current customers toward the firm which could motivate them to help the firm's business by referring friends.[5] In the gift treatment, the firm offers current customers a gift before inviting them to refer friends. The gift treatment is motivated by the field evidence that people reciprocate a gift from others by providing goods desired by the latter (Gneezy and List, 2006; Falk, 2007; Alpizar et al., 2008; Kube et al., 2012; Gilchrist et al., 2016; Chung and Narayandas, 2017), as well as the established merchandising practice in which salespeople offer customers a free sample or service to trigger reciprocity which leads to increased sales (Cialdini, 1993). We expect that the gift treatment will make current customers feel that they are treated well by the firm, and they will, out of reciprocity, return the favor to the firm by exerting more effort to carefully screen their friends and match those who they think may be valuable to the firm (i.e., by engaging in active matching).[6] Moreover, if the gift is designed in such a way that it is more valuable to high-value current customers than to low-value current customers, it will be more likely to elicit reciprocity from the former than from the latter and hence may induce a higher proportion of high-value current customers to refer their friends, which facilitates passive matching.

[5] Reciprocity refers to "the behavioral phenomenon of people responding toward (un)kind treatment likewise, even in the absence of reputational concerns" (Kube et al., 2012), and it has long been argued to be deeply embedded in and have substantial implications for social and economic interactions (Gouldner, 1960; Cialdini, 1993; Fehr and Gachter, 2000). Similar definitions of reciprocity have also been proposed in the literature. For instance, Gouldner (1960) suggests that "a norm of reciprocity, in its universal form, makes two interrelated, minimal demands: (1) people should help those who have helped them, and (2) people should not injure those who have helped them". Cialdini (1992) describes reciprocity as the norm that "we are *obligated* to the future repayment of favors, gifts, invitations, and the like". Referring to reciprocity, Fehr and Gachter (2000) state that "People repay gifts and take revenge even in interactions with complete strangers and even if it is costly for them and *yields neither present nor future material rewards*".

[6] Schmitt et al. (2011) have also suggested that reciprocity from current customers toward the firm may lead to active matching.

In the notification treatment, the firm notifies current customers about the value that they have received from the firm before inviting them to refer friends. Possible applications of this notification treatment include notifications about customers' investment return from a financial services firm, and notifications about customers' money saved from a daily deals website. We expect that such a notification will make it salient to current customers that they have received great value from the firm, which may trigger their intention to reciprocate to the firm by engaging in active matching. In addition, since the value that customers receive from the firm is likely to be positively correlated with the value that customers generate for the firm, high-value customers are more likely to have received great value from the firm and therefore more likely to refer friends after receiving the notification. In such cases, passive matching is facilitated.[7]

The third treatment, which we term as the value-based reward treatment, explores the potential of rewarding the referring customer on the basis of the value of his referred customers. In most existing referral programs, since the referring customer is rewarded contingent only on the acquisition of the referred customer, current customers have little economic incentive to engage in active matching and bring in high-value new customers. Such an incentive structure has raised concerns among managers that referral programs could end up rewarding current customers for referring low-value new customers (Schmitt et al., 2011). In order to address this issue, we propose that the firm rewards the referring customer based on the value of his referred customers on top of the regular referral reward contingent on the acquisition of the referred customers. We expect that the value-based reward treatment, by providing current customers a direct economic incentive to engage in active matching, will motivate them to refer high-value new customers.[8] At the same time, since high-value current customers are more likely to have friends who could become high-value new cus-

[7]For instance, in the case of Certificate of Deposit, as the amount of the deposit increases, the customer receives more value from the financial institution through more interest return and the financial institution derives more value from the customer through more assets under management.

[8]In the context of employee referrals, Beaman and Magruder (2012) show that providing a contingent bonus based on the referred employee's performance can make a fraction of current employees select highly skilled new employees to refer to the firm.

tomers and generate value-based rewards for them based on homophily, they are likely to be better incentivized than low-value current customers in the presence of a value-based reward, implying that passive matching is facilitated.

While we expect the three proposed treatments to promote better matching and enhance the effectiveness of referral programs, their prospects are actually far from certain. For the gift and notification treatments, it is ex ante unclear how many current customers would reciprocate to the firm, especially given the possibility that the gift and the notification are perceived as deliberate marketing devices, in which case reciprocity can hardly be triggered (Gouldner, 1960; Cialdini, 1993). As to the value-based reward treatment, since it explicitly links the size of the reward to the value of referred customers, it could impose considerable psychological cost on referring customers as they may perceive that they are exploiting their social connections to seek economic rewards, in which case the value-based reward may backfire by making current customers less likely to refer friends (Ryu and Feick, 2007; Gneezy et al., 2011). Moreover, the value-based reward may make referring customers less persuasive when advocating for the firm (Barasch et al., 2016). Consequently, whether and to what extent each of these three treatments can promote better matching and ultimately increase the total value of referred customers need to be empirically investigated.[9]

We collaborated with a leading Chinese online financial services firm to test these three treatments. The firm offers customers financial deposit services including a flexible deposit and an assortment of fixed deposits in which they can invest money and earn interest. The flexible deposit has a floating interest rate and allows customers to make investments for any duration and withdraw part or all of the investments at any time, whereas each fixed deposit has a fixed interest rate and a fixed investment duration. As part of its customer acquisition strategy, the firm runs a referral program that rewards current customers with an investment coupon that can be used to reimburse a small proportion of their next investment

[9] In this essay, we focus on understanding the impact of each of the three treatments (i.e., gift, notification, and value-based reward) on the effectiveness of referral programs rather than comparing among the three treatments. This is because these treatments take distinct forms and may operate through different mechanisms, and hence are not directly comparable ex ante.

for bringing in a friend to open an account at the firm. We conducted two randomized field experiments at the firm, testing the gift and notification treatments in the first experiment and the value-based reward treatment in the second experiment.[10]

In the first experiment, we included 93,288 current customers in a two-week campaign, and randomly assigned them to a control condition and two treatment conditions that implemented the gift treatment and the notification treatment, respectively. The regular referral program was implemented in all three conditions, and the only difference across conditions was the way by which customers were approached for the campaign. Specifically, at the beginning of the campaign, customers in the control condition received a text message inviting them to refer friends; customers in the gift condition received a gift - an investment coupon that could be used to raise the interest rate of their next investment and was valid during the campaign - and a text message inviting them to refer friends; and customers in the notification condition received a text message notifying them about their total investment return and inviting them to refer friends.

In the second experiment, we included 120,258 current customers in a 30-day campaign, and randomly assigned them to a control condition and two treatment conditions that operationalized the value-based reward treatment differently. During the campaign, the regular referral program was implemented in the control condition, while each of the two treatment conditions augmented the regular referral program with a value-based reward. In particular, a customer in the first treatment condition would receive a cash reward based on the number of his referred customers whose total investments in fixed deposits made during the campaign were above a pre-determined threshold, and a customer in the second treatment condition would receive a cash reward based on the total investments in fixed deposits made by his referred customers during the campaign. By testing two different value-based rewards, we intended to obtain a more thorough understanding of the impact of the value-based reward treatment on customer referrals.

[10] For logistical reasons, we were unable to test all three treatments in a single field experiment.

Analyzing data from the two experiments, we find that all three proposed treatments had an economically substantial and statistically significant impact on the effectiveness of the focal referral program. In particular, the gift and notification treatments increased the total value of referred customers by more than 200% in the first experiment, while both operationalizations of the value-based reward treatment increased the total value of referred customers by more than 100% in the second experiment.[11] We also find that the effects of all three treatments on the total value of referred customers were primarily driven by the positive value differential between referred customers acquired in the treatment conditions and those acquired in the control conditions, i.e., the three treatments helped the referral program acquire higher-value new customers as expected, but their impact on the number of referred customers turns out to be modest in magnitude and largely insignificant. One interesting implication of the latter finding is that, if we were to assess the effectiveness of the focal referral program using the number of successful referrals as has been done in the literature, we would erroneously conclude that these treatments are ineffective.

To derive deeper insights on the workings of the three proposed treatments, we use data from the experiments to explore the mechanisms underlying these treatments. For the gift and notification treatments, while they are designed to elicit reciprocity from current customers toward the firm which could in turn promote better matching, they may also motivate customer referrals by serving as a value signal to current customers: The gift and notification treatments may signal to current customers that the firm is able to create great value for its customers, which could enhance their perception about the value that their friends would receive once referred to the firm; since customers tend to use referrals to help friends make better choices or manage friends' impressions of them (Ryu and Feick, 2007; Kornish and Li, 2010), such an enhanced value perception is likely to motivate customer referrals. Moreover, with either reciprocity or value signaling being at work, both the gift and notification treatments could make the referral program acquire higher-value new

[11] We note that the findings are not directly comparable between experiments since the two experiments were conducted at different times and on different samples of customers.

customers as observed in the data by motivating active matching or facilitating passive matching. We conduct a series of analyses to investigate the mechanisms underlying the gift and notification treatments; the findings, however, are not encouraging. Specifically, the lack of power in most analyses prevents us from obtaining sufficient statistical evidence to draw conclusions regarding the workings of the gift and notification treatments. Similar to the gift and notification treatments, the value-based reward treatment could make the referral program acquire higher-value new customers as observed in the data by motivating active matching or facilitating passive matching. We find evidence suggesting that the second operationalization facilitated passive matching, but we do not have sufficient statistical evidence to reach a conclusion regarding the first operationalization.

Drawing on these findings, this essay contributes to the literature on customer referral programs. As we mentioned earlier, extant work empirically investigating the impact of the design and communication of referral programs has not taken into account the value of referred customers when assessing the effectiveness of referral programs (Ryu and Feick, 2007; Bapna et al., 2016; Jung et al., 2017), which is a significant limitation from both the theoretical and managerial perspectives.[12] On the other hand, while research on referral programs and customer value has proposed and identified better matching as a key mechanism through which referral programs are able to acquire high-value new customers (Schmitt et al., 2011; Van den Bulte et al., 2018), whether firms can proactively leverage this mechanism to further enhance the effectiveness of referral programs has not been investigated. We contribute to the literature by proposing three distinct treatments aimed at promoting better matching and conducting field experiments to test whether these treatments can increase the total value of referred customers, which we use to measure the effectiveness of referral programs. We find that all three treatments substantially increased the total value of referred customers, and the effects of the treatments were primarily driven by the acqui-

[12] Kumar et al. (2010) measure the effectiveness of referral programs using customer referral value (CRV) which takes into account the value of referred customers. However, their focus is on studying optimal customer targeting for a given referral marketing campaign instead of investigating the impact of the design and communication of referral programs on their effectiveness.

sition of higher-value new customers rather than the acquisition of more new customers. Therefore, our work answers the call by Godes et al. (2005) and Godes (2011) to investigate how firms can proactively leverage peer influence to enhance business performance by documenting evidence suggesting that firms can increase the total value of referred customers by promoting better matching.

The remainder of the essay is organized as follows. In Section 2.2, we describe the research setting for our study, including the design and implementation of the two experiments. We analyze data from the first experiment in Section 2.3 and data from the second experiment in Section 2.4, respectively, to investigate the effects of the three treatments and explore the mechanisms through which they operate. We discuss limitations of our work and suggest a few directions for future research in Section 2.5.

2.2. Empirical Setting

For our study, we conducted two randomized field experiments in collaboration with a leading Chinese online financial services firm. The firm offers customers financial deposit services accessible via both a website and a mobile application, in which they can invest money and earn interest. There are two types of financial deposits: a flexible deposit and an assortment of fixed deposits. In the flexible deposit, customers can make investments for any duration and withdraw part or all of the investments at any time. The flexible deposit has a floating interest rate (i.e., the interest rate varies daily) that is always lower than the interest rates of the fixed deposits. Each fixed deposit has a fixed interest rate and a fixed investment duration. In fixed deposits, customers receive both the principal and the interest of an investment at maturity, and, if they decide to withdraw the investment prior to its maturity, a transaction fee is incurred. Fixed deposits vary in terms of duration (from 3 months to 24 months) and interest rate (from 7% to 10.5% annually), with a longer-duration deposit having a higher interest rate. Both durations and interest rates are comparable to those offered by competing firms.

The firm runs a referral program as part of its customer acquisition strategy. Similar to referral programs of many competing firms, this referral program rewards current customers an investment coupon for referring a friend to open an account at the firm. Specifically, a current customer can share a referral link to a friend via either WeChat (a popular Chinese mobile messaging application), social media, or email, and, if the friend opens an account using the referral link, the referring customer receives a one-time coupon. When applied to an investment in a fixed deposit, this reward coupon reimburses a customer for 0.5% of the investment for up to 20 RMB.[13] For example, if a customer applies the reward coupon to a fixed-deposit investment of 3,000 RMB, he will receive a reimbursement of 15 RMB at the time of investing and the coupon becomes nullified. A customer can neither apply more than one reward coupon nor apply a reward coupon together with any other coupons to the same investment. On the other hand, any new customer, irrespective of the acquisition method, receives a welcome coupon that reimburses 0.5% of her fixed-deposit investments for up to 50 RMB and can be applied to multiple investments until the total reimbursements reach 50 RMB. Both the reward coupon for referring customers and the welcome coupon for new customers expire in 30 days.

We chose to collaborate with this firm for two main reasons. First, since financial investment decisions are relatively private and can be risky, customers tend to be cautious when it comes to referring friends to a firm providing financial investment services. Hence, conducting experiments at the focal Chinese online financial services firm could provide a stringent test for the proposed treatments. Second, the Chinese online financial services industry is highly competitive and new customer acquisition is challenging. Given that online advertising, the primary customer acquisition channel, has become increasingly expensive, the more cost-effective referral programs have started to gain prominence in firms' customer acquisition strategy. Testing the proposed treatments at the focal firm could serve as a first step in generating much-needed insights that firms in this industry can use to enhance the effectiveness of referral programs.

[13] RMB is the Chinese currency. 1 U.S. dollar approximately equals 7 RMB.

2.2.1. The First Experiment

The first experiment was conducted in December, 2016. Right before the start of the experiment, we selected a random sample of 93,288 customers from the firm's current customers who satisfied two criteria: (1) had at least 500 RMB invested in all deposits combined at the time of sampling and (2) had earned between 100 RMB and 10,000 RMB in interest since account opening. These two criteria were imposed to ensure that customers included in the experiment were in an active relationship with the firm and had received a nontrivial return on their investments. We randomly assigned these customers to a control condition and two treatment conditions, with the control condition including 30,977 customers, the first treatment condition including 31,241 customers, and the second treatment condition including 31,070 customers.

The experiment involved a two-week campaign inviting customers to participate in the referral program. The regular referral program was implemented in all three conditions, and the conditions only differed in how customers were approached for the campaign. In the control condition, customers received a text message encouraging them to refer friends at the beginning of the campaign. Specifically, the text message stated "Dear customer, please invite your friends to invest with us, who provide reliable and high-return financial deposit services!", and included a link to the referral program page in the mobile app.[14] Customers received the same text message again one week later as a reminder.

In the first treatment condition, customers received a one-time 1% interest-raising coupon for their future investment in fixed deposits as a gift and a text message explaining the coupon and encouraging them to refer friends at the beginning of the campaign. The gift coupon was valid for two weeks, i.e., it would expire at the end of the campaign. The text message stated "Dear customer, you have received a 1% interest-raising coupon as a gift for being a valued customer. Please feel free to use it. Also please invite your friends to invest with us, who provide reliable and high-return financial deposit services!", and included a

[14]The text quoted here and those quoted in the following are translated from Chinese.

link to the referral program page in the mobile app. One week later, customers received a text message reminding them about the gift coupon and the invitation to refer friends. In the following, we refer to this treatment condition as the gift condition.

We make several observations regarding the 1% interest-raising coupon. First, as the name suggests, this coupon could be used to raise the annual interest rate of a future fixed-deposit investment by 1 percentage point. Since the economic value of this coupon was proportional to the amount of the investment, it was likely to be more valuable to high-value customers who, compared to low-value customers, were likely to make larger investments during the campaign. Second, this coupon was familiar to customers, as interest-raising coupons with comparable raise in interest rate are sent to all customers multiple times a year. This coupon differed from other interest-raising coupons in that it was framed as a gift as opposed to a promotion. Third, since this coupon and the referral reward coupon could not be applied together to the same investment, there was a substitution between the two.[15] In particular, if a customer found the 1% interest-raising coupon more valuable, it would effectively void the value of the referral reward coupon and consequently remove the economic incentive of referring friends.

In the second treatment condition, customers received a text message notifying them about their total investment return and encouraging them to refer friends at the beginning of the campaign. Specifically, customers in this condition were divided into four subgroups based on their total investment return: Those whose total investment return was at least 100 RMB and less than 500 RMB were included in the first subgroup; those whose total investment return was at least 500 RMB and less than 1,000 RMB were included in the second subgroup; those whose total investment return was at least 1,000 RMB and less

[15]Which coupon was more valuable to a customer depended on the customer's time preferences and the amount and duration of his investment. The economic value of the 1% interest-raising coupon was proportional to the investment duration and would be received by the customer at maturity, whereas the economic value of the referral reward coupon was not contingent on the investment duration and would be received by the customer at the time of investing. In addition, there was no upper limit on the economic value that the 1% interest-raising coupon could generate as the investment amount increased, in contrast to the upper limit of 20 RMB for the referral reward coupon.

than 5,000 RMB were included in the third subgroup; and those whose total investment return was at least 5,000 RMB and no greater than 10,000 RMB were included in the fourth subgroup. At the beginning of the campaign, customers in the first subgroup received a text message that stated "Dear customer, do you notice that we have helped you earn at least 100 RMB on your investments? Please invite your friends to invest with us, who provide reliable and high-return financial deposit services!", and included a link to the referral program page in the mobile app. Customers in the three other subgroups received the same text message except that 100 RMB was replaced by 500 RMB, 1,000 RMB, and 5,000 RMB, respectively.[16] Customers in each subgroup received the same text message again one week later as a reminder. In the following, we refer to this treatment condition as the notification condition.

2.2.2. The Second Experiment

The second experiment was conducted in April and May, 2017. Right before the start of the experiment, we selected a random sample of 120,258 customers from the firm's current customers who had at least 500 RMB invested in all deposits combined at the time of sampling. We randomly assigned these customers to a control condition and two treatment conditions, with the control condition including 40,076 customers, the first treatment condition including 40,145 customers, and the second treatment condition including 40,037 customers.

The experiment involved a 30-day campaign inviting customers to participate in the referral program. In the control condition, the regular referral program was implemented and customers received a text message encouraging them to refer friends each week. The text message stated "Invite friends to invest with us and earn investment coupons!", and included a link to the referral program page in the mobile app.

[16]We chose to divide customers into subgroups and send a unified notification to customers in each subgroup instead of sending each customer a personalized notification about his total investment return given managers' concern that the latter approach might be perceived as intrusive by customers. If a customer would like to access the exact value of his total investment return, he could log into his account anytime via either the website or the mobile app in which such information is prominently displayed.

In the first treatment condition, a value-based reward was added to the referral program. For each customer, the value-based reward was specified as follows: (1) new customers who used his referral link to open an account during the campaign (i.e., referred customers) were identified, (2) the total investments in fixed deposits made by each referred customer during the campaign were calculated, and (3) for each referred customer whose total investments were at least 10,000 RMB, the referring customer would receive 50 RMB in cash at the end of the campaign. For example, if a referring customer had three referred customers, the first one investing 30,000 RMB in fixed deposits, the second one investing 10,000 RMB in fixed deposits, and the third one investing 5,000 RMB in fixed deposits, the referring customer would receive a cash reward of 100 RMB. A detailed description of this reward was shown on the referral program page. Customers also received a text message encouraging them to refer friends each week. The text message stated "Invite friends to invest with us and earn investment coupons and cash rewards!", and included a link to the referral program page in the mobile app. In the following, we refer to this treatment condition as the first value-based reward condition.

In the second treatment condition, a different value-based reward was added to the referral program. For each customer, the value-based reward was specified as follows: (1) referred customers were identified, (2) the total investments in fixed deposits made by all referred customers during the campaign were calculated, and (3) for every 10,000 RMB in the total investments, the referring customer would receive 50 RMB in cash at the end of the campaign. For example, if the referred customers invested a total of 45,000 RMB in fixed deposits, their referring customer would receive a cash reward of 200 RMB. Other than the reward, the second treatment condition was identical to the first value-based reward condition, including the text message. In the following, we refer to this treatment condition as the second value-based reward condition.[17]

[17] We note that the second value-based reward was more generous than the first value-based reward in the sense that, for any profile of referred customers, the referring customer would receive the same or more cash reward under the second reward compared to the first reward. One main difference between these two rewards was that, a referred customer's contribution to the reward for the referring customer was capped at 50 RMB under the first reward, whereas her contribution would keep increasing as her fixed-deposit investment

2.2.3. Data

For each current customer in the two field experiments, we observe the following variables that were collected right before the experiments: total investment return since account opening, investment amount in all financial deposits combined, tenure since account opening, recency of the last investment, and whether the customer had successfully referred any friend to open an account. For each referral made by a customer, it is recorded in the data if and only if it is used by the receiving friend to open an account at the firm. Ideally, it would be helpful if we can observe each referral irrespective of whether it is accepted by its receiver or not. However, since a customer can use a variety of channels including WeChat, various social media platforms, and email that the firm cannot track to share referral links to friends, his sharing of a referral link is not recorded in the data; it is only when the receiver of the referral link uses it to open an account that we can observe this referral and map the referred new customer to the referring customer. Once a new customer opens an account, we observe her investment behavior in all financial deposits.

2.3. Empirical Findings: The First Experiment

In this section, we analyze data from the first experiment to investigate the effects of the gift and notification treatments. First, we verify that the random assignment of customers was valid. Then, we estimate the aggregate effects of the gift and notification treatments on the effectiveness of the referral program. Finally, we derive deeper insights on the workings of these two treatments by exploring the mechanisms through which they operate.

2.3.1. Randomization Check

We assess the validity of the randomization by comparing customers across conditions with respect to their behaviors prior to the experiment. With a valid random assignment, we should observe little or no systematic differences in customers' pre-experiment variables

increased under the second reward. In this experiment, we tested two different value-based rewards in order to obtain a more thorough understanding of the impact of the value-based reward treatment. The optimal design of value-based reward is an interesting question that we leave for future research.

across conditions. The details of the randomization check are reported in Table 1.

Insert Table 1 here.

For each pair of experimental conditions and each pre-experiment variable, we test whether the two conditions have the same population mean on this variable using a two-sided Welch's t-test. To apply Welch's t-test, we assume that the sample in each condition consists of i.i.d. draws from an underlying population; on the other hand, Welch's t-test does not require the variances of the two populations to be equal, and given the large sample size in each condition the populations do not have to follow a normal distribution (Lumley et al., 2002). From Table 1, it is evident that customers' pre-experiment variables are well balanced across conditions, confirming the validity of the random assignment in the first experiment.

2.3.2. Aggregate Effects of the Treatments

We measure the effectiveness of the referral program for each current customer by assessing the total value of his referred customers, which is operationalized as follows. First, we identify new customers who opened an account at the firm using his referral link during the two-week campaign period (i.e., his referred customers). Second, we assess the value of each referred customer; since we do not have access to the margin information of the financial deposits, we proxy for the value of a referred customer using the amount of investments she made in all financial deposits combined during the 90-day observational period starting from the beginning of the experiment. Finally, the total value of a current customer's referred customers is calculated as the sum of the value of his referred customers, i.e., we proxy for the total value of a current customer's referred customers using the total amount of investments made by all his referred customers during the 90-day observational period.[18] While the total value of referred customers is our key outcome variable, we also consider three other outcome variables in order to obtain a more thorough understanding of the impact of the gift and notification treatments on customer referrals: (1) whether a current

[18] In the following, we use the terms "value" and "investments" interchangeably when referring to the value of a referred customer and the total value of a current customer's referred customers.

customer has acquired any referred customers during the experiment (Yes/No), (2) the number of a current customer's referred customers, and (3) the value of a referred customer assessed by the amount of investments she made during the 90-day observational period. We note that all outcome variables except the value of a referred customer are defined on each current customer, and for these outcome variables the sample size in each condition is the number of current customers in that condition: 30,977 in the control condition, 31,241 in the gift condition, and 31,070 in the notification condition. On the other hand, the value of a referred customer is defined on each referred customer, and the sample size in each condition for this outcome variable is the number of referred customers acquired in that condition: 83 in the control condition, 101 in the gift condition, and 93 in the notification condition.

Given the randomization, the average treatment effect of either the provision of the gift coupon or the notification on the total value of referred customers can be identified via a direct mean comparison between the control condition and the corresponding treatment condition. We report the result of the mean comparisons based on two-sided Welch's t-tests in the first row of Table 2.

Insert Table 2 here.

Similar to Section 2.3.1, we assume that the sample in each condition consists of i.i.d. draws from an underlying population in order to apply Welch's t-test; on the other hand, the validity of the test does not require the populations to have equal variances or be normally distributed given the large sample sizes. We find that both the gift and notification treatments had a substantial impact on the total investments of referred customers. In particular, the gift treatment and the notification treatment on average increased the total investments of a current customer's referred customers by 26.58 RMB and 25.23 RMB, respectively. Both treatment effects amount to more than 200% lift over the baseline value of 12.47 RMB, and are statistically significant with $p < 0.01$ for the gift treatment and $p < 0.05$ for the notification treatment.

So what drove the positive effects of these two treatments on the total investments of referred customers? Were they driven by the acquisition of more referred customers or by the acquisition of higher-value referred customers? To answer the question, we examine the mean comparisons with respect to the three other outcome variables, which are also based on two-sided Welch's t-tests and are reported in Table 2. From the second and third rows of Table 2, we find that both the gift and notification treatments increased the incidence of having acquired referred customers and the number of referred customers by less than 30% and 25% relative to the control condition, respectively, and none of these effects is significant at the $p < 0.05$ level. On the other hand, the last row of Table 2 shows that referred customers acquired in both treatment conditions on average invested at least 150% more than those acquired in the control condition did, and both differences are significant (ps < 0.05).[19] Therefore, the data suggest that both the gift and notification treatments enhanced the effectiveness of the referral program primarily through the acquisition of higher-investment (and hence higher-value) referred customers rather than the acquisition of more referred customers.[20]

One caveat for the mean comparisons with respect to the investments of each referred customer is that it is unclear whether the sample sizes in the three experimental conditions

[19] Unlike the three other outcome variables, investments of a referred customer is defined on referred customers who were not observational units being randomized in the experiment. Therefore, the mean comparisons with respect to investments of each referred customer are purely descriptive and do not have a causal interpretation.

[20] Since we simultaneously conduct multiple mean comparisons, multiple testing is a concern. We consider two approaches to address this concern, including the Holm procedure and the Benjamini-Hochberg procedure. The Holm procedure controls the family-wise error rate (FWER), and it does not impose any assumption on the dependency among tests and is more powerful than the Bonferroni correction (Holm, 1979). In the Holm procedure, when we set FWER at 0.05, none of the 8 comparisons is significant; on the other hand, when we set FWER at 0.1, the comparisons between the control and gift conditions with respect to the total investments of referred customers and investments of each referred customer are significant. The Benjamini-Hochberg procedure controls the false discovery rate (FDR) under a positive dependence assumption among tests called PRDS that is likely to hold in our setting; compared to procedures controlling FWER, the Benjamini-Hochberg procedure provides less stringent control of Type-I errors but has greater power (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001). When we set FDR at 0.05, the comparisons between the control and gift conditions with respect to the total investments of referred customers and investments of each referred customer are significant; on the other hand, when we set FDR at 0.1, in addition to the two comparisons that are significant when FDR is set at 0.05, the comparisons between the control and notification conditions with respect to the total investments of referred customers and investments of each referred customer are also significant.

(83, 101, and 93) are sufficient to warrant the relaxation of the normality assumption on the underlying populations when we apply Welch's t-test. Given that the majority of referred customers did not make any investment, we consider the following alternative test procedure to compare the investments of each referred customer across conditions: Between the control condition and a treatment condition, we first test whether there is a difference between the proportions of referred customers with a positive investment amount (i.e., positive observations), and then test whether there is a difference between the two positive subsamples. We report the findings based on this alternative test procedure in Table 3.

Insert Table 3 here.

We compare the proportions of positive observations across conditions using the test for equal proportions, and find that the proportions are not significantly different between the control condition and either of the gift and notification conditions (ps > 0.1). On the other hand, as the Shapiro-Wilk tests indicate that the positive subsamples in all three conditions are not normally distributed (ps < 0.001), we compare the positive subsamples across conditions using the Wilcoxon-Mann-Whitney test. We find that the positive subsamples in both the gift and notification conditions have a higher average investment amount than that in the control condition, and the difference between the control and gift conditions is significant ($p < 0.05$) whereas the difference between the control and notification conditions is marginally significant ($p < 0.1$). Taken together, the comparisons with respect to the value of each referred customer based on this alternative test procedure are consistent with our earlier findings based on Welch's t-tests.[21]

2.3.3. Regression Analyses of Aggregate Effects

In this section, we estimate the aggregate effects of the gift and notification treatments using regression models. When the outcome variable is total investments of a current customer's

[21] In Appendix, we report comparisons with respect to the total investments and number of a current customer's referred customers based on a similar alternative test procedure, and comparisons with respect to the incidence of having acquired referred customers based on the test for equal proportions.

referred customers, we consider the following linear model:

$$Y_i = \alpha + \beta_g T_{g,i} + \beta_n T_{n,i} + \theta^\top X_i + \epsilon_i, \tag{2.1}$$

where i refers to the i-th current customer, Y_i denotes the total investments of the i-th current customer's referred customers (in RMB), $T_{g,i}$ is the dummy variable for the gift treatment, $T_{n,i}$ is the dummy variable for the notification treatment, X_i is a vector containing pre-experiment variables, and ϵ_i is the idiosyncratic term. We estimate this model using ordinary least squares (OLS) and report the results in the first column of Table 4.

Insert Table 4 here.

The OLS estimation of Model (2.1) suggests that both the gift and notification treatments significantly increased the total investments of referred customers, and the estimated effect sizes are very close to those obtained from mean comparisons. As a robustness check, we consider multiple alternative regression specifications. First, since the total investments of referred customers are right-skewed, we estimate Model (2.1) with a log-transformed dependent variable using OLS; given that a positive constant needs to be added to total investments of referred customers before taking the logarithm and there is no single "optimal" choice of such a constant, we estimate the log-linear model with four different choices of the additive constant: 1, 10, 100, and 1,000. Second, we estimate a tobit specification of Model (2.1) as the total investments of referred customers are zero for most current customers. We report results from the alternative specifications in Table 4. The findings from these specifications are not encouraging: While the coefficients of both treatments are positive in all specifications, most of them are not significant at the $p < 0.05$ level. Therefore, the positive effects of the gift and notification treatments on total investments of referred customers are not robust under the alternative specifications being considered.

