MIS Quarterly

RESEARCH ARTICLE

# MONETIZING FREEMIUM COMMUNITIES: DOES PAYING FOR PREMIUM INCREASE SOCIAL ENGAGEMENT?[1]

**Ravi Bapna**
Carlson School of Management, University of Minnesota,
Minneapolis, MN 55455 U.S.A. {rbapna@umn.edu}

**Jui Ramaprasad**
Desautels Faculty of Management, McGill University,
Montréal, QC H4A 1G5 CANADA {jui.ramaprasad@mcgill.ca}

**Akhmed Umyarov**
Carlson School of Management, University of Minnesota,
Minneapolis, MN 55455 U.S.A. {akhmed.umyarov@gmail.com}

*Making sustainable profits from a baseline zero price and motivating free consumers to convert to premium subscribers is a continuing challenge for all freemium communities. Prior research has causally established that social engagement (Oestreicher-Singer and Zalmanson 2013) and peer influence (Bapna and Umyarov 2015) are two important drivers of users converting to premium subscribers in such communities. In this paper, we flip the perspective of prior research and ask whether the decision to pay for a premium subscription causes users to become more socially engaged. In the context of the Last.fm music listening freemium social community, we establish, using a novel 41-month-long panel dataset, a look-ahead propensity score matching (LA-PSM) procedure coupled with a difference-in-difference estimator of the treatment effect, that payment for premium leads to more social engagement. Specifically, we find that paying for premium leads to an increase in both content-related and community-related social engagement. Free users who convert to premium listen to 287.2% more songs, create 1.92% more playlists, exhibit a 2.01% increase in the number of forum posts made, and gain 15.77% more friends. Thus, premium subscribers create value not only for themselves by consuming more content, but also for the community and site by organizing more content and adding more friends, who are subsequently engaged by the social diffusion emerging from the focal user's activities.*

**Keywords**: Freemium, social engagement, monetization, premium subscription

[1]Bin Gu was the accepting senior editor for this paper. Gal Oestreicher-Singer served as the associate editor. Authors are listed alphabetically and contributed equally to this article.

The appendices for this paper are located in the "Online Supplements" section of the *MIS Quarterly*'s website (http://www.misq.org).

## Introduction and Background

The freemium business model is a widely used model[2] in the monetization strategies of modern software applications, mobile applications, and games. Freemium, coined by venture capitalist Fred Wilson and popularized by Chris Anderson, is defined by the simultaneous presence of a free version along with a paid premium version. The former offers a basic product or set of site features that provide a satisfying experience for the majority of users, while the latter adds enhanced features, which can come in many shapes and sizes. Anderson (2009) has a typology for different variations of freemium. Traditionally, the freemium model was employed in such contexts as software offerings, where users receive a reduced functionality version of the software for free and can get an enhanced version for a price, for example, Adobe Photoshop CS 5 (Niculescu and Wu 2010). The freemium business model has also been incorporated into online social communities such as LinkedIn, okCupid.com, Spotify, and Last.fm, the context of this study. Last.fm allows free users to listen to songs with advertisements, while premium users have an uninterrupted listening experience. The focus of this research is on improving our understanding of the link between premium subscription in freemium social communities and subsequent social engagement. We operationalize social engagement along two dimensions, content and community, capturing key activities such as number of songs listened to and the number of playlists created for the former dimension, and involvement in community forums and adding friends for the latter dimension. Thus, our overarching research question is: What is the effect of premium subscription adoption on content-related engagement and community-related engagement in the context of the freemium business model?

Prior research has linked social engagement and peer influence to increased premium subscriptions. Oestreicher-Singer and Zalmanson (2013) demonstrate that engaging socially in the freemium community leads to a higher willingness-to-pay for premium. Bapna and Umyarov (2015), by virtue of a large-scale randomized trial, demonstrate the presence of causal peer effects where paid subscribers influence their non-subscriber friends to pay on Last.fm. While prior literature has focused on drivers of adoption of premium subscription, in this paper we flip the question and ask what the impact of premium adoption is on the overall monetization of the community. Specifically, we ask whether users who adopt the premium service increase their social engagement on the site, and thereby increase the vitality of both the free and premium channels of monetization. It is critical to note that, in our context, the features that enable social engagement are equally available for both free and premium versions. The only difference in these two versions is the lack of advertising in the latter. Thus, our context is different from other products (e.g., technical software such as Stata or Adobe) that offer a significantly different set of features for premium users. If users were to become more socially engaged after converting to premium, say by creating more playlists or adding more friends to the network, they would create value not only for themselves by consuming more content, but also for the community and site. For instance, by adding more friends, who are subsequently engaged by the social diffusion emerging from the focal user's activities, they would bring an enhanced audience to advertisers. Further, premium users would be engaging and influencing friends who are likely to be free users, which in turn could increase the propensity of the new free users to convert to premium. Thus, premium users potentially have a customer lifetime value that goes beyond the net present value of their monthly subscriptions. Understanding and capturing this network value can have major implications for the site and their monetization strategies. The interlinkage between the free and premium streams of monetization, particularly through social engagement, has not yet been considered in the extant literature, and is the focus of this study.

We are motivated to ask whether premium adoption increases social engagement by multiple streams of literature that link payment for products or services to subsequent engagement with that product or service. First, established work in social psychology suggests that payment for a product or service induces a change in behavior with respect to the usage of the product by a desire to extract value. For instance, equity theory states that individuals seek to achieve a balance between inputs, such as contributions such as time, effort, and, in our case, money, and outputs, such as status, recognition, and, in our case, social engagement (Adams 1965). Similar predictions are made by theories of cognitive dissonance (Festinger 1957) that posit that individuals seek to achieve balance between inputs and outputs in order to minimize dissonance. In the same vein, work looking at the impact of pricing on the efficacy of a product finds that simply discounting the price of a performance drink reduces the performance of the subjects for both mental and physical tasks (Shiv et al. 2005). Further, theories of commitment as conceptualized by Bateman et al. (2011) and by marketing scholars such as Beatty and Kahle (1988) posit user engagement creates commitment and loyalty, which encourages payment as well as—critical for this study—active usage. Drawing on these results, we would predict an increase in social engagement by users who adopt and pay for the premium subscription.

The role of commitment has also been considered by economists who predict enhanced engagement or usage in

[2]Anderson (2009) called it "the business model of the 21st century."

correspondence with higher payment, as it is rational from a utility maximization perspective. For instance, users paying a higher fee structure for , say, a gym membership may do so as a commitment device to themselves to visit the gym with more frequency (Vigna and Malmendier 2006). Interestingly, the empirical evidence does not conform to this last prediction. Vigna and Malmendier (2006) found that consumers who sign up for an annual contract do not necessarily live up to their own commitment and, in fact, end up paying $17 per visit when they would have been able to avail of a $10 per visit pricing option. They explain this departure from rational expectations as overconfidence about future attendance and use on the part of those who opt for the annual contract. If, in our context, users had similar overconfidence about their level of engagement when becoming premium, we would expect to see a negative linkage between premium adoption and social engagement.

