Exploring Peer Effects Associated with User Retention in a Socially Connected Business

Yijun Chen
Imperial College Business School, Imperial College London, yijun.chen@imperial.ac.uk

Nitin Mehta
Rotman School of Management, University of Toronto , nitin.mehta@rotman.utoronto.ca

We investigate peer effects associated with a user's churn and subsequent retention of other users in the network. We use data from an upscale gym with information on exogenous churns of users that allows us to identify peer effects. We estimate a structural model of users' plan choice and usage decisions that allows for bidirectionality in peer effects, direct/indirect peer effects, heterogeneity in the degree of centrality of users, and dynamics in peer effects. Our main results are: (i) Unlike the churn of a low or medium centrality user, the churn of a high centrality user has a strong negative cascade effect on the usage and subscriptions of other users in the network over time. (ii) To stop this negative cascade effect and increase the retention in the rest of the network, the firm's decision on which of the central churner's peers to target with promotions depends on the trade-off between two forces: the individual level impact of the targeted promotion vs. the positive spillover effect to the rest of the network that can increase over time because of carryover effects of usage and subsequent social multiplier effects (that stem from bi-directionality in peer effects). Thus, retention strategies focused on short-term outcomes should prioritize targeting users with low centrality, while strategies aimed at long-term outcomes should prioritize targeting users with high centrality.

Key words: Peer Effect, Social Network, Customer Retention Management

1. Introduction

A product or service is socially connected if a consumer gets additional utility from consuming it with her peers such as colleagues, families and friends, as opposed to consuming it in isolation.[1] Socially connected products have become increasingly popular amongst consumers, and firms have actively responded to the trend. For example, Nike hosts an online social network, Nike Training Club, for users training together and sharing workouts with their peers. Peloton offers expensive in-door treadmills, stationary bikes, and a subscription service that connects users with their peers in live training sessions. A recent article in Harvard Business Review highlights how firms can build social features into their products and grow communities of consumers around their products as "a superior business model".[2]

[1] An article in Harvard Business Review (`https://hbr.org/2013/10/the-age-of-social-products`) uses the term "social products" to refer to products that are consumed by consumers together.

[2] Source: `https://hbr.org/2013/10/the-age-of-social-products`

In the context of socially connected products, peer effect is defined as the impact of the user's consumption on the consumption of her peers.[3] Users of socially connected products are typically part of a rich social network that has four important properties. First, the peer effects between two users are bidirectional. Bidirectionality means that the consumption decision of one user affects and is affected by the consumption decision of her peer. Further, both users recognize this and internalize these effects when making their consumption decision (Hartmann et al. 2008). Bidirectionality results in a social multiplier effect (Manski 1993) that intensifies the effect of a marketing intervention on the focal user. For instance, in the case of Peloton, if a focal user increases her consumption because of a targeted marketing intervention, she may influence her peers to increase their consumption as well, who will in turn influence the focal user to increase her consumption. This will create a positive feedback loop which will result in both the focal user and her peers to further increase their usage, till a steady state is achieved. As a result, the increase in consumption of the focal user in equilibrium will be larger than the increase in her consumption if there were no peer effects. Second, there is heterogeneity in terms of the degree of centrality of users in the network, where the degree of centrality of a user refers to the number of peers of that user (Nair et al. 2010, Iyengar et al. 2015). An increase in usage of a central user may lead to a larger overall increase in usage of her immediate peers as compared to an increase in usage of a peripheral user. Third, peer effects can be direct as well as indirect. That is, an increase in usage of a focal user can influence the usage of not only her immediate peers (direct effect) but also her peers' peers (indirect effect) through the effect on usage of the direct peers, and so on (Goldsmith-Pinkham and Imbens 2013). Thus, a change in usage of one consumer can potentially spread throughout the network. Fourth, an increase in the usage of a focal user in a given period can increase the consumption of her peers not just in that period, but also in future periods. Thus, the spillover effect of a one time change in consumption of a focal user can persist over time (Iyengar et al. 2015).

The above discussion suggests that when users are connected in a rich social network, businesses can benefit from offering a socially connected product. However, these businesses can also be at risk when some key users, who have multiple strong connections, leave the network. If that happens, then the bidirectional peer effects between the churner and her direct peers which previously helped strengthen the network, will immediately turn negative. This will lead to a reduction in consumption of not only the churner's direct peers but also her direct peers' peers and so on. Thus, depending on the position of the initial churner(s), negative effects can potentially spread through the network over time. A well-known example of this is MySpace, which was once the largest social networking site in the world. It lost over forty million unique visitors in one month in April 2008, and the

[3] In our paper, we will use 'user' and 'consumer' interchangeably, and also 'usage' and 'consumption' interchangeably.

attrition continued to exacerbate over time. As of Jan 2018, Myspace was ranked 4,153 in terms of total Web traffic, and 1,657 in the U.S.[4] An important reason for Myspace's decline is that Facebook had a better product. But what greatly accelerated the decline of MySpace is peer effects associated with churn since users were not able to see their friends anymore on MySpace (Tucker 2018).

This sets the stage for what we hope to achieve. Our goal is to study peer effects associated with a user's churn and the subsequent retention of users in a socially connected business. The specific research questions that we ask are: (i) How does the churn of a focal user impact the usage and retention of other users in the network over time, and how does that impact vary with respect to the position of the focal user in the network? (ii) What retention strategies can the firm devise to offset the negative impact of a user's churn on the usage and retention of other users in the network? Specifically, which of the direct peers of the focal user should the firm target with promotions to stop the negative impact of the focal user's churn?

To address these questions, we use data from an upscale gym in a mid-west city in the U.S. In this gym, users usually develop strong relationships with other members as they frequently exercise together in small group fitness classes, typically ranging from 3 to 8 participants. On average, there are 213 adult users each month in the gym, and they form a rich social network. Our data consists of membership plan choices and usage decisions of every active adult user in each month over a period of 7 years. The membership plan choice in a given month entails whether or not to subscribe to any plan during that month; and if the user chooses to subscribe, then whether she should subscribe to a basic plan or a premium plan. For each user who has purchased a monthly plan, we observe the number of gym visits in that month and the specific group classes she attended each day. Our dataset also has information on which users churned and their self-reported reasons for doing so. A significant number of these cases involve users moving to another city and having long-term illness or pregnancy. We use these cases to identify peer effects since these can be treated as churns that are exogenous to unobserved factors that can affect gym visits and plan choices of their peers.

Given the data, we use concepts from graph theory to characterize the network structure of all users in the gym. Amongst any two users, we operationalize their peer relationship at the gym (i.e., whether they are direct peers, and the strength of their relationship if they are direct peers) as a function of the number of prior joint visits between the two users to the same group classes in the gym. The rationale for this operationalization follows from our reduced form analysis, which we discuss in section 3.1.

Given the network structure, we then estimate a structural model of users' plan choice decisions (in each month, whether they should purchase a membership, and if so, which plan they should

[4] Source: `https://en.wikipedia.org/wiki/Myspace`.

choose) and usage decisions (how many times they should visit the gym in each month). There are several important features of our model. First, we model a user's usage utility in each period as a function of (a) contemporaneous usage of all her direct peers moderated by the strength of peer relationship with each of those direct peers, and (b) carryover effects in usage. Second, to identify the peer effects, we use exogenous churns in our data as instruments for visits of peers. Third, our model allows for bi-directionality of peer effects since the utility of any user i is modeled as a function of the usage of her direct peer j and vice versa. Fourth, as we will explain later in the paper, our model captures dynamics in spillover effects (that stem from an exogenous change in consumption of a user) through (i) carryover effects of usage in the utilities, and (ii) subsequent social multiplier effects that stem from bi-directionality in peer effects. The greater the magnitudes of (i) and (ii), the greater is the persistence of spillover effects over time.

Our key empirical results are as follows: (i) Peer effects within the network are sizable and statistically significant. These peer effects generate a social multiplier effect of 1.63. That is, if there was a marketing intervention that resulted in everyone in isolation increasing their gym visits by one, then through bidirectional peer effects, all users on average will further increase their visits by an additional 0.63 gym visits. (ii) There is significant heterogeneity in users' intrinsic usage levels at the gym. We find that there are 3×2 segments of users based on their intrinsic usage levels at the gym (high, medium, and low usage) and their carry-over effects from the last period (positive and negative carryover effects). (iii) 83% of users show strong positive carryover effects of usage i.e., lack of usage in the last period has a significant negative impact on usage in the present period.

Given the estimates of the structural model, we conduct two counterfactuals to address the two research questions. In the first counterfactual, we simulate the impact of a focal user's churn on the usage and subscription (retention) behavior of other users over time. We consider a network of 228 users and the initial churn of 3 types of focal users: focal users with high degree of centrality (85 direct peers), with medium degree of centrality (25 direct peers), and with low degree of centrality (2 direct peers). Our main findings are: (i) The churn of a high centrality user has a much higher spillover effect on the rest of the network than the churn of a medium or a low centrality user. Six months post churn, the churn of a low, medium, and high centrality users leads to a spillover loss of \$863, \$7,454, and \$66,926 in revenues to the gym respectively. (ii) When a low or a medium centrality user churns, most of the spillover effects happen in the same month as that of the churn, and after that, the subscriptions of other users stabilize. On the other hand, we see a strong negative cascade effect if a high centrality user churns - the average un-subscription rate of other users increases by 0.5% in 1 month, and by 7.2% in 6 months. The reason for the cascade effect follows from the carry over effects of usage and the subsequent negative social multiplier effects. This shows that if the

negative effect from a high centrality churner is not stopped early on, it can aggravate over time, making the problem much harder to deal with in the future.

Since the churn of a high centrality user can lead to a cascade effect, in the second counterfactual, we explore which of the direct peers of a high centrality churner should a firm target with promotions to help retain users in the network. We find that the conventional wisdom in the customer retention literature that firms should target promotions to customers with the highest lift in subscription probability or customer lifetime value (CLV) (Ascarza 2018, Lemmens and Gupta 2020) may not hold true in the context of a social network. And in the context of social networks, we find that the conventional wisdom in the product adoption literature that one should target a high centrality individual may not necessarily hold true when it comes to retention management. Our analysis shows that when it comes to retention management there are trade-offs involved in deciding which peers to target. On one hand, targeting high centrality peers can improve overall network retention through spillover effects – where these spillover effects increase over time because of carryover effects of usage and subsequent social multiplier effects. On the other hand, these high centrality peers, when targeted with a promotion, may not be the ones who experience the highest lift in CLV or subscription probability. In fact, we find a negative correlation between a user's network centrality and the lift in her subscription probability/CLV when targeted with a promotion. Thus, if we consider the short term effect of a targeted promotion, the firm is better off targeting peers with the highest increase in the subscription probability. On the other hand, since the spillover effects increase over time, we find that if we consider a long-term effect, then targeting a high centrality peer yields the best return for the gym.

2. Related Literature

Our research lies at the intersection of three streams of research: peer effects in the context of adoption, peer effects in the context of retention, and customer relationship management (CRM). We discuss these as follows.

The first stream of literature consists of papers that have studied peer effects in the context of adoption and information diffusion through the network. This stream has been widely studied in the literature; thus we will only discuss a small set of papers. Tucker (2008) studies the adoption of network communication technologies and finds that managers who are central in communication networks significantly influence adoption decisions, while ordinary workers have little impact. Iyengar et al. (2011) analyze product adoption in a physician network and show that social contagion occurs through network ties. Banerjee et al. (2013) show that the centrality of the first individuals informed about a microfinance program predicts eventual adoption in Indian village networks thus targeting high centrality individuals is a key strategy for maximizing the spread and adoption of

microfinance. Kumar and Sudhir (2025) study the effectiveness of different seeding strategies for word-of-mouth (WOM) marketing and product adoption. They find that selecting a random friend of a randomly chosen individual as a seed—thus leveraging higher-degree nodes without needing full network data—can significantly enhance WOM and adoption.

Our paper differs from the papers in the first stream in an important aspect. While these papers study peer effects in the context of adoption, we do so in the context of retention. The insights gained from studying consumer adoption differ significantly from those obtained from studying consumer retention in a social network in two respects. First, when dealing with adoption, the objective of targeting is to maximize diffusion or uptake across the network. In that case, spillovers from the targeted user to her direct/indirect peers matter the most, implying that the firm should target a high-centrality user with promotions. On the other hand, when dealing with retention, the objective of targeting is to minimize churn/maximize the lifetime value of the network. In that case, what matters is not just spillovers from the targeted user to other users, but also the increase in purchase probability of the targeted user. Thus, the decision on whom to target depends on the trade-off between the two forces: the increase in purchase probability of the targeted user and spillovers from the targeted user to other users. Since our research shows a negative correlation between a user's degree of centrality and the increase in her purchase probability when targeted with a promotion, it follows that targeting high (low) centrality users would yield a smaller (larger) increase in the purchase probability of the targeted user but a larger (smaller) spillover effect. Second, when it comes to adoption, targeting decisions remain relatively consistent regardless of whether the firm optimizes for short-term or long-term outcomes. This is not so for retention. In the case of retention, we find that the spillover effects are small initially but can amplify over time due to carry-over effects of usage and subsequent social multiplier effects, implying that the relative magnitudes of the two forces in the trade-off discussed earlier can flip when we consider short-term vs. long-term impact of targeting. Thus, targeting strategies focused on short-term (long-term) outcomes should prioritize targeting users with low (high) degree of centrality.

The second stream of literature consists of papers that have studied peer effects in the context of retention. To the best of our knowledge, no prior paper has explicitly studied peer effects in the context of retention. Thus, in this stream, we will discuss those papers that have studied peer effects in the context of usage, as usage can be considered as a proxy for retention. Nair et al. (2010) examine prescription drug usage and find that physicians' prescribing behavior is significantly influenced by opinion leaders within their reference group. Park and Manchanda (2015) study peer effects in casino gambling consumption and find that peer influence varies across individuals and is asymmetric within peer pairs. Therefore, casino managers can improve the effectiveness of their marketing by targeting more influential individuals in peer groups and by mitigating negative externalities on

non-targeted peers. Uetake and Yang (2019) investigate peer effects in the usage of a weight loss program and find that influence differs between average and top performers. Hartmann (2010) develops a model for the estimation and analysis of demand in the context of social interactions and finds significant social interaction effects in the purchasing decisions of golfers. Chen et al. (2021) propose a scalable structural discrete choice model that incorporates social influence within large-scale social networks. Their findings show that accounting for individual-specific parameters and the timing of marketing interventions significantly enhances the effectiveness of targeting influential consumers in social networks. Toker-Yildiz et al. (2017) find that online social interactions strongly drive repeat usage behavior and are more effective than monetary incentives when both are present. Iyengar et al. (2015) show that social contagion affects both the trial and the repeat prescribing of a new drug by physicians; thus marketers should design different strategies for trial versus repeat prescriptions—targeting high centrality, high-volume prescribers for trial, and focusing on immediate colleagues for repeat prescriptions.

Our paper differs from papers in the second stream in two respects. First, while the papers in the second stream have examined peer effects in the context of usage, we have examined peer effects in the context of both usage and churn. Even though usage can be a proxy for retention, it is important to have information on both usage and churn when studying retention management. This is for two reasons: (i) Low usage does not necessarily imply a high churn risk, as some users could be stable in their usage over time but do not consume much. Without churn data, it is difficult to distinguish between these low-usage stable users and other low-usage users who are indeed at risk for churn. (ii) There are cases, such as the one we have in our data, where firms care more about churn than usage. This would be when users subscribe to a service plan by paying a fee and do not need to pay for usage if it stays within the plan limit.

Second, compared to papers in the second stream, we examine peer effects in relatively more elaborate network structures where (i) there are direct and indirect peers, (ii) heterogeneity in the degree of centrality of users, (iii) dynamics in spillover effects, and (iv) bidirectionality in peer effects. Except for Chen et al. (2021), all the prior papers examined peer effects in relatively simple networks in which there are no indirect peers, and there are no dynamics in spillover effects. While papers such as Hartmann (2010) and Yang et al. (2006) allowed (explicitly or implicitly) for bidirectional peer effects, the networks they considered did not have heterogeneity in the degree of centrality of users.[5] Similarly, while Iyengar et al. (2015) and Toker-Yildiz et al. (2017) consider networks that allow for heterogeneity in the degree of centrality, they do not have bidirectional peer effects. Finally, while Chen et al. (2021) studied peer effects in a more elaborate social network, they did

[5] On the other hand, Nair et al. (2010) implicitly allowed for bidirectional peer effects, and the network they considered had heterogeneity in the degree of centrality of users

not consider dynamics in spillover effects. Allowing for all the aforementioned network ingredients (i) -(iv), which are present in more elaborate network structures, is crucial to address our research questions to uncover the fundamental trade-off we discussed earlier. For instance, if we do not allow for heterogeneity in users' degree of centrality, then we cannot address our research question on whether we should target high or low-centrality users to increase retention in the network, and we will not be able to uncover the fundamental trade-offs discussed earlier regarding whether to target low vs. high centrality users. If we do not allow for dynamics in spillover effects or if we do not allow for indirect peer effects, then we will not be able to answer our research question on how the churn of a user impacts the consumption of other direct and indirect peers over time, and we will not be able to uncover how the decision on which user to target would change when we consider short-term vs. long-term impact of targeting.

The last stream consists of papers in the CRM literature that have investigated proactive retention strategies, in terms of which customers should the firm target with promotions to improve overall retention.[6] Most papers in this area have investigated retention strategies in businesses that are not socially connected, which obviates the need to consider peer effects. For instance, Ascarza (2018) suggests that firms should target those customers with discounts who have a high lift in purchase probability after receiving discounts, and Lemmens and Gupta (2020) suggests that businesses should prioritize targeting customers with a high lift in customer lifetime value after receiving discounts. There are few exceptions that have investigated peer effects of targeted promotions, (e.g. Ascarza et al. 2017, Nitzan and Libai 2011, Godinho de Matos et al. 2018). These papers find that targeting customers has spillover effects on their direct peers that are not targeted. However, these papers do not investigate the implication of these spillover effects on which consumers should be targeted to improve overall retention and profitability. We extend the findings in this literature by showing that the conventional wisdom that firms should target promotions to customers with high churn risk, or those with high lift in purchase probability/CLV only works when we are dealing with traditional businesses that do not have social features/user networks. In the context of a socially connected business with strong peer effects, we show that the firm's decision on whether it should target users with the high lift in subscription probability/CLV or the users with high degree of centrality depends on the trade-off between two forces: individual level impact of the targeted promotion vs. the spillover effect to the rest of the network that can increase over time.

To summarize, our paper makes the following contribution to the literature. This is one of the first papers that has investigated the research question on which users a firm should target with promotions in the context of retention management in a rich social network. Such a network is

[6] Customer relationship management (CRM) in marketing research involves gathering and analyzing customer data to understand behavior and preferences, informing marketing strategies to enhance customer satisfaction and loyalty.

characterized by users having direct and indirect peers, heterogeneity in the degree of centrality of users, bidirectionality in peer effects, and dynamics in spillover effects. In such network structures, the decision on whom to target to maximize retention is not straightforward. It depends on the trade-off between two forces that we discussed earlier, the relative magnitudes of which can flip when we consider the short-term vs. long-term impact of targeting. We find that if we consider the short-term (long-term) impact of a targeted promotion, it is best to target users with low (high) degrees of centrality, and if we consider the long-term effect of the targeted promotion, we find that it is best to target users with high degrees of centrality.

The insights from our research can be generalized to other settings featuring similar network characteristics—specifically, where networks are: (i) medium-sized, (ii) dense and (iii) have strong bidirectional peer interactions. Examples of such settings include, but are not limited to, team-based gaming communities, online study groups within educational platforms, collaborative workplace cohorts, and various hobbyist clubs or interest-based groups.

3. Data, Social Network, and Reduced Form Evidence

Our data comes from an upscale gym in a Mid-West city in the U.S. It offers a variety of exercise sessions in group fitness and personal training format in categories such as sports training, pilates, and yoga. This gym is a typical socially connected business with a close-knit community. The gym offers a large number of group fitness classes. These group fitness classes are popular amongst consumers, especially the adults who mostly go these classes only. Most of the group fitness classes have between 3-8 users. The data consists of 10,967 users' plan choices and gym visits from 2012 to 2018 for seven years. For each consumer, we observe which days she visited the gym, the time of entry, and which group fitness classes she attended on those days.

There are two types of memberships offered by the gym: contractual and month-to-month. The contractual memberships have a duration of either 3, 6, or 12 months. Under month-to-month memberships, users can purchase two different plans: the basic and the premium plan. Basic plans include basic group-based fitness training and premium plans offer advanced classes such as Pilates in addition to basic group-based fitness trainings. For month-to-month memberships, a consumer makes the subscription decision at the beginning of each month, which includes the decision on whether to purchase the monthly membership and if so, which plan to choose. The consumer does not automatically get enrolled in the same plan in a given month as the plan she has in the last month. Instead, the consumer has to re-enroll herself in a plan each month. The gym has a flexible policy on canceling memberships and offers refunds even before the membership duration expires. In this study, we focus on consumers with month-to-month memberships for two reasons. First, the month-to-month memberships account for 88% of all purchases and 76% of users in the data. Second,

behavioral factors such as overestimation of future gym visits and switching costs (DellaVigna and Malmendier 2006) are less likely to exist in month-to-month memberships, as consumers make purchase decisions only based on the anticipated gym visits in each month.