When the outcome variable is the incidence of having acquired referred customers, we estimate a linear probability model (i.e., Model (2.1) with Y_i denoting the dummy variable for whether the i-th current customer has acquired referred customers) and a probit specification of Model (2.1). On the other hand, when the outcome variable is the number of referred customers, we estimate Model (2.1) with Y_i denoting the number of the i-th current customer's referred customers and a negative binomial model with a quadratic variance function (i.e., an NB2 model).[22] Results are summarized in Table 5.

Insert Table 5 here.

From Table 5, we find that the effect sizes of both treatments on the incidence and number of referred customers from the OLS estimations are very close to those obtained from mean comparisons. In terms of statistical significance, the linear probability and probit models indicate that neither treatment had an impact on the incidence of having acquired referred customers that is significant at the $p < 0.05$ level, and the linear and NB2 models show that neither treatment effect on the number of referred customers is significant at the $p < 0.05$ level; both findings are consistent with those from mean comparisons.

Finally, for investments of each referred customer, the outcome variable defined on referred customers, we consider the following linear model:

$$Z_j = \tilde{\alpha} + \tilde{\beta}_g T_{g,j} + \tilde{\beta}_n T_{n,j} + \varepsilon_j, \tag{2.2}$$

where j refers to the j-th referred customer, Z_j denotes the j-th referred customer's total investments (in RMB), $T_{g,j}$ is the dummy variable for whether the j-th referred customer was acquired in the gift condition, $T_{n,j}$ is the dummy variable for whether the j-th referred customer was acquired in the notification condition, and ε_j is the idiosyncratic term. We note that since referred customers were not randomly assigned to the experimental conditions, we do not intend to give a causal interpretation for the parameters $\tilde{\beta}_g$ and $\tilde{\beta}_n$; instead, these parameters merely serve to capture the differences between the average investment

[22]Given that an overdispersion test rejects the null hypothesis of equidispersion in favor of the alternative hypothesis of overdispersion ($p < 0.01$), we choose to estimate an NB2 model instead of a Poisson model.

amount of referred customers acquired in the control condition and that of referred customers acquired in the gift and notification conditions. We estimate this model using OLS and report the results in the first column of Table 6.

Insert Table 6 here.

Confirming our previous finding based on mean comparisons, the OLS estimation of Model (2.2) suggests that referred customers acquired in both treatment conditions on average invested at least 7,000 RMB more than those acquired in the control condition did, and both differences are significant (ps < 0.05). Similar to the case of total investments of referred customers, we assess the robustness of this finding using alternative regression specifications, including Model (2.2) with a log-transformed dependent variable with four different choices of the additive constant: 1, 10, 100, and 1,000, and a tobit specification of Model (2.2). We report results from these alternative specifications in Table 6. We find that, while the coefficients for both treatment condition dummy variables are always positive, only one turns out to be marginally significant at the $p < 0.1$ level and none is significant at the $p < 0.05$ level. Consequently, the positive differences between the two treatment conditions and the control condition in terms of investments of each referred customer are not robust under the alternative specifications being considered.

2.3.4. Exploring How Treatments Promoted Better Matching

In the previous sections, we find that both the gift and notification treatments enhanced the effectiveness of the referral program primarily through the acquisition of higher-investment referred customers, or, in other words, by promoting better matching between referred customers and the firm.[23] Research on referral programs and customer value has proposed active matching and passive matching as two distinct matching processes through which better matching operates (Schmitt et al., 2011; Van den Bulte et al., 2018); correspondingly,

[23] We emphasize that the finding that referred customers acquired in the gift and notification conditions on average invested more than those acquired in the control condition did is statistically significant under Welch's t-test, the two-stage alternative test, and the OLS estimation of the linear model; on the other hand, it is insignificant under the log-linear and tobit models.

there are two potential channels through which a treatment could promote better matching: (1) by motivating active matching, in which the treatment motivates current customers to exert greater effort to screen their friends and refer good matches to the firm, and (2) by facilitating passive matching, in which the treatment induces a higher proportion of high-value current customers to refer their friends who, by homophily, are more likely to be good matches than friends of low-value current customers.

In our setting, both the gift and notification treatments could impact customer referrals by eliciting reciprocity from and/or serving as a value signal to current customers, both of which could motivate active matching. Specifically, offering current customers a gift coupon before inviting them to refer friends may elicit reciprocity from current customers toward the firm (Gneezy and List, 2006; Falk, 2007; Alpizar et al., 2008; Kube et al., 2012; Gilchrist et al., 2016; Chung and Narayandas, 2017), which could in turn make them feel obliged to exert effort to screen their friends carefully and match those who they think may be valuable to the firm (Schmitt et al., 2011). The provision of the gift interest-raising coupon may also signal to current customers that the actual interest rates of the financial deposits could be higher than the listed interest rates, which may enhance their perception about the value that their friends would receive once referred to the firm; since customers tend to use referrals to help friends make better choices or manage friends' impressions of them (Ryu and Feick, 2007; Kornish and Li, 2010), such an enhanced value perception is likely to motivate current customers to engage in active matching. As to the notification treatment, a text message notifying current customers about their total investment return may make it salient to them that they have received great value from the firm, which could also trigger reciprocity from and/or serve as a value signal to them, and in turn motivate active matching.

In addition to motivating active matching, both the gift and notification treatments could also facilitate passive matching. Given that high-value current customers are likely to have a better match with the firm and hence invest more in the fixed deposits during the campaign period compared to low-value current customers, the gift coupon is likely to be more valuable to high-value current customers because its value is proportional to the amount of the fixed-deposit investment to which it is applied.[24] As a result, the gift treatment is more likely to elicit reciprocity from high-value current customers assuming reciprocity is the mechanism through which it impacts customer referrals, and it is likely to serve as a stronger signal to high-value current customers assuming value signaling is the underlying mechanism. In either case, the gift treatment is likely to motivate a higher proportion of high-value current customers to refer their friends and facilitate passive matching. On the other hand, since higher-value current customers are likely to have received more investment return from the firm, notifying current customers about their total investment return is more likely to induce reciprocity from and/or serves as a stronger value signal to high-value current customers. Consequently, high-value current customers are more likely to refer friends after receiving the notification compared to low-value current customers and passive matching is facilitated.

In this section, we seek to shed light on how the gift and notification treatments promoted better matching by investigating whether the data are consistent with the treatments motivating active matching and/or facilitating passive matching. We start with the implication of a treatment motivating active matching. Assuming there are two identical current customers, A and B, and A is assigned to the control condition while B is assigned to a treatment condition. If the treatment indeed motivates active matching and both A and B have acquired referred customers, we expect that the average value of B's referred customers is higher than the average value of A's referred customers. To examine whether data from the experiment are consistent with such an implication, we estimate Model (2.1) on the sample of current customers who have acquired referred customers during the experiment.[25] Here, the dependent variable Y_i denotes the average investment amount of the i-th

[24] In Appendix, we show that customers of higher value, operationalized as those with a higher total investment return and those with a larger current investment amount, invested more in the fixed deposits during the experiment when offered the gift coupon.

[25] For simplicity, we refer to current customers who have acquired referred customers during the experiment as referrers hereafter.

referrer's referred customers. Since referrers are likely to be different across conditions, we try to partially address this selection issue by controlling for their observable characteristics X_i, including log total investment return, log investment amount, tenure, time since last investment, and whether had successful referrals before the experiment. We report the results of the OLS estimation in the first column of the upper section of Table 7.

Insert Table 7 here.

We find that both treatments are estimated to have increased the average investments of a referrer's referred customers by at least 7,000 RMB and both effects are significant (ps < 0.05), which is consistent with both treatments motivating active matching. One caveat of this regression analysis is that controlling for the five referrers' observable characteristics is unlikely to fully address the selection issue. As a robustness check, we include a richer set of controls by adding the interaction of each pair of the five characteristics (i.e., a total of ten interactions) to the model and redo the analysis. We report the results of the OLS estimation in the first column of the lower section of Table 7. It is evident that the inclusion of the interactions only has a limited impact on the estimated effect sizes of both treatments and both effects remain significant (ps < 0.05). Therefore, our finding is still consistent with both treatments motivating active matching after adopting a richer set of controls.

Similar to Section 2.3.3, we also assess the robustness of this finding using multiple alternative regression specifications, including log-linear models with four different choices of the additive constant (1, 10, 100, and 1,000) and a tobit model. All models are estimated both with and without controlling for the pairwise interactions of the five referrers' characteristics. We report results from these alternative specifications in Table 7. Clearly, the functional form assumption has a considerable impact on our finding: While the coefficients of both treatments are positive in all specifications, only one turns out to be marginally significant at the $p < 0.1$ level and none is significant at the $p < 0.05$ level. Consequently, results from the log-linear and tobit models suggest that our earlier finding based on the linear models is not robust; under the alternative regression specifications, we do not have sufficient statistical evidence suggesting that the average investments of referred customers are higher for referrers in the two treatment conditions than for those in the control condition.

Now we consider the implication of a treatment facilitating passive matching. If a treatment indeed facilitates passive matching, we expect that, compared to the control condition, it induces a higher proportion of high-value current customers to refer friends and hence leads to a higher average value of referrers. To investigate whether data from the experiment are consistent with such an implication, we operationalize the value of a current customer using two observable characteristics, total investment return and current investment amount, and compare referrers in the two treatment conditions and those in the control condition with respect to these two characteristics. We report the results of the comparisons in Table 8.

Insert Table 8 here.

Using Welch's t-test, we find that referrers in the gift and notification conditions are not significantly different from those in the control condition both in terms of total investment return and current investment amount (ps > 0.1). One caveat of applying Welch's t-test in the current setting is that, as the Shapiro-Wilk tests indicate that all samples are non-normal (ps < 0.001), it is unclear whether the sample sizes (76, 99, and 83) are sufficient to warrant the relaxation of the normality assumption on the underlying populations when we apply Welch's t-test. As a robustness check, we compare referrers across conditions using the Wilcoxon-Mann-Whitney test which does not impose the normality assumption. Again, we find that the differences between the control condition and the two treatment conditions are insignificant (ps > 0.1).[26] While we do not find significant differences between the control

[26]We note that, by operationalizing the value of a current customer using his total investment return and current investment amount, we implicitly assume that referrers with a higher total investment return and those with a larger current investment amount are likely to acquire referred customers of higher value. To assess the validity of this assumption, we calculate the correlation between each of these two characteristics and the average investments of referred customers on the sample of all referrers. Between total investment return and average investments of referred customers, Pearson's $r = 0.15$ ($p < 0.05$) and Spearman's $\rho = 0.10$ ($p = 0.10$); between current investment amount and average investments of referred customers, Pearson's $r = 0.15$ ($p < 0.05$) and Spearman's $\rho = 0.05$ ($p > 0.1$). We interpret this finding as suggestive evidence supporting our assumption, with the caveat that when using Spearman's ρ we cannot reject the

condition and the two treatment conditions in terms of referrers' total investment return and current investment amount, we emphasize that referrers in the control condition and those in the two treatment conditions could differ with respect to some unobservable characteristics that are correlated with the value of referred customers through homophily. As a result, we cannot rule out the possibility of the gift and notification treatments facilitating passive matching based on the data we have.

2.3.5. Heterogeneous Effects of the Treatments

In this section, we investigate how the effects of the gift and notification treatments vary across current customers. By exploring customer heterogeneity in treatment effects, we aim to shed light on the mechanisms through which the two treatments impacted customer referrals. Specifically, we are interested in answering the following question: Did the gift and notification treatments enhance the effectiveness of the referral program by eliciting reciprocity from current customers or serving as a value signal to current customers? Understanding customer heterogeneity in treatment effects is also managerially relevant, as it could provide firms guidance on identifying current customers on whom the treatments are likely to be most effective and hence may serve as good targets for future implementations of these treatments. In the following, we consider customer characteristics that may moderate the effects of the gift and notification treatments on the total investments of referred customers (our key outcome variable), the incidence of having acquired referred customers, and the number of referred customers.

Customer Value

As discussed in Section 2.3.4, the gift and notification treatments are likely to elicit stronger reciprocity from higher-value current customers if they work by inducing reciprocity from current customers toward the firm, and they are likely to serve as a stronger signal to higher-value current customers if value signaling is the underlying mechanism. With either

null hypothesis of zero association.

mechanism being at work, high-value current customers are likely to be better motivated by the treatments to refer friends, especially those who are good matches, to the firm compared to low-value current customers. Therefore, we propose the following hypotheses regarding the moderating role of customer value:

H1: The treatments have a larger impact on current customers of higher value in terms of the total investments of referred customers.

H2: The treatments have a larger impact on current customers of higher value in terms of the incidence of having acquired referred customers.

H3: The treatments have a larger impact on current customers of higher value in terms of the number of referred customers.

Investment Recency

If the gift and notification treatments work by inducing reciprocity from current customers toward the firm, we predict that both treatments have a larger impact on customers whose last investment was more recent. This is because customers whose last investment was more recent are likely to have a stronger relationship with the firm at the time of the experiment compared to those whose last investment was less recent (Netzer et al., 2008; Kumar et al., 2010), and hence are more likely to feel the obligation to reciprocate to the firm by referring friends (especially those who are good matches) after receiving either the gift or the notification. On the other hand, if value signaling is at work, we predict that both treatments have a smaller impact on customers whose last investment was more recent based on two institutional details. First, the real-time information about total investment return is prominently displayed in a customer's account. Second, the firm sends all customers interest-raising coupons as a promotion multiple times a year, and information about current and past interest-raising coupons is available in the coupon section of a customer's account. Since before making an investment, a customer needs to log into his account and is likely to browse the coupon section to check if any coupon can be applied to the investment, these two institutional details imply that a customer who makes an investment is likely to have accessed information about his total investment return and information about promotional interest-raising coupons, which, similar to the gift coupon, may remind him that the actual interest rates of the financial deposits could be higher than the listed interest rates. Compared to customers whose last investment was less recent, those whose last investment was more recent are likely to have accessed a more up-to-date version of such information at a more recent time, in which case the signaling value of the gift coupon and the notification is smaller and hence a smaller impact on customer referrals. In sum, we propose the following hypotheses assuming that the gift and notification treatments work by eliciting reciprocity from current customers:

H4a: The treatments have a larger impact on current customers whose last investment was more recent in terms of the total investments of referred customers.

H5a: The treatments have a larger impact on current customers whose last investment was more recent in terms of the incidence of having acquired referred customers.

H6a: The treatments have a larger impact on current customers whose last investment was more recent in terms of the number of referred customers.

In contrast, we propose the following competing hypotheses assuming that the gift and notification treatments work by serving as a value signal to current customers:

H4b: The treatments have a smaller impact on current customers whose last investment was more recent in terms of the total investments of referred customers.

H5b: The treatments have a smaller impact on current customers whose last investment was more recent in terms of the incidence of having acquired referred customers.

H6b: The treatments have a smaller impact on current customers whose last investment was more recent in terms of the number of referred customers.

Past Referral Behavior

Customers who had made successful referrals in the past are likely to have a more committed relationship with the firm (Verhoef et al., 2002); if the gift and notification treatments work by inducing reciprocity, these customers are more likely to feel the obligation to reciprocate to the firm by referring friends (especially those who are good matches) after receiving either the gift or the notification than those who had not made successful referrals in the past. Therefore, we propose the following hypotheses regarding the moderating role of past referral behavior assuming that the gift and notification treatments work by eliciting reciprocity from current customers:

H7: The treatments have a larger impact on current customers who had made successful referrals before the experiment in terms of the total investments of referred customers.

H8: The treatments have a larger impact on current customers who had made successful referrals before the experiment in terms of the incidence of having acquired referred customers.

H9: The treatments have a larger impact on current customers who had made successful referrals before the experiment in terms of the number of referred customers.

Empirical Testing of the Hypotheses

In this section, we empirically test our proposed hypotheses using regression analyses. For hypotheses pertaining to the total investments of referred customers, our key outcome variable, we consider the following linear model:

$$Y_i = \alpha + \beta_g T_{g,i} + \beta_n T_{n,i} + \theta^\top X_i + \gamma_g^\top T_{g,i} X_i + \gamma_n^\top T_{n,i} X_i + \epsilon_i, \tag{2.3}$$

where i refers to the i-th current customer, Y_i denotes the total investments of the i-th current customer's referred customers (in RMB), $T_{g,i}$ is the dummy variable for the gift treatment, $T_{n,i}$ is the dummy variable for the notification treatment, X_i is a vector containing potential moderators that we are going to test, and ϵ_i is the idiosyncratic term. In order to test H1 (i.e., the moderating role of customer value), we adopt three different operationalizations of customer value, including log total investment return, log investment amount, and a single factor between log total investment return and log investment amount which we term as the value factor (Iyengar and Park, 2018).[27] Corresponding to the three operationalizations of customer value, we consider three specifications for X_i: In the first specification, X_i includes the i-th current customer's log total investment return, time since last investment, and whether he had made successful referrals in the past; the second and third specifications replace log total investment return with log investment amount and the value factor, respectively. We estimate Model (2.3) with different specifications of X_i using OLS and report the results in the first three columns of Table 9.

Insert Table 9 here.

Across different specifications of X_i, we find that the interactions between both treatments and all three operationalizations of customer value are all positive. Specifically, a 10% increase in total investment return increases the effect of the gift treatment by 0.94 RMB and the effect of the notification treatment by 1.98 RMB; a 10% increase in current investment amount increases the effect of the gift treatment by 1.38 RMB and the effect of the notification treatment by 1.75 RMB; and increasing the value factor by 0.1 increases the effect of the gift treatment by 1.35 RMB and the effect of the notification treatment by 2.19 RMB. We also find negative interactions between both treatments and time since last investment, where making the last investment one day less recent decreases the effect of the gift treatment by 0.01-0.07 RMB and the effect of the notification treatment by 0.02-0.09 RMB. Finally, the interactions between both treatments and the dummy variable of having made referrals before the experiment (i.e., the past referrals dummy) are all positive; we find

[27] We obtain the factor scores by conducting a factor analysis on log total investment return and log investment amount, which reveals a single factor accounting for 77.45% of the total variance in the two variables.

that the effect of the gift treatment is between 119.41 to 121.20 RMB larger and the effect of the notification treatment is between 18.78 to 23.43 RMB larger for current customers who had made successful referrals compared to those who had not. An inspection of the statistical significance of these interactions, however, indicates that the lack of power is an issue: Few interactions are significant at the $p < 0.05$ level, which include the interaction of the notification and log total investment return in the first specification, the interaction of the gift and log investment amount in the second specification, and the interaction of the gift and the past referrals dummy in all three specifications.

To assess the robustness of the findings from the linear models, we estimate log-linear and tobit variants of Model (2.3) and report results in Tables 9, 10, and 11.

Insert Tables 10 and 11 here.

In all log-linear and tobit models, we find that the interactions between both treatments and all three operationalizations of customer value are positive but none is significant at the $p < 0.05$ level. This finding, together with the one from the linear models, suggests that we do not have sufficient statistical evidence to reach a conclusion regarding H1. We also find that the signs of the interactions between both treatments and time since last investment are not consistent across models; most of these interactions are not significant at the $p < 0.05$ level, with the exception of the interaction between the gift and time since last investment in the tobit models, which is positive and significant (ps < 0.05). Hence, we are unable to make claims regarding H4a and H4b given these largely insignificant interactions with inconsistent signs across models. Finally, the interaction between the gift (notification) treatment and the past referrals dummy is positive (negative) in all log-linear and tobit models, but none is significant at the $p < 0.05$ level. Taken together, these insignificant results and our earlier finding from the linear models do not provide us sufficient statistical evidence to draw conclusions regarding H7.

For hypotheses pertaining to the incidence of having acquired referred customers, we estimate linear probability models (i.e., Model (2.3) with Y_i denoting the dummy variable for whether the i-th current customer has acquired referred customers) and probit models. Results are summarized in Table 12.

Insert Table 12 here.

From Table 12, we find that the interactions between both treatments and all three operationalizations of customer value are positive but none is significant at the $p < 0.05$ level based on both the linear probability and probit models; as a result, we are unable to make claims regarding H2. In the probit models, the interaction between the gift and time since last investment is positive and significant (ps < 0.05), which is consistent with H5b and inconsistent with H5a for the gift treatment; however, this finding is not robust under the linear probability models as the interaction becomes insignificant. On the other hand, the interaction between the notification and time since last investment is not significant at the $p < 0.05$ level in both the linear probability and probit models, indicating that we do not have sufficient statistical evidence to reach a conclusion regarding H5a and H5b for the notification treatment. Finally, while the interaction of the gift and the past referrals dummy changes sign across models and is never significant at the $p < 0.05$ level, the interaction of the notification and the past referrals dummy is negative and significant in both the linear probability and probit models (ps < 0.05). Therefore, our finding is inconsistent with H8 for the notification treatment.

For hypotheses pertaining to the number of referred customers, we estimate Model (2.3) with Y_i denoting the number of the i-th current customer's referred customers and NB2 models.[28] Results are summarized in Table 13.

Insert Table 13 here.

We make the following observations from Table 13. First, the interactions between both

[28]Given that an overdispersion test rejects the null hypothesis of equidispersion in favor of the alternative hypothesis of overdispersion for all specifications of X_i (ps < 0.01), we choose to estimate NB2 models instead of Poisson models.

treatments and all three operationalizations of customer value are positive in both the linear and NB2 models, but only one - the interaction between the notification and log total investment return in the NB2 model - is significant at the $p < 0.05$ level. Therefore, we do not have sufficient statistical evidence to reach a conclusion regarding H3. Second, the interaction between the gift and time since last investment is positive and significant (ps < 0.05) in the NB2 models, which is consistent with H6b and inconsistent with H6a for the gift treatment, but this finding is not robust under the linear models. On the other hand, the interaction between the notification and time since last investment is not significant at the $p < 0.05$ level in both the linear and NB2 models. Finally, while the interaction of the gift and the past referrals dummy is negative but not significant at the $p < 0.05$ level in both the linear and NB2 models, the interaction of the notification and the past referrals dummy is negative and significant in the NB2 models (ps < 0.05). This finding is inconsistent with H9 for the notification treatment, but it is not robust under the linear models as the interaction becomes not significant at the $p < 0.05$ level.

To summarize, the lack of statistical power is a serious issue in the empirical testing of our proposed hypotheses. Among all hypotheses, we only have sufficient statistical evidence to reach a conclusion regarding H8 for the notification treatment, where the data are inconsistent with H8 for the notification treatment. We fail to obtain statistically significant interactions consistent across different models that would allow us to draw sharp conclusions regarding all other hypotheses, and hence are largely unable to shed light on the mechanisms through which the gift and notification treatments impacted customer referrals.

2.4. Empirical Findings: The Second Experiment

In this section, we analyze data from the second experiment to investigate the effects of the two operationalizations of the value-based reward treatment. Similar to Section 2.3, we first verify that the random assignment of customers was valid. We then estimate the aggregate effects of the two value-based rewards on the effectiveness of the referral program. Finally, we shed light on the workings of these two value-based rewards by exploring the mechanisms through which they operate.

2.4.1. Randomization Check

We assess the validity of the randomization by comparing customers across conditions with respect to their behaviors prior to the experiment. Again, with a valid random assignment, we should observe little or no systematic differences in customers' pre-experiment variables across conditions. The details of the randomization check are reported in Table 14.

Insert Table 14 here.

Similar to Section 2.3.1, we use a two-sided Welch's t-test to examine whether two experimental conditions have the same population mean on a pre-experiment variable. From Table 14, it is evident that customers' pre-experiment variables are well balanced across conditions, confirming the validity of the random assignment in the second experiment.

2.4.2. Aggregate Effects of the Treatments

We measure the effectiveness of the referral program for each current customer using the total value of his referred customers, which is operationalized similarly as in Section 2.3.2. Specifically, we identify new customers who opened an account at the firm using a current customer's referral link during the 30-day campaign period (i.e., his referred customers), and proxy for the total value of a current customer's referred customers using the total amount of investments made by all his referred customers in all financial deposits combined during the 90-day observational period starting from the beginning of the experiment. In addition to the total value of referred customers, we also consider three other outcome variables in order to obtain a more thorough understanding of the impact of the value-based rewards on customer referrals: (1) whether a current customer has acquired any referred customers during the experiment (Yes/No), (2) the number of a current customer's referred customers, and (3) the value of a referred customer assessed by the amount of investments she made during the 90-day observational period. We note that all outcome variables except the value of a referred customer are defined on each current customer, and for these outcome variables the sample size in each condition is the number of current customers in that condition: 40,076 in the control condition, 40,145 in the first value-based reward condition, and 40,037 in the second value-based reward condition. On the other hand, the value of a referred customer is defined on each referred customer, and the sample size in each condition for this outcome variable is the number of referred customers acquired in that condition: 263 in the control condition, 300 in the first value-based reward condition, and 328 in the second value-based reward condition.

The randomization allows us to identify the average treatment effect of each value-based reward on the total value of a current customer's referred customers via a direct mean comparison between the control condition and the corresponding treatment condition. We report the result of the mean comparisons based on two-sided Welch's t-tests in the first row of Table 15.

Insert Table 15 here.

Again, we assume that the sample in each condition consists of i.i.d. draws from an underlying population in order to apply Welch's t-test; on the other hand, the validity of the test does not require the populations to have equal variances or be normally distributed given the large sample sizes. It is clear that both value-based rewards had a substantial impact on the total investments of referred customers. In particular, the first and second value-based rewards on average increased the total investments of a current customer's referred customers by 127.84 RMB and 99.93 RMB, respectively. Both treatment effects amount to more than 100% lift over the baseline value of 92.97 RMB and are statistically significant (ps < 0.01).

We also examine the mean comparisons with respect to the three other outcome variables using two-sided Welch's t-test in order to understand which of the two, the acquisition of more referred customers or the acquisition of higher-value referred customers, was the main driver of the positive effects of the two value-based rewards on the total investments of referred customers. From the second and third rows of Table 15, we find that only the second value-based reward had a positive effect on the incidence of having acquired referred customers that is close to being significant ($p = 0.053$) and a positive and significant effect on the number of referred customers ($p < 0.05$), whereas the effects of the first value-based reward on these two outcome variables are insignificant. The effect sizes of the second value-based reward, however, are modest: It only increased the incidence of having acquired referred customers and the number of referred customers by less than 25% relative to the control condition. On the other hand, the last row of Table 15 shows that referred customers acquired in the first value-based reward condition and those acquired in the second value-based reward condition on average invested at least 100% and 65% more than those acquired in the control condition did, respectively, and both differences are significant (ps < 0.05). Therefore, similar to the findings in the first experiment, the data suggest that both value-based rewards enhanced the effectiveness of the referral program primarily through the acquisition of higher-investment (and hence higher-value) referred customers rather than the acquisition of more referred customers.[29]

Similar to Section 2.3.2, one caveat for the mean comparisons with respect to the investments of each referred customer is that it is unclear whether the sample sizes in the three experimental conditions (263, 300, and 328) are sufficient to warrant the relaxation of the normality assumption on the underlying populations when we apply Welch's t-test. We

[29]Similar to Section 2.3.2, we adopt the Holm procedure and the Benjamini-Hochberg procedure to address the issue of multiple testing. In the Holm procedure, when we set the family-wise error rate (FWER) at 0.05, the comparisons between the control and both value-based reward conditions with respect to the total investments of referred customers are significant; on the other hand, when we set FWER at 0.1, in addition to the two comparisons that are significant when FWER is set at 0.05, the comparison between the control and first value-based reward conditions with respect to investments of each referred customer and the comparison between the control and second value-based reward conditions with respect to the number of referred customers are also significant. In the Benjamini-Hochberg procedure, when we set FDR at 0.05, the comparisons between the control and first value-based reward conditions with respect to the total investments of referred customers and investments of each referred customer are significant, and the comparisons between the control and second value-based reward conditions with respect to the total investments of referred customers, the number of referred customers, and investments of each referred customer are significant as well; on the other hand, when we set FDR at 0.1, in addition to the 5 comparisons that are significant when FDR is set at 0.05, the comparison between the control and second value-based reward conditions with respect to the incidence of having acquired referred customers is also significant.

adopt a similar alternative test procedure to compare the investments of each referred customer across conditions in which, between the control condition and a treatment condition, we first test whether there is a difference between the proportions of referred customers with a positive investment amount (i.e., positive observations), and then test whether there is a difference between the two positive subsamples. We report the findings based on this alternative test procedure in Table 16.

Insert Table 16 here.

We compare the proportions of positive observations across conditions using the test for equal proportions, and find that the proportions are not significantly different between the control condition and either value-based reward condition (ps > 0.1). On the other hand, as the Shapiro-Wilk tests indicate that the positive subsamples in all three conditions are not normally distributed (ps < 0.001), we compare the positive subsamples across conditions using the Wilcoxon-Mann-Whitney test. We find that the positive subsamples in both value-based reward conditions have a higher average investment amount than that in the control condition, and the difference between the control and first value-based reward conditions is marginally significant ($p < 0.1$) whereas the difference between the control and second value-based reward conditions is significant ($p < 0.05$). Taken together, the comparisons with respect to the value of each referred customer based on this alternative test procedure are consistent with our earlier findings based on Welch's t-tests.[30]

2.4.3. Regression Analyses of Aggregate Effects

In this section, we estimate the aggregate effects of both value-based reward treatments using regression models. When the outcome variable is total investments of a current customer's referred customers, we consider the following linear model:

$$Y_i = \alpha + \beta_{fr} T_{fr,i} + \beta_{sr} T_{sr,i} + \theta^\top X_i + \epsilon_i, \tag{2.4}$$

[30]In Appendix, we report comparisons with respect to the total investments and number of a current customer's referred customers based on a similar alternative test procedure, and comparisons with respect to the incidence of having acquired referred customers based on the test for equal proportions.

where i refers to the i-th current customer, Y_i denotes the total investments of the i-th current customer's referred customers, $T_{fr,i}$ is the dummy variable for the first value-based reward treatment, $T_{sr,i}$ is the dummy variable for the second value-based reward treatment, X_i is a vector containing pre-experiment variables, and ϵ_i is the idiosyncratic term. We estimate this model using OLS and report the results in the first column of Table 17.