As the extant literature described above postulates multiple mechanisms and competing outcomes for the relationship between payment and social engagement, we believe that a crucial first step is to treat this linkage as an empirical question and examine the impact of premium adoption on engagement in a freemium platform. We expect future research to focus on teasing out the specific mechanism(s) driving this linkage.

Using an enhanced matching-based treatment effect estimation procedure described in detail in the "Methodology" section, we establish that payment for premium leads to more social engagement along both content and community dimensions. Our results show that, due to premium adoption, premium users listen to 287.2% more songs, create 1.92% more playlists, exhibit a 2.01% increase in the number of forum posts made, and gain 15.77% more friends. Our results are robust to multiple alternative explanations. For instance, one could theoretically argue that premium gives the users better technology (i.e., special features such as access to music on mobile devices). It could be the case that this better "hammer" could be the sole driver of the differences we observe in engagement between the treatment and control group. This would naturally break down the theoretical motivation of cognitive dissonance and equity theory playing a role in the increased engagement after paying for premium. To rule this out, we randomly selected 2,000 users from the entire active population of more than one million in the Last.fm social network and gifted 1,000 of them the premium subscription (treatment group) while the other 1,000 did not get this subscription (control group). We found no significant increase in engagement for the gifted users who now had access to the better technology. Engagement increases only among the users who self-select into adopting a premium subscription.

Our work contributes to the information systems literature in multiple ways. Given that there is scant research on the ecosystem of monetizing freemium, we conceptualize the notion of the network value of users based on their social engagement. As mentioned earlier, premium users, by virtue of becoming more socially engaged, create value not just for themselves but for others in the network. We also link the monetization of freemium to the recent advances at the intersection of behavioral economics and marketing (Shiv et al. 2005). Our work also contributes to the causal modeling of key questions in online social communities, where we seek to understand fundamental and complex human behavior playing out in the rapidly growing online social graph. Specifically, we develop a new matching technique, look-ahead propensity score matching (LA-PSM), and demonstrate via a simulation that this technique is superior to both traditional propensity score matching and instrumental variable regression in the context of rare events involving economic decisions.

The remainder of this paper is organized as follows. First, we review the related literature. We then provide institutional details and describe our data. In the subsequent section, we present our methodology, which compares our LA-PSM approach with the standard PSM approach as well as the coarsened exact matching (CEM) method. We then present our results and, finally, our conclusions with a discussion and directions for future research.

## Literature Review

### *Freemium Models*

Freemium models bring together a variety of features that have been well studied in the IS literature. Although there are multiple implementations of the freemium model (Anderson 2009), we focus on freemium social communities, which, according to Niculescu and Wu (2010), can be categorized as feature-limited freemium. For instance, communities such as LinkedIn, Match.com, and Last.fm offer a basic version of their site with limited functionality for free, and a premium version with enhanced features. Freemium social communities are two-sided markets subject to the laws of platform competition (Parker and Van Alstyne 2005). A key idea in such markets is the presence of a money-side and subsidy-side and that these sides are linked via cross-side network externalities. Free users on the freemium site are the subsidy-side, linked typically to advertisers who are the money-side.

Unlike traditional two-sided platforms such as Facebook and Google, which only have free users, freemium sites have premium users who often get additional features by paying for a subscription. This aspect of freemium draws from the market segmentation via versioning literature (Bhargava and Choudhary 2001;Raghunathan 2000) as well as the product sampling of digital experience goods literature (Chellappa and Shivendu 2005). The free version and premium version are vertically differentiated offerings, and the key question in such settings is how to achieve optimal second-degree price discrimination without cannibalization. Online social communities can be thought of as digital, experience goods with the free version allowing users to sample a limited set of features as well as socially engage with the community. Through sampling, users learn about the value of the different features of the site and can choose to start climbing a ladder of participation (Li and Bernoff 2011), which has been linked to the willingness-to-pay for a premium subscription (Oestreicher-Singer and Zalmanson 2013).

In summary, freemium sites are two-sided platforms with versioning, and include embedded social communities. These three components draw from three different sets of literature each of which has their own set of optimal decisions. The platform literature describes well the literature on drawing a critical mass of free users that make the site attractive to advertisers. The versioning literature informs decisions on the features offered at each tier to achieve the optimal degree of vertical differentiation and segmentation. This is critical for giving users the sampling experience necessary to convert them from free to premium. W believe social engagement is critical to attract, maintain and retain both free and premium users and ensure the future viability of the site (Bateman et al. 2011). We discuss this stream of literature in more detail next.

## Online Social Engagement

Social engagement, in the form of user generated content, social interactions, and content consumption is an important aspect of freemium online social communities. We draw upon the large body of IS literature that examines antecedents and consequences of social engagement (Oestreicher-Singer and Zalmanson 2013). This entails embracing social features as a means of stimulating consumption of content as well as prolonging users' length of stay and engagement. Prior research has shown that individuals are motivated to participate in and contribute to online communities for reasons such as altruism (Andreoni 1989), reciprocity (Rabin 1993), status-seeking, and social identity/conformity. Relevant to our work is the characterization of commitment by Bateman et al. (2011). This characterization includes both content consumption and user engagement. Joyce and Kraut (2006) and Lampe and Johnston (2005) show that engagement is driven by reciprocity, providing others what they have provided to you, and feedback respectively. As has been noted by marketing scholars, ensuring user engagement creates commitment and loyalty, which encourages payment (Beatty and Kahle 1988) as well as active usage and user retention. Shmargad (2016) studies the relationship between social ties and a consumer's willingness-to-pay for continuing to use a social media service previously available free of charge. He finds that strong ties increase the likelihood of paying for the service more than weak ties. The next sub-section motivates our examination of the impact of premium subscription on social engagement.

## Payment and Value

In examining the link between payment and social engagement, we look to the literature that links payment for a product or service to a change in behavior driven by the desire to extract value. Both equity theory (Adams 1965) and theories of cognitive dissonance (Festinger 1957) posit that individuals seek to achieve balance between inputs and outputs in order to minimize dissonance. In our context, individuals paying for premium can be considered to be providing an explicit monetary input, as put forth in equity theory. Thus, it is natural to expect that they seek higher levels of output, that is, increased utility from the site. In freemium social communities this utility can come from higher levels of social engagement, by listening to more music and engaging and interacting with other users. As per the theory of cognitive dissonance (Festinger 1957), by using the services, the premium user is able to justify for herself why she paid for the service in the first place. Such justification reduces the cognitive burden of having made two incompatible decisions: paying for premium service and not using the premium service.