In the data sample that we use for the analysis in sections 3.1 and 3.2, we exclude users with less than 6 gym visits or less than 3 months of subscriptions during their entire observed consumption history. These users are likely to be those who try the gym under a free trial and are thus less likely to form peer relationships with other users. Our sample consists of 2,118 users over seven years. These users visit the gym at least once after subscribing to the month-to-month membership. Table 2 presents summary statistics on users' monthly gym visits and subscriptions. Users visit the gym an average of 7.21 times per month, with a monthly subscription rate of 0.84. On average amongst subscribers, 78.8% of users subscribe to the basic plan, and the rest subscribe to the premium plan each month. The average monthly price weighted by the observed frequency of purchases is \$76.5 for the basic plan and \$348.5 for the premium plan. The prices of both plans vary over time - the number of price changes is roughly 3 to 4 times each year. The standard deviation of prices is \$18.9 for the basic plan and \$175.0 for the premium plan over the seven years of our data. Table 3 shows the plan choice probabilities in the current month conditional on plan choices in the prior month. The switching probability is 3.2% from basic to premium and 11.9% reversely.

Our dataset also has information on which users have churned and their self-reported reasons for doing so – where "churn" in our context refers to a user leaving the gym for a long time (with a majority of users never coming back to the gym during the span of the data. On the other hand, "un-subscription" refers to not subscribing to any monthly membership plan in a month. A user who unsubscribes in a given month would subscribe again in future months).[7] There are a total of 1,314 churns that are observed in the full data. We find 86 (6.5%) of these users' churns stem from a range of personal and individual circumstances, including childcare needs, caring for ill family members, personal health issues, and both domestic and international relocation. Given the personal nature of these reasons, it is unlikely that they correlate with unobserved factors affecting the subscription or visit patterns of other users. Additionally, the relatively stable percentage of these churns due to relocation over the years suggests that these are not influenced by external economic shifts or city-wide events such as major employer exit that might cause a mass exodus in any specific year. Thus, these 86 churns can be considered exogenous to unobserved factors that can affect the gym visits and plan choices of their peers. These churns are spread out evenly throughout the 7 year period. In Table 1, we have summarized the reasons for these churns.

[7] There is a potential concern that the user may have the incentive to misreport the reason for monetary reasons. For instance, the user may report that she is moving out of town to get the gym to refund some of the membership fees. We believe this incentive to misreport is low because the gym has a flexible policy in terms of allowing members to cancel their memberships and get refunds even before the membership duration expires. Thus, regardless of their self-reported reason for the churn, the members get their refund when they leave.

Table 1 Details of Reasons for Exogenous Churn

Category	Examples	Percentage
Family	Childcare, Sick family	37%
Personal Illness	Surgery, cancer, general medical issues	5%
Personal Injury	Broken arm/wrist, back injury	13%
Pregnancy	Expecting, maternity	4%
Relocation	International(Spain, Australia, German, etc)	39%
	Domestic (New York, Atlanta, California, Chicago, etc)	

We utilize these exogenous churn cases to identify peer effects, as they can be considered independent of unobserved factors that affect gym usage and subscription choices of other users. In contrast, other reasons for churn such as a favorite trainer leaving the gym, or a user switching to another gym, are likely to be endogenous to usage and subscription decisions of other users. The rest of this section is arranged as follows. In section 3.1, we discuss how we operationalize peer relationships amongst users in the gym based on their joint visits to group classes over time. Based on the operationalization of the peer relationships, we then construct the social network of users in the gym in section 3.2. This constructed social network will then be used as an input into the structural model in section 4.

Table 2 Summary Statistics: Monthly Visit and Subscription

Variable	Min	Max	Mean	SD
Monthly Visit	0	71	7.21	8.06
Monthly Subscription	0	1	0.84	0.36

Table 3 Conditional Monthly Plan Choice Probabilities

Prior month/Current Month	No plan	Basic	Premium
No plan	55.7%	38.4%	5.9%
Basic	10.5%	86.1%	3.2%
Premium	5.9%	11.9%	82.1%

3.1. Operationalization of Peer Relationships at the Gym

The operationalization of peer relationships at the gym entails defining peer relationships in the context of small group fitness training, characterizing the direct peers of each user, and the strengths of the relationships between all pairs of direct peers. In this section, we will discuss how we infer both aspects of peer relationships at the gym. To do so we will first operationalize the latent peer relationship strength between any two users, regardless of whether they are direct peers or not; and given that, we will then characterize whether those two users are direct peers or not, and what is the strength of their peer relationship if they are direct peers.

According to information provided by the boutique gym that specializes in class-based training, users join the gym primarily to attend fitness classes led by trainers rather than doing individual solo workouts. As we mentioned earlier, these group classes are small, with typically 3-8 members in each group. This intimate class setting fosters a shared experience among attendees. Previous research in psychology and fitness training (e.g. Wayment and McDonald 2017, Nikolajsen et al.

2021, Crane et al. 2017) has shown that users may develop a close bonding with each other from the shared experiences in small group settings, and that the close bonding may in turn lead to benefits that include heightened motivation and increased adherence to fitness routines. This suggests that even if two users may not be peers before they joined a group fitness class, the greater the number of joint visits between the two users in small group fitness classes, the greater will be the bonding between the two users, and the greater will be impact of one user's attendance (or churn) in the group class on the other user's participation.

We thus define peer relationships (in the context of the small group classes at the gym) as the bonding that two users have formed with each other through shared experiences in the group classes at the gym. We define the latent peer relationship strength from user j to user i, m_{ijt} as the extent to which user i's visits to the gym are affected by user j's contemporaneous visits to the gym. We operationalize the latent peer relationship strength, m_{ijt}, as follows. We first define g_{ijt} as the total number of user j's joint visits with user i to the same small group classes prior to period t. Given this we define m_{ijt} as the share of user j's joint visits in the same group classes with user i amongst all of user i's joint visits in group classes with all members at the gym prior to period t, which results in $m_{ijt} = \frac{g_{ijt}}{\Sigma_{k=1,k\neq i}^{N} g_{ikt}}$. Note that in our data, each time a user attends a gym class, we observe the specific class they attend, including the timestamp, and the room. Because we have attendance records for all gym users, we can identify which users are in the same class at any given time. Therefore, in our operationalization of the latent peer relationships, a joint gym visit is not simply defined as two users being at the gym simultaneously; rather, it is based on those users attending the same group fitness training classes, which typically consist of 3-8 participants per session.[8]

There are four reasons for this operationalization. First, as discussed earlier, frequent shared experiences in group classes could foster a social environment where users may know each other and become peers inside the gym; and the more frequently the users share the same group classes, the more likely they are to become peers the gym. Therefore, even if users initially attended classes together because of reasons such as homophily (i.e., similar exercise preferences or schedules) or co-incidental effects, the repeated shared experiences can lead to the formation of peer relationships over time. And since bonding from shared experiences may require time and multiple interactions, we have defined m_{ijt} on a continuous scale from 0 to 1, where higher values indicate stronger

[8] This distinction is important. In our study, we operationalize peer relationships based on users' joint attendance to the same group sessions in the past - where joint attendance to a "group session" means participating in the same structured activity (e.g., yoga or pilates) in the same physical room during the same designated time slot. To construct the network, we measure the number of times two users have attended the same session together under these conditions. This operationalization captures direct shared exercise experience in a small group setting of 3-8 participants. This is in contrast to measuring simultaneous gym visits, which would only indicate that two users were present somewhere in the gym facility during overlapping time periods and could be in entirely different group sessions without any meaningful interaction.

connections based on the number of joint visits to group classes. This method allows us to capture the varying intensity of peer relationships based on the number of shared experiences and addresses the concern about the potential weaker connections among less frequent gym attendees.

Second, we utilize external information on peer relationships outside the gym for a subset of users to show that our operationalization of peer relationships can reasonably proxy peer relationships outside the gym. For 4% of our users, we have information that indicates whether any two users are real-world peers, through referrals (user who joined the gym upon recommendation from an existing member) or whether they are coworkers. We calculated the latent peer relationship strengths, m_{ijt}, between these real-world peers and found that in 93% of all observed periods, these real-world peers have latent peer relationship strengths in the highest and second-highest quartiles of all latent peer relationship strengths. This suggests that real world peers have an increased likelihood of having a high frequency of joint visits to the same group classes, which in turn suggests that it is reasonable to use past joint visits to the same group classes as a proxy for peer relationships at the gym. The details of the analysis are provided in section A of the Web Appendix.

The third rationale follows form the reduced form analysis which we have discussed in section B of the Web Appendix. In this reduced from analysis, we show that our operationalization of the latent peer relationship strength, m_{ijt}, significantly moderates the impact of other user j's visits to the gym on the number of visits of the focal user i to the gym - i.e., the greater the value of m_{ijt}, and thus the greater will be the impact of user j's usage (or number of visits to the gym) on the usage decisions of user i. Specifically, in the regression, we split m_{ijt} into four quartiles. We find that while the effect of user j's gym visits on visits of the focal user i is insignificant when m_{ijt} falls in the bottom two quartiles, it is positive and significant when m_{ijt} falls in the upper two quartiles (with the effect being the highest for the uppermost quartile). Now, if user i's and j's past joint visits to the same classes were merely due to homophily or co-incidental effects and they did not foster bonding/peer relationships, then m_{ijt} would not positively moderate the impact of user j's gym visits on user i's gym visits.

Finally, operationalizing peer relationships between any two users i and j in period t (and thereby the network structure in period t) based on their past joint visits could run into endogeneity issues when modeling user i's visit decision. To see why that is so, note that in the presence of peer effects between users i and j, user i's visit decision in period t will be a function of user j's concurrent visit decision and the latent peer relationship strength, m_{ijt}. If we operationalize m_{ijt} based on the past joint visits of users i and j, it would imply that m_{ijt} will be a function of the lagged dependent variables, namely, user i's past visits in periods $s = 1..t-1$. Now if the errors in the visit decisions are correlated over time, it could result in the peer relationship strength, m_{ijt}, to be correlated with the error in user i's visit decision in period t, thereby leading to endogeneity. We have addressed this

issue in section C of the Web Appendix, where we show using simulated data that the correlation between m_{ijt} and the error in the user i's visit decision in period t is less than 0.01 even when the serial correlation between the errors is as high as 0.95. This suggests that it is reasonable to assume that the peer relationship strength between any two users in period t is exogenous to the users' decisions in period t, conditional on all other observed covariates in period t.

3.2. Network Structure at the Gym

The findings of the regression discussed in section 3.1 suggest that we can use the median of the latent peer relationships as a threshold to define the two elements of the peer relationship between users i and j (whether any two users i and j are direct peers, and the strength of their peer relationship). User j will be user i's direct peer in period t (i.e., user j's actions will directly impact user i) if m_{ijt} lies above the median, otherwise, they will be indirect peers. Regarding the strength of the peer relationship using m_{ijt}, if user j is user i's indirect peer, then the strength of the relationship from user j to user i in period t will be 0; however, if user j is user i's direct peer, then the strength of the relationship from user j to user i in period t will be $m_{ijt} = \frac{g_{ijt}}{\Sigma_{k=1,k\neq i}^{N} g_{ikt}}$, which is the share of user j's prior joint visits to the same class with user i amongst all of user i's joint visits in the same classes with all i's peers till period t. Given this operationalization of the two elements of the peer relationship between any two users i and j, we next use the concept of adjacency matrix $\mathbf{\Omega}_t$ from graph theory (Boccaletti et al. 2006) to succinctly capture both elements of the peer relationships amongst all pairs of users in the network. For a social network with n_t users, $\mathbf{\Omega}_t$ is a $n_t \times n_t$ square matrix. All the diagonal elements in this matrix are 0, and the off-diagonal element in row i and column j, Ω_{ijt}, captures the strength of the relationship from user j to user i in period t (i.e., in the context of the peer effect of user j's actions on usage decision of user i). If user j is not user i's direct peer (i.e., if user j's actions do not directly impact user i), then Ω_{ijt} takes a value of 0. And if user j is user i's direct peer, then Ω_{ijt} will take the value of $m_{ijt} = \frac{g_{ijt}}{\Sigma_{k=1,k\neq i}^{N} g_{ikt}}$.

The adjacency matrix $\mathbf{\Omega}_t$ summarizes the entire network structure. We will use it as an input into our structural model in section 4, where we will model the usage utility of user i as a function of the usage of each of her direct peers j weighted by Ω_{ijt}, which is the peer relationship strength from user j to user i. $\mathbf{\Omega}_t$ has three important properties. First, the connection strengths between users i and j can be asymmetric. That is, Ω_{ijt} need not be the same as Ω_{jit}, which means that the impact of user j on user i's behavior can be different from the impact of user i on user j's behavior. This can be understood by considering two users i and j, where user i frequently visits the gym, but user j does not. User i is likely to have joint classes with users other than j, while user j is less likely to have many joint classes with users other than i. Thus, user i, who is a frequent user with many peer relationships, will have a stronger impact on the behavior of user j, who is an infrequent user with

peer relationships than vice versa. Second, $\mathbf{\Omega}_t$ captures heterogeneity in peer relationships across users with different degrees of centrality. A high centrality user i will have more non-zero connection strengths with the other users than a low centrality user j. That is, there will be more non-zero Ω_{ikt} in row i than Ω_{jkt} in row j of the adjacency matrix $\mathbf{\Omega}_t$. Third, $\mathbf{\Omega}_t$ captures the indirect peer relationship. If j is not i's direct peer ($\Omega_{ijt} = 0$), but is direct peer to k ($\Omega_{kjt} > 0$), and k is i's direct peer ($\Omega_{ikt} > 0$), then j is i's indirect peer.

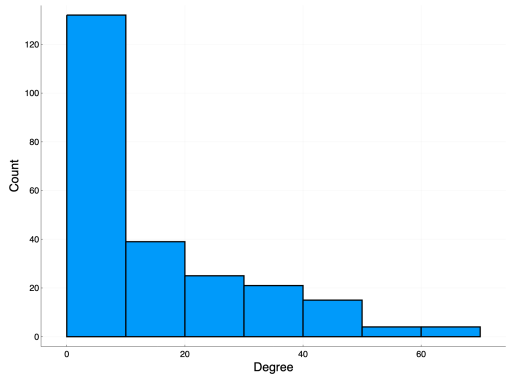
In the remaining part of this study, we focus on adult users and exclude teenagers for three reasons. First, adult users at the gym are much larger in number as compared to teenagers and they generate 75% of the revenue for the gym. Adult users also attend different classes than teenagers - adults usually go to non-sport group classes, while teenagers mostly go to sports training sessions. Second, teenagers and their peers are often teammates on sports teams, where visit and purchase decisions are likely made by their coaches or team managers rather than by the teenagers themselves. Additionally, they are observed to collectively join and leave the gym, likely influenced by common, unobservable factors and interdependent decision-making. Due to the lack of exogenous shocks, it is not feasible to identify peer effects among teenagers. Third, adults and teenagers have separate social networks with no connections between them, primarily due to their distinct gym class preferences. Therefore, it is reasonable to assume that there is no peer effect between teenagers and adults; thus omitting teenagers from the analysis is unlikely to bias the estimates of peer effects among adults. We have 1491 adult users in total over the period from 2012 to 2018, and we will use this sample of adult users for the structural model estimation.

We next present some key metrics used in graph theory to characterize the network of adult users in the gym. The metrics are size, diameter, and degree centrality.[9] On average, there are 213 users in the adult network each month. The diameter refers to the length of the longest path inside the network. The average diameter in the adult network is 6.5, which means that on average, a user could reach any other user in the network through indirect "friend of my friend" relationships of fewer than 7 users. The degree centrality represents the number of direct peers a user has, and it captures the peer influence of that user in the network. Figure 1 shows that the distribution of the degrees of adult users (in March 2013) is skewed. A few users have many direct peers. For example, the long right tail of the figure shows that only 8 out of 240 users have more than 50 peers in the gym. Those users have significantly more peers than the others, and we refer to them as "high centrality users". Those on the left side of the distribution who have a relatively low degree are referred to as "low centrality users". The rest of the users in the middle are referred to as "medium

[9] The adjacency matrix embeds the necessary information to calculate the above metrics. For instance, the degree centrality of a user can be constructed by counting the number of non-zero elements of the corresponding column. Paths can be inferred from the multiplication of the adjacency matrix. The shortest path will be the one with the largest product. One can calculate the diameter of the network in a similar way.

centrality users".[10] The key network metrics in the constructed adult network are similar to the metrics of real-world social networks reported in the prior literature (Watts and Strogatz 1998).

Figure 1 Network Degree Distribution



Note: Figure presents the degree distribution of all nodes of the adult network in the gym at March 2013.

Before we proceed, it is important to point out the following regarding the network structure that we specified in this section and the structural model that we will discuss next. In our study, we specify the network structure in any period t in a reduced form manner using the adjacency matrix, $\mathbf{\Omega}_t$. The adjacency matrix will then enter the structural model, which will allow us to capture how the user makes gym visits and plan choice/subscription decisions under peer influence. Thus, in our structural model, we do not jointly model the endogenous network formation process (in terms of how users form peer networks over time) along with users' subscription and visit decisions in each period. Instead, we only model users' visit and subscription decisions in each period t conditional on the observed network structure, $\mathbf{\Omega}_t$, that we have specified in this section. Said differently, in our structural model, we assume that the observed network structure in period t is exogenous to the users' decisions in period t, conditional on all other observed covariates in period t. As we will discuss later in section 5.1, our assumption of the exogeneity of the network structure in period t to the users' decision visits in that period is not a source of concern. Similar to what we saw in the reduced form regression discussed in section 3.1 (where we showed that the latent peer relationship strengths in period t can be considered exogenous to the users' visit decisions in period t), we will discuss in section 5.1 why it is reasonable to assume that the peer relationship strengths between all users in period t are exogenous to the users' decisions in period t.

4. Model

In this section, we specify how a user makes gym visits and plan subscription decisions under peer influence. Denote N_t as the set of adults in the network, n_t as the number of adults in the network,

[10] The cutoff for low, medium, and high centrality is 50%, 90% percentile of the degree.

and $\boldsymbol{P}_t$ as the vector of plan prices in period t. For user i, $\boldsymbol{X}_{it}$ is a vector of individual-time-specific characteristics, and $\boldsymbol{X}_t$ is the collection of $\boldsymbol{X}_{it}$ for all users. The adjacency matrix $\boldsymbol{\Omega}_t$, is a $n_t \times n_t$ matrix, which represents the network structure in period t. Its operationalization follows from the discussion in section 3.2.

The timeline of the decision process is the following: at the beginning of each period (month) t, a user i makes her plan choice $D_{it} \in \{H, L, 0\}$, where H refers to the premium plan, L to the basic plan, and 0 to the un-subscription choice (but the user can subscribe again in the next period). We assume that at the beginning of period t, the network structure $\boldsymbol{\Omega}_t$, is public information, which means that every user knows her direct and indirect peers. This is a reasonable assumption in our context because the users of the gym have been members for many years and would know of their peers' peers through frequent social interactions. We assume that users are myopic. This is for two reasons. First, the subscription is month-to-month rather than annual. Thus, users are not locked in for multiple months when they subscribe. Second, the gym does not automatically renew a user's subscription every month. Instead, the user has to make a subscription decision at the beginning of each month. This requires users to actively decide monthly whether to subscribe and, if so, which plan to select. Thus, it is not the case where the user has to make a decision once when she subscribes to a plan, and then does not incur any cost (because of automatic renewals) till she changes her subscription decision. Our data shows that on average 13.9% of basic subscribers and 17.9% of premium subscribers change their plan each month.

The plan choice depends on the user's expectation of how often she will attend gym classes in the month. Given the plan choice of either H or L, the user decides on the number of gym visits during the month. A user's plan choice and visit decision also depend on the number of gym visits of her peers, as shown in the reduced form results in section 3. Since the same would apply to her peers, it follows that both the user and her peer's decisions are affected by the plan choices and the number of gym visits of each other. Thus, in our model, each user makes the plan choice and gym visit decisions based on her expectations of her peers' decisions in the same period. We assume that users have rational expectations, which means that in equilibrium, each user can accurately predict the expected visit frequency and subscription probabilities of his/her peers. This assumption is reasonable in our context because peers in this study are users with frequent and consistent joint gym visits, so users would typically have good knowledge of their peers' gym visits and subscription preferences through repeated interactions.

The rest of this section is arranged as follows. In section 4.1, we specify a user's visit decision to the gym and the plan choice decision under peer influence. In section 4.2, we specify the equilibrium plan choices and visits across all users assuming rational expectations. In section 4.3, we specify the user heterogeneity. In section 4.4, we discuss how our model captures the key features of peer effects in socially connected businesses that we discussed in section 1.