Insert Table 17 here.

The OLS estimation of Model (2.4) suggests that both value-based rewards significantly increased the total investments of referred customers (ps < 0.01), and the estimated effect sizes are very close to those obtained from mean comparisons. Similar to Section 2.3.3, we assess the robustness of this finding using multiple alternative regression specifications, including Model (2.4) with a log-transformed dependent variable with four different choices of the additive constant: 1, 10, 100, and 1,000, and a tobit specification of Model (2.4). Results from these alternative specifications are reported in Table 17. We find that the coefficients of both value-based reward treatments are positive and significant (ps < 0.05) in all alternative specifications being considered, indicating that the finding of the positive effects of both value-based reward treatments on total investments of referred customers is robust across different regression specifications.

When the outcome variable is the incidence of having acquired referred customers, we estimate a linear probability model (i.e., Model (2.4) with Y_i denoting the dummy variable for whether the i-th current customer has acquired referred customers) and a probit specification of Model (2.4). On the other hand, when the outcome variable is the number of referred customers, we estimate Model (2.4) with Y_i denoting the number of the i-th current customer's referred customers and an NB2 specification of Model (2.4).[31] Results are summarized in Table 18.

Insert Table 18 here.

[31] Given that an overdispersion test rejects the null hypothesis of equidispersion in favor of the alternative hypothesis of overdispersion ($p < 0.01$), we choose to estimate an NB2 model instead of a Poisson model.

From Table 18, we find that the effect sizes of both treatments on the incidence and number of referred customers based on the OLS estimations are very close to those obtained from mean comparisons. In terms of statistical significance, the linear probability and probit models indicate that neither treatment had an impact on the incidence of having acquired referred customers that is significant at the $p < 0.05$ level, and the linear and NB2 models show that only the second value-based reward treatment had a significant impact on the number of referred customers (ps < 0.05); both findings are consistent with those from mean comparisons.

For investments of each referred customer, the outcome variable defined on referred customers, we consider the following linear model:

$$Z_j = \tilde{\alpha} + \tilde{\beta}_{fr} T_{fr,j} + \tilde{\beta}_{sr} T_{sr,j} + \varepsilon_j, \tag{2.5}$$

where j refers to the j-th referred customer, Z_j denotes the j-th referred customer's total investments (in RMB), $T_{fr,j}$ is the dummy variable for whether the j-th referred customer was acquired in the first value-based reward condition, $T_{sr,j}$ is the dummy variable for whether the j-th referred customer was acquired in the second value-based reward condition, and ε_j is the idiosyncratic term. Similar to Section 2.3.3, we do not intend to give a causal interpretation for the parameters $\tilde{\beta}_{fr}$ and $\tilde{\beta}_{sr}$, which merely serve to capture the differences between the average investment amount of referred customers acquired in the control condition and that of referred customers acquired in the two value-based reward conditions. We estimate this model using OLS and report the results in the first column of Table 19.

Insert Table 19 here.

Confirming our previous finding based on mean comparisons, the OLS estimation of Model (2.5) suggests that referred customers acquired in both treatment conditions on average invested at least 9,000 RMB more than those acquired in the control condition did, and both differences are significant (ps < 0.05). As a robustness check, we consider multiple alternative regression specifications, including Model (2.5) with a log-transformed dependent variable with four different choices of the additive constant: 1, 10, 100, and 1,000, and a tobit specification of Model (2.5). We report results from these alternative specifications in Table 19. While the coefficients for both treatment condition dummy variables are always positive, most of them are not significant at the $p < 0.1$ level. The largely insignificant coefficient estimates are likely due to a lack of power; since each of the two treatment conditions implemented an operationalization of the value-based reward treatment, we can seek to obtain greater statistical power by combining the two treatment conditions into a single composite value-based reward condition. We redo all analyses in Table 19 after combining the two treatment conditions, and report the results in Table 20.

Insert Table 20 here.

We find that the coefficient of the value-based reward condition dummy variable is significant in the linear, tobit, and one log-linear specifications (ps < 0.05), marginally significant in two log-linear specifications (ps < 0.1), and insignificant in one log-linear specification. Hence, combining the two original treatment conditions into one composite value-based reward condition indeed provides us more power, and the finding that referred customers acquired by current customers who were offered a value-based reward on average invested more than those acquired by current customers not offered one seems to be reasonably robust across different specifications.

2.4.4. Exploring How Treatments Promoted Better Matching

In the previous sections, we find that both value-based reward treatments enhanced the effectiveness of the referral program primarily through the acquisition of higher-investment referred customers. Since the value-based rewards offer current customers an economic incentive to bring in high-value referred customers, they could promote better matching by motivating current customers to exert greater effort to screen their friends and refer good matches to the firm (i.e., motivating active matching). On the other hand, homophily implies that, compared to low-value current customers, high-value current customers are more likely to have friends who could become high-value new customers and whose referrals could earn them value-based rewards; that is, high-value current customers are likely to be better incentivized than low-value current customers when offered a value-based reward, in which case passive matching is facilitated.

In this section, we seek to shed light on how the value-based reward treatments promoted better matching by investigating whether the data are consistent with the treatments motivating active matching and/or facilitating passive matching. Similar to Section 2.3.4, we first examine whether the average investments of referred customers are higher for referrers in the two treatment conditions than for those in the control condition after controlling for referrers' observable characteristics. Specifically, we estimate Model (2.4) on the sample of referrers. Here, the dependent variable Y_i denotes the average investment amount of the i-th referrer's referred customers, and the vector X_i includes the i-th referrer's log total investment return, log investment amount, tenure, time since last investment, and whether he had successful referrals before the experiment. We report results of the OLS estimation in the first column of the upper section of Table 21.

Insert Table 21 here.

The first value-based reward treatment is estimated to have increased the average investments of a referrer's referred customers by more than 12,000 RMB and the effect is significant ($p < 0.05$), whereas the effect of the second value-based reward treatment is considerably smaller and insignificant. After controlling for the pairwise interactions of the five characteristics, the effect of the first value-based reward treatment becomes marginally significant ($p < 0.1$) and the effect of the second value-based reward treatment remains insignificant, as shown in the first column of the lower section of Table 21. We also assess the robustness of the findings from linear models using multiple alternative regression specifications, including log-linear models with four different choices of the additive constant (1, 10, 100, and

1,000) and a tobit model; all models are estimated both with and without controlling for the pairwise interactions of the five referrers' characteristics. We report results from these alternative specifications in Table 21. While the coefficients of both treatments are positive in all specifications, all are insignificant (ps > 0.1) with the exception of the coefficient of the first value-based reward treatment in the two tobit specifications. Therefore, across the regression specifications being considered, we do not observe a consistent pattern that the average investments of referred customers are significantly higher for referrers in the two treatment conditions compared to those in the control condition, after controlling for referrers' observable characteristics. Finally, we redo all analyses in Table 21 after combining the two value-based reward conditions into a single composite value-based reward condition and report the results in Table 22.

Insert Table 22 here.

The coefficient of the composite value-based reward treatment is positive in all specifications; however, only one turns out to be marginally significant at the $p < 0.1$ level and none is significant at the $p < 0.05$ level. Hence, we still do not have sufficient statistical evidence suggesting that the average investments of referred customers are higher for referrers who were offered a value-based reward compared to those not offered one, after controlling for referrers' observable characteristics.

We now investigate whether referrers in the two value-based reward conditions on average have a higher total investment return and a larger current investment amount than those in the control condition. Results of the comparisons are summarized in Table 23.

Insert Table 23 here.

We find that referrers in both value-based reward conditions have a higher total investment return and a larger current investment amount than those in the control condition, but only the differences between the control and second value-based reward conditions are significant based on Welch's t-tests (ps < 0.05). Again, as the Shapiro-Wilk tests indicate that all samples are non-normal (ps < 0.001), it is unclear whether the sample sizes (241, 260, and 285) are sufficient to warrant the relaxation of the normality assumption on the underlying populations when we use Welch's t-test. As a robustness check, we compare referrers across conditions using the Wilcoxon-Mann-Whitney test which does not impose the normality assumption. Findings from the Wilcoxon-Mann-Whitney tests are largely consistent with those based on Welch's t-tests, with the exception that the difference between the control and first value-based reward conditions in terms of referrers' total investment return is now marginally significant ($p < 0.1$). Taken together, our findings are consistent with the second value-based reward treatment facilitating passive matching; on the other hand, we do not observe a consistent pattern that referrers in the first value-based reward condition on average have a significantly higher value than those in the control condition.[32] Finally, we redo all analyses in Table 23 after combining the two value-based reward conditions into a single composite value-based reward condition, and report the results in Table 24.

Insert Table 24 here.

We find that referrers in the composite value-based reward condition on average have a higher total investment return and a larger current investment amount than those in the control condition; the difference in terms of total investment return is significant based on both Welch's t-test and the Wilcoxon-Mann-Whitney test (ps < 0.05), whereas the difference in terms of current investment amount is marginally significant based on Welch's t-test ($p < 0.1$) and significant based on the Wilcoxon-Mann-Whitney test ($p < 0.05$). Therefore, the data are consistent with the provision of a value-based reward facilitating passive matching.

[32]Similar to Section 2.3.4, we implicitly assume that referrers with a higher total investment return and those with a larger current investment amount are likely to acquire referred customers of higher value. To assess the validity of this assumption, we calculate the correlation between each of these two characteristics and the average investments of referred customers on the sample of all referrers. Between total investment return and average investments of referred customers, Pearson's $r = 0.15$ ($p < 0.01$) and Spearman's $\rho = 0.16$ ($p < 0.01$); between current investment amount and average investments of referred customers, Pearson's $r = 0.18$ ($p < 0.01$) and Spearman's $\rho = 0.15$ ($p < 0.01$). Hence, our assumption is supported by the data.

2.4.5. Heterogeneous Effects of the Treatments

In this section, we investigate how the effects of the two value-based reward treatments vary across current customers. Understanding customer heterogeneity in treatment effects could shed light on the workings of value-based rewards and provide firms guidance on identifying current customers on whom value-based rewards are likely to be most effective. In the following, we consider customer characteristics that may moderate the effects of the value-based reward treatments on the total investments of referred customers (our key outcome variable), the incidence of having acquired referred customers, and the number of referred customers.

Customer Value

As discussed in Section 2.4.4, high-value current customers are likely to be better incentivized to refer friends (especially those who are good matches) when offered a value-based reward than low-value current customers, because homophily implies that the former are more likely to have friends who could become high-value new customers and generate value-based rewards for them. Therefore, we propose the following hypotheses regarding the moderating role of customer value:

H10: The treatments have a larger impact on current customers of higher value in terms of the total investments of referred customers.

H11: The treatments have a larger impact on current customers of higher value in terms of the incidence of having acquired referred customers.

H12: The treatments have a larger impact on current customers of higher value in terms of the number of referred customers.

Empirical Testing of the Hypotheses

In this section, we empirically test our proposed hypotheses using regression analyses. For H10, we consider the following linear model:

$$Y_i = \alpha + \beta_{fr} T_{fr,i} + \beta_{sr} T_{sr,i} + \theta X_i + \gamma_{fr} T_{fr,i} X_i + \gamma_{sr} T_{sr,i} X_i + \epsilon_i, \tag{2.6}$$

where i refers to the i-th current customer, Y_i denotes the total investments of the i-th current customer's referred customers (in RMB), $T_{fr,i}$ is the dummy variable for the first value-based reward treatment, $T_{sr,i}$ is the dummy variable for the second value-based reward treatment, X_i is a variable that operationalizes the value of the i-th current customer, and ϵ_i is the idiosyncratic term. Similar to Section 2.3.5, we adopt three different operationalizations of customer value, including log total investment return, log investment amount, and a single factor between log total investment return and log investment amount which we term as the value factor.[33] We estimate Model (2.6) with different specifications of X_i using OLS and report the results in the first three columns of Table 25.

Insert Table 25 here.

We find that the interactions between both treatments and all three operationalizations of customer value are all positive. Specifically, a 10% increase in total investment return increases the effects of the first and second value-based reward treatments by 4.14 and 4.78 RMB, respectively; a 10% increase in current investment amount increases the effects of the two treatments by 3.67 and 7.02 RMB, respectively; and increasing the value factor by 0.1 increases the effects of the two treatments by 5.77 and 8.54 RMB, respectively. An inspection of the statistical significance of these interactions indicates that, while the interactions between the first reward and the three operationalizations of customer value are not significant at the $p < 0.05$ level, the interactions between the second reward and the three

[33] We obtain the factor scores by conducting a factor analysis on log total investment return and log investment amount, which reveals a single factor accounting for 90.23% of the total variance in the two variables.

operationalizations of customer value are significant (ps < 0.05). We also estimate Model (2.6) after combining the two value-based reward conditions into a single composite value-based reward condition, and report the results in the last three columns of Table 25. We find that the interactions between the composite reward and the three operationalizations of customer value are positive, and two out of three are significant at the $p < 0.05$ level.

To assess the robustness of the findings from the linear models, we estimate log-linear and tobit variants of Model (2.6) and report results in Tables 26, 27, 28, 29, and 30.

Insert Tables 26, 27, 28, 29, and 30 here.

Across all log-linear and tobit models, a consistent pattern of results emerges: The interactions between the first reward and the three operationalizations of customer value are always positive but never significant at the $p < 0.05$ level; on the other hand, the interactions between the second reward and the three operationalizations of customer value are always positive and significant (ps < 0.05). We also find that the interactions between the composite reward and the three operationalizations of customer value are always positive and mostly significant at the $p < 0.05$ level. These findings, together with the one from the linear models, provide strong support for H10 for the second value-based reward treatment, whereas we do not have sufficient statistical evidence to reach a conclusion regarding H10 for the first value-based reward treatment.[34] In addition, when treating the two value-based rewards as a single composite reward, the data also support H10.

To test H11, we estimate linear probability models (i.e., Model (2.6) with Y_i denoting the dummy variable for whether the i-th current customer has acquired referred customers) and probit models. Results are summarized in Tables 31 and 32.

[34]The finding that the interactions between the second reward and operationalizations of customer value have larger magnitudes and smaller p-values than the interactions between the first reward and operationalizations of customer value is expected. This is because while a referred customer's contribution to the reward for the referring customer is capped at 50 RMB under the first reward, her contribution can go beyond 50 RMB as her fixed-deposit investments increase under the second reward; consequently, the second reward could be more motivating for current customers to refer friends, especially those who are good matches. The potential difference in motivation is likely to lead to a larger difference in total investments of referred customers for current customers of higher value, who as homophily implies are more likely to have friends who are good matches that they can refer to the firm.

Insert Tables 31 and 32 here.

We find that, in both the linear probability and probit models, the interactions between the first reward and the three operationalizations of customer value are always positive but never significant at the $p < 0.05$ level, and the interactions between the second reward and the three operationalizations of customer value are always positive and significant (ps < 0.05). We also find that the interactions between the composite reward and the three operationalizations of customer value are always positive and significant (ps < 0.05). Therefore, the data provide strong support for H11 for the second value-based reward treatment, whereas we do not have enough power to draw conclusions regarding H11 for the first value-based reward treatment. Moreover, when treating the two value-based rewards as a single composite reward, the data also support H11.

To test H12, we estimate Model (2.6) with Y_i denoting the number of the i-th current customer's referred customers and NB2 models.[35] Results are summarized in Tables 33 and 34.

Insert Tables 33 and 34 here.

We find that, in both the linear and NB2 models, the interactions between the first reward and the three operationalizations of customer value are always positive but mostly not significant at the $p < 0.05$ level, and the interactions between the second reward and the three operationalizations of customer value are always positive and significant (ps < 0.05). We also find that the interactions between the composite reward and the three operationalizations of customer value are always positive and significant (ps < 0.05). Therefore, the data provide strong support for H12 for the second value-based reward treatment, whereas we do not have enough power to draw conclusions regarding H12 for the first value-based reward treatment. When treating the two value-based rewards as a single composite reward, the data also support H12.

[35] Given that an overdispersion test rejects the null hypothesis of equidispersion in favor of the alternative hypothesis of overdispersion for all specifications of X_i (ps < 0.01), we choose to estimate NB2 models instead of Poisson models.

2.5. Conclusions and Future Research

Enhancing the effectiveness of referral programs is a key challenge facing firms that use referral programs to acquire new customers. In this essay, we propose three treatments aimed at enhancing the effectiveness of referral programs by promoting better matching, including (1) offering current customers a gift before inviting them to refer friends, (2) notifying current customers about the value that they have received from the firm before inviting them to refer friends, and (3) rewarding referring customers based on the value of their referred customers. We test the effectiveness of these three treatments by conducting two field experiments in collaboration with a leading Chinese online financial services firm. We find that all three treatments substantially enhanced the effectiveness of the focal referral program, which is measured for each current customer as the total value of his referred customers. We also find that the enhancement was primarily driven by the acquisition of higher-value new customers rather than the acquisition of more new customers. Moreover, we conduct a series of analyses to explore the mechanisms underlying these treatments. From the analyses, we find evidence suggesting that the second operationalization of the value-based reward treatment facilitated passive matching and had a larger impact on current customers of higher value; on the other hand, we do not have sufficient evidence to draw definitive conclusions regarding the workings of the other treatments.

Our research has several limitations. First, the lack of statistical power in many analyses conducted for the first experiment limits our ability to draw conclusions regarding the workings of the gift and notification treatments. There are a few options for alleviating this power issue that future research may adopt. For instance, researchers could increase the statistical power of a future experiment testing the gift and notification treatments by having a larger sample size in the experiment. Researchers may also consider ways that could potentially make the treatments more effective (i.e., increase their effect sizes), such as making the gift coupon more generous, using gifts that are more valuable than interest-raising coupons from customers' perspective, and making the notification more salient to customers. Moreover, researchers may explore contexts other than online financial services to test the gift and notification treatments. In our data, we find that only a very small proportion of current customers have made successful referrals in the experiments, and a likely explanation is that customers tend to be very cautious when referring friends in our setting since financial investment decisions are private and can be risky. If researchers use a context in which customers are more likely to make referrals, then there could be more power to test the gift and notification treatments.

Second, we have only considered one operationalization of the gift treatment, which is an interest-raising coupon. Other types of gifts may work better, and it is interesting to understand the relationship between the features of a gift and its impact on customer referrals. For instance, is a gift "closer" (i.e., more relevant) to the focal product/service more likely to elicit reciprocity from current customers toward the firm than a gift "more distant" (i.e., less relevant) to the focal product/service, or the other way around? A priori, it can be the case that a "closer" gift is likely to be of higher value to current customers and hence is more likely to induce reciprocity. On the other hand, it is also possible that a "closer" gift is more likely to be perceived by current customers as a deliberate marketing device instead of a genuine kind treatment, in which case reciprocity can hardly be triggered. It will be interesting to test such predictions in future experiments.

Third, we have only considered two value-based rewards with simple structures in this essay. As extensions, we can consider developing value-based rewards based on theoretical models of customer referrals and testing these value-based rewards in diverse contexts. We can also consider other reward structures that are likely to promote better matching. For example, Schmitt et al. (2011) suggest that firms may want to make the referral reward a function of the value of the referring customer. It is interesting to understand whether rewarding a referring customer based on his own value is more or less effective than rewarding him based on his referred customers' value. More broadly, investigating the design of the referral reward structure is likely to generate interesting insights on customer referrals and also prove useful for firms to enhance the effectiveness of their referral programs.

Finally, we are unable to conduct direct tests on the mechanisms underlying the treatments based on our data. To address this issue, researchers may consider augmenting field experimental data on customers' referral behaviors with post-experiment surveys and/or lab experiments in which they could directly measure the mechanisms at work. For example, researchers could test the gift and notification treatments in a lab experiment, and, after the treatments are administered, ask subjects to state their intention to reciprocate to and their value perception of the focal firm. Such information could shed additional light on how the treatments impact customer referrals.

Table 1: Randomization Check: The First Experiment

	Control		Gift		Notification		C vs. G	C vs. N	G vs. N
Total Investment Return (in RMB)	2,443.38	(13.94)	2,439.09	(13.91)	2,450.22	(13.90)	$p = 0.827$	$p = 0.729$	$p = 0.571$
Investment Amount (in RMB)	37,682.66	(330.30)	38,051.21	(319.42)	37,941.05	(318.91)	$p = 0.423$	$p = 0.574$	$p = 0.807$
Tenure (in Days)	556.46	(1.56)	555.10	(1.57)	555.91	(1.56)	$p = 0.538$	$p = 0.804$	$p = 0.712$
Time since Last Investment (in Days)	98.30	(0.64)	97.64	(0.64)	97.07	(0.63)	$p = 0.465$	$p = 0.171$	$p = 0.527$
Had Successful Referrals (Yes/No)	0.1307	(0.0019)	0.1294	(0.0019)	0.1293	(0.0019)	$p = 0.614$	$p = 0.599$	$p = 0.982$
Sample Size	30,977		31,241		31,070				

Notes. The columns "Control", "Gift", and "Notification" report means and standard errors (in parentheses) of pre-experiment variables in the control condition, the gift condition, and the notification condition, respectively. The column "C vs. G" reports p-values of two-sided Welch's t-tests comparing the control condition and the gift condition on pre-experiment variables. The columns "C vs. N" and "G vs. N" report the counterparts in the comparison between the control condition and the notification condition and those in the comparison between the gift condition and the notification condition, respectively.

Table 2: Mean Comparisons: The First Experiment

				Control vs. Gift			Control vs. Notification		
	Control	Gift	Notification	Difference	SE	p-Value	Difference	SE	p-Value
Total Investments of RCs (in RMB)	12.47	39.05	37.71	26.58	9.84	0.007	25.23	11.51	0.028
Has Acquired RCs (Yes/No)	0.0025	0.0032	0.0027	0.0007	0.0004	0.092	0.0002	0.0004	0.591
Number of RCs	0.0027	0.0032	0.0030	0.0006	0.0005	0.230	0.0003	0.0005	0.513
Investments of Each RC (in RMB)	4,655.49	12,078.99	12,597.48	7,423.51	2,841.12	0.010	7,942.00	3,636.92	0.031

Notes. RC denotes Referred Customer. The columns "Control", "Gift", and "Notification" report means of outcome variables in the control condition, the gift condition, and the notification condition, respectively. The columns under "Control vs. Gift" report the comparisons between the control and gift conditions and those under "Control vs. Notification" report the comparisons between the control and notification conditions, all based on two-sided Welch's t-tests. We note that the first three outcome variables are defined on each current customer, with the sample size in the control, gift, and notification conditions being 30,977, 31,241, and 31,070, respectively; on the other hand, the last outcome variable is defined on each referred customer, with the sample size in the control, gift, and notification conditions being 83, 101, and 93, respectively.

Table 3: An Alternative Test Comparing Investments of Each Referred Customer: The First Experiment

	Control	Gift	Notification	C vs. G	C vs. N
Full Sample of All RCs					
Proportion of RCs with Positive Investments	40.96%	45.54%	41.94%	$p = 0.635$	$p = 1.000$
Sample Size	83	101	93		
Subsample of RCs with Positive Investments					
Average Investments (in RMB)	11,364.87	26,521.27	30,040.15	$p = 0.039$	$p = 0.057$
Subsample Size	34	46	39		

Notes. RC denotes Referred Customer. In the section "Full Sample of All RCs", the columns "Control", "Gift", and "Notification" report the proportion of referred customers with a positive investment amount in each condition, and the columns "C vs. G" and "C vs. N" report p-values of the test for equal proportions comparing the control condition and the gift and notification conditions. In the section "Subsample of RCs with Positive Investments", the columns "Control", "Gift", and "Notification" report the average investment amount of referred customers with positive investments in each condition, and the columns "C vs. G" and "C vs. N" report p-values of the Wilcoxon-Mann-Whitney test comparing the control condition and the gift and notification conditions.

Table 4: Regression Analyses on Total Investments of Referred Customers (TIRC): The First Experiment

Dependent Variable	TIRC	log(TIRC+10^0)	log(TIRC+10^1)	log(TIRC+10^2)	log(TIRC+10^3)	TIRC
Model	OLS	OLS	OLS	OLS	OLS	Tobit
Gift Treatment	26.44^{**} (9.84)	4.13 (2.63)	$3.28^{\dagger}$ (1.98)	$2.45^{\dagger}$ (1.35)	1.61^{*} (0.76)	9.72 (6.33)
Notification Treatment	25.05^{*} (11.47)	1.75 (2.50)	1.54 (1.88)	1.32 (1.28)	1.00 (0.73)	5.00 (6.51)
Log Total Investment Return	0.67 (4.69)	−1.92 (1.29)	−1.40 (0.98)	−0.89 (0.67)	−0.42 (0.38)	−4.29 (3.25)
Log Investment Amount	13.42^{*} (5.98)	4.33^{**} (1.13)	3.28^{**} (0.86)	2.23^{**} (0.60)	1.25^{**} (0.35)	10.50^{**} (3.16)
Tenure	−12.30 (18.52)	−5.32 (4.54)	−4.24 (3.43)	−3.10 (2.34)	−1.88 (1.33)	−13.98 (11.95)
Time Since Last Investment	-51.88^{*} (22.41)	-23.46^{**} (6.14)	-17.45^{**} (4.62)	-11.49^{**} (3.13)	-5.95^{**} (1.75)	-154.97^{**} (47.82)
Had Successful Referrals	72.08^{**} (25.87)	24.32^{**} (5.18)	18.30^{**} (3.93)	12.34^{**} (2.70)	6.77^{**} (1.56)	35.76^{**} (6.26)
Intercept	3.23 (4.37)	6.40^{**} (1.63)	$2{,}307.23^{**}$ (1.22)	$4{,}608.10^{**}$ (0.82)	$6{,}909.14^{**}$ (0.44)	-287.47^{**} (24.66)
R^2	5.88×10^{-4}	1.18×10^{-3}	1.18×10^{-3}	1.16×10^{-3}	1.09×10^{-3}	
Log-Likelihood						−2,147.43
Observations	93,288	93,288	93,288	93,288	93,288	93,288

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. TIRC denotes Total Investments of Referred Customers. We report robust standard errors in parentheses for all linear models estimated using OLS, and report (default) standard errors in parentheses for the tobit model (Cameron and Trivedi, 2010). All continuous independent variables have been mean-centered, and tenure and time since last investment are in the unit of 1,000 days. For ease of exposition, the coefficients and standard errors have been multiplied by 1,000 in all linear models with a log-transformed dependent variable, and the coefficients and standard errors have been divided by 1,000 in the tobit model.

Table 5: Regression Analyses on Incidence and Number of Referred Customers: The First Experiment

Dependent Variable	Has Acquired Referred Customers		Number of Referred Customers	
Model	OLS	Probit	OLS	NB2
Gift Treatment	$0.0007^{\dagger}$ (0.0004)	$0.0884^{\dagger}$ (0.0516)	0.0005 (0.0005)	0.2000 (0.1584)
Notification Treatment	0.0002 (0.0004)	0.0348 (0.0532)	0.0003 (0.0005)	0.1317 (0.1610)
Log Total Investment Return	-0.0009^{**} (0.0002)	-0.1035^{**} (0.0261)	-0.0009^{**} (0.0002)	-0.2718^{**} (0.0790)
Log Investment Amount	0.0012^{**} (0.0002)	0.1521^{**} (0.0249)	0.0013^{**} (0.0002)	0.4443^{**} (0.0759)
Tenure	−0.0000 (0.0008)	−0.0379 (0.0949)	−0.0002 (0.0008)	−0.2201 (0.2915)
Time Since Last Investment	-0.0065^{**} (0.0010)	-2.1585^{**} (0.4008)	-0.0072^{**} (0.0011)	-7.6806^{**} (1.3463)
Had Successful Referrals	0.0043^{**} (0.0008)	0.3566^{**} (0.0481)	0.0049^{**} (0.0009)	1.0665^{**} (0.1427)
Intercept	0.0019^{**} (0.0003)	-3.0158^{**} (0.0456)	0.0021^{**} (0.0003)	-6.6304^{**} (0.1454)
R^2	2.04×10^{-3}		1.97×10^{-3}	
Log-Likelihood		−1,673.73		−1,740.12
Observations	93,288	93,288	93,288	93,288

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. We report robust standard errors in parentheses for all linear models estimated using OLS, and report (default) standard errors in parentheses for the probit and NB2 (negative binomial model with a quadratic variance function) models (Cameron and Trivedi, 2010). All continuous independent variables have been mean-centered, and tenure and time since last investment are in the unit of 1,000 days.

Table 6: Regression Analyses on Investments of Each Referred Customer (IERC): The First Experiment

Dependent Variable	IERC	$\log(\text{IERC}+10^0)$	$\log(\text{IERC}+10^1)$	$\log(\text{IERC}+10^2)$	$\log(\text{IERC}+10^3)$	IERC
Model	OLS	OLS	OLS	OLS	OLS	Tobit
Gift Condition	7,423.51** (2,855.72)	0.69 (0.68)	0.58 (0.51)	0.48 (0.35)	0.36† (0.21)	12,582.58 (7,832.04)
Notification Condition	7,942.00* (3,656.83)	0.37 (0.69)	0.35 (0.52)	0.32 (0.36)	0.26 (0.21)	11,360.25 (8,016.69)
Intercept	4,655.49** (1,063.78)	3.55** (0.48)	4.91** (0.36)	6.29** (0.24)	7.75** (0.13)	−22,315.93** (6,243.86)
R^2	1.86×10^{-2}	3.61×10^{-3}	4.48×10^{-3}	6.34×10^{-3}	1.04×10^{-2}	
Log-Likelihood						−1,535.30
Observations	277	277	277	277	277	277

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. IERC denotes Investments of Each Referred Customer. We report robust standard errors in parentheses for all linear models estimated using OLS, and report (default) standard errors in parentheses for the tobit model (Cameron and Trivedi, 2010).