Further, advances in the marketing literature that link price signals on the actual efficacy of the product (Shiv et al. 2005) also provide a theoretical basis for linking paying for premium to increased social engagement. Shiv et al. (2005) find that the mechanism through which price relates to actual product efficacy is that the price triggers certain response expectancies, the expectations associated with the intrinsic quality of the particular product. Price discounts, for instance, may trigger beliefs that the product's quality is inferior and therefore lead to lower levels of efficacy in the actual usage of the product. Shiv et al. find this effect using a randomized experiment with a discounted energy drink, Sobe Adrenaline

Rush, where subjects were tasked with solving puzzles. In this experiment, 125 participants were randomly assigned using a 2×2 design, where the price was either regular or discounted and the expectancy settings[3] were manipulated to be high or low. They find that the discounted price leads to a lower number of puzzles solved in both the expectancy settings and that this difference in the number of puzzles solved (between the regular and discounted price) was stronger in the high expectancy strength settings. Together, prior research suggests that price—in our context, going from free to paying for premium—not only impacts perceptions of quality but also actual behavior—in our context, social engagement—through the desire to reduce cognitive dissonance and through the role of expectancies. To the best of our knowledge, this link between payment and subsequent social engagement has not been considered in the IS literature. If this is established, further research can delve into separating out the underlying mechanisms of the link. The next section provides institutional details for our context and describes our data.

## Institutional Details and Data

### Institutional Details

We study the freemium community on Last.fm, the context of several prior social network studies in Information Systems (Bapna and Umyarov 2015; Garg et al. 2011; Oestreicher-Singer and Zalmanson 2013). These studies examine peer-influence-based music consumption, online social engagement, and peer influence in premium subscription. Last.fm is a music listening social network, where free users and paid users simultaneously interact in a social community. Free users of the Last.fm website can listen to the online radio interrupted by commercials, while paid premium subscribers of the Last.fm website enjoy a continuous commercial-free music listening experience, a prestigious black "Subscriber" icon next to their user avatars that is visible to everyone on Last.fm as a sign of status, have the ability to listen to the online radio on a mobile phone, and have access to additional colorful music statistical charts. Last.fm, like other freemium communities, employs numerous social features (Parameswaran and Whinston 2007), such as a friendship social network feature that allows website users to become listed as online friends with another website user. Being an online friend with someone typically gives certain benefits: friends can easily share information amongst themselves and exert peer influence on each other. Online friends can affect each other's music choices while sharing their own music listening experiences, can listen to friend's "recommended radio," and can review friend's "loved tracks." They can also love tracks, make new friends, and interact with other users through giving them a "shout" and posting on their "wall," similar to Facebook. Appendix A provides a snapshot of a typical Last.fm user's page.

### Data

Our dataset was collected by our custom multi-threaded, Amazon Cloud-based crawler. It represents a panel of approximately 3.9 million users that make up the largest connected component of the Last.fm network. These 3.9 million users form approximately 23 million friendship pairs, and have been tracked consistently as a panel since May 2011, with updates roughly every 3 weeks. These dynamic updates provided us with fresh snapshots of the entire social network containing the list of friends and premium subscription status for every user. In addition to this information, we have tracked self-reported demographic information and website-reported social activity information. For every snapshot at time *t*, we have collected the following data for each user:

- $Subscriber_{i,t}$. Binary variable indicating whether user *i* is a premium subscriber at time *t*.
- $Age_{i,t}$. Self-reported age of user *i* at time *t*. Age distribution was truncated to the interval between 18 and 80 in order to eliminate outlier data points that are likely fake.
- $FriendCount_{i,t}$. Total count of number of friends of user *i* at time *t*.
- $Gender_i$. Self-reported gender of user *i*. Binary variable.
- $LastFMCountry_{i,t}$. Binary variable. If user *i*'s self-reported country is USA, Germany, or UK, then LastfmCountry = 1 for this user, otherwise 0. This variable is important because Last.fm subscription rules were slightly different between those countries and the rest of the world during the period that we studied.[4]

[3]Subjects in the high-expectancy strength conditions were shown the following statements, designed to trigger the process of response expectancy: "I feel that SoBe is 'very bad' (1)/'very good' (7) at improving concentration," and "I feel that SoBe is 'very bad' (1)/'very good' (7) at improving mental performance." Subjects in the low expectancy condition were not shown these statements (Shiv et al. 2005).

[4]Even though the premium subscription costs approximately the same amount for every country, the subscription is more valuable for people outside the USA, Germany, and the UK. Several Last.fm services that were normally free for users in the USA, Germany, and the UK require premium subscription for the rest of the world because of music licensing contracts.

- $\text{Playlists}_{i,t}$. Total count of playlists made by user *i* by time *t* on Last.fm.
- $\text{Posts}_{i,t}$. Total cumulative count of forum posts made by user *i* by time *t*.
- $\text{SongsListened}_{i,t}$. Total cumulative count of all songs ever listened and reported to Last.fm by user *i* by time *t*. If a user listened to the same song twice, the song would be counted twice.[5]
- $\text{Tenure}_{i,t}$. The tenure of a user *i* on the website by time *t* (measured in "days since registration").

### Social Engagement Constructs

Following Oestreicher-Singer and Zalmanson (2013), we operationalize social engagement along two dimensions, content and community, capturing key activities such as number of songs listened to and number of playlists created for the former dimension, and involvement in community forums and adding friends for the latter dimension. We begin by presenting summary statistics of our dataset.

### Summary Statistics

The descriptive summary statistics for 3,924,064 Last.fm users are displayed in Table 1. This table provides a breakdown of statistics for free users versus existing subscribers during September 2011. In this particular month, there were 42,996 premium subscribers while 3,881,068 users remained free users in the social network on this website.

A casual glance at the summary statistics suggests the presence of differences in social engagement between premium and free users in terms of listening, forum posts, playlists and number of friends, providing motivation to look into this in a causal fashion. It is reasonable to expect that the decision to become a premium user and the various behaviors that comprise our construct of social engagement are endogenous and therefore simultaneously determined. Further, it is also reasonable to expect that other unobserved characteristics would lead to significant omitted variable bias in any standard empirical specification that does not take endogeneity concerns into account. By just looking at the demographics of premium subscribers as compared to free users, we see that they are older, more likely to be male, and less likely to be from a Last.fm country. This suggests that a specific type of user selects into purchasing premium subscriptions. This also makes it difficult to draw causal conclusions from such statistics.

In the next section we describe our novel empirical approach, which mitigates these issues and allows us to draw causal inferences about the relationship between payment and social engagement.