4.1. Usage and Plan Choice Decisions

We first specify a user's visit decision (i.e., number of visits to the gym in a period[11]) conditional on the plan choice and then specify her plan choice decision. We specify a user i's utility of visiting the gym in period t conditional on choosing the plan H or L as

$$U_{it}^{V}(\mathit{Visit}_{it}^{d}|D_{it}=d,\boldsymbol{X}_t,\boldsymbol{\Omega}_t)=(g^{d}(\boldsymbol{X}_{it})+\beta\Sigma_{j\in N_t,j\neq i}\Omega_{ijt}\,\mathit{Visit}_{jt}+e_{it})\,\mathit{Visit}_{it}^{d}-\frac{1}{2}(\mathit{Visit}_{it}^{d})^2 \quad (1)$$

In equation (1), Visit_{it}^{d} represents the number of gym visits of user i in period t conditional on the chosen plan d, Visit_{jt} represents the number of gym visits of user j, and the term $g^{d}(\boldsymbol{X}_{it})$ is a function of individual and time specific variables that affect the user i's preference to visit the gym at period t. We will provide details of it later in section 4.3. The negative quadratic term $-\frac{1}{2}(\mathit{Visit}_{it}^{d})^2$ captures the diminishing marginal utility as the user increases her gym visits. The coefficient of this term is normalized to $-\frac{1}{2}$ for model identification. The term $\Sigma_{j\in N_t,j\neq i}\Omega_{ijt}\,\mathit{Visit}_{jt}$ captures the peer effects from gym visits of direct peers of user i in period t. In this term, the gym visits of user i's direct peers are weighted by their respective connection strengths with the user in period t, Ω_{ijt}. Note that the elements of the adjacency matrix $\boldsymbol{\Omega}_t$ for all periods $t=1\cdots T$ in the estimation sample are estimated outside the structural model based on the discussion in section 3.2. If $\beta>0$, peers' visits will have a positive impact on her visit utility. By weighting the peer visits by the connection strength Ω_{ijt}, our model captures heterogeneity in peer effects in a reduced form manner. It is important to clarify that in equation (1), the parameter β accounts for the unilateral effect of the peer's gym attendance on a user's utility from gym visits. However, when calculating the equilibrium (more details in section 4.2), we account for the reciprocal nature of peer effects. This includes not only the influence a focal user exerts on her peers but also the influence her peers exert on her, and how users internalize peer effect in both directions. Consequently, the equilibrium reflects the bidirectional peer influence within the network.

The term e_{it} is the random error that represents the unobserved factors that affect user i's usage of the gym in period t. We assume these errors are distributed with mean 0 and are IID across users and over time.[12] When deciding on the number of gym visits in period t, we assume that user i knows her own usage error e_{it}, other users' preference variables $\boldsymbol{X}_{jt}$, and the network structure $\boldsymbol{\Omega}_t$. However, the user i in period t does not know the usage errors e_{jt} of her peers. Instead, she only knows the distribution of those errors. The reason for making this assumption is that even

[11] Note that we define the peer relationship strength based on the joint visits of users i and j in the same gym classes in the past. However, the user's action that we are modeling is the user's visit to the gym, and not the user's visit to a specific class in the gym. In section D of the web appendix, we have discussed why we do that and why this is not a source of concern.

[12] The independence assumption over time implies that the user i cannot use the last period's error e_{it-1} to update her belief of e_{it}. This helps greatly simplify the expression for the expected gym visits in period t when the user makes the subscription decision.

though the community at the gym is fairly close-knit, it is unlikely that each user will have complete information about all the factors affecting her peers and their peers' peers visit decisions.

Since users are myopic in their decision-making, we can derive the optimal number of visits to the gym in period t conditional on having chosen a plan by solving the first-order condition in equation (1) as

$$Visit_{it}^{d*} = g^d(\boldsymbol{X}_{it}) + \beta\Sigma_{j\in N_t, j\neq i}\Omega_{ijt}\, Visit_{jt} + e_{it} \tag{2}$$

We next specify the user's plan choice decision, which she makes at the beginning of period t before her usage in that period is realized. Since the usage errors e_{it} are only realized after the user makes her plan choice decision in period t, the user i would only know the distribution of her own usage error e_{it}, as well as the usage errors e_{jt} of other users when she makes her plan choice decision. Thus when making the plan choice decision, a user forms expectations on the number of gym visits she will make in that period, as well as the number of gym visits her direct peers will make in that period. Based on these assumptions, we specify the user i's expected utility of subscribing plan d at the beginning of period t as her expected visit utility deducted by the disutility from paying the price as follows:

$$U_{it}^d(D_{it} = d|\boldsymbol{X}_t, \boldsymbol{\Omega}_t) = E[(g^d(\boldsymbol{X}_{it}) + \beta\Sigma_{j\in N_t, j\neq i}\Omega_{ijt}\, Visit_{jt} + e_{it})\, Visit^d it - \frac{1}{2}(\, Visit_{it}^d)^2] - \alpha_i P_t^d + \tilde{\epsilon}_{it}^d \tag{3}$$

where P_t^d is the price of the plan, and $\tilde{\epsilon}_{it}^d$ is the plan choice error that captures the impact of unobserved factors related to the dis-utility of purchasing plan d in period t. We allow the price coefficient α_i to differ among users (more on that in section 4.3). Replacing $Visit_{it}^d$ by $Visit_{it}^{d*}$ in equation (2) and using the formula $E(x^2) = (E(x))^2 + var(x)$, we can simplify equation (3) as

$$U_{it}^d(D_{it} = d|\boldsymbol{X}_t, \boldsymbol{\Omega}_t) = \frac{1}{2}(g^d(\boldsymbol{X}_{it}) + \beta\Sigma_{j\in N_t, j\neq i}\Omega_{ijt}E(\, Visit_{jt}))^2 + \frac{1}{2}\overline{var}^d - \alpha P_t^d + \epsilon_{it}^d \tag{4}$$

In the last term of equation (4), $\epsilon_{it}^d = \frac{1}{2}(var(\, Visit_{it}^{d*}) - \overline{var}^d) + \tilde{\epsilon}_{it}^d$, where $var(\, Visit_{it}^{d*}) = var(\beta\Sigma_{j\in N_t, j\neq i}\Omega_{ijt}\, Visit_{jt} + e_{it})$ is the variance of i's optimal number of gym visit under plan d, and $\overline{var}^d$ is the average variance of gym visits under plan d across all users. If the user does not subscribe to any plan, her expected utility is

$$U_{it}^0(D_{it} = 0|\boldsymbol{X}_t, \boldsymbol{\Omega}_t) = \epsilon_{it}^0 \tag{5}$$

We assume that the plan choice errors ϵ_{it}^d and ϵ_{it}^0, are IID with each other, across periods and across users. Furthermore, we assume that they are independent from the usage errors e_{it} in equation (2).[13]

[13] To test for the robustness of this assumption, we have done a simulation exercise in which we first simulated a dataset using our structural model with different values of the correlations in the visit and plan choice errors (corr =0.2, 0.4, 0.6), and then estimated our model in which we assume that this correlation is zero. The estimation results show that our model was able to accurately recover the original parameters for all simulated datasets. This suggests that it is reasonable to assume that the visit and plan choice errors are not correlated. The details of this exercise are given in section E of the web appendix.

This assumption is made for computational ease and it also follows from the fact that the visit shocks e_{it} and the plan shocks ϵ_{it}^d capture different facets of the consumer utility - while the visit shocks e_{it} are related to the utility of consuming gym classes, the purchase shocks captured in ϵ_{it}^d are related to the disutility of having the membership. We allow users to have imperfect information about the plan-specific errors of their peers. Thus we assume that when making a plan choice decision in period t, while user i knows the value of her plan shocks ϵ_{it}^d, she only knows the distribution of the plan choice errors of her peers, but not the exact values. Assuming that the plan choice errors ϵ_{it}^d and ϵ_{it}^0 have the standard Gumbel distribution, we get the following expression for the probability of user i subscribing plan d in period t:

$$Pr(D_{it} = d|\boldsymbol{X}_t, \boldsymbol{\Omega}_t) = \frac{exp(\frac{1}{2}(g^d(\boldsymbol{X}_{it}) + \beta\Sigma_{j\in N_t, j\neq i}\Omega_{ijt}E(\,Visit_{jt}))^2 + \frac{1}{2}\overline{var}^d - \alpha P_t^d)}{1 + \Sigma_{d\prime=\{H,L\}}exp(\frac{1}{2}(g^{d\prime}(\boldsymbol{X}_{it}) + \beta\Sigma_{j\in N_t, j\neq i}\Omega_{ijt}E(\,Visit_{jt}))^2 + \frac{1}{2}\overline{var}^{d\prime} - \alpha P_t^{d\prime})} \tag{6}$$

There are two points to note regarding equation (6). First, the plan choice probability, $Pr(D_{it} = d|\boldsymbol{X}_t, \boldsymbol{\Omega}_t)$ does not contain the usage errors e_{it}. The reason is that the user does not know the usage errors e_{it} when she makes the subscription decision. Moreover, the plan choice errors ϵ_{it}^d and ϵ_{it}^0 are independent of the usage errors e_{it} – thus, the knowledge of the plan choice errors ϵ_{it}^d and ϵ_{it}^0 when making the plan choice decision in period t does not provide any information about the usage errors e_{it}. Second, the plan choice probability of user i in equation (6) is a function of the expected number of visits of her direct peers j, $E(\,Visit_{jt})$, which itself is a function of the subscription and visit decisions of user i. Similarly in equation (2), the number of gym visits of user i in period t are a function of the visits of her direct peers. This leads to a simultaneous set of equations. We next discuss how we solve the simultaneous system of equations in (2) and (6) to get the optimal plan choices and visit decisions of all users in equilibrium.

4.2. Rational Expectations in Equilibrium

We assume that each user is rational in their plan choice and visit decisions, making optimal choices based on rational expectations of their own visits. We assume that individual and time-specific variables X_{it} and the network structure $\boldsymbol{\Omega}_t$ are public information known to all users. However, the visit error e_{it} and the choice error ϵ_{it} are private information. Although a user doesn't know the exact realizations of their peers' visit and choice errors, they are assumed to know the distribution of these errors. Given this framework, we employ the concept of a Bayesian Nash Equilibrium. Each user makes optimal decisions based on rational expectations about the visit behaviors of their direct peers. Users form unbiased beliefs about the expected visit frequencies and subscription probabilities of their peers at equilibrium. The rational expectations assumption is needed for identification - without this assumption, we would not be able to identify what expectations users form about the activities of their peers.

Letting $Pr(D_{it}=H|\boldsymbol{X}_t,\boldsymbol{\Omega}_t)\equiv\Phi_{it}^H$ and $Pr(D_{it}=L|\boldsymbol{X}_t,\boldsymbol{\Omega}_t)\equiv\Phi_{it}^L$, the following three conditions characterize the equilibrium expectations:

$$\Phi_{it}^H=\frac{exp(\frac{1}{2}(g^H(\boldsymbol{X}_{it})+\beta\Sigma_{j\in N_t,j\neq i}\Omega_{ijt}E(\,Visit_{jt}))^2+\frac{1}{2}\overline{var}^H-\alpha P_t^H)}{1+\Sigma_{d\prime=\{H,L\}}exp(\frac{1}{2}(g^{d\prime}(\boldsymbol{X}_{it})+\beta\Sigma_{j\in N_t,j\neq i}\Omega_{ijt}E(\,Visit_{jt}))^2+\frac{1}{2}\overline{var}^{d\prime}-\alpha P_t^{d\prime})} \tag{7a}$$

$$\Phi_{it}^L=\frac{exp(\frac{1}{2}(g^L(\boldsymbol{X}_{it})+\beta\Sigma_{j\in N_t,j\neq i}\Omega_{ijt}E(\,Visit_{jt}))^2+\frac{1}{2}\overline{var}^L-\alpha P_t^L)}{1+\Sigma_{d\prime=\{H,L\}}exp(\frac{1}{2}(g^{d\prime}(\boldsymbol{X}_{it})+\beta\Sigma_{j\in N_t,j\neq i}\Omega_{ijt}E(\,Visit_{jt}))^2+\frac{1}{2}\overline{var}^{d\prime}-\alpha P_t^{d\prime})} \tag{7b}$$

$$\begin{aligned}E(\,Visit_{it})=&\Phi_{it}^H(g^H(\boldsymbol{X}_{it})+\beta\Sigma_{j\in N_t,j\neq i}\Omega_{ijt}E(\,Visit_{jt}))\\&+\Phi_{it}^L(g^L(\boldsymbol{X}_{it})+\beta\Sigma_{j\in N_t,j\neq i}\Omega_{ijt}E(\,Visit_{jt}))\end{aligned} \tag{7c}$$

Equation (7c) represents a set of $i=1,\cdots,n_t$ simultaneous equations, in which each focal user i' s visit decision to the gym is a function of the user i's plan choice probabilities for the premium and basic plans, and the expected gym visits of all of user i's peers, where $j\neq i$. This means that the visit decisions are modeled as an equilibrium, in which each focal user i forms rational expectation of all her peers' gym visits $j\neq i$, where the focal user's visit decision to the gym is the sum of the plan choice probability of each plan (H,L) times the expected visit under the plan. Note that since the subscription probabilities do not contain the usage errors e_{it}, we have $E(e_{it}|D_{it}=d)=0$ under the zero mean assumption. This gives us $E(\,Visit_{it}|D_{it}=d)=g^d(\boldsymbol{X}_{it})+\beta\Sigma_{j\in N_t,j\neq i}\Omega_{ijt}E(\,Visit_{jt})$ in equation (7c). Equations (7a) and (7b) represent a system of equations across all focal users $i=1,\cdots,n_t$, which specify each focal user i's plan choice probabilities as functions of expected peers' gym visits under utility maximization of focal user i. It means that conditional on the rational expectations of all peers' gym visit $E(\,Visit_{jt})$, the optimal plan choice of focal user i is achieved when the total utility from purchasing the plan (H or L) and visiting the gym under that plan is maximized. Jointly, the equilibrium outcome is the collection of the plan choice probabilities and the expected visits of all users $i=1,\cdots,n_t$, which are determined by the $3*n_t$ system of equations in (7a), (7b), and (7c) across all users $i=1,\cdots,n_t$.

Let $\boldsymbol{EY}_t$ be the collection of $(\Phi_{it}^H,\Phi_{it}^L,E(\,Visit_{it}))$ of every user, we can write equations (7a) to (7c) in a general form as follows.

$$\boldsymbol{EY}_t=F(\boldsymbol{EY}_t) \tag{8}$$

When the absolute value of β is less than 1 ($|\beta|<1$), there exists a unique equilibrium. Conversely, if the absolute value exceeds 1 ($|\beta|>1$), multiple equilibria emerge, leading to an unrealistic scenario where peer influence becomes so dominant that all users would collectively subscribe or unsubscribe, contradicting the observed data patterns. To align with empirical evidence, we thus constrain β to the range $\|\beta|<1$ in which the equilibrium is unique. This restriction allows us to

apply the contraction mapping method for calculating each user's plan choice probabilities (Φ_{it}^H, Φ_{it}^L) and the expected gym visits ($E(Visit_{it})$). In the next section, we discuss how this approach facilitates our model's estimation. Furthermore, in section F of the Web Appendix, we not only prove the theoretical existence and uniqueness of the equilibrium under this parameter constraint but also employ numerical simulations to validate it empirically. We also run simulations to compute equilibrium outcomes when β falls outside the constraint range to demonstrate that the resultant equilibrium behaviors significantly diverge from the data observations.

4.3. Heterogeneity in Usage Preferences

In this section, we specify the function $g^d(\boldsymbol{X}_{it})$ in the usage utility in equation (1). As discussed earlier, this function allows for the heterogeneity in exercise "states" across users and the evolution in exercise states over time. Individual users exhibit different gym visit behaviors in our data. Part of the heterogeneity across users and over time could be explained by demographics and time-fixed effects. Individuals may also differ in their time-invariant preferences and their time-varying exercise states, both of which are unobserved to researchers. We specify $g^d(\boldsymbol{X}_{it})$ as follows.

$$g^d(\boldsymbol{X}_{it}) = \boldsymbol{\gamma}_1 \boldsymbol{R}_{it} + \delta^H \mathbf{1}\{D_{it} = H\} + S_{it} \tag{9}$$

where $X_{it} = (\boldsymbol{R}_{it}, D_{it}, S_{it})$. In this specification, $\boldsymbol{R}_{it}$ is a vector of user i's age, gender, year, and season indicators at time t, D_{it} is the plan choice, and δ^H represents the incremental utility of the premium plan relative to the basic plan. The first two components on the right side of the equation capture user heterogeneity due to observed factors in data. The last component S_{it} represents the unobserved exercise state that is heterogeneous across individuals and varies over time. We next discuss the specification of the unobserved exercise state, S_{it}.

Researchers have typically used latent class segmentation or random coefficients to model unobserved heterogeneity. These methods are difficult to adopt in our context because of social interactions. To understand why that is so, consider our gym setting, where we have 200 users in the network. Since users interact with each other over time and get to know each other's preferences, the unobserved preference of any given user would be unobserved by the researcher but would be observed by other users in the network. Now consider the case where the researcher uses latent class segmentation with two latent types: H and L (where H means high preference and L means low preference). Since each user i knows all other users' types, each user i in the network would make her usage and subscription decisions in equilibrium based on her complete knowledge of the types of all other 199 users directly or indirectly connected to her. In contrast, the researcher only knows probabilistically whether each user is H or L type. Thus, the researcher would have to model usage and subscription decisions of each user i conditional on the unobserved preference type of not just

that user i, but also conditional on the unobserved preference types of all other 199 users who are directly or indirectly connected to user i. This means that the researcher would have to first condition on the types of all users $j = 1..200$ in the network and then compute the visit and subscription decisions of all users in equilibrium for that particular combination of types in the network. Since each user can be of two types, we will get 2^{200} different combinations of types in the network. This means that we would have to compute the visits and subscription decisions in equilibrium for 2^{200} combinations of types of networks. This makes it computationally infeasible.

To allow for heterogeneity in S_{it}, we use a different approach. We employ a clustering algorithm that deterministically assigns each user to a specific segment based on their observed gym visit patterns. The clustering algorithm analyzes users' historical gym behavior data (including metrics like average monthly visits and visit correlation patterns). It assigns each user to exactly one of the identified segments with certainty. This deterministic segmentation approach differs from probabilistic latent class models but serves a similar function - it allows us to account for systematic differences in preferences and behaviors across different types of users. It resolves the challenge discussed in the previous paragraph by deterministically assigning each user to a specific segment based on their observed behaviors. This eliminates the need to compute a large number of equilibria.

The clustering is done as follows. First, users are segmented based on visit levels, using average monthly visits and the number of active months as inputs for the clustering algorithm. The number of clusters is determined by the average silhouette score of all users, which measures how similar an object is to its own cluster compared to other clusters. For each user i, represented by a vector capturing multiple dimensions of gym visit behaviors, the silhouette score $s(i)$ is defined as: $s(i) = \frac{b(i)-a(i)}{\max\{a(i),b(i)\}}$, where $a(i)$ is the average Euclidean distance between user i and all other users within the same cluster, and $b(i)$ is the lowest average distance between user i and users in any other cluster. We gradually increase the number of clusters, compute the average silhouette score for each, and select the optimal number of clusters where the silhouette score is at its maximum. This yields three clusters: high, medium, and low visit levels. Next, we segment users based on visit trends and use the correlation between gym visits in the past and current months, the average and absolute difference in visits between the two months, as inputs. This yields two segments. The first one shows a positive correlation of gym visits over consecutive months, indicating that users have a consistent trend of either increasing or decreasing visits to the gym. We call this segment the "consistent" type. The other segment exhibits a negative correlation of gym visits between two consecutive months, suggesting that users switch between an active state, under which they visit the gym frequently, and an inactive state, under which they reduce visits or unsubscribe. We call it the "on-and-off" type. Combining the visit levels and visit trends segmentations gives us six user segments. Table 4 shows summary statistics of these segments. The "Consistent-Low" segment is the largest one with

approximately half of the users. A user in this segment only visits the gym 3.6 times in a period and the average correlation between visits in two consecutive months is 0.12. The second and third largest segments are "Consistent-Medium" and "Consistent-High" users, respectively. In the latter segment, a user exercises in the gym almost every day. The next three segments are the "on-and-off" types with about 15% of users altogether. The average number of monthly visits of these segments is comparable with the "consistent" type, but the average correlations between monthly visits are significantly negative.[14]

Table 4 Gym Behaviors Clustering Result

Segment	Percentage	Avg Visit	Avg Correlation
Consistent-Low	46.6%	3.60	0.12
Consistent-Medium	26.7%	12.2	0.12
Consistent-High	12.0%	27.4	0.00
On-and-off-Low	7.1%	6.17	-0.05
On-and-off-Medium	6.8%	14.0	-0.10
On-and-off-High	0.8%	25.8	-0.27

Note: Each row represents one gym behavior type; columns 2,3,4 represent the percentage of users, the average visit, and the average correlations between consecutive visits in each cluster.