Table 7: Regression Analyses on Average Investments of Referred Customers (AIRC): The First Experiment

Dependent Variable	AIRC	log(AIRC+10^0)	log(AIRC+10^1)	log(AIRC+10^2)	log(AIRC+10^3)	AIRC
Model	OLS	OLS	OLS	OLS	OLS	Tobit
Gift Treatment	7,528.74* (3,050.64)	0.41 (0.70)	0.39 (0.53)	0.36 (0.36)	0.31 (0.21)	10,353.08 (7,825.51)
Notification Treatment	8,896.33* (4,016.39)	0.39 (0.76)	0.38 (0.57)	0.36 (0.39)	0.30 (0.23)	12,333.52 (8,194.57)
Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Characteristics: Interactions	No	No	No	No	No	No
R^2	5.28×10^{-2}	3.52×10^{-2}	3.70×10^{-2}	4.01×10^{-2}	4.58×10^{-2}	
Log-Likelihood						−1,497.80
Observations	258	258	258	258	258	258
Dependent Variable	AIRC	log(AIRC+10^0)	log(AIRC+10^1)	log(AIRC+10^2)	log(AIRC+10^3)	AIRC
Model	OLS	OLS	OLS	OLS	OLS	Tobit
Gift Treatment	6,846.71* (3,018.40)	0.46 (0.74)	0.42 (0.56)	0.37 (0.38)	0.30 (0.22)	9,643.96 (7,820.44)
Notification Treatment	9,323.15* (3,892.36)	0.66 (0.79)	0.59 (0.60)	0.51 (0.41)	0.38 (0.23)	13,757.56† (8,169.17)
Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Characteristics: Interactions	Yes	Yes	Yes	Yes	Yes	Yes
R^2	7.78×10^{-2}	6.50×10^{-2}	6.79×10^{-2}	7.30×10^{-2}	8.00×10^{-2}	
Log-Likelihood						−1,493.97
Observations	258	258	258	258	258	258

Notes. $^\dagger p < .1$. $^* p < .05$. $^{**} p < .01$. We report robust standard errors in parentheses for all linear models estimated using OLS, and report (default) standard errors in parentheses for the tobit models. Characteristics include log total investment return, log investment amount, tenure, time since last investment, and whether had successful referrals prior to the experiment. Characteristics: Interactions include the interaction of each pair of these five characteristics (i.e., all ten interactions).

Table 8: Who Have Acquired Referred Customers in the First Experiment?

	Control	Gift	Notification	C vs. G	C vs. N
Total Investment Return (in RMB)	2,245.22 (258.05)	2,602.51 (263.70)	2,893.65 (295.85)		
Welch's t-test				$p = 0.334$	$p = 0.101$
Wilcoxon-Mann-Whitney test				$p = 0.653$	$p = 0.120$
Investment Amount (in RMB)	65,740.79 (9,939.21)	65,717.55 (7,656.28)	80,795.44 (10,813.78)		
Welch's t-test				$p = 0.999$	$p = 0.307$
Wilcoxon-Mann-Whitney test				$p = 0.728$	$p = 0.252$
Sample Size	76	99	83		

Notes. The columns "Control", "Gift", and "Notification" report means and standard errors (in parentheses) of observable characteristics on the samples of current customers who have acquired referred customers during the experiment (i.e., referrers) in the control condition, the gift condition, and the notification condition, respectively. The columns "C vs. G" and "C vs. N" report p-values of two-sided Welch's t-tests and the Wilcoxon-Mann-Whitney tests comparing the control and gift conditions and p-values of those comparing the control and notification conditions, respectively.

Table 9: Heterogeneous Effects on Total Investments of Referred Customers (TIRC): The First Experiment

Dependent Variable	TIRC			$\log(\text{TIRC}+10^0)$		
Model	OLS	OLS	OLS	OLS	OLS	OLS
Gift	11.05 (6.83)	10.79 (6.86)	11.12 (6.90)	1.73 (2.26)	1.68 (2.29)	1.75 (2.29)
Notification	22.62* (10.95)	22.01* (10.79)	22.51* (10.90)	3.26 (2.42)	3.13 (2.43)	3.21 (2.44)
Log Total Investment Return (Log Return)	−2.39 (2.58)			−1.41 (1.45)		
Log Investment Amount (Log Investment)		3.70† (2.11)			2.08† (1.22)	
Value Factor			0.73 (2.20)			0.37 (1.32)
Time Since Last Investment (Time)	−61.87** (17.67)	−40.84** (15.60)	−56.47** (16.66)	−49.48** (9.84)	−37.59** (9.75)	−46.50** (10.11)
Had Successful Referrals (Had Referrals)	25.26 (16.09)	23.59 (15.62)	24.05 (15.87)	21.70** (7.99)	20.74** (7.91)	21.01** (7.93)
Gift × Log Return	9.38 (8.10)			1.72 (2.30)		
Gift × Log Investment		13.78* (6.45)			2.73 (2.04)	
Gift × Value Factor			13.53† (7.26)			2.59 (2.19)
Gift × Time	−65.36 (47.84)	−10.89 (40.83)	−34.36 (39.87)	16.52 (15.99)	27.13† (15.57)	22.32 (15.77)
Gift × Had Referrals	120.21* (59.29)	121.20* (60.45)	119.41* (59.69)	19.00 (13.52)	19.13 (13.50)	18.81 (13.46)
Notification × Log Return	19.78* (8.43)			2.80 (2.04)		
Notification × Log Investment		17.53 (12.34)			1.47 (1.89)	
Notification × Value Factor			21.93† (11.40)			2.52 (2.02)
Notification × Time	−85.66 (58.94)	−24.77 (47.21)	−39.69 (46.97)	1.03 (14.07)	4.57 (13.44)	5.53 (13.88)
Notification × Had Referrals	18.78 (51.55)	23.43 (51.28)	19.44 (50.47)	−11.52 (10.99)	−10.67 (10.89)	−11.23 (10.91)
Intercept	9.21** (2.85)	9.45** (2.88)	9.37** (2.88)	6.71** (1.51)	6.85** (1.53)	6.81** (1.52)
R^2	6.75×10^{-4}	7.78×10^{-4}	7.48×10^{-4}	1.10×10^{-3}	1.28×10^{-3}	1.16×10^{-3}
Observations	93,288	93,288	93,288	93,288	93,288	93,288

Notes. $^\dagger p < .1$. $^* p < .05$. $^{**} p < .01$. TIRC denotes Total Investments of Referred Customers. We report robust standard errors in parentheses for linear and log-linear models estimated using OLS. All continuous independent variables have been mean-centered, and time since last investment is in the unit of 1,000 days. For ease of exposition, the coefficients and standard errors have been multiplied by 1,000 in log-linear models.

Table 10: Heterogeneous Effects on TIRC Continued: The First Experiment

Dependent Variable	log(TIRC+10^1)			log(TIRC+10^2)		
Model	OLS	OLS	OLS	OLS	OLS	OLS
Gift	1.33 (1.70)	1.29 (1.72)	1.34 (1.72)	0.95 (1.15)	0.92 (1.17)	0.96 (1.16)
Notification	2.57 (1.83)	2.46 (1.84)	2.53 (1.84)	1.87 (1.25)	1.80 (1.26)	1.85 (1.26)
Log Total Investment Return (Log Return)	−1.09 (1.07)			−0.76 (0.71)		
Log Investment Amount (Log Investment)		$1.53^{\dagger}$ (0.90)			$1.00^{\dagger}$ (0.59)	
Value Factor			0.24 (0.97)			0.12 (0.63)
Time Since Last Investment (Time)	-36.43^{**} (7.32)	-27.62^{**} (7.20)	-34.27^{**} (7.49)	-23.54^{**} (4.85)	-17.73^{**} (4.70)	-22.17^{**} (4.92)
Had Successful Referrals (Had Referrals)	15.67^{**} (5.88)	14.94^{*} (5.82)	15.15^{**} (5.84)	9.77^{*} (3.84)	9.27^{*} (3.80)	9.42^{*} (3.81)
Gift × Log Return	1.35 (1.73)			0.97 (1.17)		
Gift × Log Investment		2.18 (1.54)			1.61 (1.05)	
Gift × Value Factor			2.06 (1.65)			1.51 (1.11)
Gift × Time	11.30 (11.95)	$19.81^{\dagger}$ (11.64)	15.94 (11.77)	6.19 (8.03)	12.52 (7.81)	9.61 (7.88)
Gift × Had Referrals	15.44 (10.22)	15.54 (10.21)	15.29 (10.17)	$11.79^{\dagger}$ (7.00)	$11.85^{\dagger}$ (6.99)	$11.67^{\dagger}$ (6.97)
Notification × Log Return	2.23 (1.53)			1.66 (1.03)		
Notification × Log Investment		1.23 (1.43)			0.97 (0.99)	
Notification × Value Factor			2.04 (1.52)			1.55 (1.03)
Notification × Time	−0.34 (10.62)	2.75 (10.03)	3.36 (10.39)	−1.61 (7.22)	0.96 (6.69)	1.25 (6.96)
Notification × Had Referrals	−7.81 (8.25)	−7.15 (8.18)	−7.60 (8.19)	−4.25 (5.58)	−3.76 (5.53)	−4.09 (5.53)
Intercept	$2,307.56^{**}$ (1.13)	$2,307.66^{**}$ (1.14)	$2,307.63^{**}$ (1.14)	$4,608.42^{**}$ (0.75)	$4,608.49^{**}$ (0.76)	$4,608.47^{**}$ (0.76)
R^2	1.10×10^{-3}	1.28×10^{-3}	1.16×10^{-3}	1.09×10^{-3}	1.28×10^{-3}	1.15×10^{-3}
Observations	93,288	93,288	93,288	93,288	93,288	93,288

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. TIRC denotes Total Investments of Referred Customers. We report robust standard errors in parentheses for log-linear models estimated using OLS. All continuous independent variables have been mean-centered, and time since last investment is in the unit of 1,000 days. For ease of exposition, the coefficients and standard errors have been multiplied by 1,000 in log-linear models.

Table 11: Heterogeneous Effects on TIRC Continued: The First Experiment

Dependent Variable	log(TIRC+10^3)						TIRC					
Model	OLS		OLS		OLS		Tobit		Tobit		Tobit	
Gift	0.62	(0.64)	0.60	(0.65)	0.63	(0.65)	27.20*	(12.64)	24.07†	(12.39)	25.58*	(12.39)
Notification	1.18†	(0.71)	1.14	(0.71)	1.17	(0.71)	23.41†	(13.25)	20.87	(12.99)	21.49†	(13.02)
Log Total Investment Return (Log Return)	−0.41	(0.37)					−4.13	(4.03)				
Log Investment Amount (Log Investment)			0.54†	(0.31)					5.70	(4.18)		
Value Factor					0.07	(0.33)					0.54	(4.21)
Time Since Last Investment (Time)	−11.72**	(2.58)	−8.59**	(2.42)	−10.98**	(2.56)	−476.36**	(147.15)	−422.60**	(143.90)	−458.11**	(145.37)
Had Successful Referrals (Had Referrals)	4.61*	(2.02)	4.34*	(1.99)	4.42*	(2.00)	40.11**	(10.83)	36.85**	(10.55)	37.98**	(10.66)
Gift × Log Return	0.55	(0.66)					4.25	(5.33)				
Gift × Log Investment			0.98†	(0.59)					4.08	(5.50)		
Gift × Value Factor					0.89	(0.62)					4.93	(5.56)
Gift × Time	1.83	(4.43)	5.73	(4.26)	3.89	(4.29)	374.06*	(154.82)	374.48*	(153.24)	376.97*	(153.70)
Gift × Had Referrals	7.79†	(4.02)	7.81†	(4.02)	7.71†	(4.00)	7.10	(13.45)	8.33	(13.20)	7.18	(13.30)
Notification × Log Return	1.06†	(0.56)					7.58	(5.60)				
Notification × Log Investment			0.66	(0.58)					3.41	(5.69)		
Notification × Value Factor					1.01†	(0.59)					6.70	(5.82)
Notification × Time	−2.33	(4.05)	−0.48	(3.60)	−0.42	(3.77)	210.47	(168.77)	205.79	(167.09)	216.98	(167.50)
Notification × Had Referrals	−1.39	(3.13)	−1.09	(3.11)	−1.30	(3.10)	−23.38	(14.98)	−20.56	(14.74)	−22.16	(14.84)
Intercept	6,909.41**	(0.40)	6,909.45**	(0.41)	6,909.44**	(0.41)	−303.84**	(27.84)	−300.22**	(27.40)	−301.17**	(27.52)
R^2	1.05×10^{-3}		1.24×10^{-3}		1.12×10^{-3}							
Log-Likelihood							−2,151.61		−2,144.73		−2,149.67	
Observations	93,288		93,288		93,288		93,288		93,288		93,288	

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. TIRC denotes Total Investments of Referred Customers. We report robust standard errors in parentheses for log-linear models estimated using OLS, and report (default) standard errors in parentheses for tobit models. All continuous independent variables have been mean-centered, and time since last investment is in the unit of 1,000 days. For ease of exposition, the coefficients and standard errors have been multiplied by 1,000 in log-linear models, and the coefficients and standard errors have been divided by 1,000 in tobit models.

Table 12: Heterogeneous Effects on Incidence of Having Acquired Referred Customers: The First Experiment

Dependent Variable	Has Acquired Referred Customers											
Model	OLS		OLS		OLS		Probit		Probit		Probit	
Gift	0.0006	(0.0004)	0.0006	(0.0004)	0.0006	(0.0004)	0.2196**	(0.0852)	0.2052*	(0.0844)	0.2121*	(0.0838)
Notification	0.0007†	(0.0004)	0.0007	(0.0004)	0.0007†	(0.0004)	0.1790*	(0.0893)	0.1464	(0.0893)	0.1572†	(0.0885)
Log Total Investment Return (Log Return)	−0.0004†	(0.0002)					−0.0627*	(0.0320)				
Log Investment Amount (Log Investment)			0.0005*	(0.0002)					0.0702*	(0.0332)		
Value Factor					0.0000	(0.0002)					−0.0004	(0.0334)
Time Since Last Investment (Time)	−0.0120**	(0.0016)	−0.0089**	(0.0017)	−0.0113**	(0.0017)	−4.3556**	(0.9108)	−3.8111**	(0.9118)	−4.2219**	(0.9103)
Had Successful Referrals (Had Referrals)	0.0053**	(0.0013)	0.0050**	(0.0013)	0.0051**	(0.0013)	0.4552**	(0.0843)	0.4152**	(0.0832)	0.4286**	(0.0835)
Gift × Log Return	0.0001	(0.0004)					0.0262	(0.0427)				
Gift × Log Investment			0.0003	(0.0003)					0.0163	(0.0442)		
Gift × Value Factor					0.0002	(0.0004)					0.0254	(0.0446)
Gift × Time	0.0008	(0.0027)	0.0018	(0.0027)	0.0013	(0.0027)	2.2070*	(1.0585)	2.2052*	(1.0619)	2.2367*	(1.0587)
Gift × Had Referrals	0.0005	(0.0020)	0.0005	(0.0020)	0.0005	(0.0020)	−0.0432	(0.1136)	−0.0331	(0.1124)	−0.0408	(0.1127)
Notification × Log Return	0.0006†	(0.0003)					0.0849†	(0.0446)				
Notification × Log Investment			0.0004	(0.0003)					0.0539	(0.0464)		
Notification × Value Factor					0.0006†	(0.0003)					0.0834†	(0.0468)
Notification × Time	0.0006	(0.0024)	0.0019	(0.0024)	0.0018	(0.0025)	1.3032	(1.1485)	1.4518	(1.1502)	1.4690	(1.1477)
Notification × Had Referrals	−0.0036*	(0.0017)	−0.0035*	(0.0017)	−0.0036*	(0.0017)	−0.2966*	(0.1245)	−0.2681*	(0.1240)	−0.2851*	(0.1239)
Intercept	0.0018**	(0.0003)	0.0018**	(0.0003)	0.0018**	(0.0003)	−3.0968**	(0.0715)	−3.0885**	(0.0704)	−3.0828**	(0.0703)
R^2	1.62×10^{-3}		1.90×10^{-3}		1.65×10^{-3}							
Log-Likelihood							−1,693.40		−1,681.50		−1,692.66	
Observations	93,288		93,288		93,288		93,288		93,288		93,288	

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. We report robust standard errors in parentheses for linear probability models estimated using OLS, and report (default) standard errors in parentheses for probit models. All continuous independent variables have been mean-centered, and time since last investment is in the unit of 1,000 days.

Table 13: Heterogeneous Effects on Number of Referred Customers: The First Experiment

Dependent Variable	Number of Referred Customers					
Model	OLS	OLS	OLS	NB2	NB2	NB2
Gift	0.0006 (0.0004)	0.0006 (0.0004)	0.0006 (0.0004)	0.7520** (0.2883)	0.6937* (0.2837)	0.7155* (0.2835)
Notification	0.0008† (0.0004)	0.0008† (0.0004)	0.0008† (0.0004)	0.6334* (0.3033)	0.5025† (0.3015)	0.5495† (0.3002)
Log Total Investment Return (Log Return)	−0.0005† (0.0003)			−0.1956* (0.0983)		
Log Investment Amount (Log Investment)		0.0006** (0.0002)			0.2324* (0.1028)	
Value Factor			0.0001 (0.0003)			0.0033 (0.1033)
Time Since Last Investment (Time)	−0.0133** (0.0019)	−0.0096** (0.0018)	−0.0124** (0.0020)	−14.9286** (3.0827)	−12.9469** (3.0524)	−14.4437** (3.0719)
Had Successful Referrals (Had Referrals)	0.0064** (0.0017)	0.0061** (0.0017)	0.0062** (0.0017)	1.4381** (0.2517)	1.3012** (0.2462)	1.3529** (0.2486)
Gift × Log Return	0.0002 (0.0004)			0.1060 (0.1323)		
Gift × Log Investment		0.0001 (0.0004)			0.0132 (0.1376)	
Gift × Value Factor			0.0002 (0.0004)			0.0754 (0.1387)
Gift × Time	0.0018 (0.0029)	0.0020 (0.0029)	0.0021 (0.0029)	7.8816* (3.5726)	7.3985* (3.5561)	7.8486* (3.5672)
Gift × Had Referrals	−0.0007 (0.0022)	−0.0007 (0.0022)	−0.0007 (0.0022)	−0.2891 (0.3440)	−0.2374 (0.3374)	−0.2706 (0.3407)
Notification × Log Return	0.0008† (0.0004)			0.2872* (0.1368)		
Notification × Log Investment		0.0005 (0.0004)			0.1743 (0.1428)	
Notification × Value Factor			0.0007† (0.0004)			0.2813† (0.1444)
Notification × Time	0.0003 (0.0028)	0.0016 (0.0026)	0.0017 (0.0028)	4.3625 (3.8637)	4.6634 (3.8408)	4.9276 (3.8516)
Notification × Had Referrals	−0.0041† (0.0023)	−0.0039† (0.0022)	−0.0040† (0.0023)	−0.8781* (0.3660)	−0.7787* (0.3620)	−0.8402* (0.3638)
Intercept	0.0018** (0.0003)	0.0019** (0.0003)	0.0019** (0.0003)	−6.9557** (0.2452)	−6.9201** (0.2394)	−6.9055** (0.2404)
R^2	1.55×10^{-3}	1.85×10^{-3}	1.61×10^{-3}			
Log-Likelihood				−1,758.75	−1,746.16	−1,757.16
Observations	93,288	93,288	93,288	93,288	93,288	93,288

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. We report robust standard errors in parentheses for linear models estimated using OLS, and report (default) standard errors in parentheses for NB2 models. All continuous independent variables have been mean-centered, and time since last investment is in the unit of 1,000 days.

Table 14: Randomization Check: The Second Experiment

	Control	First Reward	Second Reward	C vs. F	C vs. S	F vs. S
Total Investment Return (in RMB)	9,239.12 (125.94)	9,412.78 (131.48)	9,425.45 (130.09)	$p = 0.340$	$p = 0.303$	$p = 0.945$
Investment Amount (in RMB)	72,213.53 (913.36)	72,492.67 (875.08)	73,013.00 (861.17)	$p = 0.825$	$p = 0.524$	$p = 0.672$
Tenure (in Days)	601.56 (1.79)	601.62 (1.80)	604.53 (1.80)	$p = 0.979$	$p = 0.242$	$p = 0.253$
Time since Last Investment (in Days)	92.38 (0.58)	91.83 (0.57)	92.34 (0.58)	$p = 0.496$	$p = 0.962$	$p = 0.527$
Had Successful Referrals (Yes/No)	0.1110 (0.0016)	0.1124 (0.0016)	0.1119 (0.0016)	$p = 0.535$	$p = 0.683$	$p = 0.833$
Sample Size	40,076	40,145	40,037			

Notes. The columns "Control", "First Reward", and "Second Reward" report means and standard errors (in parentheses) of pre-experiment variables in the control condition, the first value-based reward condition, and the second value-based reward condition, respectively. The column "C vs. F" reports p-values of two-sided Welch's t-tests comparing the control condition and the first value-based reward condition on pre-experiment variables. The columns "C vs. S" and "F vs. S" report the counterparts in the comparison between the control condition and the second value-based reward condition and those in the comparison between the first and second value-based reward conditions, respectively.

Table 15: Mean Comparisons: The Second Experiment

				Control vs. First Reward			Control vs. Second Reward		
	Control	FR	SR	Difference	SE	p-Value	Difference	SE	p-Value
Total Investments of RCs (in RMB)	92.97	220.81	192.90	127.84	46.99	0.007	99.93	35.10	0.004
Has Acquired RCs (Yes/No)	0.0060	0.0065	0.0071	0.0005	0.0006	0.405	0.0011	0.0006	0.053
Number of RCs	0.0066	0.0075	0.0082	0.0009	0.0007	0.172	0.0016	0.0007	0.016
Investments of Each RC (in RMB)	14,166.89	29,548.51	23,546.02	15,381.62	6,129.20	0.012	9,379.12	4,313.32	0.030

Notes. RC denotes Referred Customer. The columns "Control", "FR", and "SR" report means of outcome variables in the control condition, the first value-based reward condition, and the second value-based reward condition, respectively. The columns under "Control vs. First Reward" report the comparisons between the control and first value-based reward conditions and those under "Control vs. Second Reward" report the comparisons between the control and second value-based reward conditions, all based on two-sided Welch's t-tests. We note that the first three outcome variables are defined on each current customer, with the sample size in the control, first value-based reward, and second value-based reward conditions being 40,076, 40,145, and 40,037, respectively; on the other hand, the last outcome variable is defined on each referred customer, with the sample size in the control, first value-based reward, and second value-based reward conditions being 263, 300, and 328, respectively.

Table 16: An Alternative Test Comparing Investments of Each Referred Customer: The Second Experiment

	Control	First Reward	Second Reward	C vs. F	C vs. S
Full Sample of All RCs					
Proportion of RCs with Positive Investments	37.26%	40.67%	40.85%	$p = 0.460$	$p = 0.422$
Sample Size	263	300	328		
Subsample of RCs with Positive Investments					
Average Investments (in RMB)	38,019.32	72,660.28	57,635.03	$p = 0.052$	$p = 0.028$
Subsample Size	98	122	134		

Notes. RC denotes Referred Customer. In the section "Full Sample of All RCs", the columns "Control", "First Reward", and "Second Reward" report the proportion of referred customers with a positive investment amount in each condition, and the columns "C vs. F" and "C vs. S" report p-values of the test for equal proportions comparing the control condition and the two value-based reward conditions. In the section "Subsample of RCs with Positive Investments", the columns "Control", "First Reward", and "Second Reward" report the average investment amount of referred customers with positive investments in each condition, and the columns "C vs. F" and "C vs. S" report p-values of the Wilcoxon-Mann-Whitney test comparing the control condition and the two value-based reward conditions.

Table 17: Regression Analyses on Total Investments of Referred Customers (TIRC): The Second Experiment

Dependent Variable	TIRC	$\log(\text{TIRC}+10^0)$	$\log(\text{TIRC}+10^1)$	$\log(\text{TIRC}+10^2)$	$\log(\text{TIRC}+10^3)$	TIRC
Model	OLS	OLS	OLS	OLS	OLS	Tobit
First Reward	127.26^{**} (46.95)	7.31^{*} (3.57)	5.90^{*} (2.75)	4.46^{*} (1.95)	2.94^{*} (1.21)	21.26^{*} (10.15)
Second Reward	98.42^{**} (35.05)	9.77^{**} (3.66)	7.73^{**} (2.82)	5.72^{**} (2.00)	3.76^{**} (1.24)	24.95^{*} (10.08)
Log Total Investment Return	$55.53^{\dagger}$ (30.90)	0.58 (1.63)	0.70 (1.26)	0.78 (0.90)	0.73 (0.57)	0.49 (5.57)
Log Investment Amount	48.71 (33.87)	8.29^{**} (1.71)	6.39^{**} (1.32)	4.52^{**} (0.95)	2.75^{**} (0.60)	22.16^{**} (5.84)
Tenure	$-216.78^{\dagger}$ (118.08)	-18.46^{**} (6.37)	-14.74^{**} (4.96)	-10.91^{**} (3.58)	-6.95^{**} (2.29)	-48.96^{**} (16.96)
Time Since Last Investment	-128.82 (96.14)	-39.83^{**} (9.38)	-29.39^{**} (7.18)	-19.12^{**} (5.03)	-9.89^{**} (3.02)	-288.80^{**} (65.72)
Had Successful Referrals	-73.00 (45.50)	24.64^{**} (6.55)	17.94^{**} (5.00)	11.44^{**} (3.49)	5.59^{**} (2.08)	47.60^{**} (10.72)
Intercept	101.82^{**} (20.56)	19.31^{**} (2.34)	$2{,}317.23^{**}$ (1.79)	$4{,}615.22^{**}$ (1.26)	$6{,}913.51^{**}$ (0.77)	-654.27^{**} (32.44)
R^2	7.74×10^{-4}	1.43×10^{-3}	1.44×10^{-3}	1.45×10^{-3}	1.43×10^{-3}	
Log-Likelihood						$-6{,}402.59$
Observations	120,258	120,258	120,258	120,258	120,258	120,258

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. TIRC denotes Total Investments of Referred Customers. We report robust standard errors in parentheses for all linear models estimated using OLS, and report (default) standard errors in parentheses for the tobit model. All continuous independent variables have been mean-centered, and tenure and time since last investment are in the unit of 1,000 days. For ease of exposition, the coefficients and standard errors have been multiplied by 1,000 in all linear models with a log-transformed dependent variable, and the coefficients and standard errors have been divided by 1,000 in the tobit model.

Table 18: Regression Analyses on Incidence and Number of Referred Customers: The Second Experiment

Dependent Variable	Has Acquired Referred Customers		Number of Referred Customers	
Model	OLS	Probit	OLS	NB2
First Reward	0.0004 (0.0006)	0.0323 (0.0325)	0.0009 (0.0007)	0.1500 (0.0940)
Second Reward	$0.0011^{\dagger}$ (0.0006)	$0.0623^{\dagger}$ (0.0320)	0.0016^{*} (0.0007)	0.2214^{*} (0.0928)
Log Total Investment Return	-0.0007^{**} (0.0002)	-0.0517^{**} (0.0181)	-0.0009^{**} (0.0003)	-0.1652^{**} (0.0528)
Log Investment Amount	0.0021^{**} (0.0003)	0.1267^{**} (0.0190)	0.0025^{**} (0.0003)	0.3797^{**} (0.0555)
Tenure	-0.0028^{**} (0.0009)	-0.1611^{**} (0.0547)	-0.0032^{**} (0.0012)	-0.4324^{**} (0.1576)
Time Since Last Investment	-0.0108^{**} (0.0015)	-1.4915^{**} (0.2136)	-0.0126^{**} (0.0017)	-4.7587^{**} (0.6562)
Had Successful Referrals	0.0093^{**} (0.0011)	0.3678^{**} (0.0331)	0.0107^{**} (0.0013)	0.9778^{**} (0.0948)
Intercept	0.0050^{**} (0.0004)	-2.6531^{**} (0.0260)	0.0054^{**} (0.0004)	-5.4779^{**} (0.0774)
R^2	3.29×10^{-3}		3.06×10^{-3}	
Log-Likelihood		−4,530.65		−4,855.71
Observations	120,258	120,258	120,258	120,258

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. We report robust standard errors in parentheses for all linear models estimated using OLS, and report (default) standard errors in parentheses for the probit and NB2 models. All continuous independent variables have been mean-centered, and tenure and time since last investment are in the unit of 1,000 days.

Table 19: Regression Analyses on Investments of Each Referred Customer (IERC): The Second Experiment

Dependent Variable	IERC	log(IERC+10^0)	log(IERC+10^1)	log(IERC+10^2)	log(IERC+10^3)	IERC
Model	OLS	OLS	OLS	OLS	OLS	Tobit
First Reward	15,381.62* (6,139.72)	0.54 (0.41)	0.46 (0.31)	0.37 (0.23)	0.27† (0.14)	27,545.96* (12,987.92)
Second Reward	9,379.12* (4,320.55)	0.53 (0.40)	0.45 (0.31)	0.36 (0.22)	0.27† (0.14)	19,690.42 (12,767.43)
Intercept	14,166.89** (2,687.99)	3.47** (0.29)	4.91** (0.22)	6.37** (0.16)	7.90** (0.10)	−69,303.10** (10,298.73)
R^2	7.47×10^{-3}	2.50×10^{-3}	2.97×10^{-3}	3.75×10^{-3}	4.90×10^{-3}	
Log-Likelihood						−4,955.25
Observations	891	891	891	891	891	891

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. IERC denotes Investments of Each Referred Customer. We report robust standard errors in parentheses for all linear models estimated using OLS, and report (default) standard errors in parentheses for the tobit model.

Table 20: Regression Analyses on IERC Continued: The Second Experiment

Dependent Variable	IERC	$\log(\text{IERC}+10^0)$	$\log(\text{IERC}+10^1)$	$\log(\text{IERC}+10^2)$	$\log(\text{IERC}+10^3)$	IERC
Model	OLS	OLS	OLS	OLS	OLS	Tobit
Value-Based Reward	12,246.56** (4,157.06)	0.53 (0.35)	0.45† (0.27)	0.37† (0.19)	0.27* (0.12)	23,458.84* (11,379.43)
Intercept	14,166.89** (2,687.99)	3.47** (0.29)	4.91** (0.22)	6.37** (0.16)	7.90** (0.10)	−69,366.62** (10,306.08)
R^2	6.21×10^{-3}	2.50×10^{-3}	2.97×10^{-3}	3.75×10^{-3}	4.90×10^{-3}	
Log-Likelihood						−4,955.46
Observations	891	891	891	891	891	891

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. IERC denotes Investments of Each Referred Customer. We report robust standard errors in parentheses for all linear models estimated using OLS, and report (default) standard errors in parentheses for the tobit model.