## Methodology

The ideal method to understand causal relationships is through running randomized trials (Aral and Walker 2011), which has been employed recently in IS (Bapna et al. 2016; Bapna and Umyarov 2015; Molitor et al. 2016). However, random assignment of premium adoption is infeasible for us given that our goal is to study the causal behavioral changes resulting from the unprovoked, naturally occurring decision to adopt. In the absence of random assignment, a common approach toward causal inferences is to simulate exogenous adoption through naturally occurring experiments or instrumental variables. Given that adoption of premium subscription is a rare event (only 1.1% of Last.fm users are premium subscribers at any given moment), finding an instrument among the observed variables with enough statistical power is inherently difficult. Thus, the IV approach does not work in our context. To surmount these challenges, we utilize a combination of matching techniques, in the spirit of, but substantially building upon, propensity score matching (Aral et al. 2009; Mithas and Krishnan 2009; Rosenbaum and Rubin 1983), and difference-in-difference (DID) analysis to estimate the treatment effect. The DID analysis allows us to see the difference between the matched treatment and control groups' change in engagement pre- and post-adoption of the premium subscription. Layering the DID analysis on top of the matching-based techniques for causal inference allows us to further account for unobservable covariates and is an approach that has been used recently in IS research (Goh et al. 2013; Rishika et al. 2013; Xu et al. 2016). In addition, we also rely heavily on falsification tests to ensure robust results. We expand on this three-stage process—matching, DID, and falsification—in the remainder of this section.

Standard propensity score matching is a two-step approach where first a predictive model is built to predict the likelihood of a user being treated based on the user's own observable characteristics. This model gives us the propensity-to-be-treated scores that allow us to take the observed treatment

[5]The maximum number of songsListened is capped at 1,000,000. There were a small number of user accounts that exceeded that limit. We have manually verified several of these users and they appeared to be legitimate and long-standing human user accounts. These users appeared to have used some custom software to upload a list of listened songs from their past. In order to avoid the effect of this small number of extreme outliers, we capped the SongsListened variable at 1,000,000.

**Table 1. User Summary**

| Subscriber | Variable | Mean | St. Dev. | Min. | Median | Max. |
|---|---|---|---|---|---|---|
| 0 | Age | 24.73 | 6.46 | 18 | 23 | 80 |
| 1 | Age | 30.56 | 9.24 | 18 | 28 | 80 |
| 0 | FriendCount | 12.25 | 49.88 | 1 | 3 | 19,340 |
| 1 | FriendCount | 31.99 | 115.70 | 1 | 9 | 9,856 |
| 0 | Gender (Male = 1) | 0.61 | 0.49 | 0 | 1 | 1 |
| 1 | Gender (Male = 1) | 0.74 | 0.44 | 0 | 1 | 1 |
| 0 | LastFMCountry | 0.37 | 0.48 | 0 | 0 | 1 |
| 1 | LastFMCountry | 0.33 | 0.47 | 0 | 0 | 1 |
| 0 | Playlists | 0.50 | 4.29 | 0 | 0 | 7,023 |
| 1 | Playlists | 1.45 | 7.84 | 0 | 1 | 1,114 |
| 0 | Posts | 3.41 | 89.36 | 0 | 0 | 66,689 |
| 1 | Posts | 25.43 | 447.48 | 0 | 0 | 52,626 |
| 0 | SongsListened | 12,415.28 | 23,851.68 | 0 | 3,413 | 1,000,000 |
| 1 | SongsListened | 29,810.70 | 42,879.58 | 0 | 15,933 | 1,000,000 |
| 0 | Tenure | 1,144.10 | 579.15 | 0 | 1,120 | 3,254 |
| 1 | Tenure | 1,195.09 | 620.63 | 6 | 1,132 | 3,253 |

group of adopters and select a matched control group from the observed non-adopters. The treatment group and the control group are then compared in terms of the differences in the outcome variables.

A key limitation of PSM is that it can only account for observed and observable covariates (Pearl 2009; Shadish et al. 2002), which limits the matching procedure and may potentially lead to estimation bias. For instance, consider an economic decision such as the decision to adopt a premium subscription on Last.fm. An important but unobserved variable here can be the variable that describes whether the person has impatient character and, thus, is not tolerant to advertisements.

This "impatient character" variable is likely to be linked to both the selection into treatment (purchase of premium subscription) and the outcome of interest (say, listening to new songs). However, because this variable is unobserved, regular PSM methods will fail to control for it during matching and, thus, may produce biased results.

In our paper, we overcome this limitation of regular PSM by introducing a novel matching approach called look-ahead propensity score matching (LA-PSM), detailed in the next subsection. LA-PSM suggests that we match existing adopters to users who are currently non-adopters *but who will become adopters in the future* and who, therefore, share the unobserved time-constant characteristics that are generally causing users to become adopters. We subsequently use a difference-in-difference (DID) estimator on this matched sample in order to estimate the treatment effect.

### Look-Ahead Propensity Score Matching (LA-PSM)

Traditional propensity score matching (PSM) is frequently used as a starting point for developing a control group in observational studies with the goal of trying to determine a causal inference about the effect of the treatment. While the traditional PSM method is well-known, it suffers from the major limitation that it can only match users based on their observed characteristics. However, given the context of an economic decision, it is natural to believe that there is a range of unobserved variables that are different for users who self-select themselves into the treatment group and into the control group, such as distaste for advertisements on webpages, household income level, etc. In this study, we employ the LA-PSM in order to account not just for the observed characteristics in our matching procedure, but also for time-invariant, unobserved characteristics, by taking advantage of our long-term panel data. In our suggested LA-PSM approach, we match current adopters to current non-adopters based on propensity score (just as a regular PSM) but we pick the control group only from the non-adopters who will become adopters in the near future.

Specifically, we focus on a group of users who switched from free to premium in the period between September 2011 and June 2012. We match each of these users in the treatment group to a user who remained a free user in this 10-month period, but who actually adopted premium between June 2012 and January 2015. That is, the matching was not only based on the observable characteristics, but also involved looking ahead and ensuring that the matched control user also did actually switch from free to premium but later, in the subsequent 2.5 year period. Because both the treatment and the matched control group ultimately adopt the premium subscription (only at different times), we account not just for the observed characteristics via propensity-score matching, but also for the unobserved time-invariant characteristics linked to the user's intrinsic propensity to adopt. The matching procedure has successfully found a match for 10,372 treatment users and these users were subsequently compared both before the treatment and after the treatment.

Given the methodological challenges with not being able to use randomized experiments, and the difficulties of finding an instrument with enough statistical power in the context of rare events, this novel modification of PSM takes advantage of our large, longitudinal dataset and allows us to causally identify our treatment effect. While large cross-sectional samples are well known to be desirable for PSM (Shadish et al. 2002) as that increases the likelihood of finding matched treatment-control pairs, we demonstrate that having a significant time dimension to this data allows us to look ahead and use revealed future behaviors to control for the time-invariant unobservables. We believe that this is a methodological contribution to causal inference and is particularly relevant to IS researchers who often have access to such data. For the benefit of future researchers, we demonstrate the effectiveness of LA-PSM as compared to standard PSM in the simulation described in the next section.