Based on this segmentation, we specify the unobserved exercise state, S_{it} for each segment as the sum of two parts. The first part varies over time, and it captures the exercise state that is influenced by past exercises. We represent it as a product of a time-varying state variable $\tilde{S}_{it}$ and a time-invariant visit level variable λ_i. The second part is a constant λ_i^C that captures the exercise state that is not influenced by past exercises.

$$S_{it} = \tilde{S}_{it} \cdot \lambda_i + \lambda_i^C \tag{10}$$

In the equation (10), λ_i and λ_i^C are the segment-specific parameters based on the visit level. The state variable $\tilde{S}_{it}$ is restricted within $[0, 1]$, and how it varies over time represents the segmentation based on the visit trend. For the consistent type, the specification of $\tilde{S}_{it}$ is:

$$\tilde{S}_{it} = \gamma_2^{CN} \tilde{S}_{it-1} + (1 - \gamma_2^{CN})[\mathbf{1}\{Visit_{it-1} > Visit_{it-2}\} + \tilde{S}_{it-1} \cdot \mathbf{1}\{Visit_{it-1} = Visit_{it-2}\}] \tag{11}$$

where γ_2^{CN} is a parameter that is restricted between $[0, 1]$. It represents the carry-over effect from the exercise state in the previous month to the current month. $\mathbf{1}\{Visit_{it-1} > Visit_{it-2}\}$ is an indicator variable that equals 1 if the number of gym visits increases in the prior month, and 0 otherwise. When the number of gym visits increases, the exercise state of the consistent user will further increase by $(1 - \gamma_2^{CN})$. When the number of gym visits remains the same, i.e., when $\mathbf{1}\{Visit_{it-1} = Visit_{it-2}\}$ equals 1, the exercise state will remain unchanged. The exercise state will decrease when the visits

[14] Please see section G the Web Appendix for additional analysis that supports the use of cluster analysis to segment the users.

decrease. Under this specification, consistent users can have an increasing, constant, or decreasing trend of gym visits across months. For on-and-off type, the specification of $\tilde{S}_{it}$ is:

$$\tilde{S}_{it} = \gamma_2^{ON}(1-\tilde{S}_{it-1}) + (1-\gamma_2^{ON}) \cdot \mathbf{1}\{Visit_{it-1} < Visit_{it-2} \; or \; Visit_{it-1} = 0\} \tag{12}$$

In this equation, γ_2^{ON} is the carry-over effect parameter that is also restricted between $[0,1]$. It is multiplied by $(1-\tilde{S}_{it-1})$, implying that the higher the exercise state was in the past month, the lower the exercise state will be in the current month. The state will increase by $(1-\gamma_2^{ON})$ under two conditions: visit decreases or no gym visits in the prior month. Under this specification, after an "off" period with reduced visits, a user's exercise state will increase, switching to the "on" state, resulting in increased visits.[15]

In equations (10) - (12), the parameters λ's and γ's and the state variables $\tilde{S}_{it}$ will be estimated together with all the other parameters in the structural model. Details are discussed later in section 5.

Our clustering approach effectively categorizes individuals based on their exercise levels and trends. By incorporating the unobserved preferences of users in S_{it}, we can account for correlated preferences between users i and j (i.e. homophily). Specifically, the parameter λ_i^C in equation (10) represents an individual's time-invariant exercise preference. When users i and j have high values of λ_i^C and λ_j^C, their exercise preferences become positively correlated. This correlation in preferences, in turn, increases the probability of joint gym visits. Additionally, equation (10) includes a time-varying component $\tilde{S}_{it} \cdot \lambda_i$, which allows exercise preferences to change over time. Finally note that while incorporating S_{it} may not completely control for correlated preferences, this is not a concern for identifying peer effects. This is because identifying the peer effects parameter β relies on the exogenous churns (details in section 5.1).

4.4. Model Properties

In this section, we discuss how our model captures the key properties of a socially connected business that we discussed in section 1: direct and indirect peer effects, dynamics in spillover effects, heterogeneity in peer effects, and bi-directionality of peer effects.

In equations (7a)-(7c), since the peer relationship strength between user i and user j is positive (i.e. $\Omega_{ijt} > 0$) only if users i and j are direct peers, the plan choice and visit decisions of user i are connected with the decisions of her direct peers, whose decisions are also connected with their own direct peers. This implies that user i's decisions will also affect her indirect peers through her direct peers. The peer effect in our model is bidirectional because we jointly capture user i's decision under her peer j's influence and user j's decision under i's influence in the simultaneous

[15] Because the carry-over effect parameters γ_2^{CN} and γ_2^{ON} are restricted between $[0,1]$, as long as the initial state is restricted $[0,1]$, then as per equations (11) and (12), all the subsequent states will be bounded between 0 and 1.

system of equations (8). To see how this results in a social multiplier effect, suppose that because of an external shock, user i's gym visits were to increase by one visit in the absence of peer effects. Equation (2) shows that the gym visits of user j, who sees i as a peer, will increase by $\beta\Omega_{jit}$. If i also sees j as a peer, the increase will further increase i's gym visits by another factor of $\beta^2\Omega_{ijt}\Omega_{jit}$. This will go on until it converges, which means both i and j internalize the effects from each other. As j's gym visits increase, her other direct peers will also be influenced, leading to more multiplicative increases among users in the whole network. Therefore, in equilibrium, the total gym visits of user i can be significantly higher than the original one visit increase from user i in isolation.

The dynamics in the spillover effects stem from the carryover effects of usage in the utilities and the subsequent social multiplier effects. As a result of the carry-over effects, when a user increases his usage in a given period because of an exogenous shock, it increases the number of gym visits of her peers not only in period t but also in subsequent periods. To see why that is so, note that when a focal user increases her usage in period t, it increases the number of gym visits of her peers in period t because of the contemporaneous peer effects, β. Now since most users belong to the consistent type, an increase in gym visits of focal user j's peers in period t will in turn increase their exercise state in period $t+1$ due to carry-over effects in their utilities. As a result, there will be an increase in the usage of user j's peers in periods $t+1$ and onwards. This increase in usage utilities of user j's peers gets further compounded by the positive social multiplier effect that we discussed in the previous paragraph. Thus, these two factors (carryover effects and the social multiplier effects) can lead to persistence in spillover effects over time - and the greater the magnitudes of these two factors, the greater will be the persistence in spillover effects.

Finally, regarding heterogeneity in peer effects, recall from section 3.2 that we allow for the connection strengths to vary across different pairs of users. This in turn implies that peer effects will also vary across different pairs of users. Note that in our model, we have assumed that (i) heterogeneity in peer effects only stems from heterogeneity in the strength of peer relationships and not from heterogeneity in the coefficient of the peer effect, β, and (ii) the strength of the peer relationship between the two users is determined by their share of prior joint visits to the same group classes. We have already discussed the rationale for assumption (ii) in section 3.1. Regarding assumption (i), the reason why we do not model unobserved heterogeneity in β is that the curse of dimensionality arising from social interactions makes it computationally infeasible to model heterogeneity in peer effects using typical latent class and random coefficients models.

5. Model Estimation, Identification and Results

We estimate the structural model on the adult users over 7 years from 2012 to 2018. The parameters that we estimate are: (i) peer effect coefficient β, (ii) price sensitivity α, (iii) incremental utility

of the premium plan relative to the basic plan δ^H, (iv) coefficients of year and season indicators, and user's demographics in the usage utility, $\boldsymbol{\gamma}_1$, (v) the coefficients $\{\lambda, \lambda^C\}$ of the variables in the unobserved exercise state S_{it}, (vi) the carry-over effect coefficients γ_2 in the time-varying unobserved exercise state $\tilde{S}_{it}$. We allow for the exercise level parameters $\{\lambda, \lambda^C\}$, and price sensitivities α, to differ across the visit level segments. To capture heterogeneity in visit patterns, we allow the carry-over effect coefficient γ_2 to differ across consistent and on-off types. We assume the rest of the parameters, namely the peer effect coefficient β, the premium plan surplus δ^H, and the coefficients of age, gender, year, and season in the usage utility to be the same across all segments.

In estimating the peer effect coefficient β, addressing the identification challenge inherent in peer effect models is crucial. There are three factors that can confound the estimates of the peer effects and outline our strategies for addressing each. First, the reflection problem, as described by Manski (1993), arises from the simultaneity of behaviors among users and their peers. To tackle the reflection problem, we first employ a simultaneous equation system (7), with equilibrium conditions (8) capturing the simultaneity of peer influence. Given the simultaneity, the peer effect coefficient is identified from the exogenous churn. Essentially, the exogenous churn, such as a user relocating, acts as an exogenous shock to the network. This shock creates a differential impact across users: users closely connected to the churner experience a stronger effect than those less closely connected. This spatial pattern in how the exogenous shock propagates through the network - with stronger effects on close peers and weaker effects on distant peers - helps isolate the true peer effect from other confounding factors that arise from simultaneity. Second, homophily refers to the tendency of individuals with similar exercise behaviors and preferences to form peer connections. We address the homophily by employing a clustering method to control for unobserved individual heterogeneity in exercise preferences that explains peer relationships (see section 4.3). Third, co-incidental effects emerge from unobserved time-varying factors that are common across users, which is different from homophily. Recall from our discussion in section 4.1 that we do not model the co-incidental effects since assume the visit errors are IID across users. However, this should not bias the estimates of the peer effects since the identification of peer effects in our model stems from the exogenous instrumental variables. To verify this, in section H of the web appendix, we simulated two data sets using our structural model in which we assumed the visit errors to be correlated across users with correlations as 0.2 and 0.5. Following that we estimated our model on both datasets in which we assumed the visit errors are IID across users. We found that our model is able to accurately recover the parameters. This suggests that it is reasonable to assume the visit errors are IID across users.

We use GMM to estimate the parameters. We start with the equations (7a)-(7c) to get $\boldsymbol{EY}_t = (\Phi^H_{it}, \Phi^L_{it}, E(\mathit{Visit}_{it}))$, which is an $n_t \times 3$ matrix of predicted probabilities of plan subscriptions and number of visits across all users. Let $\boldsymbol{Y}_{it} = (1\{D_{it} = H\}, 1\{D_{it} = L\}, \mathit{Visit}_{it})$ be the observed plan

choice and visits. Thus, for each user i in period t, we get the prediction error, $\boldsymbol{\xi}_{it} = \boldsymbol{Y}_{it} - \boldsymbol{EY}_{it}$, as a 1×3 vector. For GMM estimation, our task is to find a set of instrumental variables $\boldsymbol{Z}_{it}$, with the dimension larger or equal to the number of parameters, so that we can construct the moment conditions $E(\boldsymbol{\xi}_{it}|\boldsymbol{Z}_{it}) = \mathbf{0}$. Estimating our model presents a challenge because the left-hand side (LHS) of equations (7a)-(7c) involves user i's decision probabilities, $\boldsymbol{EY}_{it}$, while the right-hand side (RHS) includes the decision probabilities of user i's peers j, denoted as $\boldsymbol{EY}_{jt}$. This setup leads to two key issues. The first issue is that the decision probabilities $\boldsymbol{EY}_{it}$ do not have a closed-form expression, making direct computation infeasible. To over come this issue, we numerically compute $\boldsymbol{EY}_t$ in equation (8) across all n_t users using contraction mapping, conditional on a given set of parameters. The second issue arises because the peer decision probabilities $\boldsymbol{EY}_{jt}$ on the RHS of equations (7a)-(7c) depend on the parameters we aim to estimate. This dependence introduces endogeneity, as $\boldsymbol{EY}_{jt}$ becomes correlated with the prediction error term $\boldsymbol{\xi}_{it}$ for user i. This endogeneity implies that, for each focal user i, we need to find instruments $\boldsymbol{M}_{jt}$ for $\boldsymbol{EY}_{jt}$ for all of i's direct peers j. This instrument must be such that it is correlated with user j's expected number of visits under peer influence and satisfies $E(\boldsymbol{\xi}_{it}|\boldsymbol{M}_{jt}) = \mathbf{0}$. To obtain such valid instruments for identifying peer effects, we leverage the observed exogenous churns in our data. In the next section, we will discuss the construction of these instruments, the identification strategy, and the estimation procedure.

5.1. Instrumental Variables, Identification and Estimation Procedure

Instrumental Variables for Identifying Peer Effects. For each direct peer j of the focal user i, we need a valid instrument M_{jt} that help identify the peer effect coefficient β. To get such valid instruments, we make use of the observed exogenous churns in our data. Since the churn is exogenous, it will be independent of the prediction error. Next note that when a user leaves the gym, it impacts the usage and subscription decisions of her direct peers - the stronger the connection strength between the churner and the direct peer, the greater will be the impact on that peer. For indirect peers, the churn also impacts their utility through the change in usage and subscription of the direct peers of the churner. Thus, the closer the distance between a user and the churner in the network, the stronger will be the impact on the user. The spatial characteristics of the network in terms of the distances between the churner and her direct and indirect peers will determine the variation in the impact of the exogenous churn across these peers.

To construct the instrument M_{jt} for the direct peer j of the focal user i, we use the following procedure. Assume user k is the potential churner, for user $k \neq i$ and $k \neq j$ we calculate the shortest distance between j and k. If the shortest distance is 1, (i.e., user j is user k's direct peer with $\Omega_{jkt} > 0$), we construct the intermediate variable $M_{jkt} = \mathbf{1}\{Churn_{kt}\}\Omega_{jkt}$, where Ω_{jkt} is the connection strength from k to j, and $\mathbf{1}\{Churn_{kt}\}$ is an indicator variable that is equal to 1 if k churns

exogenously in period t and 0 otherwise. If the distance from user j to user k is 2 (i.e. $\Omega_{jkt}=0$ but there is at least one path $k \to l \to j$ between k and j, as such $\Omega_{lkt}\Omega_{jlt}>0$), we choose the path with the largest $\Omega_{lkt}\Omega_{jlt}$, implying the "shortest" path, and $M_{jkt}=\mathbf{1}\{Churn_{kt}\}\Omega_{lkt}\Omega_{jlt}$

If the distance from k to j is larger than 2, we follow the above method and construct M_{jkt} by multiplying the Ω's in the shortest path. We repeat the above procedure for every user k in the network who is neither user i nor user j and then sum up M_{jkt} to construct the following instrument variable M_{jt} for peer j as $M_{jt}=\Sigma_{k\in N_t,k\neq i,j}M_{jkt}$

This way, we construct the instrumental variable M_{jt} for every direct peer $j \in N_t$ of focal user i. Since the peer effect term $\Sigma_{j\in N_t,j\neq i}\Omega_{ijt}E(\,Visit_{jt})$ on the RHS of equations (7a)-(7c) for focal user i is the average of all i's direct peer j's expected visit weighted by the peer relationship strength between i and j, we construct an instrumental variable IV_{it}^1 for the peer effect term by taking the relationship strength weighted average of the instrument M_{jt} for all i's direct peers as $IV_{it}^1=\Sigma_{j\in N_t}\Omega_{ijt}M_{jt}$

If peer effects are positive, the variable IV_{it}^1 is positively correlated with $\Sigma_{j\in N_t,j\neq i}\Omega_{ijt}E(\,Visit_{jt})$ in equations (7a) to (7c), thus the IV_{it}^1 satisfies the relevance condition. Since we use only exogenous churns, $\mathbf{1}\{Churn_{kt}\}$ in the instruments is independent of prediction errors. Regarding the exogeneity of the peer relationship strength terms Ω_{ijt} in the instruments, it is not as straightforward. This is because Ω_{jkt} is a function of the lagged dependent variables, namely, user i's past visits ($Visit_{is}$, for $s=1,...,t1$). Thus, if the visit errors in equation (2), are correlated over time, then the terms Ω_{jkt} will be correlated with the current visit error e_{it}. This would imply that the terms m_{ijt} and m_{jkt} could be endogenous if the errors e_{it} are serially correlated. To mitigate the concern that errors could be correlated over time for reasons such as users developing a habit for exercise, we model habit formation and heterogeneity in gym visit behavior in the term $g^d(\boldsymbol{X}_{it})$. Thus we would not expect the serial correlation in the errors to be large. To further assess the validity of our method, we have done a robustness test in section C of the Web Appendix in which we first simulated 4 datasets using our structural model with different values of the serial correlations in the errors e_{it} at 0.2, 0.4, 0.6, 0.95, and then estimated our model in which we assume that the serial correlation in e_{it} is zero. The estimation results show that our model was able to accurately recover the original parameters for all simulated datasets. This suggests that IV_{it}^1 satisfies the exclusion restriction and serves as a valid instrument.

Before we proceed, it is important to point out that the instruments we discussed above to identify the peer effects parameters are constructed from 86 exogenous churns in our data. These churns are distributed evenly throughout the time-span of the data, with about 12 churns per year (see Section I of the Web Appendix). There is also a spatial pattern in how exogenous churn impacts the network - the impact is stronger for users closer to the churn source (direct peers) and weaker for those further away (indirect peers). Thus, we used this spatial pattern when we constructed the

instruments above for the churner's direct and indirect peers. Doing so helps capture the attenuated impact of exogenous churn throughout the network.

Next note that peer users' visits can be impacted by not only churns in the current month but also those in the past months. We thus repeat the above procedures for churns in month $t-1$ to construct another instrument IV_{it}^2. We repeat the exercise until $t-5$. Our final set of instruments for $\Sigma_{j\in N_t, j\neq i}\Omega_{ijt}E(\mathit{Visit}_{jt})$ in equations (7a) to (7c) are $\boldsymbol{IV}_{it}=\{IV_{it}^1, IV_{it}^2, \cdots, IV_{it}^6\}$. For the first stage regression, the first stage (incremental) F-statistic of $\boldsymbol{IV}_{it}$ is 27.7, indicating that the constructed IV satisfies the relevance condition. In addition to $\boldsymbol{IV}_{it}$, we also include individual characteristics $\boldsymbol{X}_{it}$ defined earlier in section 4.3 which includes the age, gender, year, season indicators and plan prices $\boldsymbol{P}_t=\{P_t^H, P_t^L\}$ as instruments. Thus, we have the instrument vector for the focal user i in period t as $\boldsymbol{Z}_{it}=(\boldsymbol{X}_{it}, \boldsymbol{P}_t, \boldsymbol{IV}_{it})$.

Identification. The set of instrumental variables $\boldsymbol{Z}_{it}=(\boldsymbol{X}_{it}, \boldsymbol{P}_t, \boldsymbol{IV}_{it})$ inform us about the model identification. As discussed earlier, the identification of the peer effect parameter *beta* comes from the differential impact of the exogenous churns on the usage and subscription decisions across close and distant peers of the churners. The variation in $\boldsymbol{X}_{it}$ (which includes gender, age, seasons, and years) across users and/or time helps identify the parameters in $g^d(\boldsymbol{X}_{it})$. The price coefficient α is identified by the variation in the plan prices over time. To mitigate the concern of price endogeneity that can stem from seasonal demand shocks, we include year and season fixed effects in $\boldsymbol{X}_{it}$. To mitigate the price endogeneity that may be caused by personalized pricing, we use the average price paid for the plan weighted by the observed frequency of purchases by all the users each month across all plans. This way, there is no concern about price endogeneity at the individual level and the identification comes from the variations in average plan choices over time.

Estimation Procedure The evolution of exercise states in equations (11) and (12) comes with an initial value problem since we do not know S_{it} in the first period in our data. To address this problem, we use the first $T_0=6$ months as initialization periods. Since the unobserved time-varying state, $\tilde{S}_{i,1}$, takes a value between 0 and 1, we draw the unobserved time-varying state at the beginning of the initialization sample, $\tilde{S}_{i,-T_0}$, from a uniform distribution. Then conditional on the model parameters, we let $\tilde{S}_{it}$ evolve based on the gym visits in each period observed in the initialization sample until the end of the initialization sample, which is the first period in the estimation sample. As the model parameters change during the estimation, we accordingly adjust the unobserved time-varying state at the beginning of the estimation sample $\tilde{S}_{i,1}$.

There are two stages in our estimation procedure. We start with a set of trial model parameters $\boldsymbol{\Theta}^0$. Conditional on the parameters, we use contraction mapping in the inner stage to compute $\boldsymbol{EY}_{it}(\boldsymbol{\Theta}^0)=(\Phi_{it}^H(\boldsymbol{\Theta}^0), \Phi_{it}^L(\boldsymbol{\Theta}^0), E(\mathit{Visit}_{it})(\boldsymbol{\Theta}^0))$ for every active user in every month from equation (8). We then calculate $\boldsymbol{\xi}_{it}(\boldsymbol{\Theta}^0)=\boldsymbol{Y}_{it}-\boldsymbol{EY}_{it}(\boldsymbol{\Theta}^0)$, and based on that, we calculate the GMM criterion function $J(\boldsymbol{\Theta})=\left[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\boldsymbol{Z}_{it}'\boldsymbol{\xi}_{it}(\boldsymbol{\Theta})\right]' W \left[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\boldsymbol{Z}_{it}'\boldsymbol{\xi}_{it}(\boldsymbol{\Theta})\right]$.