Table 21: Regression Analyses on Average Investments of Referred Customers (AIRC): The Second Experiment

Dependent Variable	AIRC	$\log(\text{AIRC}+10^0)$	$\log(\text{AIRC}+10^1)$	$\log(\text{AIRC}+10^2)$	$\log(\text{AIRC}+10^3)$	AIRC
Model	OLS	OLS	OLS	OLS	OLS	Tobit
First Reward	12,239.28* (6,204.93)	0.58 (0.44)	0.47 (0.34)	0.36 (0.24)	0.23 (0.15)	24,594.68† (12,612.29)
Second Reward	2,053.12 (4,563.88)	0.41 (0.42)	0.32 (0.33)	0.23 (0.23)	0.16 (0.15)	8,814.09 (12,511.24)
Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Characteristics: Interactions	No	No	No	No	No	No
R^2	8.09×10^{-2}	4.59×10^{-2}	5.40×10^{-2}	6.68×10^{-2}	8.57×10^{-2}	
Log-Likelihood						−4,737.91
Observations	786	786	786	786	786	786
Dependent Variable	AIRC	$\log(\text{AIRC}+10^0)$	$\log(\text{AIRC}+10^1)$	$\log(\text{AIRC}+10^2)$	$\log(\text{AIRC}+10^3)$	AIRC
Model	OLS	OLS	OLS	OLS	OLS	Tobit
First Reward	12,451.33† (6,369.94)	0.56 (0.44)	0.46 (0.34)	0.35 (0.24)	0.23 (0.15)	25,116.83* (12,320.97)
Second Reward	791.10 (4,571.81)	0.31 (0.43)	0.24 (0.33)	0.18 (0.23)	0.12 (0.15)	5,795.87 (12,188.45)
Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Characteristics: Interactions	Yes	Yes	Yes	Yes	Yes	Yes
R^2	1.27×10^{-1}	7.34×10^{-2}	8.18×10^{-2}	9.59×10^{-2}	1.18×10^{-1}	
Log-Likelihood						−4,720.98
Observations	786	786	786	786	786	786

Notes. $^\dagger p < .1$. $^* p < .05$. $^{**} p < .01$. We report robust standard errors in parentheses for all linear models estimated using OLS, and report (default) standard errors in parentheses for the tobit models. Characteristics include log total investment return, log investment amount, tenure, time since last investment, and whether had successful referrals prior to the experiment. Characteristics: Interactions include the interaction of each pair of these five characteristics (i.e., all ten interactions).

Table 22: Regression Analyses on AIRC Continued: The Second Experiment

Dependent Variable	AIRC	log(AIRC+10^0)	log(AIRC+10^1)	log(AIRC+10^2)	log(AIRC+10^3)	AIRC
Model	OLS	OLS	OLS	OLS	OLS	Tobit
Value-Based Reward	6,962.55† (4,098.51)	0.49 (0.37)	0.39 (0.29)	0.29 (0.20)	0.19 (0.13)	16,536.41 (11,114.01)
Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Characteristics: Interactions	No	No	No	No	No	No
R^2	7.76×10^{-2}	4.57×10^{-2}	5.37×10^{-2}	6.65×10^{-2}	8.54×10^{-2}	
Log-Likelihood						−4,738.80
Observations	786	786	786	786	786	786
Dependent Variable	AIRC	log(AIRC+10^0)	log(AIRC+10^1)	log(AIRC+10^2)	log(AIRC+10^3)	AIRC
Model	OLS	OLS	OLS	OLS	OLS	Tobit
Value-Based Reward	6,371.57 (4,125.65)	0.43 (0.38)	0.35 (0.29)	0.26 (0.21)	0.17 (0.13)	15,174.15 (10,847.45)
Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Characteristics: Interactions	Yes	Yes	Yes	Yes	Yes	Yes
R^2	1.23×10^{-1}	7.29×10^{-2}	8.13×10^{-2}	9.53×10^{-2}	1.18×10^{-1}	
Log-Likelihood						−4,722.38
Observations	786	786	786	786	786	786

Notes. $^\dagger p < .1$. $^* p < .05$. $^{**} p < .01$. We report robust standard errors in parentheses for all linear models estimated using OLS, and report (default) standard errors in parentheses for the tobit models. Characteristics include log total investment return, log investment amount, tenure, time since last investment, and whether had successful referrals prior to the experiment. Characteristics: Interactions include the interaction of each pair of these five characteristics (i.e., all ten interactions).

Table 23: Who Have Acquired Referred Customers in the Second Experiment?

	Control	First Reward	Second Reward	C vs. F	C vs. S
Total Investment Return (in RMB)	8,619.57 (1,066.68)	9,230.89 (1,194.46)	14,454.24 (2,122.08)		
Welch's t-test				$p = 0.703$	$p = 0.014$
Wilcoxon-Mann-Whitney test				$p = 0.055$	$p = 0.001$
Investment Amount (in RMB)	94,739.58 (9,522.25)	104,103.73 (12,474.38)	133,068.36 (11,272.54)		
Welch's t-test				$p = 0.551$	$p = 0.010$
Wilcoxon-Mann-Whitney test				$p = 0.456$	$p = 0.001$
Sample Size	241	260	285		

Notes. The columns "Control", "First Reward", and "Second Reward" report means and standard errors (in parentheses) of observable characteristics on the samples of current customers who have acquired referred customers during the experiment (i.e., referrers) in the control condition, the first value-based reward condition, and the second value-based reward condition, respectively. The columns "C vs. F" and "C vs. S" report p-values of two-sided Welch's t-tests and the Wilcoxon-Mann-Whitney tests comparing the control and first value-based reward conditions and p-values of those comparing the control and second value-based reward conditions, respectively.

Table 24: Who Have Acquired Referred Customers in the Second Experiment (Continued)?

	Control		Value-Based Reward		C vs. V
Total Investment Return (in RMB)	8,619.57	(1,066.68)	11,962.37	(1,251.38)	
Welch's t-test					$p = 0.042$
Wilcoxon-Mann-Whitney test					$p = 0.003$
Investment Amount (in RMB)	94,739.58	(9,522.25)	119,250.37	(8,391.62)	
Welch's t-test					$p = 0.054$
Wilcoxon-Mann-Whitney test					$p = 0.019$
Sample Size	241		545		

Notes. The columns "Control" and "Value-Based Reward" report means and standard errors (in parentheses) of observable characteristics on the samples of current customers who have acquired referred customers during the experiment (i.e., referrers) in the control and composite value-based reward conditions, respectively. The column "C vs. V" reports p-values of two-sided Welch's t-tests and the Wilcoxon-Mann-Whitney tests comparing the control and composite value-based reward conditions.

Table 25: Heterogeneous Effects on Total Investments of Referred Customers (TIRC): The Second Experiment

Dependent Variable	TIRC					
Model	OLS	OLS	OLS	OLS	OLS	OLS
First Reward	127.71** (47.07)	127.70** (47.09)	127.69** (47.09)			
Second Reward	98.52** (34.99)	98.02** (34.95)	98.10** (34.96)			
Composite Reward				113.16** (32.23)	113.01** (32.24)	112.99** (32.24)
Log Total Investment Return (Log Return)	35.12** (11.15)			35.12** (11.15)		
Log Investment Amount (Log Investment)		56.43** (16.89)			56.43** (16.89)	
Value Factor			66.27** (20.24)			66.27** (20.24)
First Reward × Log Return	41.36 (27.01)					
First Reward × Log Investment		36.71 (39.50)				
First Reward × Value Factor			57.66 (47.23)			
Second Reward × Log Return	47.83* (21.03)					
Second Reward × Log Investment		70.16* (29.08)				
Second Reward × Value Factor			85.39* (35.93)			
Composite Reward × Log Return				44.56* (18.85)		
Composite Reward × Log Investment					$53.35^{\dagger}$ (27.29)	
Composite Reward × Value Factor						71.44* (32.96)
Intercept	93.32** (18.94)	93.43** (18.97)	93.42** (18.97)	93.32** (18.94)	93.43** (18.97)	93.42** (18.97)
R^2	5.39×10^{-4}	7.46×10^{-4}	6.97×10^{-4}	5.35×10^{-4}	7.29×10^{-4}	6.87×10^{-4}
Observations	120,258	120,258	120,258	120,258	120,258	120,258

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. TIRC denotes Total Investments of Referred Customers. We report robust standard errors in parentheses for linear models estimated using OLS. All continuous independent variables have been mean-centered.

Table 26: Heterogeneous Effects on TIRC Continued: The Second Experiment

Dependent Variable	log(TIRC+10^0)					
Model	OLS	OLS	OLS	OLS	OLS	OLS
First Reward	7.40* (3.58)	7.39* (3.58)	7.39* (3.58)			
Second Reward	9.78** (3.65)	9.70** (3.65)	9.72** (3.65)			
Composite Reward				8.60** (3.03)	8.57** (3.03)	8.58** (3.03)
Log Total Investment Return (Log Return)	2.32* (1.09)			2.32* (1.09)		
Log Investment Amount (Log Investment)		6.15** (1.40)			6.15** (1.40)	
Value Factor			6.02** (1.75)			6.02** (1.75)
First Reward × Log Return	2.63 (1.61)					
First Reward × Log Investment		1.65 (2.10)				
First Reward × Value Factor			3.20 (2.59)			
Second Reward × Log Return	6.43** (1.82)					
Second Reward × Log Investment		8.33** (2.34)				
Second Reward × Value Factor			10.74** (2.93)			
Composite Reward × Log Return				4.53** (1.44)		
Composite Reward × Log Investment					4.98** (1.86)	
Composite Reward × Value Factor						6.96** (2.32)
Intercept	22.00** (2.31)	22.03** (2.31)	22.02** (2.31)	22.00** (2.31)	22.03** (2.31)	22.02** (2.31)
R^2	5.89×10^{-4}	1.17×10^{-3}	9.32×10^{-4}	5.50×10^{-4}	1.08×10^{-3}	8.67×10^{-4}
Observations	120,258	120,258	120,258	120,258	120,258	120,258

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. TIRC denotes Total Investments of Referred Customers. We report robust standard errors in parentheses for log-linear models estimated using OLS. All continuous independent variables have been mean-centered. For ease of exposition, the coefficients and standard errors have been multiplied by 1,000 in log-linear models.

Table 27: Heterogeneous Effects on TIRC Continued: The Second Experiment

Dependent Variable	log(TIRC+10^1)					
Model	OLS	OLS	OLS	OLS	OLS	OLS
First Reward	5.97^{*} (2.76)	5.96^{*} (2.76)	5.96^{*} (2.76)			
Second Reward	7.74^{**} (2.82)	7.68^{**} (2.82)	7.69^{**} (2.82)			
Composite Reward				6.87^{**} (2.33)	6.84^{**} (2.33)	6.84^{**} (2.33)
Log Total Investment Return (Log Return)	1.94^{*} (0.84)			1.94^{*} (0.84)		
Log Investment Amount (Log Investment)		4.85^{**} (1.09)			4.85^{**} (1.09)	
Value Factor			4.83^{**} (1.36)			4.83^{**} (1.36)
First Reward × Log Return	$2.06^{\dagger}$ (1.25)					
First Reward × Log Investment		1.32 (1.64)				
First Reward × Value Factor			2.52 (2.02)			
Second Reward × Log Return	4.97^{**} (1.42)					
Second Reward × Log Investment		6.54^{**} (1.83)				
Second Reward × Value Factor			8.37^{**} (2.29)			
Composite Reward × Log Return				3.51^{**} (1.12)		
Composite Reward × Log Investment					3.92^{**} (1.45)	
Composite Reward × Value Factor						5.44^{**} (1.81)
Intercept	$2{,}319.19^{**}$ (1.77)	$2{,}319.21^{**}$ (1.77)	$2{,}319.21^{**}$ (1.77)	$2{,}319.19^{**}$ (1.77)	$2{,}319.21^{**}$ (1.77)	$2{,}319.21^{**}$ (1.77)
R^2	6.24×10^{-4}	1.21×10^{-3}	9.76×10^{-4}	5.86×10^{-4}	1.13×10^{-3}	9.11×10^{-4}
Observations	120,258	120,258	120,258	120,258	120,258	120,258

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. TIRC denotes Total Investments of Referred Customers. We report robust standard errors in parentheses for log-linear models estimated using OLS. All continuous independent variables have been mean-centered. For ease of exposition, the coefficients and standard errors have been multiplied by 1,000 in log-linear models.

Table 28: Heterogeneous Effects on TIRC Continued: The Second Experiment

Dependent Variable	log(TIRC+10^2)					
Model	OLS	OLS	OLS	OLS	OLS	OLS
First Reward	4.50^{*} (1.96)	4.50^{*} (1.96)	4.50^{*} (1.96)			
Second Reward	5.73^{**} (2.00)	5.68^{**} (2.00)	5.69^{**} (2.00)			
Composite Reward				5.12^{**} (1.65)	5.10^{**} (1.65)	5.11^{**} (1.65)
Log Total Investment Return (Log Return)	1.54^{*} (0.60)			1.54^{*} (0.60)		
Log Investment Amount (Log Investment)		3.54^{**} (0.78)			3.54^{**} (0.78)	
Value Factor			3.63^{**} (0.98)			3.63^{**} (0.98)
First Reward × Log Return	$1.49^{\dagger}$ (0.89)					
First Reward × Log Investment		0.99 (1.19)				
First Reward × Value Factor			1.85 (1.47)			
Second Reward × Log Return	3.52^{**} (1.03)					
Second Reward × Log Investment		4.75^{**} (1.34)				
Second Reward × Value Factor			6.01^{**} (1.67)			
Composite Reward × Log Return				2.50^{**} (0.80)		
Composite Reward × Log Investment					2.87^{**} (1.05)	
Composite Reward × Value Factor						3.93^{**} (1.31)
Intercept	$4{,}616.47^{**}$ (1.24)	$4{,}616.48^{**}$ (1.24)	$4{,}616.47^{**}$ (1.24)	$4{,}616.47^{**}$ (1.24)	$4,616.48^{**}$ (1.24)	$4{,}616.47^{**}$ (1.24)
R^2	6.77×10^{-4}	1.27×10^{-3}	1.04×10^{-3}	6.41×10^{-4}	1.19×10^{-3}	9.77×10^{-4}
Observations	120,258	120,258	120,258	120,258	120,258	120,258

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. TIRC denotes Total Investments of Referred Customers. We report robust standard errors in parentheses for log-linear models estimated using OLS. All continuous independent variables have been mean-centered. For ease of exposition, the coefficients and standard errors have been multiplied by 1,000 in log-linear models.

Table 29: Heterogeneous Effects on TIRC Continued: The Second Experiment

Dependent Variable	log(TIRC+10^3)					
Model	OLS	OLS	OLS	OLS	OLS	OLS
First Reward	2.96^{*} (1.21)	2.96^{*} (1.21)	2.96^{*} (1.21)			
Second Reward	3.77^{**} (1.24)	3.74^{**} (1.24)	3.75^{**} (1.24)			
Composite Reward				3.37^{**} (1.01)	3.36^{**} (1.01)	3.36^{**} (1.01)
Log Total Investment Return (Log Return)	1.10^{**} (0.37)			1.10^{**} (0.37)		
Log Investment Amount (Log Investment)		2.25^{**} (0.50)			2.25^{**} (0.50)	
Value Factor			2.40^{**} (0.62)			2.40^{**} (0.62)
First Reward × Log Return	$0.94^{\dagger}$ (0.57)					
First Reward × Log Investment		0.67 (0.77)				
First Reward × Value Factor			1.20 (0.95)			
Second Reward × Log Return	2.13^{**} (0.66)					
Second Reward × Log Investment		3.01^{**} (0.86)				
Second Reward × Value Factor			3.73^{**} (1.08)			
Composite Reward × Log Return				1.54^{**} (0.51)		
Composite Reward × Log Investment					1.84^{**} (0.68)	
Composite Reward × Value Factor						2.46^{**} (0.84)
Intercept	$6{,}914.12^{**}$ (0.75)	$6{,}914.13^{**}$ (0.75)	$6{,}914.13^{**}$ (0.75)	$6{,}914.12^{**}$ (0.75)	$6{,}914.13^{**}$ (0.75)	$6{,}914.13^{**}$ (0.75)
R^2	7.49×10^{-4}	1.34×10^{-3}	1.12×10^{-3}	7.15×10^{-4}	1.25×10^{-3}	1.06×10^{-3}
Observations	120,258	120,258	120,258	120,258	120,258	120,258

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. TIRC denotes Total Investments of Referred Customers. We report robust standard errors in parentheses for log-linear models estimated using OLS. All continuous independent variables have been mean-centered. For ease of exposition, the coefficients and standard errors have been multiplied by 1,000 in log-linear models.

Table 30: Heterogeneous Effects on TIRC Continued: The Second Experiment

Dependent Variable	TIRC					
Model	Tobit	Tobit	Tobit	Tobit	Tobit	Tobit
First Reward	$17.45^{\dagger}$ (10.26)	$19.96^{\dagger}$ (10.68)	$17.88^{\dagger}$ (10.45)			
Second Reward	15.10 (10.51)	9.11 (11.35)	10.64 (10.93)			
Composite Reward				$16.74^{\dagger}$ (9.07)	$15.78^{\dagger}$ (9.59)	15.12 (9.30)
Log Total Investment Return (Log Return)	$6.50^{\dagger}$ (3.79)			$6.50^{\dagger}$ (3.79)		
Log Investment Amount (Log Investment)		19.35^{**} (4.60)			19.37^{**} (4.60)	
Value Factor			18.22^{**} (5.80)			18.22^{**} (5.80)
First Reward × Log Return	6.34 (5.16)					
First Reward × Log Investment		0.53 (6.07)				
First Reward × Value Factor			5.42 (7.78)			
Second Reward × Log Return	15.09^{**} (5.27)					
Second Reward × Log Investment		16.57^{**} (6.30)				
Second Reward × Value Factor			22.98^{**} (8.00)			
Composite Reward × Log Return				10.72^{*} (4.56)		
Composite Reward × Log Investment					8.36 (5.41)	
Composite Reward × Value Factor						14.12^{*} (6.91)
Intercept	-634.94^{**} (31.39)	-632.95^{**} (31.29)	-632.33^{**} (31.24)	-634.88^{**} (31.39)	-633.64^{**} (31.33)	-632.62^{**} (31.26)
Log-Likelihood	−6,462.12	−6,428.96	−6,442.97	−6,463.73	−6,432.70	−6,445.74
Observations	120,258	120,258	120,258	120,258	120,258	120,258

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. TIRC denotes Total Investments of Referred Customers. We report (default) standard errors in parentheses for tobit models. All continuous independent variables have been mean-centered. For ease of exposition, the coefficients and standard errors have been divided by 1,000 in tobit models.

Table 31: Heterogeneous Effects on Incidence of Having Acquired Referred Customers: The Second Experiment

Dependent Variable	Has Acquired Referred Customers					
Model	OLS	OLS	OLS	OLS	OLS	OLS
First Reward	0.0005 (0.0006)	0.0005 (0.0006)	0.0005 (0.0006)			
Second Reward	$0.0011^{\dagger}$ (0.0006)	$0.0011^{\dagger}$ (0.0006)	$0.0011^{\dagger}$ (0.0006)			
Composite Reward				0.0008 (0.0005)	0.0008 (0.0005)	0.0008 (0.0005)
Log Total Investment Return (Log Return)	0.0003 (0.0002)			0.0003 (0.0002)		
Log Investment Amount (Log Investment)		0.0012^{**} (0.0002)			0.0012^{**} (0.0002)	
Value Factor			0.0011^{**} (0.0003)			0.0011^{**} (0.0003)
First Reward × Log Return	$0.0005^{\dagger}$ (0.0002)					
First Reward × Log Investment		0.0002 (0.0003)				
First Reward × Value Factor			0.0005 (0.0004)			
Second Reward × Log Return	0.0009^{**} (0.0003)					
Second Reward × Log Investment		0.0012^{**} (0.0003)				
Second Reward × Value Factor			0.0015^{**} (0.0004)			
Composite Reward × Log Return				0.0007^{**} (0.0002)		
Composite Reward × Log Investment					0.0007^{**} (0.0003)	
Composite Reward × Value Factor						0.0010^{**} (0.0003)
Intercept	0.0060^{**} (0.0004)	0.0060^{**} (0.0004)	0.0060^{**} (0.0004)	0.0060^{**} (0.0004)	0.0060^{**} (0.0004)	0.0060^{**} (0.0004)
R^2	4.72×10^{-4}	1.49×10^{-3}	9.91×10^{-4}	4.39×10^{-4}	1.41×10^{-3}	9.34×10^{-4}
Observations	120,258	120,258	120,258	120,258	120,258	120,258

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. We report robust standard errors in parentheses for linear probability models estimated using OLS. All continuous independent variables have been mean-centered.

Table 32: Heterogeneous Effects on Incidence of Having Acquired Referred Customers Continued: The Second Experiment

Dependent Variable	Has Acquired Referred Customers					
Model	Probit	Probit	Probit	Probit	Probit	Probit
First Reward	0.0183 (0.0320)	0.0208 (0.0333)	0.0165 (0.0326)			
Second Reward	0.0393 (0.0319)	0.0173 (0.0341)	0.0249 (0.0329)			
Composite Reward				0.0296 (0.0277)	0.0210 (0.0291)	0.0220 (0.0283)
Log Total Investment Return (Log Return)	0.0164 (0.0113)			0.0164 (0.0113)		
Log Investment Amount (Log Investment)		0.0768** (0.0137)			0.0768** (0.0137)	
Value Factor			0.0647** (0.0174)			0.0647** (0.0174)
First Reward × Log Return	0.0268† (0.0159)					
First Reward × Log Investment		0.0089 (0.0191)				
First Reward × Value Factor			0.0279 (0.0243)			
Second Reward × Log Return	0.0495** (0.0159)					
Second Reward × Log Investment		0.0564** (0.0193)				
Second Reward × Value Factor			0.0770** (0.0244)			
Composite Reward × Log Return				0.0384** (0.0138)		
Composite Reward × Log Investment					0.0329* (0.0167)	
Composite Reward × Value Factor						0.0530* (0.0211)
Intercept	−2.5126** (0.0227)	−2.5324** (0.0238)	−2.5202** (0.0231)	−2.5126** (0.0227)	−2.5324** (0.0238)	−2.5202** (0.0231)
Log-Likelihood	−4,709.01	−4,646.59	−4,677.54	−4,710.56	−4,650.09	−4,680.05
Observations	120,258	120,258	120,258	120,258	120,258	120,258

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. We report (default) standard errors in parentheses for probit models. All continuous independent variables have been mean-centered.

Table 33: Heterogeneous Effects on Number of Referred Customers: The Second Experiment

Dependent Variable	Number of Referred Customers					
Model	OLS	OLS	OLS	OLS	OLS	OLS
First Reward	0.0009 (0.0007)	0.0009 (0.0007)	0.0009 (0.0007)			
Second Reward	0.0016* (0.0007)	0.0016* (0.0007)	0.0016* (0.0007)			
Composite Reward				0.0013* (0.0006)	0.0012* (0.0006)	0.0013* (0.0006)
Log Total Investment Return (Log Return)	0.0002 (0.0002)			0.0002 (0.0002)		
Log Investment Amount (Log Investment)		0.0013** (0.0002)			0.0013** (0.0002)	
Value Factor			0.0011** (0.0003)			0.0011** (0.0003)
First Reward × Log Return	0.0006* (0.0003)					
First Reward × Log Investment		0.0005 (0.0004)				
First Reward × Value Factor			0.0008† (0.0004)			
Second Reward × Log Return	0.0013** (0.0003)					
Second Reward × Log Investment		0.0016** (0.0004)				
Second Reward × Value Factor			0.0021** (0.0005)			
Composite Reward × Log Return				0.0009** (0.0003)		
Composite Reward × Log Investment					0.0010** (0.0003)	
Composite Reward × Value Factor						0.0014** (0.0004)
Intercept	0.0066** (0.0004)	0.0066** (0.0004)	0.0066** (0.0004)	0.0066** (0.0004)	0.0066** (0.0004)	0.0066** (0.0004)
R^2	4.60×10^{-4}	1.43×10^{-3}	9.51×10^{-4}	4.21×10^{-4}	1.36×10^{-3}	8.91×10^{-4}
Observations	120,258	120,258	120,258	120,258	120,258	120,258

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. We report robust standard errors in parentheses for linear models estimated using OLS. All continuous independent variables have been mean-centered.

Table 34: Heterogeneous Effects on Number of Referred Customers Continued: The Second Experiment

Dependent Variable	Number of Referred Customers					
Model	NB2	NB2	NB2	NB2	NB2	NB2
First Reward	0.1051 (0.0951)	0.1008 (0.0990)	0.0931 (0.0967)			
Second Reward	0.1480 (0.0952)	0.0679 (0.1018)	0.0954 (0.0982)			
Composite Reward				0.1296 (0.0829)	0.0902 (0.0870)	0.0985 (0.0845)
Log Total Investment Return (Log Return)	0.0346 (0.0339)			0.0346 (0.0340)		
Log Investment Amount (Log Investment)		0.2149** (0.0409)			0.2149** (0.0409)	
Value Factor			0.1716** (0.0519)			0.1716** (0.0520)
First Reward × Log Return	0.0842† (0.0474)					
First Reward × Log Investment		0.0431 (0.0566)				
First Reward × Value Factor			0.0990 (0.0722)			
Second Reward × Log Return	0.1593** (0.0475)					
Second Reward × Log Investment		0.1766** (0.0574)				
Second Reward × Value Factor			0.2442** (0.0727)			
Composite Reward × Log Return				0.1236** (0.0413)		
Composite Reward × Log Investment					0.1111* (0.0496)	
Composite Reward × Value Factor						0.1744** (0.0631)
Intercept	−5.0284** (0.0684)	−5.0902** (0.0712)	−5.0507** (0.0692)	−5.0284** (0.0684)	−5.0902** (0.0712)	−5.0507** (0.0693)
Log-Likelihood	−5,031.00	−4,969.43	−5,000.59	−5,032.62	−4,972.38	−5,002.86
Observations	120,258	120,258	120,258	120,258	120,258	120,258

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. We report (default) standard errors in parentheses for NB2 models. All continuous independent variables have been mean-centered.

CHAPTER 3 : A Low-Dimension Learning Approach to Modeling Consumer Heterogeneity in Choice-Based Conjoint Estimation

3.1. Introduction

Conjoint analysis, and choice-based conjoint (CBC) in particular, has been widely used by both researchers and practitioners to assess how consumers with heterogeneous preferences value different product or service attributes (Wittink and Cattin, 1989; Green and Srinivasan, 1990; Huber, 2004). The understanding of consumers' heterogeneous preferences plays a central role in a variety of marketing decisions, such as pricing, targeted promotions, differentiated product offerings, and market segmentation (Allenby and Rossi, 1998). In most conjoint applications, researchers are faced with the challenge that short questionnaires are adopted due to concerns over response rates and response quality, and, as a result, the amount of information elicited from each consumer is limited (Lenk et al., 1996). Given the scarcity of information from each consumer, adequate modeling of consumer heterogeneity becomes critical for conjoint estimation.

Modeling consumer heterogeneity entails pooling information across consumers; for each consumer, her information is used to help estimating other consumers' partworths while other consumers' information also contributes to the estimation of her own partworths. Hence, identifying effective information pooling mechanisms is the key for accurate estimation of the individual-level partworths. In the marketing literature, researchers have primarily investigated three distinct information pooling mechanisms, each of which has been implemented by multiple models for recovering consumer heterogeneity. First, in the hierarchical Bayes (HB) model with a normal population distribution (Lenk et al., 1996; Rossi et al., 1996) and the convex optimization model of Evgeniou et al. (2007), information sharing is induced by shrinking the individual-level partworths toward the population mean. Second, the HB normal component mixture model (Allenby et al., 1998) and the sparse learning model of Chen et al. (2017) pool information across consumers by recovering segments in the population and shrinking the individual-level partworths toward their respective segment means. Finally, the finite mixture model (Kamakura and Russell, 1989) and the HB model with a Dirichlet process prior (Ansari and Mela, 2003; Kim et al., 2004) approximate the individual-level partworths using discrete points. As these models have all demonstrated strong empirical performance in modeling consumer heterogeneity, the effectiveness of their underlying information pooling mechanisms in addressing the challenge of limited individual-level information is evident.

In this essay, we propose an innovative low-dimension learning model for recovering consumers' heterogeneous partworths in CBC, built upon an information pooling mechanism that is *distinct* from but also *compatible* with the three well-established mechanisms in the marketing literature. We assume that most variations in consumers' heterogeneous partworths, viewed as vectors in an Euclidean space, are along a small number of orthogonal directions; consequently, consumers' partworth vectors all reside near some *low-dimensional affine subspace* of the Euclidean space (James et al., 2017).[36] This assumption is likely to hold when consumers have more substantial preference variations over a small number of attributes compared to the other attributes. Since there is a good low-dimensional affine subspace approximation for consumers' heterogeneous partworths, a natural mechanism for pooling information across consumers is to shrink the individual-level partworths toward this low-dimensional affine subspace, which is also inferred from the data. Our model implements such a low-dimension information pooling mechanism using a convex optimization framework in which both the distance between each partworth vector and the affine subspace as well as the dimension of the affine subspace are penalized.

We note that our low-dimension information pooling mechanism may lead to effective information sharing across consumers that cannot be induced by the three information pooling mechanisms in the literature, and vice versa. Therefore, coupling the low-dimension mechanism with any of the extant information pooling mechanisms could potentially give

[36] Examples of affine subspace of an Euclidean space include lines and planes that do not necessarily contain the origin of the space. We will provide a formal definition of affine subspace in Section 3.2.

rise to more effective information sharing across consumers than any single mechanism. In other words, we view our low-dimension mechanism and the extant information pooling mechanisms as *complements* rather than substitutes in addressing the challenge posed by the scarcity of information from each consumer. To illustrate this perspective, in our low-dimension learning model, we choose to *also* shrink the individual-level partworths toward the population mean besides shrinking them toward a low-dimensional affine subspace by incorporating the convex regularization of Evgeniou et al. (2007), which is straightforward within our convex optimization framework. Similar to extant machine learning-based models for recovering consumer heterogeneity (Evgeniou et al., 2007; Chen et al., 2017), we determine the amount of each type of shrinkage using cross-validation (Vapnik, 1998; Hastie et al., 2016).