## *Effectiveness of LA-PSM Versus PSM: Simulation Analysis*

In this subsection we use a simulation approach to demonstrate how the LA-PSM method is able to outperform the traditional PSM approach in dealing with time-invariant unobservable characteristics that strongly impact both the decision to adopt and the outcome. In particular, we synthetically generate four scenarios where we create a number of simulated users and for each of them, we simulate (1) premium adoption time and (2) outcome variable of interest (say, number of songs listened).

For all of these scenarios, the data is generated such that the null hypothesis is true and there is no true effect of premium adoption on the variable of interest. However, as we go from the first scenario to the fourth one, we gradually make the data generation assumptions less and less ideal (as described below) in order to discover the points where our estimation methods start failing.

We apply three different methods to these four scenarios: (1) a traditional PSM method, (2) our LA-PSM method, and (3) a simulated randomized experiment. Keeping in mind that there is not actually a true treatment effect in any scenario, ideally each estimation method should report the p-values that are insignificant. In other words, if we see an extremely low p-value for some method in some scenario, we can conclude that the method fails for this scenario.

In order to describe the scenarios below, we will introduce some notation: each user is represented by a simulated variable $x$, representing users' observed characteristics, variable $y$, the outcome of interest, variable $z$, an omitted variable that may or may not impact $y$ depending on the scenario and, finally, variable $t$, representing adoption time. Our four scenarios and the corresponding predictions of the estimated treatment effects are

1. *Independence Scenario:* adoption time $t$ is independent of any user characteristics while outcome variable $y$ only depends on the observed user characteristics $x$. In this nearly perfect scenario, we would expect that all three estimation methods (PSM, LA-PSM, and the randomized experiment) would correctly demonstrate the lack of true significant treatment effect.

2. *Perfect observed controls:* adoption time $t$ and outcome variable $y$ both depend on the observed user characteristics $x$ and do not depend on any unobserved characteristics $z$. This scenario illustrates the case when we have good control variables and we don't have any important omitted variables. Again, we would expect that all three estimation methods (PSM, LA-PSM, and the randomized experiment) should be able to discover the lack of true treatment effect.

3. *Omitted time-invariant unobserved variable:* both the adoption time $t$ and outcome variable $y$ depend on the observed user characteristics $x$ *as well as on the unobserved* time-constant user characteristics $z$. For this scenario, we predict that regular PSM will fail to account for the omitted-variable $z$ and, thus, will detect a signifi-

**Table 2. Estimated p-Values When No True Effect Exists**

| | Case 1 | Case 2 | Case 3 | Case 4 |
|---|---|---|---|---|
| Regular PSM | 0.6423 | 0.7539 | 0 | 0 |
| LA PSM | 0.4208 | 0.7971 | 0.2138 | 0 |
| Randomized Experiment | 0.6905 | 0.3835 | 0.2671 | 0.1711 |

cant treatment effect when there is none. However, we expect that LA-PSM will successfully account for the presence of unobserved $z$ since LA-PSM does matching between current adopters and future adopters (who operate in the same unobserved "regime"). We expect that LA-PSM reports insignificant results when there is no true effect. Randomized experiment should also be able to show the lack of significant differences.

4. *Omitted time-variant unobserved variable:* this is the worst-case scenario where both the adoption time $t$ as well as outcome variable $y$ depend on observed $x$ as well as on the *unobserved and time-varying* user characteristics $z$. In this scenario, we expect that all of our matching methods would incorrectly signal the presence of treatment effect even when there is no true effect. The randomized experiment, being the gold standard of causal inference, is the only method in our list that should still correctly detect the lack of actual treatment effect.

Table 2 shows the p-values from the simulation of each of the scenarios described above, and the results are consistent with the aforementioned predictions. It is evident, in column 3, that when there is an unobserved time-invariant variable, PSM gets "tricked" into picking up a significant treatment effect, while LA-PSM accounts for the fact that adopters operate in the unknown but different regime than non-adopters.

In addition to the simulation tests above, we do a similar falsification test based on our actual data. This falsification test is described in detail in the "Results" section while the results of this test for regular PSM, LA-PSM, and CEM are presented in Appendix B.

As can be seen from the results in Appendix B, both the regular PSM and CEM methods fail the falsification test since they report a significant treatment effect even before the treatment has actually been applied. On the other hand (as can be seen from later Table 5), LA-PSM successfully detects the lack of treatment effects before the treatment has actually been applied. We discuss the details of this falsification test in the "Results" section.

### *Implementation of LA-PSM: Static and Dynamic Matching*

Our LA-PSM approach can be implemented in both a static and dynamic manner. In the static implementation, the researcher chooses a treatment window, and matches adopters in that window with future adopters in a second window of appropriate duration that immediately follows the chosen treatment window. Static LA-PSM is our main inference approach. However, it can be argued that there is some subjectivity around the choice of the duration of the treatment window. Hence, we also implement a dynamic LA-PSM method, wherein it is not necessary to *a priori* lock in a treatment duration, and look for matching adopters in the subsequent period of equivalent duration. Rather, we propose implementing a dynamic approach such that as a naturally occurring adoption happens, we can theoretically consider any future adopter as a candidate to be matched. Dynamic LA-PSM works as follows: (1) select all adopters at time period $t$, (2) attempt to match them to a good control group selected among the non-adopters at time $t$ who would necessarily adopt in the future (just after time $t$), (3) once the match at time period $t$ is achieved and treatment and control groups are constructed, we shift our attention to the adopters at time period $t+1$ and repeat the process. In other words, dynamic LA-PSM, attempts to match current adopters to future adopters at every time period $t$. We discuss dynamic LA-PSM in more detail later.

## Results

Before we report our main analysis, we compare the effectiveness of our static LA-PSM matching by comparing the matched treatment and the control group prior to treatment. We ensure that the matching was successful so as there is no difference between the groups prior to treatment in terms of the observed variables using the statistical difference metric employed in d'Agostino (1998). We do this first for the static demographics and cumulative social engagement measures over a three-week period during September 2011, our pre-

**Table 3. Match Comparison Tests for Static LA-PSM**

| Variable | Value |
|---|---|
| Age | 0.0200 |
| LastFMCountry | -0.0073 |
| Gender (Male = 1) | -0.0124 |
| Tenure | -0.0183 |

treatment time period. As demonstrated by the match comparison tests reported in Table 3 all differences are below the standard threshold of 0.1 as suggested by d'Agostino. Therefore, our treatment group T and control group C are similar along observable characteristics prior to treatment.