In the outer stage, we use the differential evolution method to search for the optimal parameter vector $\boldsymbol{\Theta}^*$ that minimizes the criterion function. To enhance the efficiency of the estimator, the weighting matrix W is computed using an iterative approach. We start with the identity matrix and use the residuals to estimate the covariance matrix of the moment conditions, iteratively updating the weighting matrix W until convergence is achieved.

5.2. Estimation Results

The results on model fit are discussed in section J of the Web Appendix. The results for the adult network are presented in Table 5, showing a significant peer effect coefficient $\beta = 0.18$ at $p < 0.01$. To understand the magnitude of this peer effect, we examine the social multiplier. Consider a scenario where a marketing intervention motivates each active user in the network to increase their gym attendance by one visit per month. Calculating the new equilibrium under peer effect, the average gym visits per user increases by 1.63 visits per month. This indicates that alongside each user's individual one-visit increase, there's an additional 0.63 visits explained by peer influences. It's important to note that in the individual utility equation (1), β is specified to account for the unilateral effect of a peer j's gym attendance on a user i's value derived from gym visits. However, in calculating the equilibrium, we consider the reciprocal nature of these influences—how users affect each other. Thus, the calculated increase of 0.63 visits in the equilibrium reflects the bidirectional peer influence within the network.

Table 5 Model Estimation Result

Key Parameter	Estimate	Fixed Effect	Estimate
Peer Effect Coefficient β	0.18 (0.01)	Middle Age (Reference: Young)	1.76 (0.15)
Premium Plan δ^H	3.33 (0.06)	Old Age	1.68 (0.33)
Exercise State Level High λ_H	32.07 (0.36)	Male (Reference: Female)	-0.18 (0.10)
Exercise State Level Medium λ_M	7.99 (0.40)	Year 2013 (Reference: 2012)	-1.49 (0.49)
Exercise State Level Low λ_L	6.63 (0.45)	Year 2014	-1.45 (0.25)
Exercise State Level High Constant λ_H^C	4.45 (0.33)	Year 2015	-1.56 (0.15)
Exercise State Level Medium Constant λ_M^C	3.77 (0.41)	Year 2016	-1.12 (0.09)
Price Coefficient Level High α^H	-0.27 (0.03)	Year 2017	-0.89 (0.11)
Price Coefficient Level Medium α^M	-0.17 (0.01)	Year 2018	-2.08 (0.15)
Price Coefficient Level Low α^L	-0.26 (0.01)	Season Summer (Reference: Spring)	-0.36 (0.06)
Carry-over Effect (Consistent) γ_2^{CN}	0.56 (0.02)	Season Autumn	-0.02 (0.07)
Carry-over Effect (On-and-off) γ_2^{ON}	0.56 (0.09)	Season Winter	-0.12 (0.11)
Visit Variance (Premium) $\overline{var}^H$	22.0 (0.33)		
Visit Variance (Basic) $\overline{var}^L$	10.2 (0.32)		

Note: Standard errors are presented in parentheses; young age: age < 25; middle age: $25 \leq$ age < 55; old age: age ≥ 55

The price coefficients α of users in each of the three visit level segments H, M, and L are significantly negative. Medium-level users ($\alpha_M = -0.17$) are less price sensitive than high ($\alpha_H = -0.27$) and low-level users ($\alpha_L = -0.26$), implying that they are more responsive to price discounts (in terms of both subscription and upgrading to a premium membership) as compared to low-level and high-level users. The premium plan coefficient is positive and significant ($\delta^H = 3.33$), indicating that a user gets higher usage utility by going to classes covered by the premium plan. In terms of visits,

this translates to 3.33 more gym visits in a month if a user switches from the basic plan to the premium plan.

The exercise state level parameters λ's in equation (10) increase from 6.63 to 7.99 and further to 32.07 as we go from the low to high user segments. The higher the λ, the higher the utility a user gains per visit at all levels of the states between 0 and 1, and the more visits the user will have. The carry-over parameters γ's in equations (11) and (12) are both 0.56 for consistent and on-and-off segments. Even at similar values, the carry-over parameters would generate different visiting patterns for the two segments. For all the consistent segments (85.3% of all the users), the estimate of 0.56 indicates a strong state dependence: a user's exercise state $\tilde{S}_{it}$ will decrease from the maximum 1.0 to 0.56 if she does not visit the gym for 1 month and will further decrease to 0.03 in 6 months if she does not return to the gym. For the smaller on-and-off segment (14.7%), assuming that a user was in a high exercise state of 0.9 last month, and visited the gym more frequently than the month before. From equation (12), the exercise state will drop to 0.056 this month. In the next month, her exercise state will be high again, jumping back to $0.56 \times (1 - 0.056) + 0.44 = 0.96$. Therefore, gym visits of on-and-off users fluctuate every month. The estimated parameters for demographics, year, and season fixed effects are smaller in magnitudes than the exercise level parameters. This indicates that after controlling for the exercise state types, demographics and time factors are less important in explaining heterogeneity in visits across individuals and over time. We also estimated a positive correlation (0.57) between users' intrinsic visit levels, as captured by the exercise state level parameters, and their degree centrality in the network. This indicates that users with more peer connections also have higher intrinsic visit utilities.

6. Counterfactuals

We conduct two counterfactuals to address the two research questions that we discussed in section 1. In the first counterfactual, we address the research question of how does churn of a focal user impacts the usage and subscription behavior of other users in the network over time, and how that impact varies with the position of the focal user in the network. In the second counterfactual, we address the second research question, namely, what retention strategies can the firm devise to offset the negative impact of a focal user's churn on the usage of other users in the network? Specifically, which of the direct peers of the churner should the firm target with promotions to mitigate the negative impact of the churn of the focal user on the rest of the network?

6.1. Impact of Churn on Subscription and Usage of Peers over Time

In this counterfactual, we simulate the subscription and usage behavior of users in the network over time following an exogenous churn of the focal user. We select from data a month in which there are 228 active users in the network, and we simulate the subscription/usage behavior of these users

across four different scenarios: no churn, churn of a focal user with low degree of centrality, the churn of a focal user with medium degree of centrality, and churn of a focal user with high degree of centrality. The low centrality focal user that we consider has 2 peers, the medium centrality focal user has 25 peers, and the high centrality focal user has 85 peers.

In the simulations, we not only have to simulate the visit and subscription decisions of all users in equilibrium in each period $t = 1, \cdots, T$, but also how the network structure $\mathbf{\Omega}_t$ evolves over time. Simulating the users' decisions and the evolution of the network structure over periods $t = 1, \cdots, T$ entails two steps in each time period t: (i) given the network structure (or the peer relationship strengths between all pairs of users) in period t, we would need to simulate subscription decisions and the gym visit decisions to the gym of all users in period t, and (ii) given the network structure in period t and the simulated gym visit and subscription decisions of all users prior to period t, we would need to specify how the network structure evolves from period t to $t+1$. Note that the simulations of the gym visit and subscription decisions in (i) follow directly from the specification of equilibrium conditions in our structural model in equations (7a)-(7c). However, the specification of the evolution of the network structure in (ii) does not follow from our structural model. This is because in the structural model, we only model the visit and subscription decisions of all users in each period t conditional on the observed network structure in that period t. Thus, we use a reduced form approach to specify the evolution of the network, which is grounded in the data and aligns with the prior literature on social networks that has focused on how the network structure evolves with the presence/absence of shared activities between people (Granovetter 1973, Krackhardt et al. 2003, Burt 2002, Suitor et al. 1997). In section K of the web appendix, we have explained steps (i) and (ii). To simulate the gym visits and subscriptions of all users in the network, we randomly draw $100 * T$ usage errors as well as plan choice errors. We introduce a burnout period of 12 months for the visits and purchases to stabilize before introducing the churns. We then calculate the average unsubscription rate and the average gym visits of each user in each month for up to 6 months.

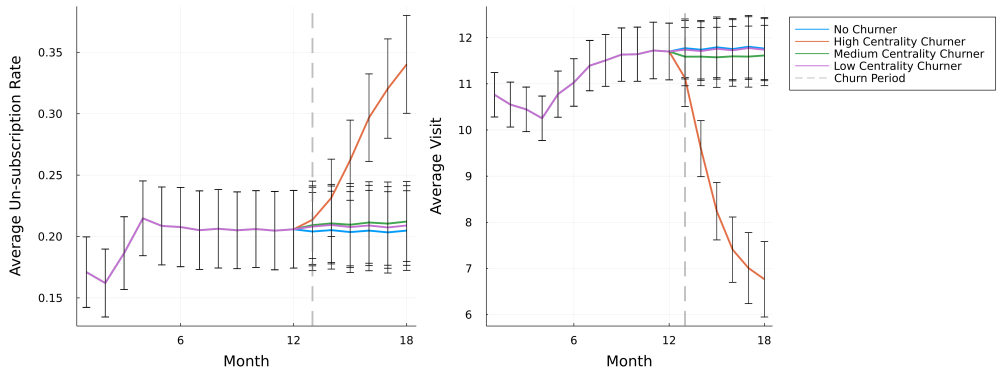
The results are shown in Figure 2 and Table 6. There are two main results here. First, the churn of a high centrality user has a much higher impact on the rest of the users than a medium or a low centrality user. Six months post churn, a low centrality churner leads to a revenue loss of \$1,217 to the gym (where \$354 is the revenue loss that stems from the user's own subscriptions and \$863 is the revenue loss that stems from the spillover effect on others' subscriptions), a medium centrality churner leads to a revenue loss of \$10,129 (where \$2,675 is the revenue loss that stems from the user's own subscriptions and \$7,454 is the revenue loss that stems from the spillover effect), and a high centrality churner leads to a revenue loss of \$69,342 (where \$2,677 is the revenue loss that stems from the user's own subscriptions and \$66,925 is the revenue loss that stems from the spillover effect). Second, unlike a low or a medium centrality churner, the negative impact of a high centrality

churner on usage and subscriptions increases over time. For the low and medium centrality churners, most of the changes in visits and un-subscriptions of the other users happen in the same month as that of the churn, and after that, the visits and subscriptions stabilize. On the other hand, for the case of the high centrality churner, the average visits of other users decrease by 0.7 in 1 month, and by 3.4 in 6 months. Similarly, the un-subscriptions increase by 0.5% in 1 month, and by 7.2% in 6 months. Thus we see a cascade effect when a high centrality user churns.

The second result on the cascade effect when a high centrality user leaves follows from the dynamics in the networks, which in turn stems from the carry-over effects in the exercise states and the subsequent negative social multiplier effect that stems from bidirectional peer effects. When the focal user churns from the network in month t, it leads to a decrease in visits and subscriptions of the focal user's direct and indirect peers in the same month t because of peer effects. Now since most users belong to the consistent type, the reduction in gym visits of users in period t will reduce the exercise state of those users in period $t+1$, which will lead to lead to decrease in their usage utilities in periods $t+1$ and onwards. This decrease in usage utility gets further compounded by the negative social multiplier effect that stems from bidirectional peer effects. These two factors thus lead to a substantial decrease in visits and subscriptions in periods $t+1$ and onwards.

This finding sends an important message: if the negative effects from a high centrality churner are not stopped early on, they can aggravate over time, making the problem much harder to deal with in the future. In the next counterfactual, we discuss targeted promotional strategies that a firm can employ to mitigate the negative impact of the churn of a high centrality user.

Figure 2 Post Churn Comparison, Churners with Low, Medium, and High Centrality



Note: Figure represents simulated unsubscription rate and average visit per month post churn; the first 12 months represent the burnout periods; the vertical dotted line represents the month of peer churn; each series represents one churner scenario; the vertical bars represent 90% confidence interval.

Table 6 The Impact of Churn, Low vs Medium vs High Centrality Churner

Churner Scenario	No Churner	Low Centrality	Medium Centrality	High Centrality
1M Visit	11.8	11.7	11.6	11.1
6M Visit	11.8	11.7	11.3	8.4
1M Unsubscription	20.4%	20.8%	20.9%	21.3%
6M Unsubscription	20.5%	21.0%	21.1%	27.7%
6M Revenue	189858	188641	179729	120256

Note: Each column represents one churner scenario. 1M represents the first month since churn; 6M represents the first six months since churn.

6.2. Mitigating the Negative Impact of Churn

In this counterfactual, we consider the case when a user with high degree of centrality (85 peers) churns, and we explore the targeted promotional offers that the gym can offer to the direct peers of the churner to mitigate the negative impact of churn on the other users in the network.

The promotional offers that we consider are discounts for the premium plan to the direct peers of the churner. As discussed in the estimation results section 5.2, a user is expected to have 3.33 more gym visits on average if they switch from the basic plan to the premium plan. By offering discounts on the premium plan, we aim to incentivize users to subscribe to it, thereby increasing their gym visits. Since the negative effect of churn persists over time, we consider promotions that are offered over multiple months. We consider six different scenarios. Scenario 1 represents the case when there was no churn at the beginning of the first month. Scenario 2 represents the case when the high centrality churner leaves at the beginning of the first month and the gym does not target anyone with promotions. Scenario 3 represents the baseline case in which the high centrality churner leaves the gym at the beginning of the first month and the gym targets all 85 direct peers of the high centrality churner with promotional offers for three months starting from the first month. This will serve as the benchmark for the other three counterfactual scenarios (scenarios 4-6) in which will investigate which subset of the direct peers of the high centrality churner should the firm target with promotions in the first 3 months after the churn if it only had 50% of the promotional budget it had under scenario 3. In scenario 3, the promotional discount is the same for all the direct peers of the churner, but the discount is dynamically adjusted each month. We set the discount in each month (starting from month 1 and then proceeding sequentially to month 3) such that it can offset the loss in the overall number of visits during that month because of the churn.[16] We have explained the procedure for calculating the discounts in each month in section L of the Web Appendix.

Our analysis for scenario 3 shows that discount levels (i.e., the percentage price discount for the premium plan offered to all the direct peers of the churner) needed to offset the loss in the overall number of visits during that month because of the churn are 27%, 33%, 15% from the first month to the third month respectively. The total promotional budget for scenario 3, calculated as the sum of all discounts given to direct peers of the churner over the 3-month campaign, is \$25K.

[16] We start with month 1 and then proceed onwards to months 2 and above because the promotion in month t will not only positively affect the demand in month t, but it will also affect the demand in months $t+1$ and onwards

We next compare the net revenues (revenues minus promotional budget) earned by the gym across scenarios 1-3 in the short run (1 month after the churn), and in the long run (6 months after the churn, which includes the 3 months of promotion after the churn, and 3 months of no promotion after that). The net revenues earned by the gym after the first month are \$56,394 for scenario 3, \$47,347 for scenario 2, and \$55,464 for the scenario 1. The net revenues earned by the gym over six months are \$196,252 for scenario 3, \$120,256 for scenario 2 and \$189,585 for the scenario 1. Thus we see that the net revenues earned by the gym in the first month and over six months are similar across scenarios 1 and 3, implying that targeting all the direct peers of the central churner with promotions offsets the loss in revenue that stems from the central churner leaving the network. Scenario 3 will serve as the benchmark promotional case for scenarios 4-6 in which we consider different targeting strategies under a budget constraint.

To explore the effectiveness of different targeting strategies under a budget constraint, we assume the gym still offers discounts for 3 months but has half of the required budget(i.e. \$12.5K out of the \$25K in the benchmark scenario 3). We study three different targeting scenarios. In scenario 4, we rank the direct peers of the churner based on their degree of centrality and prioritize targeting those with a higher degree of centrality. Since we offer the same level of discount as calculated in scenario 3 to each of the targeted users, with half the budget, we can offer discounts to half of the direct peers (i.e. 42 out of the total 85 direct peers). In scenarios 5 to 6, we separately prioritize the direct peers with high lift in purchase probability (as in Ascarza 2018), and high lift in CLV (as in Lemmens and Gupta 2020), and offer the same level of discount to the same number of 42 targeted users. Outcomes of all scenarios are shown in Table 7.

To evaluate the effectiveness of each targeting strategy, we look at the lift in revenue across the entire network under the targeting strategy. In the short term (one month into the campaign), targeting users with high lift in purchase probability, high lift in CLV and high centrality would respectively achieve 99%, 98.7% and 95.4% of the revenue that the firm would earn when targeting all users in scenario 3. This indicates that targeting users with high purchase probability or CLV lift is more effective initially. However, in the long term (six months after the start of the campaign), targeting users with the high lift in purchase probability, high lift in CLV and high centrality would respectively achieve 74%, 76%, and 87% of the full budget targeting scenario. This demonstrates that targeting users with high centrality becomes the more effective strategy over time.

While previous studies have shown the benefits of targeting high centrality users for product adoption (e.g. Kumar and Sudhir 2025), this strategy can not be directly applied to retention management and churn prevention. Unlike product adoption scenarios—where targeting high centrality users boosts the adoption of the product in the network—retention management involves the trade-offs between two forces, the relative magnitudes of which vary over time. The first force is the

increase in CLV/lift in subscription probability of the user who is targeted with the promotion. The second force is the spillover effect of the targeted promotion to the rest of the network, which increases over time because of carryover effect and the social multiplier effect that we discussed earlier. Next, note that users with a higher lift in purchase probability/CLV are not necessarily the ones with a high degree of centrality. For instance, the number of peer connections of users that are targeted is 18 in scenario 5 (high lift in purchase probabilities), and scenario 6 (high lift in CLV), and 28 in scenario 4 (high degree centrality). This implies that targeting users with a high lift in purchase probability/CLV will not generate substantial spillover effects – it will only serve to increase the performance of those targeted users. Since the spillover effects are not that large after the first month, it follows that if one is only interested in the short term effect of targeting, then it is most profitable to target the users with the highest lift in subscription probabilities. And regarding targeting a high centrality user, even though it may not substantially increase the subscription probability/revenue from that centrality user as compared to other targeting strategies in the short run, however, it yields the highest payoff for the firm in the long run (6 months).

Table 7 Counterfactual Results, Targeting Campaign Comparison

Scenario	1M Un-Subscription	6M Un-Subscription	1M Visit	6M Visit	1M Revenue	6M Revenue
No Churn	0.204	0.204	11.8	11.8	55464	189585
Churn Without Targeting	0.213	0.277	11.1	8.3	47347	120256
Full Budget Targeting	0.210	0.223	11.8	11.6	56394	196252
High Centrality Seeding	0.211	0.231	11.5	10.8	53828	172252
High Purchase Lift Seeding	0.212	0.256	11.4	9.4	55851	145213
High CLV Seeding	0.211	0.238	11.6	10.5	55690	150365

Note: 1M represents the month of churn and the start of the campaign; 6M represents the first six months since churn.

7. Conclusions and Directions for Future Research

In this paper, we address two research questions in the context of retention management in the rich network structure: (i) How does the churn of a focal user impact the usage and retention of other users in the network over time, and how does that impact vary with respect to the position of the focal user in the network? (ii) What retention strategies can the firm devise to offset the negative impact of a user's churn on the usage and retention of other users in the network? Specifically, which of the direct peers of the focal user should the firm target with promotions to stop the negative impact of the focal user's churn?

To address these questions, we use data from an upscale gym in a mid-west city. Our dataset also has information on which users have unsubscribed their memberships for an extended period and their self-reported reasons for doing so. A significant number of these cases involve users moving to another city and having long-term illness or pregnancy. We use these cases to identify peer effects since these can be treated as churns that are exogenous to unobserved factors that can affect the

gym visits and plan choices of their peers. Given the data, we use concepts from graph theory to construct the network structure of all adult users in the gym.

Given the network structure, we then estimate a structural model of users' plan choice decisions and usage decisions, We model a user's usage utility in each period as a function of the contemporaneous usage of all her direct peers moderated by the strength of peer relationship with each of those direct peers. The users internalize the contemporaneous peer effect when making their plan choices and consumption decisions. We find that the peer effects within the network of adult users are sizable and statistically significant. These peer effects generate a social multiplier effect of 1.63. That is, if there was an external shock that resulted in everyone in isolation increasing their gym visits by one, then through bidirectional peer effects, all users on average will further increase their visits by an additional 0.63 gym visits.

Given the estimates of the structural model, we run two counterfactuals to address the two research questions. Regarding the first question, we find that as compared to a low/medium centrality user's churn, a high centrality user's churn has a much higher cumulative negative impact across her peers' subscriptions/usage, which grows over time and spreads throughout the network. Regarding the second question, we find that the conventional wisdom that firms should target promotions to customers with higher lift in purchase probability/higher lift in CLV (Ascarza 2018, Lemmens and Gupta 2020) holds true when we are dealing with a traditional business that does not have a social network, or in a socially connected business when one only considers the short term impact of the targeted promotion. On the other hand, when we are dealing with a socially connected business with strong peer effects and if we consider the long term impact of the targeted promotion, the firms should instead target promotions to users with high degree of centrality.