We compare our low-dimension learning model and a restricted version of the model in which only the low-dimension information pooling mechanism is implemented to multiple benchmark models using simulation experiments and two field data sets. In simulations, the low-dimension learning model and its restricted version overall outperform the benchmark models both in terms of parameter recovery and predictive accuracy. In particular, both demonstrate strong performance irrespective of whether their underlying assumption - the true individual-level partworths have a good low-dimensional affine subspace approximation - seems to hold or not. In the two field data sets, these two models also emerge as the overall best performing models. We also find that, across simulations and field data sets, the performance of the low-dimension learning model is very close to that of its restricted version, suggesting that the incremental value of shrinking the individual-level partworths toward the population mean can be very limited when we are already shrinking the individual-level partworths toward a low-dimensional affine subspace.

The remainder of this essay is organized as follows. In Section 3.2, we describe the setup of CBC and develop our low-dimension learning model for recovering consumers' heterogeneous partworths in CBC. We compare our low-dimension learning model and the benchmark models using simulation experiments in Section 3.3 and two field conjoint data sets in Section 3.4. We conclude in Section 3.5.

3.2. The Low-Dimension Learning Model

3.2.1. Conjoint Setup

We consider a choice-based conjoint (CBC) experiment consisting of I respondents. In the CBC experiment, respondent i makes choice decisions over J choice sets, each of which includes H conjoint profiles with p attributes.[37] We use the row vector $x_{ijh} \in \mathbb{R}^p$ to represent the h-th profile in the j-th choice set of respondent i, and use the column vector $\beta_i \in \mathbb{R}^p$ to represent the partworth vector of respondent i. We assume that respondent i chooses her most preferred profile x_{ijh*} from her j-th choice set $\left\{x_{ijh}\right\}_{h=1}^{H}$ using a standard logit model with an additive specification of the utility function. Specifically, we assume that $U_{ijh} = x_{ijh}\beta_i + \epsilon_{ijh}$, where $\left\{\epsilon_{ijh}\right\}_{i,j,h}$ are independently and identically distributed type-I extreme value random variables, and x_{ijh*} is chosen such that $U_{ijh*} = \max\left\{U_{ijh}\right\}_{h=1}^{H}$ (Train, 2009).

3.2.2. Model Development

Given data from the CBC experiment, a natural starting point for conjoint estimation is to separately estimate each respondent's partworths using the logit maximum likelihood estimator (MLE). The logit MLE of respondent i's partworths is obtained by solving the following optimization problem:

$$\begin{aligned} \min \quad & -\sum_{j=1}^{J} \log \frac{e^{x_{ijh*}\beta_i}}{\sum_{h=1}^{H} e^{x_{ijh}\beta_i}}, \\ \text{s.t.} \quad & \beta_i \in \mathbb{R}^p. \end{aligned}$$

[37] For ease of exposition, we assume that all respondents need to answer the same number of choice questions and that all choice sets contain the same number of conjoint profiles. Our model, however, can be straightforwardly extended to accommodate varying numbers of choice sets across respondents and varying numbers of conjoint profiles across choice sets.

Equivalently, we can obtain the logit MLE of all respondents' partworths by solving the following joint optimization problem:

$$\begin{aligned} \min \quad & -\sum_{i=1}^{I}\sum_{j=1}^{J}\log\frac{e^{x_{ijh*}\beta_i}}{\sum_{h=1}^{H}e^{x_{ijh}\beta_i}}, \\ \text{s.t.} \quad & \beta_i \in \mathbb{R}^p, \ \text{for } i = 1, 2, \ldots, I. \end{aligned} \tag{3.1}$$

In most conjoint studies, short questionnaires are adopted due to concerns over response rates and response quality and the amount of information elicited from each respondent is limited (Lenk et al., 1996). The scarcity of individual-level information poses a serious challenge to the separate estimation of each respondent's logit model: When the number of choice sets J is small, the individual-level logit models become excessively flexible relative to the amount of data available to estimate these models; consequently, the logit MLE tends to overfit and yield poor estimates of respondents' partworths, and can be unidentified for many respondents (Rossi and Allenby, 1993; Allenby and Rossi, 1998; James et al., 2017).

In machine learning, a well-established approach to controlling the flexibility of a model and improving its predictive accuracy and interpretability is to *regularize* the estimation of the model (Hoerl and Kennard, 1970; Tibshirani, 1996; Zou and Hastie, 2005; Hastie et al., 2016). In the regularization approach, instead of simply minimizing an empirical loss function derived from the model that measures the fit between the model parameters and the data (e.g., the minus log-likelihood function in Problem (3.1)), the researcher minimizes the sum of the empirical loss function and a regularization function. The regularization function is specified to reflect the researcher's prior belief regarding which model parameters are more appropriate for his research objectives, such as prediction and variable selection, and those deemed as more appropriate are assigned lower values. Including the regularization function in the optimization problem thus effectively limits the set of possible parameter estimates and makes model parameters that are regarded as "better" based on the researcher's prior knowledge more likely to be chosen by the estimation procedure. Classical examples of the regularization approach include ridge regression (Hoerl and Kennard, 1970) and the lasso

(Tibshirani, 1996).

In the following, we develop our low-dimension learning model as an approach to regularizing the estimation of the individual-level logit models. To motivate, let us imagine for now that we can actually observe the set of true partworths $\{\hat{\beta}_i\}_{i=1}^{I}$. We further assume that we have conducted a principal component analysis on $\{\hat{\beta}_i\}_{i=1}^{I}$ and have found that the first few principal components collectively account for most variance in the true partworths, which is likely to be the case if respondents have more substantial preference variations over a small number of attributes than over the other attributes. Under this assumption, we can construct an affine subspace, obtained by appropriately shifting the linear subspace spanned by these principal components, that is close to $\{\hat{\beta}_i\}_{i=1}^{I}$ (James et al., 2017); in other words, there exists a good low-dimensional affine subspace approximation for $\{\hat{\beta}_i\}_{i=1}^{I}$.[38] This observation suggests that, when estimating the individual-level logit models, we can consider regularizing the estimated partworths $\{\beta_i\}_{i=1}^{I}$ so that they reside in a low-dimensional affine subspace. Such a regularization is likely to induce some bias, but the cost of the bias could be well outweighed by the benefit of the reduction in the variance for the estimation, and as a result the regularization could lead to a better bias-variance trade-off and more accurate estimates (Hastie et al., 2016).

To operationalize this low-dimensional affine subspace regularization, we propose the following optimization problem:

$$\begin{aligned}\min \quad & -\sum_{i=1}^{I}\sum_{j=1}^{J}\log\frac{e^{x_{ijh*}\beta_i}}{\sum_{h=1}^{H}e^{x_{ijh}\beta_i}}+\lambda\cdot\Phi\big(\{\beta_i\}_{i=1}^{I}\big),\\ \text{s.t.} \quad & \beta_i\in\mathbb{R}^p,\ \text{for } i=1,2,\ldots,I.\end{aligned} \tag{3.2}$$

In Problem (3.2), the regularization function $\Phi\big(\{\beta_i\}_{i=1}^{I}\big)$ is defined as the minimum of the dimension of any affine subspace of $\mathbb{R}^p$ that contains all β_i's, and its inclusion in the objective function serves to encourage the partworths $\{\beta_i\}_{i=1}^{I}$ to reside in a low-dimensional

[38] A subset W of $\mathbb{R}^p$ is an affine subspace of $\mathbb{R}^p$ if there exist a vector $\rho\in\mathbb{R}^p$ and a linear subspace V of $\mathbb{R}^p$ such that $W=\{\rho+v|v\in V\}$. Intuitively, W is obtained by shifting V by ρ.

affine subspace. Moreover, the regularization parameter $\lambda \geq 0$ controls the relative strength of the two components of the objective function and hence the trade-off between fit and regularization. To make Problem (3.2) more mathematically tractable, we reformulate the regularization function $\Phi\big(\{\beta_i\}_{i=1}^{I}\big)$ through the following equations:

$$
\begin{aligned}
\Phi\big(\{\beta_i\}_{i=1}^{I}\big) \quad = \quad & \min \quad \mathrm{Dim}(W), \\
& \text{s.t.} \quad W \text{ is an affine subspace of } \mathbb{R}^p;\ \beta_i \in W,\ \text{ for } i = 1, 2, \ldots, I. \\
= \quad & \min \quad \mathrm{Dim}(V), \\
& \text{s.t.} \quad V \text{ is a linear subspace of } \mathbb{R}^p,\ \rho \in \mathbb{R}^p;\ \beta_i - \rho \in V,\ \text{ for } i = 1, 2, \ldots, I. \\
= \quad & \min \quad \mathrm{Dim}\big(\mathrm{Span}(\beta_1 - \rho, \beta_2 - \rho, \ldots, \beta_I - \rho)\big), \\
& \text{s.t.} \quad \rho \in \mathbb{R}^p. \\
= \quad & \min \quad \mathrm{Rank}\big([\beta_1 - \rho, \beta_2 - \rho, \ldots, \beta_I - \rho]\big), \\
& \text{s.t.} \quad \rho \in \mathbb{R}^p.
\end{aligned}
$$

We briefly discuss the rationale behind each of these four equations. For simplicity, we refer to the optimization problem to the right of the q-th equals sign as the q-th optimization problem. The first equation simply reflects the mathematical definition of the regularization function $\Phi\big(\{\beta_i\}_{i=1}^{I}\big)$. As to the second equation, we note that W is an affine subspace of $\mathbb{R}^p$ if and only if there exist a vector $\rho \in \mathbb{R}^p$ and a linear subspace V of $\mathbb{R}^p$ such that $W = \rho + V \triangleq \big\{\rho + v \big| v \in V\big\}$. It is straightforward to verify that, for any feasible W in the first optimization problem, any (ρ, V)-pair satisfying $W = \rho + V$ is feasible in the second optimization problem, and $\mathrm{Dim}(V) = \mathrm{Dim}(W)$; on the other hand, for any feasible (ρ, V)-pair in the second optimization problem, $W \triangleq \rho + V$ is feasible in the first optimization problem, and $\mathrm{Dim}(W) = \mathrm{Dim}(V)$. As a result, the optimal values of the first and second optimization problems are equal and hence the second equation holds. Given that the second optimization problem has two decision variables, ρ and V, a simple approach to solving this problem is to first find the optimal V for any given ρ and then plug in the optimal V as a function of ρ and optimize over ρ. It is clear that, any linear subspace V of $\mathbb{R}^p$ containing $\beta_i - \rho$ for all i also contains the linear subspace spanned by $\big\{\beta_i - \rho\big\}_{i=1}^{I}$, which we denote as $\text{Span}(\beta_1 - \rho, \beta_2 - \rho, \ldots, \beta_I - \rho)$, and therefore for any given ρ the associated optimal $V^*(\rho)$ is $\text{Span}(\beta_1 - \rho, \beta_2 - \rho, \ldots, \beta_I - \rho)$.[39] The third optimization problem is obtained once we plug $V^*(\rho)$ in the second optimization problem. Finally, a basic result in linear algebra indicates that the dimension of the linear subspace spanned by a set of vectors in $\mathbb{R}^p$ is equal to the rank of the matrix whose columns are composed of these vectors, which immediately leads to the fourth equation.

Replacing $\Phi\big(\{\beta_i\}_{i=1}^I\big)$ with the fourth optimization problem in Problem (3.2), we arrive at the following equivalent optimization problem:

$$\begin{aligned} \min \quad & -\sum_{i=1}^{I}\sum_{j=1}^{J} \log \frac{e^{x_{ijh*}\beta_i}}{\sum_{h=1}^{H} e^{x_{ijh}\beta_i}} + \lambda \cdot \text{Rank}\big([\beta_1 - \rho, \beta_2 - \rho, \ldots, \beta_I - \rho]\big), \\ \text{s.t.} \quad & \beta_i \in \mathbb{R}^p, \text{ for } i = 1, 2, \ldots, I;\ \rho \in \mathbb{R}^p. \end{aligned} \tag{3.3}$$

In the rest of this section, we make three modifications to Problem (3.3) to enhance its modeling flexibility and computational tractability. First, while we assume that there exists a good low-dimensional affine subspace approximation for the true partworths $\{\hat{\beta}_i\}_{i=1}^I$, it is unlikely that $\{\hat{\beta}_i\}_{i=1}^I$ will actually reside in a low-dimensional affine subspace. Therefore, compared to the "hard" regularization in Problem (3.3) that encourages the partworths $\{\beta_i\}_{i=1}^I$ to reside in a low-dimensional affine subspace, a "soft" regularization only encouraging $\{\beta_i\}_{i=1}^I$ to be close to a low-dimensional affine subspace is likely to better capture the preferences of respondents. To implement such a "soft" regularization, we first reformulate Problem (3.3) in the following equivalent form:

$$\begin{aligned} \min \quad & -\sum_{i=1}^{I}\sum_{j=1}^{J} \log \frac{e^{x_{ijh*}\beta_i}}{\sum_{h=1}^{H} e^{x_{ijh}\beta_i}} + \lambda \cdot \text{Rank}\big([\theta_1, \theta_2, \ldots, \theta_I]\big) + \infty \cdot \sum_{i=1}^{I} \big|\big|\beta_i - \rho - \theta_i\big|\big|_2^2, \\ \text{s.t.} \quad & \beta_i, \theta_i \in \mathbb{R}^p, \text{ for } i = 1, 2, \ldots, I;\ \rho \in \mathbb{R}^p. \end{aligned} \tag{3.4}$$

Here $\big|\big|a\big|\big|_2^2 \triangleq \sum_{t=1}^{p} a_t^2$ for $a = (a_1, a_2, \ldots, a_p) \in \mathbb{R}^p$, and θ_i's are decision variables. The equiv-

[39] Formally, $\text{Span}(\beta_1 - \rho, \beta_2 - \rho, \ldots, \beta_I - \rho) \triangleq \big\{\sum_{i=1}^{I} k_i(\beta_i - \rho) | k_i \in R, \text{ for } i = 1, 2, \ldots, I\big\}$.

alence between Problems (3.3) and (3.4) can be seen by noting that the third component of the objective function of Problem (3.4), $\infty \cdot \sum_{i=1}^{I} \big|\big|\beta_i - \rho - \theta_i\big|\big|_2^2$, enforces $\theta_i = \beta_i - \rho$ for all i in the optimal solution. To switch to the "soft" regularization, we replace ∞ by a finite regularization parameter $\tau \geq 0$ in Problem (3.4), leading to the following modified optimization problem:

$$
\begin{aligned}
\min \quad & -\sum_{i=1}^{I}\sum_{j=1}^{J} \log \frac{e^{x_{ijh*}\beta_i}}{\sum_{h=1}^{H} e^{x_{ijh}\beta_i}} + \lambda \cdot \mathrm{Rank}\big(\big[\theta_1, \theta_2, \ldots, \theta_I\big]\big) + \tau \cdot \sum_{i=1}^{I} \big|\big|\beta_i - \rho - \theta_i\big|\big|_2^2, \\
\text{s.t.} \quad & \beta_i, \theta_i \in \mathbb{R}^p, \text{ for } i = 1, 2, \ldots, I; \ \rho \in \mathbb{R}^p.
\end{aligned}
\tag{3.5}
$$

In Problem (3.5), while the second component of the objective function (i.e., the rank function) encourages $\big\{\theta_i\big\}_{i=1}^{I}$ to reside in a low-dimensional linear subspace and hence $\big\{\rho + \theta_i\big\}_{i=1}^{I}$ to reside in a low-dimensional affine subspace, the third component of the objective function, by penalizing the discrepancy between β_i and $\rho+\theta_i$, encourages $\big\{\beta_i\big\}_{i=1}^{I}$ to be close to the low-dimensional affine subspace containing $\big\{\rho + \theta_i\big\}_{i=1}^{I}$ but without being enforced to reside in a low-dimensional affine subspace. As a result, Problem (3.5) implements the "soft" regularization discussed above.

Second, we note that Problem (3.5) is computationally intractable since the rank function is neither continuous nor convex. In the optimization and machine learning literatures, an effective approach to dealing with the computational challenge posed by the rank function that has seen many successful applications is to approximate the rank of a matrix using its nuclear norm (Cai et al., 2010; Pong et al., 2010; Candès et al., 2011; Ma et al., 2011). The nuclear norm of a matrix A, which we denote as $\big|\big|A\big|\big|_*$, is defined as the sum of its singular values, and a key property of the nuclear norm is that it is the tightest convex approximation of the rank function (Fazel, 2002). We adopt such a convex relaxation approach and replace the rank function by the nuclear norm in Problem (3.5), which gives rise to the following modified optimization problem:[40]

$$
\begin{aligned}
\min \quad & -\sum_{i=1}^{I}\sum_{j=1}^{J}\log\frac{e^{x_{ijh*}\beta_i}}{\sum_{h=1}^{H}e^{x_{ijh}\beta_i}} + \lambda\cdot\big|\big|\big[\theta_1,\theta_2,\cdots,\theta_I\big]\big|\big|_* + \tau\cdot\sum_{i=1}^{I}\big|\big|\beta_i-\rho-\theta_i\big|\big|_2^2, \\
\text{s.t.} \quad & \beta_i,\theta_i\in\mathbb{R}^p,\ \text{for } i=1,2,\ldots,I;\ \rho\in\mathbb{R}^p.
\end{aligned}
\tag{3.6}
$$

Effectively, our proposed low-dimension regularization pools information across respondents by shrinking the individual-level partworths toward a low-dimensional affine subspace. Such an information pooling mechanism is distinct from but also compatible with several well-established information pooling mechanisms in the marketing literature, including (1) shrinking the individual-level partworths toward the population mean (Lenk et al., 1996; Rossi et al., 1996; Evgeniou et al., 2007), (2) recovering segments and shrinking the individual-level partworths toward their respective segment means (Allenby et al., 1998; Chen et al., 2017), and (3) approximating the individual-level partworths using discrete points (Kamakura and Russell, 1989; Ansari and Mela, 2003; Kim et al., 2004). Since these mechanisms could lead to effective information sharing across respondents that cannot be induced by the low-dimension regularization, we could potentially enhance the estimation accuracy of our model by complementing the low-dimension regularization with any of these information pooling mechanisms. In this essay, we choose to shrink the individual-level partworths toward the population mean by incorporating the convex regularization of Evgeniou et al. (2007) in our approach, which gives rise to the low-dimension learning

[40]One may attempt to solve Problem (3.5) by enumerating all possible values of the rank function. That is, one finds $\big\{\{\beta_i^t,\theta_i^t\}_{i=1}^I,\rho^t\big\}$ that minimizes the objective function subject to the constraint that $\text{Rank}\big(\big[\theta_1,\theta_2,\ldots,\theta_I\big]\big)=t$ for $t=0,1,2,\ldots,\min(p,I)$, and then compares the objective values among $\big\{\{\beta_i^t,\theta_i^t\}_{i=1}^I,\rho^t\big\}$'s and identifies the one with the lowest objective value. The problem of this solution approach is that the constraint $\text{Rank}\big(\big[\theta_1,\theta_2,\ldots,\theta_I\big]\big)=t$ is non-convex, i.e., the feasible set it implies is non-convex, in which case a local optimal solution is not guaranteed to be a global optimal solution; consequently, the optimization problem becomes computationally intractable.

(LDL) optimization problem:[41]

$$
\begin{aligned}
\min \quad & -\sum_{i=1}^{I}\sum_{j=1}^{J}\log\frac{e^{x_{ijh*}\beta_i}}{\sum_{h=1}^{H}e^{x_{ijh}\beta_i}} + \lambda\cdot\big|\big|[\theta_1,\theta_2,\cdots,\theta_I]\big|\big|_* + \tau\cdot\sum_{i=1}^{I}\big|\big|\beta_i-\rho-\theta_i\big|\big|_2^2 \\
& + \gamma\cdot\sum_{i=1}^{I}(\beta_i-\beta_0)^\top D^{-1}(\beta_i-\beta_0), \qquad (3.7)\\
\text{s.t.} \quad & \beta_i,\theta_i\in\mathbb{R}^p,\ \text{for } i=1,2,\ldots,I;\ \ \rho,\beta_0\in\mathbb{R}^p;\\
& D\in\mathbb{R}^{p\times p} \text{ is a positive semidefinite matrix scaled to have trace 1.}
\end{aligned}
$$

The last component of the objective function, $\gamma\cdot\sum_{i=1}^{I}(\beta_i-\beta_0)^\top D^{-1}(\beta_i-\beta_0)$, is the convex regularization function of Evgeniou et al. (2007). Using the first order condition, it can be shown that in the optimal solution of the LDL optimization problem we have $\beta_0=\frac{1}{I}\sum_{i=1}^{I}\beta_i$, and hence the last component indeed shrinks the individual-level partworths toward the population mean (Evgeniou et al., 2007).[42] We note that in the LDL optimization problem, the decision variables include $\{\beta_i\}_{i=1}^{I}$, β_0, D, ρ, and $\{\theta_i\}_{i=1}^{I}$, and the regularization parameters include $\lambda>0$, $\tau>0$, and $\gamma>0$.

3.2.3. The Full Model and the Solution Strategy

Using the LDL optimization problem to estimate the individual-level partworths involves two tasks: (1) selecting the regularization parameters and (2) solving for the optimal solution of the LDL optimization problem given the regularization parameters. In this section, we discuss our solution strategy for each task and present our full low-dimension learning (LDL) model. Since our solution strategy for selecting the regularization parameters relies on solving the LDL optimization problem, we first discuss the solution algorithm for the LDL optimization problem and then the selection of the regularization parameters.

[41] Our choice of incorporating the convex regularization of Evgeniou et al. (2007) in our model is motivated by its strong empirical performance in modeling consumer heterogeneity and the fact that its incorporation in our model is straightforward, only requiring the addition of a new component in the objective function and a new constraint.

[42] Readers are referred to Evgeniou et al. (2007) for more details on the convex regularization function and the associated constraint on the matrix D.

Solving the Low-Dimension Learning Optimization Problem

Given the regularization parameters $\lambda > 0$, $\tau > 0$, and $\gamma > 0$, the LDL optimization problem is a convex optimization problem. A fundamental property of convex optimization problems is that any locally optimal solution is also a globally optimal solution, implying that convex optimization problems are tractable in theory (Boyd and Vandenberghe, 2004). To solve for the optimal solution of the LDL optimization problem, we exploit its special structure and propose a block coordinate descent (BCD) algorithm (Tseng, 2001; Hong et al., 2017). In the BCD algorithm, we partition the decision variables into several blocks, and minimize the objective function with respect to the first block of decision variables while holding the other blocks fixed, and then minimize the objective function with respect to the second block of decision variables while holding the other blocks fixed, and so on; once we have minimized the objective function with respect to the last block of decision variables, we move to the first block and iterate the process. We formally state the BCD algorithm in Table 35, where we refer to the objective function of the LDL optimization problem as $\mathcal{L}\Big(\{\beta_i\}_{i=1}^I, \beta_0, D, \rho, \{\theta_i\}_{i=1}^I\Big)$ for ease of exposition.

Insert Table 35 here.

From Table 35, we note that the implementation of the BCD algorithm entails solving five subproblems, each corresponding to one block of decision variables. We discuss the solution algorithm for each subproblem below.

Solving $\{\beta_i\}_{i=1}^I$-Subproblem. The $\{\beta_i\}_{i=1}^I$-subproblem is separable in β_i's, i.e., solving this subproblem is equivalent to separately solving the following optimization problem for each i:

$$\min_{\beta_i \in \mathbb{R}^p} \quad -\sum_{j=1}^{J} \log \frac{e^{x_{ijh*}\beta_i}}{\sum_{h=1}^{H} e^{x_{ijh}\beta_i}} + \tau \cdot \big|\big|\beta_i - \rho^l - \theta_i^l\big|\big|_2^2 + \gamma \cdot (\beta_i - \beta_0^l)^\top (D^l)^{-1} (\beta_i - \beta_0^l). \tag{3.8}$$

Since Problem (3.8) is both convex and smooth, we solve it using Newton's method (Boyd and Vandenberghe, 2004).

Solving β_0-Subproblem. The β_0-subproblem is equivalent to the following optimization problem:

$$\min_{\beta_0 \in \mathbb{R}^p} \quad \sum_{i=1}^{I} (\beta_0 - \beta_i^{l+1})^\top (D^l)^{-1} (\beta_0 - \beta_i^{l+1}). \tag{3.9}$$

Problem (3.9) is both convex and quadratic, and hence can be solved by simply taking the first-order condition, which yields the optimal solution $\beta_0^{l+1} = \frac{1}{I} \sum_{i=1}^{I} \beta_i^{l+1}$.

Solving D-Subproblem. The D-subproblem is equivalent to the following optimization problem:

$$\begin{aligned} \min \quad & \sum_{i=1}^{I} (\beta_i^{l+1} - \beta_0^{l+1})^\top D^{-1} (\beta_i^{l+1} - \beta_0^{l+1}), \\ \text{s.t.} \quad & D \in \mathbb{R}^{p \times p} \text{ is a positive semidefinite matrix scaled to have trace 1.} \end{aligned} \tag{3.10}$$

Problem (3.10) admits a closed-form solution, $D^{l+1} = \frac{1}{2v} \big(\sum_{i=1}^{I} (\beta_i^{l+1} - \beta_0^{l+1})(\beta_i^{l+1} - \beta_0^{l+1})^\top \big)^{\frac{1}{2}}$, where v is selected so that D^{l+1} has trace 1 (Evgeniou et al., 2007).

Solving ρ-Subproblem. The ρ-subproblem is equivalent to the following optimization problem:

$$\min_{\rho \in \mathbb{R}^p} \quad \sum_{i=1}^{I} ||\rho - \beta_i^{l+1} + \theta_i^l||_2^2. \tag{3.11}$$

Similar to Problem (3.9), Problem (3.11) is also both convex and quadratic. Taking the first-order condition, we obtain the optimal solution $\rho^{l+1} = \frac{1}{I} \sum_{i=1}^{I} (\beta_i^{l+1} - \theta_i^l)$.

Solving $\{\theta_i\}_{i=1}^I$-Subproblem. The $\{\theta_i\}_{i=1}^I$-subproblem is equivalent to the following optimization problem:

$$\min_{\theta_i \in \mathbb{R}^p, \text{ for } i=1,2,\ldots,I} \quad \lambda \cdot \big|\big| [\theta_1, \theta_2, \cdots, \theta_I] \big|\big|_* + \tau \cdot \sum_{i=1}^{I} \big|\big| \theta_i - \beta_i^{l+1} + \rho^{l+1} \big|\big|_2^2. \tag{3.12}$$

Problem (3.12) admits a closed-form solution. To construct the optimal solution, we define the matrix $B^{l+1} \triangleq \big[\beta_1^{l+1} - \rho^{l+1}, \beta_2^{l+1} - \rho^{l+1}, \ldots, \beta_I^{l+1} - \rho^{l+1}\big]$, and conduct a singular value decomposition (SVD) on B^{l+1}. Let the SVD be $B^{l+1} = U_B \cdot \text{diag}(\sigma) \cdot V_B^\top$, where $\text{diag}(\sigma)$ is a diagonal matrix of which the diagonal vector σ consists of the singular values of B^{l+1}. The optimal solution of Problem (3.12), $\Theta^{l+1} \triangleq \big[\theta_1^{l+1}, \theta_2^{l+1}, \ldots, \theta_I^{l+1}\big] = U_B \cdot \text{diag}\big(\max(\sigma - \frac{\lambda}{2\tau}, 0)\big) \cdot V_B^\top$, where the max operator works element-wise (Cai et al., 2010; Ma et al., 2011).

The Stopping Criterion. We terminate the BCD algorithm when the gap between the objective values of two consecutive iterations is small. Specifically, we end the algorithm after the L-th iteration, where L is the smallest number satisfying the following condition:

$$\mathcal{L}\Big(\big\{\beta_i^L\big\}_{i=1}^I, \beta_0^L, D^L, \rho^L, \big\{\theta_i^L\big\}_{i=1}^I\Big) - \mathcal{L}\Big(\big\{\beta_i^{L+1}\big\}_{i=1}^I, \beta_0^{L+1}, D^{L+1}, \rho^{L+1}, \big\{\theta_i^{L+1}\big\}_{i=1}^I\Big) < \eta.$$

In the empirical applications, we set $\eta = 0.01$.[43]

Selecting the Regularization Parameters

Similar to extant machine learning-based models for recovering consumer heterogeneity (Evgeniou et al., 2007; Chen et al., 2017), we select the regularization parameters (λ, τ, γ) using cross-validation (Vapnik, 1998; Hastie et al., 2016). To this end, we specify a grid $\Lambda \subset \mathbb{R}^3$ from which the triplet (λ, τ, γ) is chosen. For each $(\lambda, \tau, \gamma) \in \Lambda$, we evaluate its cross-validation error, $CVE(\lambda, \tau, \gamma)$, which is defined as follows:[44]

[43]While the choice of $\eta = 0.01$ works well in our empirical applications, we can also consider treating η as an additional regularization parameter and selecting η endogenously using cross-validation.

[44]Our operationalization of the cross-validation error follows that adopted in Evgeniou et al. (2007) and Chen et al. (2017). Other operationalizations, which differ in how the data are divided into "calibration" and "holdout" sets, can also be used.

(1) Set $CVE(\lambda, \tau, \gamma) = 0$.

(2) For $j = 1$ to J:

(a) Divide the conjoint data into two disjoint subsets - a "calibration" set Calib_j and a "holdout" set Hold_j. Calib_j contains all conjoint data except the j-th choice set of each respondent, and Hold_j contains only the j-th choice set of each respondent.

(b) Obtain the individual-level partworth estimates $\{\beta_i^{(-j)}\}_{i=1}^{I}$ by solving the LDL optimization problem on the "calibration" set Calib_j given the regularization parameters (λ, τ, γ).

(c) For each respondent i, compute the logistic error of $\beta_i^{(-j)}$ on her j-th choice set (i.e., her only choice set in the "holdout" set Hold_j), $-\log \frac{e^{x_{ijh*}\beta_i^{(-j)}}}{\sum_{h=1}^{H} e^{x_{ijh}\beta_i^{(-j)}}}$. Let $\Delta(j)$ be the sum of the logistic errors over all I respondents.

(d) Set $CVE(\lambda, \tau, \gamma) = CVE(\lambda, \tau, \gamma) + \Delta(j)$.

For the individual-level partworth estimates obtained by solving the LDL optimization problem on the full data set given the regularization parameters (λ, τ, γ), the cross-validation error $CVE(\lambda, \tau, \gamma)$ provides an effective estimate of their *out-of-sample* predictive accuracy using only *in-sample* data, i.e., the data available to the researcher for model calibration. Therefore, we select $(\lambda^*, \tau^*, \gamma^*)$ as the minimizer of $CVE(\lambda, \tau, \gamma)$ over the grid Λ, which is expected to lead to the individual-level partworth estimates with the optimal predictive accuracy on out-of-sample data.