While Table 3 establishes that the static LA-PSM matching works (i.e., the matched treatment and control groups are similar), later we provide a falsification test that demonstrates a lack of differences between the treatment and control groups immediately before the actual treatment. This test is followed by two robustness checks to address concerns around the duration of our treatment window and pre-treatment trends. As mentioned above, these robustness checks first incorporate a dynamic matching procedure, which reduces the sensitivity of our results to the choice of the duration of our treatment window. In other words, the dynamic matching procedure radically shortens the difference in adoption times between the treatment and control group. Additionally, we rule out the existence of pre-treatment trends impacting selection into the treatment based on the relative time model developed by Greenwood and Agarwal (2015).

## *Impact on Engagement*

Having established the validity of our matching procedure, we estimate the difference-in-difference model comparing the outcomes in the treatment and in the control groups as reported in Table 4. Recall that our treatment group is made up of users who switched from free to premium in the period between September 2011 and June 2012. We match each of these users in the treatment group to a user who remained a free user in this 10-month period, but who actually adopted premium between June 2012 and January 2015. The results suggest that, after adoption of premium subscription, the users exhibit a statistically and economically significant increase in engagement along the dimensions of content-related engagement and community-related engagement on the site. Specifically, due to premium adoption, premium users listen to 287.2% more songs, create 1.92% more playlists, exhibit a 2.01% increase in the number of forum posts made, and gain 15.77% more friends.

## *Falsification Test*

We conducted several robustness checks to rule out alternative explanations. However, before describing these robustness checks we would like to concentrate on a particular falsification procedure that we consider to be the most important robustness test for our analysis. The idea of this falsification test relies on reestimating the exact same DID model with all the same data, model parameters, and constraints but with only one small change: true adoption time is shifted backward by one time period.

Recall that our estimation approach relies on matching followed by DID estimation, which is based on the time of adoption by each user (captured by variable After). This model allows for a natural and very strict falsification procedure: we can shift the real adoption time one month earlier (thus, constructing the "fake" adoption time variable), re-estimate our model with this shifted fake adoption time instead of the real adoption time, and see if our estimation still produces a significant coefficient. Since no true effect is possible before real adoption has occurred, the coefficient from this shifted model with fake adoption time should be insignificant.[6] More importantly, exactly the same model and exactly the same procedures are used in this case both for the real adoption time variable (Table 4) and for the shifted fake adoption time variable (Table 5). Comparing Tables 4 and 5, we can see that our model and identification procedure immediately detects the shift to the earlier fake time and reports that the effect disappears as can be seen by the inter-

[6] In this falsification procedure, since our time is shifted by 1 period earlier, the "after" time period in DID is pointing at the time period right before actual adoption. Thus, there is no reason whatsoever for the treatment and control groups to show any treatment effect since no actual adoption has occurred yet.

**Table 4. DID for Static LA-PSM**

| Dependent Variables | Posts (1) | Playlists (2) | SongsListened (3) | FriendCount (4) |
|---|---|---|---|---|
| Constant | 0.0320*** | 0.0103*** | 3.8206*** | 0.2823*** |
| | (0.0070) | (0.0030) | (0.0804) | (0.0167) |
| Tenure | 0.000003 | 0.00001*** | 0.0002*** | 0.0001*** |
| | (0.000002) | (0.000001) | (0.00002) | (0.000005) |
| Age | 0.0008*** | 0.00003 | 0.0539*** | 0.0035*** |
| | (0.0002) | (0.0001) | (0.0019) | (0.0004) |
| Gender (Male = 1) | 0.0015 | 0.0017 | 0.4265*** | 0.0470*** |
| | (0.0026) | (0.0011) | (0.0302) | (0.0063) |
| LastFMCountry | 0.0047** | 0.0050*** | 0.3696*** | 0.0145** |
| | (0.0024) | (0.0010) | (0.0273) | (0.0057) |
| Adopter | 0.0014 | 0.0011 | 0.0139 | 0.0088 |
| | (0.0032) | (0.0014) | (0.0373) | (0.0078) |
| After | 0.0002 | 0.0012 | 0.0281 | 0.0086 |
| | (0.0033) | (0.0014) | (0.0376) | (0.0078) |
| Adopter * After | 0.0199*** | 0.0190*** | 1.3538*** | 0.1464*** |
| | (0.0045) | (0.0020) | (0.0522) | (0.0109) |
| Observations | 40,313 | 40,313 | 40,304 | 40,313 |
| $R^2$ | 0.0027 | 0.0083 | 0.0942 | 0.0317 |

**Notes**: *p < 0.1; **p < 0.05; ***p < 0.01
Robust heteroscedasticity-consistent standard errors in parentheses

**Table 5. Falsification Test for Static LA-PSM**

| Dependent Variables | Posts (1) | Playlists (2) | SongsListened (3) | FriendCount (4) |
|---|---|---|---|---|
| Constant | 0.0232*** | 0.0013 | 4.0658*** | 0.2998*** |
| | (0.0064) | (0.0020) | (0.0886) | (0.0161) |
| Tenure | 0.000005*** | 0.000002*** | 0.0002*** | 0.0001*** |
| | (0.000002) | (0.000001) | (0.00002) | (0.000004) |
| Age | 0.0005*** | 0.0002*** | 0.0661*** | 0.0042*** |
| | (0.0002) | (0.00005) | (0.0021) | (0.0004) |
| Gender (Male = 1) | 0.0050** | 0.0005 | 0.4587*** | 0.0314*** |
| | (0.0023) | (0.0007) | (0.0322) | (0.0059) |
| LastFMCountry | 0.0016 | 0.0004 | 0.5216*** | 0.0081 |
| | (0.0021) | (0.0007) | (0.0291) | (0.0053) |
| Adopter | 0.0008 | 0.0009 | 0.0943** | 0.0022 |
| | (0.0029) | (0.0009) | (0.0406) | (0.0074) |
| After | 0.0004 | 0.0004 | 0.0009 | 0.0005 |
| | (0.0029) | (0.0009) | (0.0403) | (0.0073) |
| Adopter * After | 0.0031 | 0.0001 | 0.0247 | 0.0082 |
| | (0.0041) | (0.0013) | (0.0562) | (0.0102) |
| Observations | 37,916 | 37,916 | 37,901 | 37,917 |
| $R^2$ | 0.0015 | 0.0015 | 0.0531 | 0.0197 |

**Notes**: *p < 0.1; **p < 0.05; ***p < 0.01
Robust heteroscedasticity-consistent standard errors in parentheses

action coefficient between the adoption variable and the after adoption variable. This falsification procedure ensures that the result in Table 4 is not driven by a spurious correlation, model misspecification, inappropriate matching, software bugs in the statistical procedure, and, most importantly, the unaccounted preexisting pre-treatment trends that could have started prior to actual adoption.