Our paper suffers from some shortcomings which can be avenues for future research. First, we use a reduced form approach to specify the strengths of the peer relationships in the model equations instead of implementing an endogenous network formation model. Without modeling how users form networks endogenously, our approach does not capture how users become peers in the first place and how heterogeneity in peer relationships is generated. Second, we have used simulations to rule out the endogeneity that can result from using past joint visits of users to operationalize the network. The use of simulations is not foolproof. Thus, we cannot completely rule out the possibility of endogenous network formation. Third, we only capture peer relationships in terms of the share of prior joint visits inside the gym. Other unobserved factors could also contribute to peer relationships, such as users' interactions outside the gym. As explained in section 3.1, we find suggestive evidence supporting our operationalization of peer relationships through past joint visits. Nevertheless, we acknowledge that this remains an assumption of our model rather than a conclusively proven approach. Fourth, we assume users are myopic in their subscription choices.

However, even with month-to-month subscriptions, there could still be option value of choosing a premium subscription. If there was such an option value, then assuming a myopic model could result in a over-estimation of the user's intrinsic preference for exercising. Recall that the number of visits depends on the user's intrinsic preference for exercising, carry-over effects from past usage, and peer effects. A user would choose the basic plan over the no-purchase option if her monthly usage is high enough to offset the price of the basic plan. In our data, when users join the gym, they typically subscribe to the basic plan and then upgrade to the premium plan if their visits become high enough. Now, if the user is indeed forward-looking, she may choose the basic plan (over choosing the no-purchase option) even when her intrinsic preference for exercising is very low. The reason is that the forward-looking user knows that if she persists with her exercise in the future, her number of visits will increase because of the strengthening of peer relationships and the positive carry-over effects from past usage. As a result, she will have high utility in the future by upgrading to the premium plan. This means that if we assume such a user to be myopic, we would overestimate the user's intrinsic preference for exercising. We leave these issues for future research.

Funding and Competing Interests: All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no funding to report.

References

Ascarza, Eva. 2018. Retention futility: Targeting high-risk customers might be ineffective. *Journal of Marketing Research* **55**(1) 80–98.

Ascarza, Eva, Peter Ebbes, Oded Netzer, Matthew Danielson. 2017. Beyond the target customer: Social effects of customer relationship management campaigns. *Journal of Marketing Research* **54**(3) 347–363.

Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, Matthew O Jackson. 2013. The diffusion of microfinance. *Science* **341**(6144) 1236498.

Boccaletti, Stefano, Vito Latora, Yamir Moreno, Martin Chavez, D-U Hwang. 2006. Complex networks: Structure and dynamics. *Physics reports* **424**(4-5) 175–308.

Burt, Ronald S. 2002. Bridge decay. *Social networks* **24**(4) 333–363.

Chen, Xi, Ralf van der Lans, Michael Trusov. 2021. Efficient estimation of network games of incomplete information: Application to large online social networks. *Management Science* **67**(12) 7575–7598.

Crane, Deborah A, Jeanne M Hoffman, Maria R Reyes. 2017. Benefits of an exercise wellness program after spinal cord injury. *The journal of spinal cord medicine* **40**(2) 154–158.

DellaVigna, Stefano, Ulrike Malmendier. 2006. Paying not to go to the gym. *american economic Review* **96**(3) 694–719.

Eagle, Nathan, Alex Pentland, David Lazer. 2009. Inferring friendship network structure by using mobile phone data. *Proceedings of the national academy of sciences* **106**(36) 15274–15278.

Godinho de Matos, Miguel, Ferreira Pedro, Belo Rodrigo. 2018. Target the ego or target the group: Evidence from a randomized experiment in proactive churn management. *Marketing Science* **37**(5) 793–811.

Goldsmith-Pinkham, Paul, Guido W Imbens. 2013. Social networks and the identification of peer effects. *Journal of Business & Economic Statistics* **31**(3) 253–264.

Granovetter, Mark S. 1973. The strength of weak ties. *American journal of sociology* **78**(6) 1360–1380.

Hartmann, Wesley R. 2010. Demand estimation with social interactions and the implications for targeted marketing. *Marketing science* **29**(4) 585–601.

Hartmann, Wesley R, Puneet Manchanda, Harikesh Nair, Matthew Bothner, Peter Dodds, David Godes, Kartik Hosanagar, Catherine Tucker. 2008. Modeling social interactions: Identification, empirical methods and policy implications. *Marketing letters* **19**(3-4) 287–304.

Iyengar, Raghuram, Christophe Van den Bulte, Jae Young Lee. 2015. Social contagion in new product trial and repeat. *Marketing Science* **34**(3) 408–429.

Iyengar, Raghuram, Christophe Van den Bulte, Thomas W Valente. 2011. Opinion leadership and social contagion in new product diffusion. *Marketing Science* **30**(2) 195–212.

Krackhardt, David, N Nohria, B Eccles. 2003. The strength of strong ties. *Networks in the knowledge economy* **82**.

Kumar, Vineet, K Sudhir. 2025. Can random friends seed more buzz and adoption? leveraging the friendship paradox. *Management Science* **71**(1) 942–953.

Lee, Lung-fei, Ji Li, Xu Lin. 2014. Binary choice models with social network under heterogeneous rational expectations. *Review of Economics and Statistics* **96**(3) 402–417.

Lemmens, Aurélie, Sunil Gupta. 2020. Managing churn to maximize profits. *Marketing Science* **39**(5) 956–973.

Manski, Charles F. 1993. Identification of endogenous social effects: The reflection problem. *The review of economic studies* **60**(3) 531–542.

Nair, Harikesh S, Puneet Manchanda, Tulikaa Bhatia. 2010. Asymmetric social interactions in physician prescription behavior: The role of opinion leaders. *Journal of Marketing Research* **47**(5) 883–895.

Nikolajsen, Helene, Louise Fleng Sandal, Carsten Bogh Juhl, Jens Troelsen, Birgit Juul-Kristensen. 2021. Barriers to, and facilitators of, exercising in fitness centres among adults with and without physical disabilities: a scoping review. *International journal of environmental research and public health* **18**(14) 7341.

Nitzan, Irit, Barak Libai. 2011. Social effects on customer retention. *Journal of Marketing* **75**(6) 24–38.

Park, Hee Mok, Puneet Manchanda. 2015. When harry bet with sally: An empirical analysis of multiple peer effects in casino gambling behavior. *Marketing Science* **34**(2) 179–194.

Shahapure, Ketan Rajshekhar, Charles Nicholas. 2020. Cluster quality analysis using silhouette score. *2020 IEEE 7th international conference on data science and advanced analytics (DSAA)*. IEEE, 747–748.

Suitor, J Jill, Barry Wellman, David L Morgan. 1997. It's about time: How, why, and when networks change. *Social networks* **19**(1) 1–7.

Toker-Yildiz, Kamer, Minakshi Trivedi, Jeonghye Choi, Sue Ryung Chang. 2017. Social interactions and monetary incentives in driving consumer repeat behavior. *Journal of Marketing Research* **54**(3) 364–380.

Tucker, Catherine. 2008. Identifying formal and informal influence in technology adoption with network externalities. *Management Science* **54**(12) 2024–2038.

Tucker, Catherine. 2018. Network effects and market power: what have we learned in the last decade? *Antitrust* **32**(2) 72–79.

Uetake, Kosuke, Nathan Yang. 2019. Inspiration from the "biggest loser": Social interactions in a weight loss program. *Marketing Science* .

Watts, Duncan J, Steven H Strogatz. 1998. Collective dynamics of 'small-world'networks. *nature* **393**(6684) 440–442.

Wayment, Heidi A, Rachael L McDonald. 2017. Sharing a personal trainer: personal and social benefits of individualized, small-group training. *The Journal of Strength & Conditioning Research* **31**(11) 3137–3145.

Yang, Sha, Vishal Narayan, Henry Assael. 2006. Estimating the interdependence of television program viewership between spouses: A bayesian simultaneous equation model. *Marketing Science* **25**(4) 336–349.

Yook, Soon-Hyung, Hawoong Jeong, A-L Barabási, Yuhai Tu. 2001. Weighted evolving networks. *Physical review letters* **86**(25) 5835.

Web Appendix for "Exploring Peer Effects Associated with User Retention in Socially Connected Businesses"

A. Robustness Check of Peer Relationships Using External Peer Relationship Data

In this robustness check, we first show that individuals who consistently attend the same gym classes are likely to have a social connection outside the gym. Following that, we will discuss why we only model the users' decision to visit the gym and not a specific class in the gym.

For around 4% of the users, the gym had information about their relationships outside the gym. This information included "reference relationships" indicating who referred whom to the gym, and "employment relationships" identifying coworkers. We use this information to test whether there is any conformance between these relationships outside the gym and our measure of the peer relationships (which is based on past joint visits).

First, for these 4% of the users, we examined the hit rate of our peer inference method, which measures the proportion of periods in which the users with real world relationship (based on whether they are co-workers or were referred by the other user) are identified as peers among all the observed periods. Our analysis reveals that the hit rate is 93%, indicating a high level of precision in our peer relationship inference.

Second, Table A.1 shows that real-world peers, compared to those inferred as peers based on high latent peer relationship strengths (m_{ijt} in the top two quartiles) but lacking real-world relationship data, exhibit a 3.5% greater likelihood of having any joint visits and a 50% higher average monthly joint visit frequency across all observation periods. This higher joint visit frequency among real-world peers is expected, as their external relationships contribute to stronger social bonding and more shared gym visits. This finding underscores a strong association between higher joint visit frequency and the likelihood of individuals being real-world peers, supporting the validity of using past joint visits as a proxy for strengths of peer relationships in the gym.

Table A.1 Comparison of Real World Peers and Inferred Peers

Statistics	Real-world Peers	Inferred Peers
Likelihood of Joint Visit	0.359	0.324
Average Monthly Joint Visit	1.34	0.89
Average Percentage of Joint Visit	0.244	0.194

B. Reduced Form Analysis Supporting the Operationalization of Peer Relationship

In this reduced form analysis, we regress each focal user i's gym visits in each period[17] (month) on the gym visits of each other user $j = 1, \cdots, N, j \neq i$, and the interaction of the gym visits of each of

[17] Note that we define the peer relationship strength based on the joint visits of users i and j in the same gym classes in the past. However, the user's action that we are modeling in both the reduced form model and the structural model is the user's visit to the gym, and not the user's visit to a specific class in the gym. In section A of the web appendix, we have discussed why we do that and why this is not a source of concern.

those users j with the moderating variable m_{ijt}. To compare the effect of the gym visits of other users $j \neq i$ with different values of m_{ijt} on the focal user i's gym visit, we divide the m_{ijt} among all active users i and j into four quartiles and run the following regression:

$$\begin{aligned} Visit_{it} =& \alpha_i + Month_t + Year_t + \delta Premium_{it} + \beta_0 Visit_{it-1} + \\ & \Sigma_{j\neq i}(\beta_1 Firstm_{ijt} + \beta_2 Secondm_{ijt} + \beta_3 Thirdm_{ijt} + \beta_4 Fourthm_{ijt}) Visit_{jt} + e_{it} \end{aligned} \tag{B.1}$$

In equation (B.1), $Visit_{it}$ is the number of gym visits for user i in period t, α_i, $Month_t$ and $Year_t$ are the user i, period and year fixed effect, and $Visit_{ijt-1}$ is the lag monthly gym visit which controls for individual-level temporary visit shocks. $Premium_{it}$ is an indicator of premium plan that equals 1 if individual i subscribes to the premium plan at month t and 0 otherwise. $Visit_{jt}$ represents the visit count of user $j \neq i$ in the network in period t. $Firstm_{ijt}$ is a 0-1 indicator variable, which takes the value of 1 if m_{ijt} falls within the lowest quartile, and similarly for $Secondm_{ijt}$, $Thirdm_{ijt}$, and $Fourthm_{ijt}$ which represent the subsequent higher quartiles. The peer effect term $\Sigma_{j\neq i}(\beta_1 Firstm_{ijt} + \beta_2 Secondm_{ijt} + \beta_3 Thirdm_{ijt} + \beta_4 Fourthm_{ijt}) Visit_{jt}$ is the summation of other user j's gym visit weighted by the quartiles of m_{ijt}.[18] The variable Visit$_{jt}$ is endogenous because each user's gym visit at time t is simultaneously influenced by other users' gym visits in the same period. To estimate the peer effect coefficients β_1 to β_4, we use the exogenous churns to construct an instrumental variable for user j's gym visit $IV(Visit_{jt}) = \Sigma_k 1(Churn_{kt}) * m_{jkt}$, where $1(Churn_{kt})$ is an indicator variable that equals 1 if user k churns in period t and 0 otherwise, and m_{jkt} is the peer relationship strength between users k and j in period t.

The regression analysis above rests on the assumption that network defined through joint class attendance is exogenous, that is the term m_{ijt} in equation (B.1) and the term m_{jkt} in the instrument are exogenous. However, the terms m_{ijt} and m_{jkt} could be endogenous since they are functions of the lagged dependent variables, namely, user i's past visits ($Visit_{is}$ for $s = 1, \cdots, t-1$). Thus, if the error terms in equation (B.1) are correlated over time, then the terms m_{ijt} and m_{jkt} will be correlated with the current error term e_{it}. This implies that the terms m_{ijt} and m_{jkt} could be endogenous if the errors e_{it} are serially correlated. To address this issue, we control for individual FE and the individual-level temporary visit shocks through the regressor $Visit_{it-1}$ in equation (B.1) and it is unlikely that the visit errors in equation (B.1) will have a strong serial correlation. To further assess the validity of our analysis, we conducted a robustness test in Web Appendix C. In this test, we simulated 4 datasets using equation (B.1) with varying levels of serial correlations in the error terms (from 0 to 0.95) and re-estimated the parameters in equation (B.1) using 2SLS. The estimated coefficients β_1 to β_4 remained consistent with the true values, suggesting that our

[18] In the regression in equation (B.1), we have not included the main effects terms for $Firstm_{ijt}$, $Secondm_{ijt}$, and $Thirdm_{ijt}$. The estimates of their coefficients were not significantly different from zero.

findings are not sensitive to serial correlation in the error terms. The reason for that is because the correlation between m_{ijt} and e_{it} (and between m_{jkt} and e_{it}) was less than 0.01 even when the serial correlation between the errors was as high as 0.6.[19] This suggests that it is reasonable to assume that m_{ijt} and m_{jkt} are exogenous.

We estimate the regression with two-stage least square using the entire sample of 2,118 users over a period of seven years. The incremental F-statistic for the instrumental variable (IV) in the first stage of the regression is 21.5. This indicates that the instrument is not weak. The results of the regression in equation (B.1) are shown in Table B.2.

Table B.2 Impact of Other Users' Visit on the Focal User's Visit

Variable	Estimate	Clustered S.E.
$Visit_{it-1}$	0.32^{***}	0.006
$Premium_{it}$	3.16^{***}	0.13
$Visit_{jt} * Firstm_{ijt}$	-0.0018	0.002
$Visit_{jt} * Secondm_{ijt}$	-0.001	0.001
$Visit_{jt} * Thirdm_{ijt}$	0.025^{***}	0.002
$Visit_{jt} * Fourthm_{ijt}$	0.03^{***}	0.002

Note: Standard errors, clustered at the user-month level, are presented in parentheses. Significance level:***p<0.01, **p<0.05,*p<0.1.

Table B.2 shows that subscribing to a premium plan is positively correlated with the individual's gym visits. The impact of other user's gym visit on the focal user's gym visits is larger for users with a higher share of prior joint visits to the same classes with the focal user. Specifically, the effect of the other users with prior joint visits to the same classes falling within the first two quartiles is statistically insignificant. However, the impact of user j's gym visits on the visits of user i is positive and significant in the third and fourth quartiles. This indicates that m_{ijt} moderates the causal impact of gym visits of other users on the gym visits by the focal user. Therefore, the share of prior joint visits to the same class between any two users can reasonably proxy for the latent strength of their peer relationship.

The above regression findings also suggest that we can use the median of individual-time specific joint visits as a threshold to define peer relationships. User j will be user i's direct peer in period t (i.e., user j 's actions will directly impact user i) if the share of the number of classes that i attends with j till period t is greater than the median of i's joint visits to the same class with all the users at

[19] The reason for this can be understood as follows. Recall that $Visit_{it-1}$ represents the number of user i's visits to gym in period $t-1$, while m_{ijt} refers to the ratio of 'the number of past joint visits of user i and user j to the same classes in the gym' to 'sum of the number of past joint visits of user i and each of the users k to the same classes in the gym'. There are two points to note here: (a) $Visit_{it-1}$ enters both the numerator and the denominator of m_{ijt}. (b) While $Visit_{it-1}$ is the number of visits by user i to the gym in period t, m_{ijt} is a function of the prior joint visits of user i with each of the other users to the same classes in the gym. Both points (a) and (b) result in m_{ijt} to be a very weak function of $Visit_{it-1}$, which in turn results in the correlation between $Visit_{it-1}$ and the visit errors to be insignificant.

the gym till period t, otherwise, they will be indirect peers. In addition to capturing whether the peer relationship exists or not between two users, we also capture the strength of the peer relationship using m_{ijt}. If user j is user i's direct peer, then the strength of the relationship from user j to user i in period t will be $m_{ijt} = \frac{g_{ijt}}{\Sigma_{k=1,k\neq i}^{N} g_{ikt}}$, which is the share of user j's prior joint visits to the same class with user i amongst all of user i's joint visits in the same classes with all i's peers till period t. If j is not i's direct peer, the strength of their relationship will be 0. This approach of operationalizing whether or not two users are direct peers follows from the prior literature (e.g. Yook et al. 2001, Boccaletti et al. 2006, Eagle et al. 2009) in which direct peers are defined as users whose joint usage exceeds a certain threshold quantile of their joint usage of all other users. In our analysis, we use the median as the quantile threshold also because it yields a network with similar key metrics (degree distribution, transitivity, diameter) as the metrics of the real-world social networks reported in the prior literature (more on this later). In conclusion, we find suggestive evidence that supports our operationalization of peer relationships through past joint visits.[20]

C. Robustness Test of Intertemporal Correlation in Visit Errors

Reduced Form Model in Equation (B.1): To investigate the impact of inter-temporal correlation on the model estimates, we conducted a simulation study using four datasets. In each dataset, we simulated user visit decisions according to equation (B.1), holding several elements constant: all coefficients in regression (B.1) (including the peer effect parameters, β), and all variables other than user visits across periods (including plan choices, assumed exogenous, and fixed effects). We use the method in section K of the Web Appendix to simulate the dynamics of m_{ijt} between any two users i and j over time, capturing how visit error terms influence user visits, which in turn affects the values of m_{ijt}. The only difference across the four datasets was the degree of inter-temporal correlation in the visit equation errors, e_{it}, for the same individual i. We introduced correlations of 0, 0.2, 0.4, 0.6 and 0.95, respectively. Additionally, we assumed a constant correlation of 0.2 in the period t errors across different users j and i. We then re-estimated equation (B.1) on each simulated dataset.

Table C.3 presents the estimation results. Notably, the estimates of the peer effect parameters, β_1 to β_4, remained consistent across all four datasets, regardless of the level of inter-temporal error correlation. This consistency provides strong evidence that even in the presence of some degree of inter-temporal correlation in the errors, the estimation of the peer effect parameters remains robust and unbiased. We compute the correlation between e_{it} and m_{ijt} in the simulated datasets, and the

[20] Related to this operationalization, it is possible that one can have alternative operationalizations in which the joint visits in the distant past should be weighted less than the joint visits in the recent past. We have tried these alternative operationalizations in which we have discounted the joint visits from the distant past, and we found that our results are robust.

correlation is 0.0019, 0.0029, 0.0038, and 0.0084 when the inter-temporal correlation in the e_{it} is set as 0.2, 0.4, 0.6 and 0.95. The correlation between e_{it} and m_{jkt} is smaller than 0.001 in all scenarios. This indicates that the serial correlation in the visit error e_{it} does not induce a correlation between e_{it} and m_{ijt} or correlation between e_{it} and m_{jkt}, thus m_{ijt} and m_{jkt} satisfies the exclusion restriction.