The Low-Dimension Learning Model: A Summary

By solving the LDL optimization problem using the BCD algorithm and selecting the regularization parameters using cross-validation, our LDL model obtains the individual-level partworth estimates via the following two-step approach:

Step 1. Select the optimal regularization parameters $(\lambda^*, \tau^*, \gamma^*) = \underset{(\lambda,\tau,\gamma)\in\Lambda}{\operatorname{argmin}}\, CVE(\lambda, \tau, \gamma)$.

Step 2. Solve for the optimal solution of the LDL optimization problem given $(\lambda^*, \tau^*, \gamma^*)$; the $\{\beta_i\}_{i=1}^{I}$-component of the optimal solution is the individual-level partworth estimates of the LDL model.

3.3. Simulation Experiments

In this section, we assess the empirical performance of our LDL model using simulation experiments. We compared the LDL model to three extant models that have shown strong performance in modeling consumer heterogeneity, including (1) the HB model with a normal population distribution (Lenk et al., 1996; Rossi et al., 1996), (2) the HB normal component mixture model (Allenby et al., 1998), and (3) the convex optimization model of Evgeniou et al. (2007). We refer to these three models as the UHB (for unimodal HB), NCM, and LOG-Het models, respectively. LOG-Het can be seen as a restricted version of the LDL model, in which the information pooling mechanism of shrinking the individual-level partworths toward a low-dimensional affine subspace is shut down by setting the regularization parameters λ and τ to 0. We also tested another restricted version of the LDL model, termed as LDL-RV, in which we set $\gamma = 0$ and select λ and τ using cross-validation. Clearly, in LDL-RV the individual-level partworths are not shrunk toward the population mean and information pooling relies completely on shrinking the individual-level partworths toward a low-dimensional affine subspace.[45]

3.3.1. Data

The setup of the simulation experiments, including the data-generating process and the experimental design, largely followed the choice-based simulations in Evgeniou et al. (2007) with one modification that we will highlight below.

[45] We provide the specifications of the grid of regularization parameters for the LDL, LDL-RV, and LOG-Het models and discuss the setup and implementation of the UHB and NCM models in the appendix.

Data-Generating Process

We assumed that there were 10 binary attributes and each choice set consisted of two conjoint profiles. To generate the choice sets, we first derived a design matrix $M \in \mathbb{R}^{24\times 10}$ with 24 rows and the j-th row M_j providing the dummy coding of the j-th profile.[46] We then applied the shifting method of Bunch et al. (1996) to the design matrix M to obtain 24 choice sets, where the j-th choice set consisted of M_j and $1 - M_j$. In each data set, we generated choice data for 100 synthetic respondents. The true partworths of each respondent were drawn from a multivariate normal distribution with the mean vector $\mu = [mag, mag, \ldots, mag]$ and the covariance matrix $\Sigma = \mathrm{diag}\big(\{\sigma_t^2\}_{t=1}^{10}\big)$, where $\sigma_t^2 = het \times mag$ for $t = 1, 2, \ldots, 5$ and $\sigma_t^2 = het \times mag \times ratio$ for $t = 6, 7, \ldots, 10$. As pointed out in Evgeniou et al. (2007), the parameter *mag* controls the amount of response error and the parameter *het* controls the amount of heterogeneity. We differed from Evgeniou et al. (2007) in the introduction of the parameter $ratio \in [0, 1]$, which controls the proportion of preference variations allocated to the first 5 attributes as opposed to the last 5 attributes and hence how well the true individual-level partworths can be approximated by a low-dimensional affine subspace. For example, when $ratio = 0$, all preference variations are restricted to the first 5 attributes and the true individual-level partworths actually reside in a 5-dimensional affine subspace; on the other hand, when $ratio = 1$, the preference variations are equally allocated to all 10 attributes and all directions see the same amount of heterogeneity due to normality, and therefore the true individual-level partworths may not have a good low-dimensional affine subspace approximation. Given the true individual-level partworths, we simulated each respondent's choices using the logit model. We randomly selected J (out of the 24) choice sets for each respondent as the calibration data set, and randomly selected another 8 choice sets for each respondent as the holdout data set.

[46] The design matrix M was generated using the SAS Macro %mktex(2 2 2 2 2 2 2 2 2 2, n = 24).

Experimental Design

We experimentally manipulated four data characteristics: (1) the number of choice sets per respondent for calibration via the parameter J, (2) the amount of response error via the parameter mag, (3) the amount of heterogeneity via the parameter het, and (4) the proportion of preference variations allocated to the first 5 attributes via the parameter $ratio$. Specifically, we adopted the following $2^3 \times 3$ design:

Factor 1. J: 8 or 16;

Factor 2. mag: 0.2 or 1.2;

Factor 3. het: 1 or 3;

Factor 4. $ratio$: 0, 0.5, or 1.

The choices of levels for the first three factors were identical to those of the choice-based simulations in Evgeniou et al. (2007). We considered three levels for $ratio$: $ratio = 0$ and $ratio = 1$ led to boundary conditions in which the preference variations were either allocated to the first 5 attributes only or equally allocated to all attributes, whereas $ratio = 0.5$ led to intermediate conditions in which the preference variations allocated to the last 5 attributes were half of those allocated to the first 5 attributes (in terms of variance). By varying $ratio$, we were interested in understanding how critical the assumption that the true individual-level partworths can be well approximated by a low-dimensional affine subspace is for the LDL model to be effective in modeling consumer heterogeneity. In sum, we had a total of 24 experimental conditions; for each condition, we randomly generated 5 data sets and estimated all five models separately on each data set (Evgeniou et al., 2007; Chen et al., 2017).[47]

[47] Estimating all five models on a single data set took around 7 hours when $J = 8$ and around 13 hours when $J = 16$. As the next step, we plan to generate more data sets for each experimental condition.

Performance Measures

We compared all five models in terms of parameter recovery and predictive accuracy. Parameter recovery was assessed using the root mean squared error between the true individual-level partworths $\hat{\beta}_i$ and the estimated individual-level partworths β_i, which we denote as RMSE (Andrews et al., 2002).[48] RMSE for respondent i was defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{p}||\hat{\beta}_i - \beta_i||_2^2}. \tag{3.13}$$

Following Evgeniou et al. (2007) and Chen et al. (2017), for any given model, we computed RMSE for each respondent in each data set and report the average RMSE across respondents and data sets for each experimental condition.

Predictive accuracy was evaluated using two performance measures, the holdout sample log-likelihood, which we denote as Holdout-LL, and the holdout sample hit rate, which we denote as Holdout-HIT (Andrews et al., 2002; Evgeniou et al., 2007; Iyengar and Jedidi, 2012; Chen et al., 2017). For respondent i, Holdout-LL was defined as follows:

$$\text{Holdout-LL} = \frac{1}{\tilde{J}} \sum_{j=1}^{\tilde{J}} \log \frac{e^{\tilde{x}_{ijh*}\beta_i}}{\sum_{h=1}^{H} e^{\tilde{x}_{ijh}\beta_i}}, \tag{3.14}$$

and Holdout-HIT was defined as follows:

$$\text{Holdout-HIT} = \frac{1}{\tilde{J}} \sum_{j=1}^{\tilde{J}} \mathbf{1}\left(\tilde{x}_{ijh*}\beta_i = \max_h \left\{\tilde{x}_{ijh}\beta_i\right\}\right), \tag{3.15}$$

where $\tilde{J}$ denotes the number of holdout choice sets for each respondent, i.e., $\tilde{J} = 8$ in our study, $\left\{\tilde{x}_{ijh}\right\}_{h=1}^{H}$ denotes the j-th holdout choice set for respondent i, and the indicator function $\mathbf{1}(\cdot)$ takes value 1 when the argument is true, and 0 otherwise. Again, for any given model, we computed Holdout-LL and Holdout-HIT for each respondent in each data set and

[48] The LDL, LDL-RV, and LOG-Het models generate the point estimates β_i's directly, whereas for the UHB and NCM models we calculated β_i's using the means of the posterior distributions.

report the averages across respondents and data sets for each experimental condition.

Between Holdout-LL and Holdout-HIT, the two measures assessing predictive accuracy, the continuous Holdout-LL is more sensitive (Iyengar and Jedidi, 2012). One important difference between the two measures stems from the ways in which they penalize a model for assigning a low choice probability for the chosen profile of a choice set - the logarithmic functional form of Holdout-LL imposes a heavy penalty on such a scenario whereas the stepwise functional form of Holdout-HIT imposes a constant penalty as long as the chosen profile is not assigned with the highest choice probability.

3.3.2. Results

We first compare the five models with respect to RMSE. The average RMSEs of the models are reported in Tables 36 and 37.

Insert Tables 36 and 37 here.

We make several observations from Tables 36 and 37. First, the overall best performing models are LDL and LDL-RV, followed by UHB: LDL and LDL-RV perform best or not significantly different from best ($p > 0.05$) in 11 and 12 conditions, respectively, whereas UHB performs best or not significantly different from best ($p > 0.05$) in 8 conditions. Specifically, we find that LDL and LDL-RV show superior performance in conditions with high response error (i.e., $mag = 0.2$) and conditions with low response error and low heterogeneity (i.e., $mag = 1.2$, $het = 1$); on the other hand, UHB performs very well in conditions with low response error and high heterogeneity (i.e., $mag = 1.2$, $het = 3$). Here, the comparisons are based on paired t-tests over the same 500 respondents in each experimental condition (i.e., 100 respondents per data set $\times$ 5 data sets per condition).

Second, LDL-RV, the restricted version of LDL that only shrinks the individual-level partworths toward a low-dimensional affine subspace, outperforms LOG-Het, the restricted version of LDL that only shrinks the individual-level partworths toward the population mean.

This finding suggests that in our simulation experiments the former information pooling mechanism is more effective in recovering heterogeneous preferences than the latter. On the other hand, the average RMSEs of LDL and LDL-RV are very close, indicating that in our simulation experiments the incremental value of shrinking the individual-level partworths toward the population mean is limited when we are already shrinking the individual-level partworths toward a low-dimensional affine subspace.

Third, both LDL and LDL-RV demonstrate strong performance in conditions with $ratio = 1$, in which the true individual-level partworths are unlikely to be well approximated by a low-dimensional affine subspace. This finding is very encouraging, as it suggests that the LDL and LDL-RV models as well as the information pooling mechanism of shrinking the individual-level partworths toward a low-dimensional affine subspace can be effective in modeling consumer heterogeneity even when their underlying assumption does not seem to hold.

We conduct a regression analysis to further investigate the impact of the four data characteristics on the relative performance between LDL and other models. To this end, we consider the following regression specification:

$$\begin{aligned}\Delta_t \quad = \quad & \alpha + \beta_1 \times \mathbf{1}\big(J = 16\big)_t + \beta_2 \times \mathbf{1}\big(mag = 1.2\big)_t + \beta_3 \times \mathbf{1}\big(het = 3\big)_t \\ & + \beta_4 \times \mathbf{1}\big(ratio = 0.5\big)_t + \beta_5 \times \mathbf{1}\big(ratio = 1\big)_t + \epsilon_t,\end{aligned} \tag{3.16}$$

where the index t refers to the t-th experimental condition ($t = 1, 2, \ldots, 24$). The dependent variable Δ_t denotes the difference between the average RMSEs of LDL and one of the other four models (i.e., LDL-RV, LOG-Het, NCM, and UHB) in the t-th condition, and the independent variables are appropriately defined dummy variables for the four experimental factors.[49] We estimate this linear model using ordinary least squares (OLS) and report the results in Table 38. Note that a smaller dependent variable indicates a better relative performance for LDL.

[49]For instance, when comparing LDL and UHB, Δ_t is defined as the average RMSE of LDL less that of UHB in the t-th condition.

Insert Table 38 here.

Table 38 shows that the experimental factors have no significant impact on the relative performance between LDL and LDL-RV. Compared to LOG-Het, the relative performance of LDL improves with lower response error (i.e., $mag = 1.2$) and lower heterogeneity (i.e., $het = 1$). The performance of LDL relative to NCM and UHB is more favorable in conditions with fewer calibration choice sets (i.e., $J = 8$), higher response error (i.e., $mag = 0.2$), and lower heterogeneity (i.e., $het = 1$); moreover, compared to UHB, the relative performance of LDL improves when the preference variations are more concentrated on the first five attributes (i.e., $ratio = 0$).

Now, we compare the predictive accuracy of the five models using Holdout-LL and Holdout-HIT. Model comparisons in terms of Holdout-LL are reported in Tables 39 and 40.

Insert Tables 39 and 40 here.

We find that LDL and LDL-RV are again the overall best performing models, being either the best model or indistinguishable from the best model ($p > 0.05$) in 24 and 19 conditions, respectively. We also find that the observations regarding RMSE still hold when the performance measure is Holdout-LL, including that LDL-RV outperforms LOG-Het and is comparable to LDL, and that LDL and LDL-RV have strong performance in conditions with $ratio = 1$. Similar to the case of RMSE, we conduct a regression analysis to understand how the relative performance between LDL and other models with respect to Holdout-LL varies across experimental conditions. We use the regression specification (3.16), with the dependent variable Δ_t denoting the difference between the average Holdout-LLs of LDL and one of the other four models in the t-th condition. Results of the OLS estimation are summarized in Table 41. Here, unlike the case of RMSE, a larger dependent variable indicates a better relative performance for LDL.

Insert Table 41 here.

For Holdout-LL, we find that the experimental factors have no significant impact on the relative performance between LDL and LDL-RV. Compared to the other three models - LOG-Het, NCM, and UHB - the relative performance of LDL is more favorable in conditions with fewer calibration choice sets (i.e., $J = 8$). In addition, the performance of LDL relative to LOG-Het and UHB improves with lower response error (i.e., $mag = 1.2$) and lower heterogeneity (i.e., $het = 1$), respectively.

When the performance measure is Holdout-HIT, model comparisons are reported in Tables 42 and 43.

Insert Tables 42 and 43 here.

The performance gaps among the models in terms of Holdout-HIT are much smaller than those in terms of Holdout-LL. Specifically, LDL, LDL-RV, LOG-Het, NCM, and UHB perform best or not significantly different from best ($p > 0.05$) in 24, 22, 19, 18, 18 conditions, respectively. To investigate the impact of the experimental factors on the relative performance between LDL and other models with respect to Holdout-HIT, we conduct a regression analysis using the specification (3.16), with the dependent variable Δ_t denoting the difference between the average Holdout-HITs of LDL and one of the other four models in the t-th condition. Results of the OLS estimation are summarized in Table 44. Here, a larger dependent variable indicates a better relative performance for LDL.

Insert Table 44 here.

From Table 44, we find that the performance of LDL relative to LDL-RV is less favorable when the preference variations are evenly allocated to all attributes (i.e., $ratio = 1$). Compared to LOG-Het, the relative performance of LDL improves with fewer calibration choice sets (i.e., $J = 8$) and more even allocation of preference variations to attributes (i.e., $ratio = 0.5$ or 1). On the other hand, the experimental factors have no significant impact on the relative performance between LDL and the two HB models, NCM and UHB.

3.4. Field Data

In this section, we compare the predictive accuracy of the five models (i.e., LDL, LDL-RV, LOG-Het, NCM, and UHB) using two field CBC data sets.

3.4.1. The Hospital Data

In the first data set, a total of 200 respondents participated in a CBC study on hospitals.[50] Each respondent was shown 6 choice sets, and each choice set consisted of 4 profiles and the no-choice option was not included. There were 3 attributes describing a profile, with the first attribute having 6 levels, and the second and third attributes each having 4 levels, respectively.[51] We randomly selected 4 out of the 6 choice sets for each respondent for model calibration, and used the remaining 2 choice sets for holdout validation.

We evaluate the predictive accuracy of the models using Holdout-LL and Holdout-HIT. Similar to Section 3.3, for any given model, we computed Holdout-LL and Holdout-HIT for each respondent and report the averages across all 200 respondents. Results are reported in the upper section of Table 45.

Insert Table 45 here.

We find that LDL, LDL-RV, and UHB are the best performing models on the hospital data set, each being either the best model or indistinguishable from the best model ($p > 0.1$) both in terms of Holdout-LL and Holdout-HIT. NCM has an average Holdout-LL that is significantly lower than the best model ($p < 0.1$), whereas LOG-Het performs significantly worse than the best model for both Holdout-LL and Holdout-HIT ($p < 0.1$). Here, the comparisons are based on paired t-tests over the sample of 200 respondents.[52]

[50] We thank Rajan Sambandam from TRC Market Research for sharing the hospital data set with us.

[51] We were agnostic about the specific attributes and levels used in this study.

[52] To be more specific, when comparing two models with respect to a performance measure (i.e., either Holdout-LL or Holdout-HIT), we calculate the measure for each of the 200 respondents and obtain 200 individual-level measures for each model. We then conduct a paired t-test on the two sets of 200 individual-level performance measures.

3.4.2. The Soft Drink Data

In the second data set, 192 respondents took part in a CBC study on soft drinks.[53] Each respondent answered 22 choice questions, each of which included 8 profiles and there was no no-choice option. Three attributes were used in the study, including brand (with 6 levels), size (with 7 levels), and price (with 7 levels). We randomly selected 16 out of the 22 choice sets for each respondent for model calibration, and used the remaining 6 choice sets for holdout validation.

We again measure the predictive accuracy of the models using Holdout-LL and Holdout-HIT. Results are reported in the lower section of Table 45. We find that this time the best performing models include LDL, LDL-RV, and LOG-Het, each of which performs best or not significantly different from best ($p > 0.1$) for both Holdout-LL and Holdout-HIT. On the other hand, both UHB and NCM perform significantly worse than the best model in terms of Holdout-LL ($p < 0.1$), and NCM also has an average Holdout-HIT that is significantly lower than the best model ($p < 0.1$). We note that the comparisons are based on paired t-tests over the sample of 192 respondents.

3.4.3. Summary

Across the two field CBC data sets, LDL and LDL-RV emerge as the overall best performing models and the predictive accuracy of LDL and LDL-RV is very close, which are consistent with our findings in the simulation experiments. As the next step, we plan to compare the models using more field CBC data sets to assess the robustness of the findings based on the hospital and soft drink data sets and explore settings in which LDL and LDL-RV are likely to perform particularly well. We also plan to investigate whether and when the LDL and LDL-RV models lead to more profitable pricing strategies.

[53]The soft drink data have been previously analyzed in Evgeniou et al. (2007). We thank Olivier Toubia for sharing this data set with us.

3.5. Conclusions and Future Research

Adequate modeling of consumer heterogeneity is critical for CBC estimation since the amount of information elicited from each consumer is limited in most CBC studies. In this essay, we propose an innovative LDL model for recovering consumers' heterogeneous partworths in CBC, built upon an information pooling mechanism that shrinks the individual-level partworths toward a low-dimensional affine subspace that is also inferred from the data. Our model implements such a low-dimension information pooling mechanism using a convex optimization framework in which both the distance between each partworth vector and the affine subspace as well as the dimension of the affine subspace are penalized. In order to further enhance the effectiveness of the LDL model, we also incorporate the information pooling mechanism of shrinking the individual-level partworths toward the population mean in the LDL model.

We compare the LDL model and a restricted version of the model, LDL-RV, to benchmark models including UHB, NCM, and LOG-Het using simulation experiments and two field data sets. LDL and LDL-RV overall outperform the benchmark models both in terms of parameter recovery and predictive accuracy. We find that LDL and LDL-RV demonstrate strong performance even when their underlying assumption that the true individual-level partworths have a good low-dimensional affine subspace approximation seems unlikely to hold. We also find that the performance of LDL is very close to that of LDL-RV, suggesting that the incremental value of shrinking the individual-level partworths toward the population mean can be very limited when we are already shrinking the individual-level partworths toward a low-dimensional affine subspace.

There are a few questions that have not been answered in this essay and could serve as starting points for future research. First, why do we find in the simulation experiments that LDL and LDL-RV perform well even in situations where their underlying assumption seems unlikely to hold? Is this finding an artifact of the setup of our simulation experiments or it actually speaks to a general property of the low-dimension information pooling mechanism? Second, are there scenarios in which LDL outperforms LDL-RV? That is, when does shrinking the individual-level partworths toward the population mean noticeably enhance the low-dimension information pooling mechanism? Third, are the findings on model comparisons based on the two field CBC data sets robust? Comparing the models using more field CBC data sets should enhance our understanding of the relative performance of LDL and LDL-RV. Fourth, we have empirically demonstrated the superior performance of LDL and LDL-RV in terms of parameter recovery and predictive accuracy. Do they (and when do they) lead to more profitable targeting strategies? Finally, are there tractable ways to couple the low-dimension information pooling mechanism and information pooling mechanisms other than shrinking the individual-level partworths toward the population mean?

Table 35: The Block Coordinate Descent (BCD) Algorithm

The BCD Algorithm
1: **Initialization**: Choose $\left\{\beta_i^0\right\}_{i=1}^I, \beta_0^0, D^0, \rho^0, \left\{\theta_i^0\right\}_{i=1}^I$
2: **for** $l = 0, 1, \cdots$ until the stopping criterion is satisfied **do**
3: $\left\{\beta_i\right\}_{i=1}^I$**-Subproblem**: $\left\{\beta_i^{l+1}\right\}_{i=1}^I \leftarrow \underset{\left\{\beta_i\right\}_{i=1}^I}{\operatorname{argmin}} \mathcal{L}\Big(\left\{\beta_i\right\}_{i=1}^I, \beta_0^l, D^l, \rho^l, \left\{\theta_i^l\right\}_{i=1}^I\Big)$
4: β_0**-Subproblem**: $\beta_0^{l+1} \leftarrow \underset{\beta_0}{\operatorname{argmin}} \mathcal{L}\Big(\left\{\beta_i^{l+1}\right\}_{i=1}^I, \beta_0, D^l, \rho^l, \left\{\theta_i^l\right\}_{i=1}^I\Big)$
5: D**-Subproblem**: $D^{l+1} \leftarrow \underset{D}{\operatorname{argmin}} \mathcal{L}\Big(\left\{\beta_i^{l+1}\right\}_{i=1}^I, \beta_0^{l+1}, D, \rho^l, \left\{\theta_i^l\right\}_{i=1}^I\Big)$
6: ρ**-Subproblem**: $\rho^{l+1} \leftarrow \underset{\rho}{\operatorname{argmin}} \mathcal{L}\Big(\left\{\beta_i^{l+1}\right\}_{i=1}^I, \beta_0^{l+1}, D^{l+1}, \rho, \left\{\theta_i^l\right\}_{i=1}^I\Big)$
7: $\left\{\theta_i\right\}_{i=1}^I$**-Subproblem**: $\left\{\theta_i^{l+1}\right\}_{i=1}^I \leftarrow \underset{\left\{\theta_i\right\}_{i=1}^I}{\operatorname{argmin}} \mathcal{L}\Big(\left\{\beta_i^{l+1}\right\}_{i=1}^I, \beta_0^{l+1}, D^{l+1}, \rho^{l+1}, \left\{\theta_i\right\}_{i=1}^I\Big)$
8: **end for**
9: **return** $\Big(\left\{\beta_i^{L+1}\right\}_{i=1}^I, \beta_0^{L+1}, D^{L+1}, \rho^{L+1}, \left\{\theta_i^{L+1}\right\}_{i=1}^I\Big)$, where L is the last iteration

Table 36: Model Comparisons on RMSE: Simulation Experiments

				RMSE				
J	*mag*	*het*	*ratio*	LDL	LDL-RV	LOG-Het	NCM	UHB
8	0.2	1	0.0	**0.2995**	**0.2995**	0.3083	0.9035	0.7371
		1	0.5	0.3589	**0.3570**	0.3717	1.0007	0.7313
		1	1.0	**0.3990**	**0.3987**	0.4082	1.0641	0.8120
		3	0.0	**0.4589**	**0.4596**	0.4660	1.0465	0.7739
		3	0.5	**0.5801**	**0.5790**	0.5997	1.0769	0.7868
		3	1.0	**0.6590**	**0.6594**	0.6981	1.1157	0.7976
8	1.2	1	0.0	0.8475	**0.8231**	0.9766	1.4867	1.4877
		1	0.5	**1.0113**	1.0378	1.1282	1.4982	1.2466
		1	1.0	1.1333	**1.1240**	1.2008	1.7050	1.2247
		3	0.0	1.2586	1.2305	1.3012	1.3720	**1.1088**
		3	0.5	1.5499	1.5687	1.5621	1.8354	**1.3883**
		3	1.0	1.8662	1.8559	1.8921	1.8020	**1.6196**

Notes. Bold numbers in each experimental condition indicate best or not significantly different from best at the $p < 0.05$ level based on paired t-tests.

Table 37: Model Comparisons on RMSE Continued: Simulation Experiments

				RMSE				
J	*mag*	*het*	*ratio*	LDL	LDL-RV	LOG-Het	NCM	UHB
16	0.2	1	0.0	**0.2662**	0.2733	0.2685	0.5934	0.5442
		1	0.5	**0.3215**	**0.3222**	0.3243	0.5531	0.5449
		1	1.0	**0.3646**	0.3696	0.3754	0.5855	0.5493
		3	0.0	0.3938	**0.3892**	0.3941	0.6567	0.5888
		3	0.5	**0.4895**	0.4971	0.4976	0.6898	0.5788
		3	1.0	0.5737	**0.5679**	**0.5667**	0.7457	0.6109
16	1.2	1	0.0	**0.7661**	0.7768	0.8649	1.1038	0.9990
		1	0.5	0.8820	**0.8741**	0.9754	0.9107	**0.8776**
		1	1.0	0.9826	1.0045	1.0457	1.1701	**0.9108**
		3	0.0	1.1012	1.0517	1.0996	1.0692	**0.9781**
		3	0.5	1.4670	1.4696	1.4439	1.1923	**1.1632**
		3	1.0	1.6136	1.6245	1.6203	1.5649	**1.3231**

Notes. Bold numbers in each experimental condition indicate best or not significantly different from best at the $p < 0.05$ level based on paired t-tests.

Table 38: Regression Analysis on RMSE

Dependent Variable	LDL − LDL-RV		LDL − LOG-Het		LDL − NCM		LDL − UHB	
$\mathbf{1}(J = 16)$	−0.0023	(0.0074)	0.0197	(0.0142)	0.3226 **	(0.0597)	0.1538 *	(0.0541)
$\mathbf{1}(mag = 1.2)$	0.0038	(0.0074)	−0.0431 **	(0.0142)	0.2197 **	(0.0597)	0.2536 **	(0.0541)
$\mathbf{1}(het = 3)$	0.0072	(0.0074)	0.0405 *	(0.0142)	0.2322 **	(0.0597)	0.2772 **	(0.0541)
$\mathbf{1}(ratio = 0.5)$	−0.0167	(0.0100)	0.0056	(0.0189)	0.0929	(0.0686)	0.1461 †	(0.0769)
$\mathbf{1}(ratio = 1)$	−0.0126	(0.0091)	0.0090	(0.0158)	0.0849	(0.0787)	0.1962 *	(0.0764)
Intercept	0.0067	(0.0073)	−0.0444 *	(0.0197)	−0.7422 **	(0.0738)	−0.5705 **	(0.0946)
R^2	0.272		0.592		0.816		0.830	
Observations	24		24		24		24	

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. Robust standard errors are reported in parentheses.

Table 39: Model Comparisons on Holdout-LL: Simulation Experiments

				Holdout-LL				
J	*mag*	*het*	*ratio*	LDL	LDL-RV	LOG-Het	NCM	UHB
8	0.2	1	0.0	**-0.6624**	**-0.6629**	−0.6688	−1.1801	−0.9767
		1	0.5	**-0.6605**	**-0.6594**	−0.6709	−1.3729	−0.9637
		1	1.0	**-0.6647**	**-0.6639**	−0.6693	−1.2589	−1.0019
		3	0.0	**-0.6431**	**-0.6429**	**-0.6440**	−1.2026	−0.9493
		3	0.5	**-0.6464**	**-0.6474**	−0.6569	−1.2051	−0.9139
		3	1.0	**-0.6262**	**-0.6277**	−0.6514	−1.0642	−0.8298
8	1.2	1	0.0	**-0.4311**	−0.4347	−0.4634	−0.8352	−0.7220
		1	0.5	**-0.4502**	−0.4580	−0.4810	−0.8912	−0.7039
		1	1.0	**-0.4639**	**-0.4644**	−0.4924	−0.9863	−0.7172
		3	0.0	**-0.4737**	**-0.4751**	**-0.4757**	−0.8971	−0.7209
		3	0.5	**-0.5124**	**-0.5121**	−0.5238	−1.2984	−0.8317
		3	1.0	**-0.5041**	**-0.5054**	−0.5226	−1.0946	−0.7608

Notes. Bold numbers in each experimental condition indicate best or not significantly different from best at the $p < 0.05$ level based on paired t-tests.

Table 40: Model Comparisons on Holdout-LL Continued: Simulation Experiments

				Holdout-LL				
J	*mag*	*het*	*ratio*	LDL	LDL-RV	LOG-Het	NCM	UHB
16	0.2	1	0.0	**-0.6493**	**-0.6496**	**-0.6497**	−0.8831	−0.8470
		1	0.5	**-0.6468**	**-0.6486**	−0.6530	−0.8561	−0.8488
		1	1.0	**-0.6407**	**-0.6424**	−0.6482	−0.8508	−0.8137
		3	0.0	**-0.5993**	−0.6033	−0.6012	−0.8394	−0.7739
		3	0.5	**-0.6090**	**-0.6100**	**-0.6116**	−0.8896	−0.7555
		3	1.0	**-0.5834**	−0.5860	−0.5880	−0.7994	−0.6992
16	1.2	1	0.0	**-0.3756**	**-0.3757**	−0.3922	−0.5347	−0.5253
		1	0.5	**-0.4060**	**-0.4061**	−0.4207	−0.5890	−0.5513
		1	1.0	**-0.4308**	**-0.4278**	**-0.4337**	−0.7427	−0.5977
		3	0.0	**-0.3552**	−0.3591	**-0.3567**	−0.5607	−0.5169
		3	0.5	**-0.3980**	**-0.3977**	**-0.4022**	−0.6115	−0.5206
		3	1.0	**-0.4103**	**-0.4112**	−0.4164	−0.7953	−0.5460

Notes. Bold numbers in each experimental condition indicate best or not significantly different from best at the $p < 0.05$ level based on paired t-tests.