As we show in Appendix B, when we apply the same shifted "fake" adoption time falsification test to both regular PSM and CEM, both of these regular methods fail the falsification test as they seem to detect significant treatment effect even before the treatment has occurred. In other words, the use of LA-PSM based matching is not a luxury or an incremental improvement but a necessity in our scenario.

## Robustness

### Dynamic LA-PSM

Our first robustness check implements a dynamic LA-PSM method, which is analogous in its construction to the regular dynamic PSM model. Just like a regular dynamic PSM model, the idea behind our dynamic matching is to substitute one static matching procedure with continuous month-to-month dynamic matching.

To illustrate the dynamic LA-PSM with a specific example, consider the first several steps of the algorithm: as a first step, we populate our treatment group with all adopters who adopted in September 2011. In order to find the appropriate matches for these treatment users, we look for similar users among the universe of people who have not yet adopted as of September 2011 *but who will adopt*[7] some time between October 2011 and January 2015. Both the treatment group and the corresponding matched control group are then saved for our analysis.

The process then shifts the time window by one month and repeats itself: those users who adopted in October 2011 are collected as the treatment group and are matched to users who have not yet adopted as of October 2011 but who will adopt some time between November 2011 and January 2015. This process repeats again and again for each month until we reach the treatment group month[8] June 2012.

[7]The clause "but who will adopt in the future" is what separates dynamic LA-PSM from a regular dynamic PSM model.

[8]We used the same treatment group cut-off point for both static LA-PSM and dynamic LA-PSM to make the results comparable between the two since dynamic LA-PSM serves as a robustness check for static LA-PSM

As shown in Table 6, the results are qualitatively similar between static LA-PSM and dynamic LA-PSM. Looking at the interaction term coefficient, we observe that coefficient estimates are very close to the estimates obtained from the static LA-PSM in Table 4, and thus, we are more certain of the magnitude of the claimed effect sizes.

In order to confirm that dynamic LA-PSM is a valid estimation method in our scenario, we conduct the same falsification that we did for static LA-PSM described earlier. As can be seen from the results presented in Appendix C, dynamic LA-PSM also passes the falsification test, and therefore is a valid robustness test.

One of the advantages of the dynamic LA-PSM model over a static LA-PSM model is that it shortens the potential time difference in adoption between the users in the treatment and matched control groups. In effect, this serves as a robustness check for the sensitivity of the estimation to the choice of the time and duration of the treatment window.

### Ruling Out Pre-treatment Trends

While the core idea of the shifted time falsification tests above is to demonstrate the lack of pre-treatment trends, we also conduct and demonstrate a number of additional tests as suggested by prior literature. We additionally check for this by testing a period-by-period model based on Greenwood and Agarwal (2015), where we estimate a vector of coefficients for each relative time period and absolute time periods, and also include user-level fixed effects to control for individual-level heterogeneity.

In this equation, the absolute time controls are period dummies based on the calendar periods of three weeks. These are identical for each user in the dataset. The relative time controls, by contrast, are a series of dummies that indicate the chronological distance between the observation and the eventual, or previous, adoption of premium.

Table 7 reports the results from this model applied to our multiple social engagement measures. In this model, a lack of positive coefficients for $t-3$, $t-2$, and $t-1$ suggests that the users did not show any unusual positive spikes in activity prior to treatment. Notice that the signs for $t+0$ and further time periods are all positive, significant, and very similar to our estimated treatment effects, providing further support of our findings above. More importantly, this set of robustness checks gives us confidence in that our results are not driven by pre-treatment heterogeneity in our users.

**Table 6. Dynamic LA-PSM**

| Dependent Variables | Posts (1) | Playlists (2) | SongsListened (3) | FriendCount (4) |
|---|---|---|---|---|
| Constant | 0.0322*** | 0.0071** | 4.0502*** | 0.3109*** |
| | (0.0077) | (0.0033) | (0.0867) | (0.0184) |
| Tenure | 0.00001*** | 0.000004*** | 0.0001*** | 0.0001*** |
| | (0.000001) | (0.000001) | (0.00002) | (0.000004) |
| Age | 0.0008*** | 0.00002 | 0.0621*** | 0.0047*** |
| | (0.0001) | (0.0001) | (0.0014) | (0.0003) |
| Gender (Male = 1) | 0.0053*** | 0.0004 | 0.3976*** | 0.0472*** |
| | (0.0020) | (0.0009) | (0.0227) | (0.0048) |
| LastFMCountry | 0.0071*** | 0.0044*** | 0.4322*** | 0.0119** |
| | (0.0020) | (0.0009) | (0.0225) | (0.0048) |
| Adopter | 0.0017 | 0.0002 | 0.0214 | 0.0169*** |
| | (0.0025) | (0.0011) | (0.0283) | (0.0060) |
| After | 0.0005 | 0.0001 | 0.0036 | 0.0054 |
| | (0.0025) | (0.0011) | (0.0287) | (0.0061) |
| Adopter * After | 0.0196*** | 0.0191*** | 1.4211*** | 0.1347*** |
| | (0.0035) | (0.0015) | (0.0399) | (0.0085) |
| Observations | 69,850 | 69,850 | 69,843 | 69,850 |
| $R^2$ | 0.0036 | 0.0085 | 0.1007 | 0.0386 |

**Notes**: *p < 0.1; **p < 0.05; ***p < 0.01
Robust heteroscedasticity-consistent standard errors in parentheses

**Table 7. Relative Time Model**

| Dependent Variables | Posts (1) | Playlists (2) | SongsListened (3) | FriendCount (4) |
|---|---|---|---|---|
| t – 3 | 0.0028** | 0.00004 | 0.2046*** | 0.0099*** |
| | (0.0011) | (0.0004) | (0.0131) | (0.0025) |
| t – 2 | 0.0011 | 0.0002 | 0.2127*** | 0.0093*** |
| | (0.0011) | (0.0004) | (0.0128) | (0.0024) |
| t – 1 | 0.0011 | 0.0003 | 0.1486*** | 0.0001 |
| | (0.0011) | (0.0004) | (0.0126) | (0.0024) |
| t + 0 | 0.0203*** | 0.0197*** | 1.3990*** | 0.1329*** |
| | (0.0010) | (0.0004) | (0.0125) | (0.0024) |
| t + 1 | 0.0112*** | 0.0062*** | 1.2240*** | 0.0892*** |
| | (0.0010) | (0.0003) | (0.0122) | (0.0024) |
| t + 2 | 0.0051*** | 0.0020*** | 0.7987*** | 0.0330*** |
| | (0.0010) | (0.0004) | (0.0123) | (0.0024) |
| t + 3 | 0.0019* | 0.0017*** | 0.5892*** | 0.0109*** |
| | (0.0010) | (0.0004) | (0.0124) | (0.0024) |
| Observations | 1,028,248 | 1,028,248 | 1,027,372 | 1,051,674 |
| $R^2$ | 0.0014 | 0.0034 | 0.0617 | 0.0125 |