Table C.3 Peer Effect Parameter Estimates with Varying Levels of Inter-temporal Correlation

Parameters	True Value	Cor=0	Cor=0.2	Cor=0.4	Cor=0.6	Cor=0.95
δ	0.32^{***}	0.32^{***}	0.32^{***}	0.32^{***}	0.32^{***}	0.33^{***}
β_0	3.16^{***}	3.19^{***}	3.19^{***}	3.19^{***}	3.18^{***}	3.16^{***}
β_1	-0.0018	-0.0017	-0.0018	-0.0018	-0.0018	-0.0018
β_2	-0.001	-0.0011	-0.0011	-0.0011	-0.0011	-0.0010
β_3	0.025^{***}	0.0258^{***}	0.0259^{***}	0.026^{***}	0.026^{***}	0.026^{***}
β_4	0.03^{***}	0.0303^{***}	0.0303^{***}	0.0302^{***}	0.0302^{***}	0.0298^{***}

Note: Significance level:***p<0.01, **p<0.05,*p<0.1

Strucutral Model: We also simulated four datasets based on our structural model. In the simulation, the true parameter values are in Table C.4, which are in the similar range as our original estimates. We keep all observed characteristics $\boldsymbol{X}_{it}$ the same as they are in the data. We then introduce serial correlations in the visit choice error e_{it} at levels of 0.2,0.4,0.6 and 0.95. The network $\boldsymbol{\Omega}_t$ is computed using the method in the section K of the Web Appendix. We numerically compute the users decisions $\boldsymbol{EY}_t$ in equation (8) across all the users using contraction mapping.

Then we re-estimated our structural model on each simulated dataset. Table C.4 presents a comparison between the true parameter values used in generating the simulated datasets and the corresponding parameter estimates obtained from our model estimation. The results demonstrate that across all three scenarios, our model successfully recovered the true parameter values with minimal deviation. This finding indicates that our model exhibits robustness even in the presence of moderate serial correlation in the visit error.

Table C.4 Estimation Results with Varying levels of Serial Correlation in Visit Error

Key Parameters	True Value	Cor = 0.2	Cor = 0.4	Cor = 0.6	Cor = 0.95
Peer Effect Coefficient β	0.3	0.32	0.28	0.27	0.28
Exercise State Level High λ_H	15	17.37	17.47	18.01	18.04
Exercise State Level Medium λ_M	10	13.19	12.97	11.97	11.52
Exercise State Level Low λ_L	5	5.30	4.99	4.99	4.52
Exercise State Level High Constant λ_H^C	20	19.98	19.07	19.66	18.72
Exercise State Level Medium Constant λ_M^C	15	16.37	15.00	16.04	15.74
Price Coefficient Level High α^H	-0.2	-0.17	-0.17	-0.18	-0.15
Price Coefficient Level Medium α^M	-0.15	-0.13	-0.12	-0.13	-0.12
Price Coefficient Level Low α^L	-0.1	-0.07	-0.06	-0.08	-0.06
Carry-over Effect (Consistent) γ_2^{CN}	0.7	0.67	0.72	0.71	0.71
Carry-over Effect (On-and-off) γ_2^{ON}	0.7	0.72	0.68	0.67	0.65

D. Discussion on the Issue of not Modeling Users' Visit Decision to Each Class in the Gym

We next discuss why we model the user's decision to visit the gym in each period and not her decision to visit each specific group class in the gym in each period. The reason why we only model the user's visit decision to the gym is because there are 50+ different types of group classes, based on time of the day, day of the week and the type of exercise. Modeling the user's visit decision to each of these classes in an equilibrium framework in a structural model is computationally infeasible. As a side note, the same problem would be there in an endogenous network formation model, unless if we were to model the users' decisions on which of the 50+ gym classes to attend in each period. Having said that, not modeling the users' visit decisions to each class and instead modeling the user's visit decision to the gym is not a source of concern when it comes to estimating the reduced form or the structural models. This is because in our data, each time a user attends a gym class, we observe the specific class they attend, including the timestamp, and the room. Because we have attendance records for all gym users, we can identify which users are in the same class at any given time. Thus, we can infer the peer relationship strengths over time directly from the data. However, it does lead to an issue when we are simulating the network in the simulations. This is because in the simulations, we are updating the network structure (that is, the peer relationship strength between users) over time based on the past simulated decisions of the users. The challenge here is that from the model, we can only simulate the visits of users i and j to the gym in each period - however, recall that our operationalization of the peer relationship strength of users i and j in period $t+1$ is based on the share of prior joint visits of users i and j to the 'same classes' in the gym (and not merely just their visits to the gym). Thus, to calculate the peer relationship strength between users i and j in period $t+1$ from the simulated number of visits of users i and j in the prior periods, we need some kind of a mapping from the simulated number of visits of users i and j to the gym in a given period to the number of visits of users i and j to the 'same classes' in the gym in that same period. To deal with this challenge, in our simulations, we propose a reduced form mapping that yields the users' joint visits to the same classes in the gym using their past simulated visits to the gym. This mapping is estimated from the data and it aligns with the prior literature on social networks, which has focused on how the network structure evolves with the presence/absence of shared activities between people (Granovetter 1973, Krackhardt et al. 2003, Burt 2002, Suitor et al. 1997). Please see Appendix K for details.

E. Robustness Test of Correlation between Visit and Choice Errors

In this section, we conduct a a robustness test to assess how correlations between the plan choice error term ϵ_{it}^{d} and the visit error term e_{it} might affect our parameter estimates.

Specifically, we simulated three datasets based on our structural model, keeping all parameters in the similar range as our original estimates. In these simulations, we introduced correlations between the errors of the plan choice ϵ_{it}^d and visit e_{it} at levels of 0.2,0.4 and 0.6. The simulation method from the structural model is explained in section K of the Web Appendix.

Subsequently, we re-estimated our structural model on each simulated dataset. Crucially, during the re-estimation process, we maintained the assumption of zero correlation between the error terms (as in our original model). Table E.5 presents a comparison between the true parameter values used in generating the simulated datasets and the corresponding parameter estimates obtained from our model estimation. The results demonstrate that across all three scenarios, our model successfully recovered the true parameter values with minimal deviation. This finding indicates that our model exhibits robustness even in the presence of moderate correlation between the plan choice and visit error terms

Table E.5 Estimation Results with Varying levels of Correlation between Visit and Choice Error

Key Parameters	True Value	Cor = 0.2	Cor = 0.4	Cor = 0.6
Peer Effect Coefficient β	0.3	0.27	0.30	0.33
Exercise State Level High λ_H	15	16.58	16.27	13.70
Exercise State Level Medium λ_M	10	12.63	9.64	8.97
Exercise State Level Low λ_L	5	4.86	4.68	4.48
Exercise State Level High Constant λ_H^C	20	19.44	18.83	19.06
Exercise State Level Medium Constant λ_M^C	15	14.48	16.15	15.05
Price Coefficient Level High α^H	-0.2	-0.17	-0.16	-0.22
Price Coefficient Level Medium α^M	-0.15	-0.14	-0.13	-0.17
Price Coefficient Level Low α^L	-0.1	-0.08	-0.07	-0.09
Carry-over Effect (Consistent) γ_2^{CN}	0.7	0.72	0.69	0.67
Carry-over Effect (On-and-off) γ_2^{ON}	0.7	0.73	0.68	0.71

F. Proof of Existence and Uniqueness of Equilibrium and Numerical Simulations

The equilibrium $\boldsymbol{E(Y_t)}$ is characterized by the $3 \times n_t$ simultaneous equations in equations (7a)-(7c). Here we prove the existence and the uniqueness of the equilibrium in certain ranges of the peer effect coefficient β.

The choice probabilities $(\Phi_{it}^H, \Phi_{it}^L)$ in equations (7a) and (7b) are characterized by a discrete choice model with logit specification. Here we use $p = (p_1, p_2, \cdot, p_n)$ to represent the vector of choice probabilities $(\Phi_{it}^H, \Phi_{it}^L)$ for all the users in the network with a size of n. Then equations (7a) and (7b) can be rewritten as $p = H(p)$, where $H(p) = (F(u_1(p)), F(u_2(p)), \cdot F(u_n(p))))$, $F(u) = \frac{exp(u)}{1+exp(u)}$, and $u_i(p)$ could be decomposed to two parts: $f(X_{it})$ as a function of the individual characteristics that is free of the peer effect component and $\beta \times \Omega_i \times p$ as the weighted average of the choice probability vector p multiplied by the peer effect coefficient β.

The gradient of $H(p)$ is:

$$\frac{\partial H(p)}{\partial p'} = \lambda \begin{pmatrix} \Omega_{11} f(u_1(p)) & \cdots & \Omega_{1n} f(u_1(p)) \\ \vdots & \ddots & \vdots \\ \Omega_{n1} f(u_n(p)) & \cdots & \Omega_{nn} f(u_n(p)) \end{pmatrix}, \tag{F.1}$$

where $f(.)$ is the density function of the distribution function $F(.)$. Then we have:

$$\left\| \frac{\partial H(p)}{\partial p'} \right\|_\infty = |\lambda| \max_{i=1,\ldots,n} \sum_{j=1}^{n} |\Omega_{ij}| \cdot f(u_i(p)) \leq |\lambda| \cdot \|\mathbf{\Omega}\|_\infty \max_u f(u), \tag{F.2}$$

$$\left\| \frac{\partial H(p)}{\partial p'} \right\|_1 = |\lambda| \max_{j=1,\ldots,n} \sum_{i=1}^{n} |w_{ij}| f(u_i(p)) \leq |\lambda| \cdot \|\mathbf{\Omega}\|_1 \max_u f(u). \tag{F.3}$$

where $\|\cdot\|_\infty$ and $\|\cdot\|_1$ are the maximum row and column sum matrix norms of the square matrix. So if the following condition holds, $H(p)$ will be a contraction mapping.

$$|\lambda| \leq 1/\{\min[\|\mathbf{\Omega}\|_\infty, \|\mathbf{\Omega}\|_1] \cdot \max_u f(u)\}, \tag{F.4}$$

In logit model, the plan choice error ϵ_{it}^d and ϵ_{it}^0 follows the standard Gumbel distribution, thus the maximum value of the density function $f(u)$ is $\frac{1}{4}$. Thus if $|\beta| < 4$, $H(p)$ is a contraction mapping.[21]

Next, in equation (7c), if $|\beta| < 1$, the absolute value of coefficients in front of $E(Visit_{it})$ on the right-hand side is smaller than 1, and it is a contraction mapping for continuous variables $E(Visit_{it})$.

Jointly, the system of simultaneous equations that characterize the equilibrium is a contraction mapping given the peer effect coefficient $|\beta| < 1$. By Banach's fixed point theorem, the equilibrium exists and is unique.

To empirically test the uniqueness of the equilibrium under the parameter range restriction $|\beta| < 1$, we conducted numerical simulations. We first drew 100 sets of model parameters. The peer effect coefficient, β, was drawn subject to the restriction $|\beta| < 1$ (see Figure F.1), while other parameters were drawn from a wide range of reasonable values. For each parameter set, we computed the equilibrium using contraction mapping with 50 iterations from 50 random starting points. We then calculated the absolute difference between every pair of simulated equilibrium outcomes (visit probabilities and visit counts) obtained under the same parameter set. Finally, we averaged these differences across all individuals and parameter sets. The average difference in purchase probabilities was 0.0008, and the average difference in visits was 0.0142. Figure F.2 displays the distribution of these differences in equilibrium outcomes for both purchase probabilities and visits.

To explore the impact of violating the parameter restriction, we used the same 100 parameter sets but increased the value of the peer effect coefficient by 1, ensuring $|\beta| > 1$. All other model parameters remained unchanged. We then recomputed the equilibrium using contraction mapping. The resulting average purchase probability across all users was 0.995, and the average visit count was extremely large, deviating significantly from the observed data.

Figure F.1 Histogram of Drawn Peer Effect Coefficients

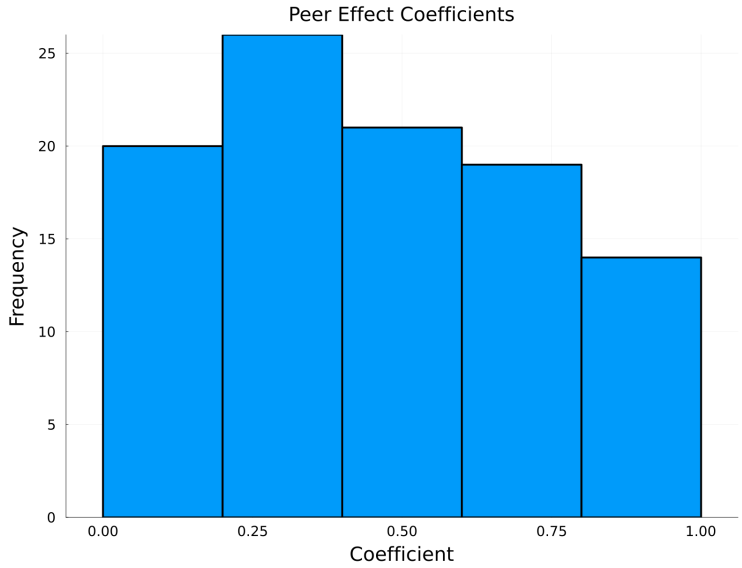
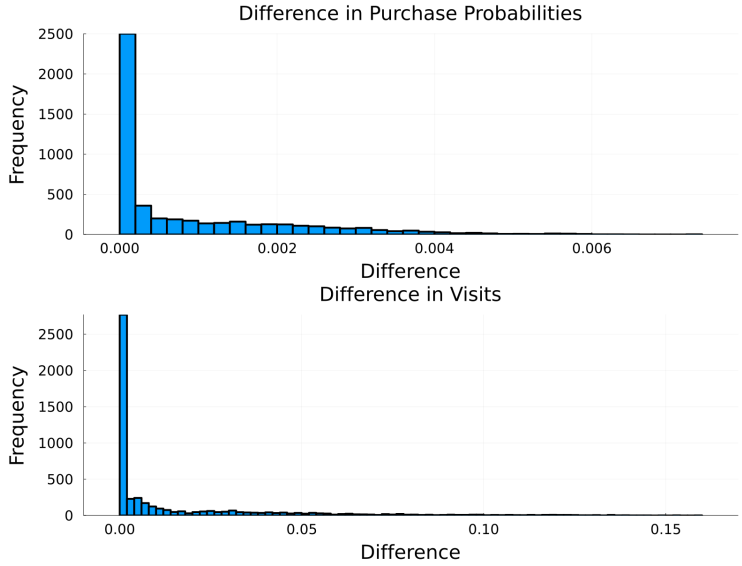


Figure F.2 Differences in Equilibrium Outcomes



G. Evidence to Support the Clustering Method

This section presents empirical evidence supporting the validity of our clustering method.

We first show that that users within the same cluster have similar preferences for specific gym class types. We calculate each user's attendance frequency for each class type (e.g., Cardio, Strength) by dividing the number of times a user attended that particular class by their total class attendance. We

[21] The details of the proof can be found in Lee et al. (2014)

then calculate the absolute difference in these attendance frequencies for pairs of users both within the same cluster and across different clusters. Table G.6 reveals that users within the same cluster have similar preferences for specific gym class types. For example, for "Cardio" class attendance, the average absolute difference in attendance frequency between users within the same cluster is 0.07, which is smaller than the average difference between users from different clusters is 0.12. This pattern holds consistently across all class types, demonstrating that users within the same cluster exhibit more similar preferences for specific exercise modalities compared to users from different clusters. This suggests that our clustering effectively captures users' heterogeneous preferences for classes. In other words, users within the same cluster are more likely to attend similar classes, indicating that our clustering approach can account for variations in class preferences that might otherwise influence the formation of peer connections.

Table G.6 Preference for Classes Within and Across Clusters

Class	Same Cluster	Different Cluster
Acceleration	0.006	0.04
Cardio	0.07	0.12
Strength	0.02	0.10
Small Group Personal Training	0.04	0.07
Cardio/Strength	0.0003	0.07

Note: Table entries represent the average absolute difference in class attendance frequencies between user pairs. The values under "Same Cluster" and "Different Cluster" reflect the mean absolute difference in these frequencies for pairs within the same cluster and across clusters, respectively. Lower values (e.g., 0.07 vs. 0.12 for Cardio) indicate greater similarity in preferences.

We validated our clustering approach using the silhouette score, a standard metric for evaluating clustering quality by measuring both how similar users are within the same cluster and how distinct a cluster is from others (Shahapure and Nicholas 2020). The silhouette score ranges from -1 to 1, where values closer to 1 indicate that users are well-matched to their own cluster and poorly matched to others. Our analysis shows silhouette scores primarily around 0.6, indicating good cluster separation and appropriate assignment of users to segments. This high positive silhouette score suggests that users within each identified segment exhibit similar gym behavior patterns that are distinctly different from users in other segments.

To validate our clustering method's ability to effectively segment consumers, we performed a reduced-form exercise. This exercise aims to show that our approach yields results comparable to directly estimating unobserved heterogeneity alongside other parameters.

In this exercise, we estimated two variants of the reduced-form regression defined in equation (B.1) of the paper:

- In Variant 1, we estimated the regression specified in equation (B.1), including individual fixed effects (FE) for all $N_1 = 1491$ users in the sample. These N_1 user FEs directly capture the heterogeneity in each user's intrinsic propensity to visit the gym. This variant represents the approach where unobserved heterogeneity is estimated directly for each user along with the other parameters.
- In Variant 2, we first grouped all users into $N_2 = 3$ clusters using the same clustering algorithm applied in our structural model. We then estimated the reduced-form regression from equation (B.1). Unlike Variant 1, where we estimated N_1 individual user FEs along with the other parameters, in Variant 2, we estimated only $N_2 = 3$ cluster FEs along with other parameters. This assumes all users within the same cluster share a common FE (the cluster FE).

Next, we compare the segmentations implied by Variants 1 and 2. For Variant 2, we already have the $N_2 = 3$ clusters as the segments. For Variant 1, after estimating the individual user FEs, we grouped these user FEs into clusters using the same clustering algorithm that we employed in Variant 2. This means that users with similar estimated user FEs in Variant 1 were placed in the same cluster. Therefore, we obtained clusters from both variants. Note that the key difference in the segmentation across the two variants is when clustering occurred: in Variant 2, clustering was performed before estimation, while in Variant 1, clustering was performed after estimating individual user FEs. Thus, Variant 2 is synonymous with what we did with our structural model, where we first clustered the users and then estimated the model.

Finally, we compared the cluster memberships between Variant 1 and Variant 2. The overlap was substantial: 95.5% of users assigned to the 'high' cluster in Variant 1 were also in the 'high' cluster in Variant 2; 95.4% of users assigned to the 'medium' cluster in Variant 1 were also in the 'medium' cluster in Variant 2 and 97.2% of users assigned to the 'low' cluster in Variant 1 were also in the 'low' cluster in Variant 2. The number of users in each cluster for Variant 1 and Variant 2 is shown in Table G.7.

This high degree of consistency between the clusters derived from the two variants demonstrates that our pre-estimation clustering approach (Variant 2), as used in the structural model, effectively groups users based on their unobserved preferences.

Table G.7 Number of Users in Variant 1 and Variant 2 Clusters

Variant 2 / Variant 1	**Low**	**Medium**	**High**
Low	895	18	0
Medium	25	439	5
High	0	3	106

H. Robustness Test of Correlation between Visit Errors Across Users

In this section, we conduct a a robustness test to assess how correlations in the visit error term e_{it} among users might affect our parameter estimates.

Specifically, we simulated two datasets based on our structural model, keeping all parameters in the similar range as our original estimates. In these simulations, we introduced correlations in e_{it} among users in each period t at levels of 0.2 and 0.5. The simulation method is explained in the section K of the Web Appendix.

Subsequently, we re-estimated our structural model on each simulated dataset. Crucially, during the re-estimation process, we maintained the assumption of zero correlation between the error terms (as in our original model) among users. Table H.8 presents a comparison between the true parameter values used in generating the simulated datasets and the corresponding parameter estimates obtained from our model estimation. The results demonstrate that across the two scenarios, our model successfully recovered the true parameter values with minimal deviation. This finding indicates that our model exhibits robustness even in the presence of moderate correlation in the visit choice errors between users.

Table H.8 Estimation Results with Varying Levels of Correlation in Visit Errors between Users

Key Parameters	True Value	Cor = 0.2	Cor = 0.5
Peer Effect Coefficient β	0.3	0.31	0.28
Exercise State Level High λ_H	15	17.87	18.04
Exercise State Level Medium λ_M	10	11.22	11.52
Exercise State Level Low λ_L	5	5.34	4.52
Exercise State Level High Constant λ_H^C	20	18.75	18.72
Exercise State Level Medium Constant λ_M^C	15	14.27	15.73
Price Coefficient Level High α^H	-0.2	-0.23	-0.15
Price Coefficient Level Medium α^M	-0.15	-0.16	-0.12
Price Coefficient Level Low α^L	-0.1	-0.07	-0.08
Carry-over Effect (Consistent) γ_2^{CN}	0.7	0.73	0.71
Carry-over Effect (On-and-off) γ_2^{ON}	0.7	0.76	0.78

I. Empirical Evidence on Exogenous Churn

We use a total of 86 exogenous churns to construct the instrumental variables for the model estimation. In this section, we provide a detailed breakdown of the reasons users have given for leaving, which supports our argument that these churns are exogenous and validates our instrumental variables.