Table 41: Regression Analysis on Holdout-LL

Dependent Variable	LDL − LDL-RV		LDL − LOG-Het		LDL − NCM		LDL − UHB	
$\mathbf{1}(J = 16)$	−0.0002	(0.0011)	−0.0094 *	(0.0037)	−0.3083 **	(0.0437)	−0.1218 **	(0.0136)
$\mathbf{1}(mag = 1.2)$	0.0003	(0.0011)	0.0074 †	(0.0037)	−0.0121	(0.0437)	−0.0199	(0.0136)
$\mathbf{1}(het = 3)$	0.0004	(0.0011)	−0.0060	(0.0037)	0.0332	(0.0437)	−0.0275 †	(0.0136)
$\mathbf{1}(ratio = 0.5)$	−0.0005	(0.0015)	0.0036	(0.0040)	0.0802	(0.0539)	−0.0103	(0.0145)
$\mathbf{1}(ratio = 1)$	−0.0011	(0.0011)	0.0045	(0.0052)	0.0656	(0.0436)	−0.0250	(0.0159)
Intercept	0.0014	(0.0010)	0.0117 *	(0.0047)	0.4865 **	(0.0462)	0.3149 **	(0.0118)
R^2	0.066		0.511		0.797		0.868	
Observations	24		24		24		24	

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. Robust standard errors are reported in parentheses.

Table 42: Model Comparisons on Holdout-HIT: Simulation Experiments

				Holdout-HIT				
J	*mag*	*het*	*ratio*	LDL	LDL-RV	LOG-Het	NCM	UHB
8	0.2	1	0.0	**0.5952**	**0.5913**	**0.5917**	**0.5815**	**0.5883**
		1	0.5	**0.6045**	**0.6062**	0.5887	0.5925	0.5905
		1	1.0	**0.6025**	**0.6028**	**0.6008**	**0.6018**	**0.5975**
		3	0.0	**0.6308**	**0.6268**	**0.6235**	**0.6250**	**0.6318**
		3	0.5	**0.6305**	**0.6335**	0.6178	**0.6230**	**0.6310**
		3	1.0	**0.6532**	**0.6528**	0.6232	**0.6515**	**0.6495**
8	1.2	1	0.0	**0.7857**	**0.7860**	**0.7867**	**0.7823**	**0.7845**
		1	0.5	**0.7770**	**0.7708**	**0.7768**	0.7668	**0.7730**
		1	1.0	**0.7702**	**0.7710**	**0.7652**	0.7602	**0.7678**
		3	0.0	**0.7702**	**0.7708**	**0.7755**	**0.7695**	**0.7682**
		3	0.5	**0.7572**	**0.7538**	0.7475	0.7355	**0.7540**
		3	1.0	**0.7590**	**0.7592**	0.7435	0.7338	0.7475

Notes. Bold numbers in each experimental condition indicate best or not significantly different from best at the $p < 0.05$ level based on paired t-tests.

Table 43: Model Comparisons on Holdout-HIT Continued: Simulation Experiments

				Holdout-HIT				
J	*mag*	*het*	*ratio*	LDL	LDL-RV	LOG-Het	NCM	UHB
16	0.2	1	0.0	**0.6215**	**0.6175**	**0.6218**	0.6082	0.6050
		1	0.5	**0.6295**	**0.6252**	**0.6248**	**0.6195**	0.6158
		1	1.0	**0.6360**	**0.6380**	**0.6290**	**0.6345**	**0.6358**
		3	0.0	**0.6835**	0.6790	**0.6857**	**0.6773**	0.6755
		3	0.5	**0.6865**	0.6823	**0.6835**	**0.6827**	**0.6855**
		3	1.0	**0.7035**	**0.7033**	**0.7030**	**0.7030**	**0.7020**
16	1.2	1	0.0	**0.8333**	**0.8320**	**0.8347**	**0.8303**	0.8287
		1	0.5	**0.8217**	**0.8223**	**0.8207**	**0.8180**	**0.8185**
		1	1.0	**0.8143**	**0.8153**	**0.8113**	**0.8103**	**0.8095**
		3	0.0	**0.8515**	**0.8512**	**0.8500**	**0.8522**	**0.8510**
		3	0.5	**0.8363**	**0.8367**	**0.8355**	**0.8337**	**0.8327**
		3	1.0	**0.8323**	**0.8317**	**0.8275**	**0.8250**	**0.8250**

Notes. Bold numbers in each experimental condition indicate best or not significantly different from best at the $p < 0.05$ level based on paired t-tests.

Table 44: Regression Analysis on Holdout-HIT

Dependent Variable	LDL − LDL-RV		LDL − LOG-Het		LDL − NCM		LDL − UHB	
$\mathbf{1}(J = 16)$	0.0004	(0.0012)	−0.0061 *	(0.0028)	−0.0048	(0.0032)	0.0010	(0.0024)
$\mathbf{1}(mag = 1.2)$	−0.0009	(0.0012)	−0.0042	(0.0028)	0.0012	(0.0032)	−0.0017	(0.0024)
$\mathbf{1}(het = 3)$	0.0000	(0.0012)	0.0033	(0.0028)	−0.0003	(0.0032)	−0.0030	(0.0024)
$\mathbf{1}(ratio = 0.5)$	−0.0006	(0.0017)	0.0057 *	(0.0025)	0.0033	(0.0035)	0.0004	(0.0027)
$\mathbf{1}(ratio = 1)$	−0.0025 **	(0.0009)	0.0082 †	(0.0039)	0.0007	(0.0042)	−0.0003	(0.0030)
Intercept	0.0024 †	(0.0012)	0.0037	(0.0026)	0.0076 †	(0.0040)	0.0067 *	(0.0029)
R^2	0.230		0.509		0.193		0.151	
Observations	24		24		24		24	

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. Robust standard errors are reported in parentheses.

Table 45: Field Conjoint Data Sets

			The Hospital Data		
	LDL	LDL-RV	LOG-Het	NCM	UHB
Holdout-LL	**-0.9286**	**-0.9381**	−0.9777	−1.0231	**-0.9547**
Holdout-HIT	**0.6325**	**0.6275**	0.6075	**0.6275**	**0.6325**
			The Soft Drink Data		
	LDL	LDL-RV	LOG-Het	NCM	UHB
Holdout-LL	**-1.3489**	**-1.3493**	**-1.3497**	−1.4043	−1.3968
Holdout-HIT	**0.5208**	**0.5226**	**0.5191**	0.5087	**0.5130**

Notes. Bold numbers in each row indicate best or not significantly different from best at the $p < 0.1$ level based on paired t-tests.

APPENDIX

A.1. Alternative Tests Comparing Outcome Variables

In this section, we compare the three outcome variables defined on each current customer - the total investments of referred customers, the number of referred customers, and the incidence of having acquired referred customers - in both experiments using alternative test procedures. Specifically, we use a test procedure similar to the one applied to compare investments of each referred customer in the text to compare the total investments and number of referred customers, and use the test for equal proportions to compare the incidence of having acquired referred customers.

A.1.1. The First Experiment

We first conduct the comparisons with respect to the total investments of a current customer's referred customers. Between the control condition and a treatment condition, we first test whether there is a difference between the proportions of current customers whose referred customers have made positive investments (i.e., positive observations), and then test whether there is a difference between the two positive subsamples. We report the findings in Table A1.

Insert Table A1 here.

We compare the proportions of positive observations across conditions using the test for equal proportions, and find that the proportions are not significantly different between the control condition and either of the gift and notification conditions (ps > 0.1). On the other hand, as the Shapiro-Wilk tests indicate that the positive subsamples in all three conditions are not normally distributed (ps < 0.001), we compare the positive subsamples across conditions using the Wilcoxon-Mann-Whitney test. We find that the positive subsamples in both the gift and notification conditions have a higher average total investment amount made by referred customers than that in the control condition, and both differences are significant (ps < 0.05).

We then compare the number of a current customer's referred customers across conditions. Between the control condition and a treatment condition, we first test whether there is a difference between the proportions of current customers who have acquired referred customers during the experiment (i.e., positive observations), and then test whether there is a difference between the two positive subsamples. We report the findings in Table A2.

Insert Table A2 here.

Again, we compare the proportions of positive observations across conditions using the test for equal proportions, and find that the proportions are not significantly different between the control condition and either of the gift and notification conditions (ps > 0.1). On the other hand, as the Shapiro-Wilk tests indicate that the positive subsamples in all three conditions are not normally distributed (ps < 0.001), we compare the positive subsamples across conditions using the Wilcoxon-Mann-Whitney test. We find that there is no significant difference between the positive subsample of the control condition and those of the gift and notification conditions (ps > 0.1).

Finally, we compare the incidence of a current customer having acquired referred customers across conditions using the test for equal proportions. In fact, we have already conducted the comparisons in Table A2, simply noting that current customers with a positive number of referred customers are by definition those who have acquired referred customers. We find that the proportions of current customers who have acquired referred customers are not significantly different between the control condition and either of the gift and notification conditions (ps > 0.1).

A.1.2. The Second Experiment

We first conduct the comparisons with respect to the total investments of a current customer's referred customers and report the findings in Table A3.

Insert Table A3 here.

Using the test for equal proportions, we find that the proportions of positive observations are not significantly different between the control and first value-based reward conditions ($p > 0.1$), whereas the second value-based reward condition has a significantly higher proportion of positive observations than the control condition ($p < 0.05$). As the Shapiro-Wilk tests indicate that the positive subsamples in all three conditions are not normally distributed (ps < 0.001), we compare the positive subsamples across conditions using the Wilcoxon-Mann-Whitney test. We find that the positive subsamples in both value-based reward conditions have a higher average total investment amount made by referred customers than that in the control condition, and the difference between the control and first value-based reward conditions is marginally significant ($p < 0.1$) whereas the difference between the control and second value-based reward conditions is significant ($p < 0.05$).

We then compare the number of a current customer's referred customers across conditions and report the findings in Table A4.

Insert Table A4 here.

We compare the proportions of positive observations across conditions using the test for equal proportions, and find that the proportions of positive observations are not significantly different between the control and first value-based reward conditions ($p > 0.1$), whereas the second value-based reward condition has a higher proportion of positive observations than the control condition that is marginally significant ($p < 0.1$). On the other hand, as the Shapiro-Wilk tests indicate that the positive subsamples in all three conditions are not normally distributed (ps < 0.001), we compare the positive subsamples across conditions using the Wilcoxon-Mann-Whitney test. We find that there is no significant difference between the positive subsample of the control condition and those of the two value-based reward conditions (ps > 0.1).

Finally, we compare the incidence of a current customer having acquired referred customers across conditions using the test for equal proportions. Again, we have already conducted the comparisons in Table A4, noting that current customers with a positive number of referred customers are by definition those who have acquired referred customers. We find that the proportions of current customers who have acquired referred customers are not significantly different between the control and first value-based reward conditions ($p > 0.1$), whereas the second value-based reward condition has a higher proportion than the control condition that is marginally significant ($p < 0.1$).

A.2. Who Were Likely to Derive More Value from the Gift?

We empirically show in this section that, when offered the gift interest-raising coupon, customers of higher value, operationalized as those with a higher total investment return and those with a larger current investment amount, invested more in the fixed deposits during the first experiment and hence were likely to derive more value from the gift coupon. To this end, we estimate two linear models on customers in the gift condition: In the first model, we regress customers' fixed-deposit investments made during the experiment on an intercept and their total investment return; in the second model, we regress the same dependent variable on an intercept and customers' current investment amount. We note that this regression analysis is purely descriptive and not intended to establish a causal relationship between the dependent and independent variables. We report the results of the OLS estimation in Table A5. For completeness, we also estimate these two models on customers in the control condition and those in the notification condition and report the results in the same table.

Insert Table A5 here.

We find that in the gift condition both total investment return and current investment amount are positively associated with fixed-deposit investments made during the experiment, indicating that, when offered the gift coupon, customers with a higher total investment return and those with a larger current investment amount invested more in the fixed deposits during the experiment, and hence were likely to have derived more value from the gift coupon. Similar results are also found in the control and notification conditions.

A.3. Grids of Regularization Parameters

In the LDL, LDL-RV, and LOG-Het models, the regularization parameters are selected from a pre-specified finite grid. We present in the following the grid used for each model.

The LDL Model.

- $\lambda \in \left\{10^{-2+\frac{v}{3}}\right\}_{v=0}^{12}$, $\tau \in \left\{10^{-3+\frac{v}{3}}\right\}_{v=0}^{12}$, $\gamma \in \left\{10^{-3+\frac{v}{3}}\right\}_{v=0}^{12}$.

The LDL-RV Model.

- $\lambda \in \left\{10^{-2+\frac{v}{3}}\right\}_{v=0}^{12}$, $\tau \in \left\{10^{-3+\frac{v}{3}}\right\}_{v=0}^{12}$.

The LOG-Het Model.

- $\gamma \in \left\{10^{-3+\frac{v}{3}}\right\}_{v=0}^{12}$.

By choosing these grids, we aim to cover a wide range of values for the regularization parameters while keeping the total number of grid points moderate so that the computational time remains reasonable.

A.4. The NCM and UHB Models

We focus on the setup and implementation of the NCM model as UHB is a special case of NCM in which the number of normal components is set to 1. We adopted the following specification of NCM:

Likelihood:

$$\mathrm{Prob}\big(x_{ijh*} \text{ is chosen}\big) = \frac{e^{x_{ijh*}\beta_i}}{\sum_{h=1}^{H} e^{x_{ijh}\beta_i}};$$

First-stage prior:

$$\beta_i \sim N\left(\mu_{\mathrm{Ind}_i}, \Sigma_{\mathrm{Ind}_i}\right),$$
$$\mathrm{Ind}_i \sim \mathrm{Multinomial(pvec)};$$

Second-stage prior:

$$\begin{aligned}
&\mathrm{pvec} \sim \mathrm{Dirichlet}(\alpha), \\
&\mu_k \sim N\left(\bar{\mu}, \Sigma_k \otimes a_\mu^{-1}\right), \quad k = 1, 2, \ldots, K, \\
&\Sigma_k \sim \mathrm{IW}(v, V).
\end{aligned}$$

We set $\alpha_k = 2$, for $k = 1, 2, \ldots, K$, $\bar{\mu} = 0$, $a_\mu = 1/8$, $v = p + 3$, and $V = vI$. For any fixed number of components K, we used the Gibbs sampler to generate draws from the posterior distribution. We executed the Gibbs sampler for 30000 iterations, using the first 15000 iterations as the burn-in period and the last 15000 iterations to obtain parameter estimates. We estimated the NCM model for $K \in \{1, 2, \ldots, 10\}$ and selected K using the deviance information criterion (Spiegelhalter et al., 2002).

Table A1: An Alternative Test Comparing Total Investments of Referred Customers: The First Experiment

	Control	Gift	Notification	C vs. G	C vs. N
Full Sample of All CCs					
Proportion of CCs with Positive Investments from RCs	0.11%	0.15%	0.12%	$p = 0.233$	$p = 0.822$
Sample Size	30,977	31,241	31,070		
Subsample of CCs with Positive Investments from RCs					
Average Total Investments of RCs (in RMB)	11,364.87	26,521.27	31,663.95	$p = 0.039$	$p = 0.021$
Subsample Size	34	46	37		

Notes. CC denotes Current Customer, and RC denotes Referred Customer. In the section "Full Sample of All CCs", the columns "Control", "Gift", and "Notification" report the proportion of current customers whose referred customers have made positive investments in each condition, and the columns "C vs. G" and "C vs. N" report p-values of the test for equal proportions comparing the control condition and the gift and notification conditions. In the section "Subsample of CCs with Positive Investments from RCs", the columns "Control", "Gift", and "Notification" report the average total investments of referred customers for current customers whose referred customers have made positive investments in each condition, and the columns "C vs. G" and "C vs. N" report p-values of the Wilcoxon-Mann-Whitney test comparing the control condition and the gift and notification conditions.

Table A2: An Alternative Test Comparing Number of Referred Customers: The First Experiment

	Control	Gift	Notification	C vs. G	C vs. N
Full Sample of All CCs					
Proportion of CCs with a Positive Number of RCs	0.25%	0.32%	0.27%	$p = 0.108$	$p = 0.647$
Sample Size	30,977	31,241	31,070		
Subsample of CCs with a Positive Number of RCs					
Average Number of RCs	1.09	1.02	1.12	$p = 0.125$	$p = 0.851$
Subsample Size	76	99	83		

Notes. CC denotes Current Customer, and RC denotes Referred Customer. In the section "Full Sample of All CCs", the columns "Control", "Gift", and "Notification" report the proportion of current customers who have acquired referred customers in each condition, and the columns "C vs. G" and "C vs. N" report p-values of the test for equal proportions comparing the control condition and the gift and notification conditions. In the section "Subsample of CCs with a Positive Number of RCs", the columns "Control", "Gift", and "Notification" report the average number of referred customers for current customers who have acquired referred customers in each condition, and the columns "C vs. G" and "C vs. N" report p-values of the Wilcoxon-Mann-Whitney test comparing the control condition and the gift and notification conditions.

Table A3: An Alternative Test Comparing Total Investments of Referred Customers: The Second Experiment

	Control	First Reward	Second Reward	C vs. F	C vs. S
Full Sample of All CCs					
Proportion of CCs with Positive Investments from RCs	0.23%	0.30%	0.32%	$p = 0.102$	$p = 0.019$
Sample Size	40,076	40,145	40,037		
Subsample of CCs with Positive Investments from RCs					
Average Total Investments of RCs (in RMB)	39,637.16	74,492.05	59,408.41	$p = 0.056$	$p = 0.031$
Subsample Size	94	119	130		

Notes. CC denotes Current Customer, and RC denotes Referred Customer. In the section "Full Sample of All CCs", the columns "Control", "First Reward", and "Second Reward" report the proportion of current customers whose referred customers have made positive investments in each condition, and the columns "C vs. F" and "C vs. S" report p-values of the test for equal proportions comparing the control condition and the two value-based reward conditions. In the section "Subsample of CCs with Positive Investments from RCs", the columns "Control", "First Reward", and "Second Reward" report the average total investments of referred customers for current customers whose referred customers have made positive investments in each condition, and the columns "C vs. F" and "C vs. S" report p-values of the Wilcoxon-Mann-Whitney test comparing the control condition and the two value-based reward conditions.

Table A4: An Alternative Test Comparing Number of Referred Customers: The Second Experiment

	Control	First Reward	Second Reward	C vs. F	C vs. S
Full Sample of All CCs					
Proportion of CCs with a Positive Number of RCs	0.60%	0.65%	0.71%	$p = 0.431$	$p = 0.058$
Sample Size	40,076	40,145	40,037		
Subsample of CCs with a Positive Number of RCs					
Average Number of RCs	1.09	1.15	1.15	$p = 0.185$	$p = 0.245$
Subsample Size	241	260	285		

Notes. CC denotes Current Customer, and RC denotes Referred Customer. In the section "Full Sample of All CCs", the columns "Control", "First Reward", and "Second Reward" report the proportion of current customers who have acquired referred customers in each condition, and the columns "C vs. F" and "C vs. S" report p-values of the test for equal proportions comparing the control condition and the two value-based reward conditions. In the section "Subsample of CCs with a Positive Number of RCs", the columns "Control", "First Reward", and "Second Reward" report the average number of referred customers for current customers who have acquired referred customers in each condition, and the columns "C vs. F" and "C vs. S" report p-values of the Wilcoxon-Mann-Whitney test comparing the control condition and the two value-based reward conditions.

Table A5: Which Current Customers Invested More in Fixed Deposits during the First Experiment?

DV: Fixed-Deposit Investments	The Gift Condition		The Control Condition		The Notification Condition	
Total Investment Return	0.50** (0.03)		0.34** (0.04)		0.35** (0.03)	
Investment Amount		0.06** (0.01)		0.04** (0.01)		0.04** (0.01)
Intercept	951.15** (78.14)	−5.19 (252.84)	612.37** (71.68)	−19.90 (188.19)	731.67** (80.51)	2.10 (230.05)
R^2	0.0101	0.0695	0.0056	0.0404	0.0056	0.0416
Observations	31,241	31,241	30,977	30,977	31,070	31,070

Notes. $^{\dagger}p < .1$. $^{*}p < .05$. $^{**}p < .01$. Results are from linear models estimated using OLS. Robust standard errors are reported in parentheses.

BIBLIOGRAPHY

G. M. Allenby and P. E. Rossi. Marketing Models of Consumer Heterogeneity. *Journal of Econometrics*, 89(1-2):57–78, 1998.

G. M. Allenby, N. Arora, and J. L. Ginter. On the Heterogeneity of Demand. *Journal of Marketing Research*, 35(3):384–389, 1998.

F. Alpizar, F. Carlsson, and O. Johansson-Stenman. Anonymity, Reciprocity, and Conformity: Evidence from Voluntary Contributions to a National Park in Costa Rica. *Journal of Public Economics*, 92(5-6):1047–1060, 2008.

R. L. Andrews, A. Ainslie, and I. S. Currim. An Empirical Comparison of Logit Choice Models with Discrete versus Continuous Representations of Heterogeneity. *Journal of Marketing Research*, 39(4):479–487, 2002.

A. Ansari and C. F. Mela. E-Customization. *Journal of Marketing Research*, 40(2):131–145, 2003.

R. Bapna, A. Gupta, J. Jung, and S. Sen. Analyzing the Impact of Incentive Structure on the Diffusion of Mobile Social Games: A Randomized Field Experiment. *Working Paper*, 2016.

A. Barasch, J. Z. Berman, and D. A. Small. When Payment Undermines the Pitch: On the Persuasiveness of Pure Motives in Fund-Raising. *Psychological Science*, 27(10):1388–1397, 2016.

L. Beaman and J. Magruder. Who Gets the Job Referral? Evidence from a Social Networks Experiment. *American Economic Review*, 102(7):3574–3593, 2012.

Y. Benjamini and Y. Hochberg. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B*, 57(1):289–300, 1995.

Y. Benjamini and D. Yekutieli. The Control of the False Discovery Rate in Multiple Testing under Dependency. *The Annals of Statistics*, 29(4):1165–1188, 2001.

B. Berman. Referral Marketing: Harnessing the Power of Your Customers. *Business Horizons*, 59(1):19–28, 2016.

S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

M. Brown, E. Setren, and G. Topa. Do Informal Referrals Lead to Better Matches? Evidence from a Firm's Employee Referral System. *Journal of Labor Economics*, 34(1):161–209, 2016.

D. S. Bunch, J. J. Louviere, and D. Anderson. A Comparison of Experimental Design

Strategies for Multinomial Logit Models: The Case of Generic Attributes. *Working Paper*, 1996.

J.-F. Cai, E. J. Candès, and Z. Shen. A Singular Value Thresholding Algorithm for Matrix Completion. *SIAM Journal on Optimization*, 20(4):1956–1982, 2010.

A. C. Cameron and P. K. Trivedi. *Microeconometrics Using Stata*. Stata Press, 2010.

E. J. Candès, X. Li, Y. Ma, and J. Wright. Robust Principal Component Analysis? *Journal of the ACM*, 58(3):11:1–11:37, 2011.

E. J. Castilla. Social Networks and Employee Performance in a Call Center. *American Journal of Sociology*, 110(5):1243–1283, 2005.

Y. Chen, R. Iyengar, and G. Iyengar. Modeling Multimodal Continuous Heterogeneity in Conjoint Analysis - A Sparse Learning Approach. *Marketing Science*, 36(1):140–156, 2017.

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

R. Cialdini. Social Motivations to Comply: Norms, Values, and Principles. In J. A. Roth and J. T. Scholz, editors, *Taxpayer Compliance*, volume 2, pages 200–227. University of Pennsylvania Press, 1992.

R. Cialdini. *Influence: The Psychology of Persuasion*. HarperCollins, 1993.

T. Evgeniou, M. Pontil, and O. Toubia. A Convex Optimization Approach to Modeling Consumer Heterogeneity in Conjoint Estimation. *Marketing Science*, 26(6):805–818, 2007.

A. Falk. Gift Exchange in the Field. *Econometrica*, 75(5):1501–1511, 2007.

M. Fazel. *Matrix Rank Minimization with Applications*. PhD thesis, Stanford University, 2002.

E. Fehr and S. Gachter. Fairness and Retaliation: The Economics of Reciprocity. *Journal of Economic Perspectives*, 14(3):159–181, 2000.

R. M. Fernandez, E. J. Castilla, and P. Moore. Social Capital at Work: Networks and Employment at a Phone Center. *American Journal of Sociology*, 105(5):1288–1356, 2000.

D. S. Gilchrist, M. Luca, and D. Malhotra. When 3+1>4: Gift Structure and Reciprocity in the Field. *Management Science*, 62(9):2639–2650, 2016.

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

U. Gneezy, S. Meier, and P. Rey-Biel. When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*, 25(4):191–210, 2011.

D. Godes. Invited Comment on "Opinion Leadership and Social Contagion in New Product Diffusion". *Marketing Science*, 30(2):224–229, 2011.

D. Godes, D. Mayzlin, Y. Chen, S. Das, C. Dellarocas, B. Pfeiffer, B. Libai, S. Sen, M. Shi, and P. Verlegh. The Firm's Management of Social Interactions. *Marketing Letters*, 16 (3-4):415–428, 2005.

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

P. E. Green and V. Srinivasan. Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice. *Journal of Marketing*, 54(4):3–19, 1990.

T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer, 2016.

A. E. Hoerl and R. W. Kennard. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 12(1):55–67, 1970.

S. Holm. A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979.

M. Hong, X. Wang, M. Razaviyayn, and Z.-Q. Luo. Iteration Complexity Analysis of Block Coordinate Descent Methods. *Mathematical Programming, Series A*, 163(1-2):85–114, 2017.

J. Huber. Conjoint Analysis: How We Got Here and Where We Are (An Update). *Sawtooth Software Research Paper Series*, pages 1–14, 2004.

R. Iyengar and K. Jedidi. A Conjoint Model of Quantity Discounts. *Marketing Science*, 31 (2):334–350, 2012.

R. Iyengar and Y.-H. Park. Shareable Coupons. *Working Paper*, 2018.

G. James, D. Witten, T. Hastie, and R. Tibshirani. *An Introduction to Statistical Learning with Applications in R.* Springer, 2017.

J. Jung, R. Bapna, J. Golden, and T. Sun. Altruism Pays! Towards Optimal Call-to-Action for Online Referral: A Randomized Field Experiment. *Working Paper*, 2017.

W. A. Kamakura and G. J. Russell. A Probabilistic Choice Model for Market Segmentation and Elasticity Structure. *Journal of Marketing Research*, 26(4):379–390, 1989.

J. G. Kim, U. Menzefricke, and F. M. Feinberg. Assessing Heterogeneity in Discrete Choice Models Using a Dirichlet Process Prior. *Review of Marketing Science*, 2(1):1–39, 2004.

L. J. Kornish and Q. Li. Optimal Referral Bonuses with Asymmetric Information: Firm-Offered and Interpersonal Incentives. *Marketing Science*, 29(1):108–121, 2010.

S. Kube, M. A. Marechal, and C. Puppe. The Currency of Reciprocity: Gift Exchange in the Workplace. *American Economic Review*, 102(4):1644–1662, 2012.

V. Kumar, J. A. Petersen, and R. P. Leone. Driving Profitability by Encouraging Customer Referrals: Who, When, and How. *Journal of Marketing*, 74(5):1–17, 2010.

P. J. Lenk, W. S. DeSarbo, P. E. Green, and M. R. Young. Hierarchical Bayes Conjoint Analysis: Recovery of Partworth Heterogeneity from Reduced Experimental Designs. *Marketing Science*, 15(2):173–191, 1996.

T. Lumley, P. Diehr, S. Emerson, and L. Chen. The Importance of the Normality Assumption in Large Public Health Data Sets. *Annual Review of Public Health*, 23:151–169, 2002.

S. Ma, D. Goldfarb, and L. Chen. Fixed Point and Bregman Iterative Methods for Matrix Rank Minimization. *Mathematical Programming, Series A*, 128(1-2):321–353, 2011.

J. D. Montgomery. Social Networks and Labor-Market Outcomes: Toward an Economic Analysis. *American Economic Review*, 81(5):1408–1418, 1991.

O. Netzer, J. M. Lattin, and V. Srinivasan. A Hidden Markov Model of Customer Relationship Dynamics. *Marketing Science*, 27(2):185–204, 2008.

A. Pallais and E. G. Sands. Why the Referential Treatment? Evidence from Field Experiments on Referrals. *Journal of Political Economy*, 124(6):1793–1828, 2016.

T. K. Pong, P. Tseng, S. Ji, and J. Ye. Trace Norm Regularization: Reformulations, Algorithms, and Multi-Task Learning. *SIAM Journal on Optimization*, 20(6):3465–3489, 2010.

P. E. Rossi and G. M. Allenby. A Bayesian Approach to Estimating Household Parameters. *Journal of Marketing Research*, 30(2):171–182, 1993.

P. E. Rossi, R. E. McCulloch, and G. M. Allenby. The Value of Purchase History Data in Target Marketing. *Marketing Science*, 15(4):321–340, 1996.

G. Ryu and L. Feick. A Penny for Your Thoughts: Referral Reward Programs and Referral Likelihood. *Journal of Marketing*, 71(1):84–94, 2007.

P. Schmitt, B. Skiera, and C. Van den Bulte. Referral Programs and Customer Value. *Journal of Marketing*, 75(1):46–59, 2011.

D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. Van der Linde. Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639, 2002.

R. Tibshirani. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 58(1):267–288, 1996.

K. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, 2009.

P. Tseng. Convergence of a Block Coordinate Descent Method for Nondifferentiable Minimization. *Journal of Optimization Theory and Applications*, 109(3):475–494, 2001.

C. Van den Bulte, E. Bayer, B. Skiera, and P. Schmitt. How Customer Referral Programs Turn Social Capital into Economic Capital. *Journal of Marketing Research*, 55(1):132–146, 2018.

V. Vapnik. *Statistical Learning Theory*. Wiley, 1998.

P. C. Verhoef, P. H. Franses, and J. C. Hoekstra. The Effect of Relational Constructs on Customer Referrals and Number of Services Purchased from a Multiservice Provider: Does Age of Relationship Matter? *Journal of the Academy of Marketing Science*, 30(3): 202–216, 2002.

D. R. Wittink and P. Cattin. Commercial Use of Conjoint Analysis: An Update. *Journal of Marketing*, 53(3):91–96, 1989.

H. Zou and T. Hastie. Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320, 2005.