Table 1 outlines the reasons for customer churn, categorizing them into specific segments such as family obligations, personal illness, injury, pregnancy, and geographical relocation. These categories encompass a variety of deeply personal and individual circumstances, ranging from childcare needs and caring for ill family members to personal health issues and relocating, both domestically and internationally. Due to the personal nature of these reasons, it is unlikely that they are connected with unobserved factors that might influence the subscription or visit patterns of the peers of those who churn. Consequently, we can consider these instances of churn as exogenous when estimating peer effects.

Table I.9 Percentages of Relocation Churns

Year	Percentage
2012	0.367
2013	0.434
2014	0.378
2015	0.35
2016	0.367
2017	0.4
2018	0.333

One might contend that if there's a significant economic change, such as a major employer exiting the city, the observed relocation churns might no longer be exogenous. This is because the peers of those who churn might also leave the city due to the same economic factors, affecting their gym memberships similarly. To address this concern, we analyzed the annual percentage of churners who relocated. The stability in this percentage, as shown in Table I.9, supports our assertion that these churns are primarily exogenous.

The data, spanning from 2012 to 2018, reveals that the percentages of individuals moving out of town due to churn fluctuate only slightly around a consistent average, ranging from 33.3% in 2018 to 43.4% in 2013. This relative stability over seven years suggests that the reasons for relocating are not significantly impacted by external economic shifts or city-wide factors that could lead to a mass exodus in any particular year. Instead, the consistency of these percentages indicates that churn related to moving is driven by individual circumstances and decisions, rather than broader economic conditions in the city. This evidence strengthens the argument that these churns can be treated as exogenous when estimating a peer effect model.

In addition, within three months following an exogenous churn, the data shows that only a minimal 3.7% of the peers associated with the churners also decide to churn immediately. Instead, the predominant immediate response among the remaining peers is a reduction in visit, not outright churn. If the churners and their peers had been influenced by the same external factors, such as a major economic shift in the city, one might expect a more pronounced mirroring of behavior, with a higher rate of churn among peers closely following the initial observed churn events.

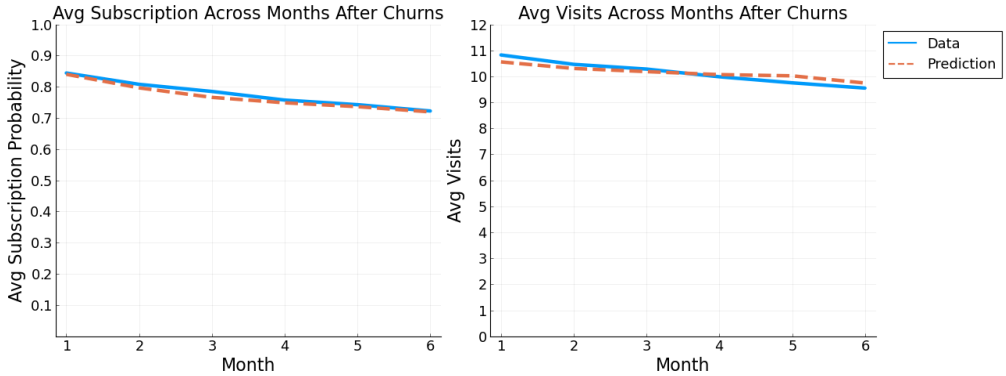
However, the fact that most peers primarily reduce their visit frequency rather than churn outright suggests a response more aligned with peer influence rather than a collective reaction to external factors. This distinction supports the idea that the original churn events are exogenous, with peer behaviors responding differently, indicating individual decisions rather than reactions to broader economic conditions.

J. Model Fit

To investigate model fit, we use the estimated coefficients to simulate the equilibrium subscriptions and gym visits and compare those with the ones in the data. In Figure J.3, we take the average

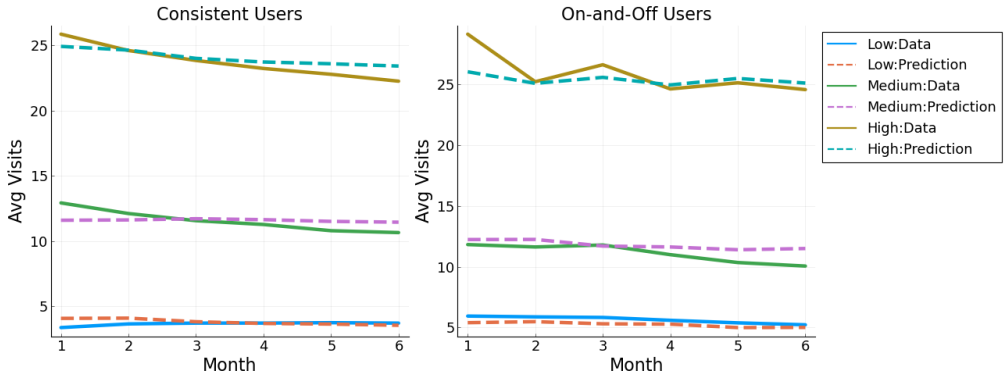
of the subscription probability and the visit count in each month after a churner leaves the gym. Overall, we see that the estimated model does a good job of recovering the observed decreasing trend of visit count and subscription probability after the churn. Figure J.4 further breaks down the visit count by consumer types. It shows that our estimated model does a good job of recovering the post-churn heterogeneous visit patterns across users. This suggests that our ex-ante segmentation method works well in capturing individual-level heterogeneity.

Figure J.3 Model Fit: Average Subscription and Visits



Note: Figure represents model fit for users subscription and visits; X-axis represents month since exogenous churn; Y-axis represents average subscription probability or average visits in each month; the solid series represents data; the dashed series represents model prediction.

Figure J.4 Model Fit: Visits By Consumer Types



Note: Figure represents average visits broken down by consumer types. The X-axis represents the month since peer churn; the Y-axis represents average visits in each month; the left figure represents consistent users; the right figure represents the on-and-off users; each series represents the average visit count per month for one user type in data or in model prediction; the solid series represents data; the dashed series represents model prediction.

K. Network Simulations

In this section, we specify how we update the network structures in the structural model simulations in section 6 and how we update the latent peer relationship strengths for the reduced form regression in section C of the Web Appendix.

Simulation for the Structural Model: In the simulations for the structural model, we not only have to simulate the visit and subscription decisions of all users in equilibrium in each period $t = 1, \cdots, T$, but also how the network structure $\mathbf{\Omega}_t$ evolves over time. Simulating the users' decisions and the evolution of the network structure over periods $t = 1, \cdots, T$ entails two steps in each time period t: (i) given the network structure (or the peer relationship strengths between all pairs of users) in period t, we would need to simulate subscription decisions and the gym visit decisions to the gym of all users in period t, and (ii) given the network structure in period t and the simulated gym visit and subscription decisions of all users prior to period t, we would need to specify how the network structure evolves from period t to $t+1$. Note that the simulations of the gym visit and subscription decisions in (i) follow directly from the specification of equilibrium conditions in our structural model in equations (7a)-(7c). However, as discussed in section 6, we use a reduced form approach for simulating the network in (ii) which captures the essence of a structural network formation model. We discuss the steps (i) and (ii) as follows.

We first discuss how we simulate the visits and subscription choices of all users in period t given the peer relationship strengths in period t. Following that, we will discuss how we calculate the peer relationship strength in period $t+1$ given the peer relationship strength in period t and the visits and subscription choices of all users in period t. Given the adjacency matrix $\mathbf{\Omega}_t$ in period t, we numerically compute the expected visit $E(Visit_{it})$ and subscription probabilities $(\Phi_{it}^H,\ \Phi_{it}^L)$ of all users $i \in N$ using the $3 * N$ system of simultaneous equations in equation (8). The actual visit $Visit_{it}$ of each user i is calculated as $Visit_{it} = E(Visit_{it}) + e_{it}$. The visit error e_{it} is drawn from the normal distribution of mean 0 and standard deviation of 1 for the structural model simulation in section 6. In section C of the web appendix, the e_{it} is drawn from an AR1 process with the specified level of inter-temporal correlation. We simulate each user's actual subscription choice using a Monte Carlo simulation based on the probabilities $(\Phi_{it}^H,\ \Phi_{it}^L)$, where a random number from a uniform distribution between 0 and 1 is drawn and evaluated against the threshold $\Phi_{it}^H,\ \Phi_{it}^H + \Phi_{it}^L$ to determine which choice the user makes.

We next discuss how we specify the peer relationship strength in period $t+1$ given the visits of all users in periods $s = 1, \cdots, t$ and the peer relationship strengths between all users in period t. Note that we do not model the endogenous network formation in a structural way. Thus to simulate the peer relationships strengths in a given period $t+1$, we use a reduced form approach, which captures the essence of structural network formation model. In the reduced form approach, we make the following assumption. We do not consider new users joining the network and new connections formed between users who are not peers in the data. This means the connection strength matrix $\mathbf{\Omega}_t$ has the same dimension across periods and $\Omega_{ijt'}$ will remain 0 for all subsequent periods t' if $\Omega_{ijt} = 0$ in the data. We next discuss how we compute the peer relationship strength between users

i and j in period $t+1$ in the simulation, given their peer relationship strengths in period t in the simulation and given the simulated number of visits of users i and j in prior periods. The challenge here is that from the model, we can only simulate the visits of users i and j to the gym in each period - however, recall that our operationalization of the peer relationship strength of users i and j in period $t+1$ is based on the share of prior joint visits of users i and j to the 'same classes' in the gym (and not merely just their visits to the gym). Thus, to calculate the peer relationship strength between users i and j in period $t+1$ from the simulated number of visits of users i and j in the prior periods, we need some kind of a mapping from the simulated number of visits of users i and j to the gym in a given period to the number of visits of users i and j to the 'same classes' in the gym in that same period.[22] We have tried two approaches here, where both approaches yielded similar answers. In what follows, we discuss the two approaches, where we have followed Approach 2 in our simulations.

Approach 1: In this approach, we assume that if users i and j had positive number of simulated visits in period t, then their shared visits to the same classes in period t will be a fraction of the minimum of the simulated visits of either users i and j in period t, where this fraction is estimated from the data.

For i and j, we assume that the shared visits between i and j to the same classes in period t, G_{ijt}, is a fraction of the minimum of the visits of either i and j if $\Omega_{ijt} > 0$, i.e., $G_{ijt} = \alpha * min(Visit_{it}, Visit_{jt})$, where α is a fraction between 0 and 1 that is estimated from the data, $Visit_{it}$ is the simulated number of user i's gym visits in period t and $Visit_{jt}$ is the simulated number of gym visits of user j in period t. To see how α is estimated from the data, recall that in the data, we observe the visits of all users to the gym in period t, and we also observe the visits of all pairs of users to the same classes in the gym in period t. Thus we can regress the observed joint visits of any pair of users i and j to the same classes in the gym in period t on the observed minimum of number of visits to the gym of users i and j in period t. In this regression, α is the coefficient of the observed minimum of number of visits to the gym of users i and j in period t. Under this assumption if $Visit_{it} > 0, Visit_{jt} > 0$, $G_{ijt} = \alpha * min(Visit_{it}, Visit_{jt})$ and if $Visit_{it} = 0$ or $Visit_{jt} = 0$, then $G_{ijt} = 0$. Then we update the total number of shared visits between i and j to the same class up to period t, $g_{ijt} = g_{ijt-1} + G_{ijt}$, where g_{ijt} is the cumulative number of visits of i and j to the same group classes from periods $s = 1$ to t. Given the g_{ijt} for each pair of user i and j in period t, we use the similar formula in section 3.1, i.e., $\Omega_{ijt+1} = \frac{g_{ijt}}{\Sigma_{k=1,k\neq i}^{N} g_{ikt}}$ to to get the value of the peer relationship strength between users i and j in period $t+1$ in the simulation. Note that by estimating α from

[22] As a side note, the same problem would be there in an endogenous network formation model, unless if we were to model the users' decisions on which of the 50+ gym classes to attend in each period.

the data, we ensure that on average, in the simulation, the network in the simulation evolves in a manner similar to what we observe in the data.

Approach 2: For i and j, we define two 0-1 indicator variables D_{it} and D_{jt} such that $D_{it}=1$ if $Visit_{it}>0$ and $D_{jt}=1$ if $Visit_{jt}>0$. If $D_{it}=1, D_{jt}=1$ and $\Omega_{ijt}>0$, we create an intermediary variable $F_{ijt}=\Omega_{ijt}\times\Delta_1$ to capture the increase in connection strength between i and j after joint gym visits in period t. Δ_1 is a strength multiplier that we would expect to be larger than 1. If $D_{it}=0$ or $D_{jt}=0$, the intermediary connection strength between i and j will decrease: $F_{ijt}=\Omega_{ijt}\times\Delta_2$, where we would expect the strength multiplier Δ_2 to be less than 1.

For each individual i, we calculate the intermediary connection strengths for all peers $j\in N_t$ and normalize across the network to get the updated connection strength between i and j in period $t+1$ in the simulation as:

$$\Omega_{ijt+1}=\frac{F_{ijt}}{\Sigma_{k\in N,k\neq i}F_{ikt}} \tag{K.1}$$

We compute the updated connection strength for every individual pair i,j in the network to get the updated network $\mathbf{\Omega}_{t+1}$. Note that after normalization of the intermediary connection strength F_{ijt} in equation (K.1), $\mathbf{\Omega}_{t+1}$ has the same property as $\mathbf{\Omega}_t$: the sum over each row, which represents the total connection strength for each individual, equals 1.

We apply the same multiplier Δ_1 for the connection strengths between individuals who continue to visit the gym together, and Δ_2 for the connection strengths between individuals who stop to visit the gym together. The values of Δ_1 and Δ_2 are estimated from the observed network dynamics in the data. To see how we estimate Δ_1 and Δ_2 from the data, note that from the data, we have the 'observed' peer relationship strengths Ω_{ijt} between all pairs of users i and j over time. As discussed in sections 3.1 and 3.2, these observed peer relationship strengths are calculated from the observed joint visits of all pairs of users i and j to the same classes over time. Also, in the data, we observe the number of visits of all users i and j to the gym in each period. Given these variables in the data, we estimate Δ_1 and Δ_2 using the following regression: $\Omega_{ijt+1}=\Delta_1*\Omega_{ijt}*I(Visit_{it}>0\,and\,Visit_{jt}>0)+\Delta_2*\Omega_{ijt}*I(Visit_{it}=0\,or\,Visit_{jt}=0)$, where $Visit_{it}$ and $Visit_{jt}$ are the observed number of visits of users i and j to the gym in period t, and Ω_{ijt} and Ω_{ijt+1} are the 'observed' peer relationships in the data in periods t and $t+1$ respectively (which as discussed in sections 3.1 and 3.2, are calculated as the share of the observed number of joint visits of users i and j to the same classes in the gym). We use the estimated $\Delta_1=1.04$ and $\Delta_2=0.95$ in the simulations. We substitute the estimated values of Δ_1 and Δ_2 into equation (K.1) to get the values of the peer relationship strengths in period $t+1$ in the simulation. Since $\Delta_1>1$ and $\Delta_2<1$, it follows that the peer relationship strength between users i and j will become stronger if both users go to the gym in the same period and weaker if either of them does not.

By estimating Δ_1 and Δ_2 from the data, we ensure that on average, in the simulation, the network in the simulation evolves in a manner similar to what we observe in the data.

We next discuss how our reduced form approach is a reasonable approximation of how the networks evolve in the real world with the presence/absence of shared activities between people. First, it is grounded in the data since the parameters of the mapping in our reduced form approach are estimated from the data. Second, it aligns with the prior literature on social networks, which has focused on how the network structure evolves with the presence/absence of shared activities between people (Granovetter 1973, Krackhardt et al. 2003, Burt 2002, Suitor et al. 1997). To see what that is so, note that our approach for specifying the network structure over time has two main properties: (i) If any two users i and j increase their attendance in the same gym classes, then their peer relationship strength will increase. (ii) If any two users i and j decrease their attendance in the same gym classes, then their peer relationship strength will decrease. We next discuss how properties (i) and (ii) align closely to those found in prior literature in social networks. Property (i) follows from the seminal network paper by Granovetter (1973) who introduces the concept that the strength of a tie is influenced by factors such as the amount of time spent together, emotional intensity, intimacy, and reciprocal services. In other words, increased joint activities enhance these factors, thereby strengthening the connection over time. Property (ii) follows from the classic paper by Burt (2002) who examines how the absence of interaction leads to the weakening of social ties over time. He emphasizes that ongoing interaction is necessary to maintain strong connections, as ties tend to decay without it. Other than these two properties, our simulation makes the following assumption: no new connections can be formed between users who were not peers (that is, their peer relationship strength was zero). This assumption is reasonable for our simulation since in our counterfactuals, we only consider a duration of 6 months. During this relatively short period of time, there will be few new peer relationships, and even if new peer relationships are formed, the strengths will be a lot weaker than the existing peer relationships. Thus, the new peer relationships will be far less important than existing relationships in influencing users' consumption and subscription decisions.

Simulation for the Reduced Form Model in Equation (B.1): The simulation method for the reduced form model is very similar to that in the structural model we discussed above. In the simulations, we not only have to simulate the visit decisions of all users in equilibrium in each period $t = 1, \cdots, T$, but also how the latent peer relationship strengths between all pairs of users, m_{ijt}, evolves over time. Note that unlike the simulation in the structural model, we do not simulate the plan choice/subscription decisions in the reduced form model. Also, unlike the simulation in the structural model in which we specified the evolution the peer relationship strengths Ω_{ijt}, in the reduced form model, we specify the evolution of the latent peer relationship strengths, m_{ijt}. Simulating the users' visit decisions and the evolution of the latent peer relationship strengths over

periods $t = 1, \cdots, T$ entails following two steps in each time period t: (i) given the latent peer relationship strengths between all pairs of users in period t, we would need to simulate the visit decisions to the gym of all users in period t, and (ii) given the latent peer relationship strengths between all pairs of users in period t and the simulated visit decisions of all users prior to period t, we would need to specify how the peer relationship strengths between all pairs of users evolve in period $t+1$. Note that the simulations of the visit decisions in (i) follow directly from the specification of the visit decisions in equation (B.1). However, specifying how the network structure evolves in (ii) does not follow directly from equation (B.1). Thus, similar to what we did in the structural model, we use a reduced form approach for specifying the evolution of the latent peer relationship strengths in (ii). We discuss these as follows.

For step (i), to simulate the gym visits of all users in the network in period t, we start with equation (B.1) for each user $i = 1, \cdots, N$. We randomly draw the usage errors (see equation (B.1)) for all N users to get a system of N linear equations (which are visit equations for all users $i = 1, \cdots, N$ with N unknowns (which are the visits of all N users in the network in that period t). Solving these equations yields the number of visits of all users in period t. For step (ii), we update the latent peer relationship strengths between all pairs of users for period $t+1$, m_{ijt+1}, given the simulated visits of all users prior to period t and the latent peer relationship strengths in period t, m_{ijt}. We tried the same two updating approaches for the latent peer relationship strength terms as in the simulation of the structural model. We got similar results for both approaches. Thus we only report the results from Approach 2. For Approach 2, we use $\Delta_1 = 1.04$ and $\Delta_2 = 0.98$, which are estimated from the observed gym visits and latent peer relationship strengths data.

L. Simulation of Benchmark Targeting Campaign

In this section, we explain how we compute the promotional discounts offered to targeted users in the benchmark targeting campaign that mitigates the negative impact from the high centrality churner in the counterfactual section 6.2.

Note that in equation (3), the loss of the churner will diminish the exercise utility of the churner's direct peer. To make up for the loss, the counterfactual strategy we explore here is to offer a discount for the premium plan to each of the churner's direct peers. With a premium discount, the subscription probability of the premium plan Φ_{it}^H in equation (7a) for the direct peers will increase, while Φ_{it}^L in equation (7b) and the un-subscription probability of the direct peers will decrease. As a result, the direct peer will visit the gym more under the premium plan, and the expected gym visits in equation (7c) will increase. Since the users' exercise state levels and the connection strengths with other users will change across months, the discount will also be dynamically adjusted.

Specifically, let i be a direct peer of the churner. In month t, we first calculate the expected visits in equation (7c) for every user, assuming the churner had not left, and compute the average

expected visits in the gym. With the churn, we replace the original premium plan price P_t^H with a trial discounted $\tilde{P}_{it}^H$ for all direct peers i of the churner and compute the average expected visit in the gym under the trial discounts. We keep lowering $\tilde{P}_{it}^H$ from the original P_t^H, until the average expected visit of all the users in the network is equal to the original level without the churn. The discount is set at $\delta_{it} = P_t^H - \tilde{P}_{it}^H$. It is offered specifically for the direct peer i and can only be redeemed if the user subscribes to the premium plan. Based on the simulated visits under the promotional discounts in the month t, we update the exercise states $\tilde{S}_{it+1}$ with equations (11) and (12), and update the network structure $\mathbf{\Omega}_{t+1}$ with the method explained in section K. Then we repeat the above process to compute the discounts δ_{it+1} in the month $t+1$. We look at the targeting campaigns with a maximum of six-month duration, i.e., discounts are offered from one month to six consecutive months since the churn.