**Notes**: *p < 0.1; **p < 0.05; ***p < 0.01
Robust heteroscedasticity-consistent standard errors in parentheses

While we use the relative time model[9] primarily for robustness purposes, Table 7 does hint at the possibility that the effects could be declining with time after adoption. This suggests that the value capture in terms of the network happens immediately. We would, however, caution against making any strong claims about fleeting effects here since there is another possible explanation: the observed coefficient can be reduced noticeably after 1 month simply because some of the users who purchased a subscription for 1 month do not proceed to buy premium subscription for a second month. In other words, at $t + 0$, every user is a premium subscriber.[10] However, by $t + 3$, some people will stop being premium subscribers. It is basically guaranteed that at least some adopters must have dropped their premium subscription after three time periods and, therefore, may have reduced the use of the website to their normal levels. In other words, even if paying users have no reduction in use in month 3, coefficient for $t + 3$ may still decline just because there are fewer paying users.

### Ruling out Similar Increase in Social Engagement in Artificially Induced Premium Adopters

Recall that the research question in which we are interested is limited to the post-adoption behavior in the universe of self-selected adopters,[11] that is, those users who choose to adopt a premium subscription on their own volition. A complementary question is whether users who chose *not to become premium* were forced to become premium anyway (e.g., through receiving a premium subscription gift given randomly to a non-subscriber) exhibit similar behaviors. In particular, one argument that we may face is that premium users become more engaged because they have more features or arguably better technology, say, the ability to listen to music on their mobile devices. Theoretically, having this additional channel of access to music could be the sole driver of the differences we observe in engagement between the treatment and control group. If this were to be the case, it would break down the theoretical motivation of cognitive dissonance and equity theory playing a role in the increased engagement after paying for premium. To rule this out, we randomly selected 2,000 users from the active user population of Last.fm social network and gifted 1,000 of them the premium subscription (treatment group) while the other 1,000 did not receive this subscription (control group). Table 8 shows the lack of effect of premium subscription on social engagement for subscribers who were gifted, as contrasted with subscribers from Table 4 who chose to purchase the premium subscription themselves. Note that, unlike our prior analysis, these 1,000 gifted premium users were selected at random among non-premium users and thus, did not undergo the conscious process of choosing to convert from free to premium. This reaffirms our main finding that users who are choosing to pay for premium experience a different treatment effect from the premium subscription compared to random active website users.

### Regular Matching Methods

In addition to the robustness tests presented above, we also ran numerous regular matching methods including but not limited to static regular PSM, dynamic regular PSM, static CEM, and dynamic CEM. All these methods report significant treatment effects similar to the ones that were obtained with LA-PSM so in the interest of space we do not repeat them.

However, all these regular methods *fail our falsification test* described earlier so these do not act as a valid robustness test. For illustration purposes, we present only a few examples of falsification test failures in Appendix B. All these matching methods fail because they produce matches that seem to be different in some important unobserved characteristics even before the actual treatment has occurred. This can be contrasted with falsification tests in Table 5 for static LA-PSM and Appendix C for dynamic LA-PSM which both pass the falsification test and show lack of differences before the treatment has occurred.

## Discussion and Conclusions

Freemium business models are the model of choice for a large variety of online social platforms, but the simultaneous presence of free and premium leads to significant monetization challenges. In order to address the challenge, this research combines the extant IS literature related to the monetization of freemium communities and challenges with online platforms, with the behavioral social psychology literature that looks at the impact payment on value extraction. While prior literature has looked at social engagement and peer influence and the impact of each on premium subscrip-

[9]Note that Table 7 includes all time periods before adoption as observations in the regression. For example, consider the observation that corresponds to time $t - 5$, that is, 5 time periods before adoption for a particular user $i$. It is clear that, for this observation, the regressor dummy variables $t - 3$, $t - 2$, $t - 1$, $t + 0$, $t + 1$, $t + 2$, and $t + 3$ are all zeros. Therefore, we do not have a dummy variable trap here since these relative time regressors do not add up to 1 for this observation. This observation will ultimately end up influencing the time fixed effect and user fixed effect.

[10]Just by definition, $t + 0$ is the dummy variable for the time period when user $i$ first buys premium subscription.

[11]A statistical term that here means *the average treatment effect on the treated* (ATET).

**Table 8. No Effect for Gifted Adopters**

| Variable | Mean(Non-gifted) | Mean(Gifted) | t-value | p-value |
|---|---|---|---|---|
| FriendCount | 0.243 | 0.120 | 1.1782 | 0.2390 |
| Playlists | 0.000 | -0.001 | 1.0000 | 0.3176 |
| Posts | 0.008 | 0.035 | -1.4907 | 0.1363 |
| SongsListened | 113.773 | 124.995 | -1.2368 | 0.2163 |

tion, this study flips the agenda and asks how premium subscription impacts behavior in terms of social engagement. In doing this, we connect social engagement, peer influence and premium subscription and shed new light on the network value of consumers who choose to pay premium and become more socially engaged. We draw upon the online communities and social engagement literature and examine the impact of payment on multiple social engagement dimensions, as well as how this impact varies across levels of user experience in the site, in the Last.fm music community.

We contribute by showing that premium users are substantially more engaged along the dimensions of content consumption, content organization, and community involvement—all features that were available to them as free users. Increased content consumption has significant value-chain impact as it increases royalties paid to artists. When premium users create more playlists (i.e., organize content), they make the music listening experience more attractive to other users. Finally, the fact that they make more friends (i.e., add new links to the social graph) and, via the natural process of social diffusion built into online social networks, engage and interact with these new friends, has not previously been considered. There is strong evidence that suggests that the focal user's sharing volume increases with audience size (Burke et al. 2009). Note that, given the overwhelming prevalence of free users on the site, the premium users' new friends are likely to be free users. The consequence of this newly formed friendship is that the activities, content consumption[12] and behaviors of the focal premium user will now be socially diffuse and shared on the free user's wall. This core feature of online social networking is likely to increase the vitality and engagement for the newly friended free user, which, for the site, maps to eyeballs for advertisers. Further, prior research has established the likelihood of driving them to convert to premium via peer influence (Bapna and Umyarov 2015). In summary, engaging and influencing friends who are (likely to be) free users increases advertising revenues, as well as the likelihood, via peer influence, that the free users will convert to premium.

[12]For instance, the new user is likely to get music recommendations on the friend's listening patterns.

The current paper offers an alternative perspective into the inner workings of freemium communities as compared to that of Oestreicher-Singer and Zalmanson (2013), and Bapna and Umyarov (2015). We believe that this opens up a wave of future questions. What profile of users adopt via the social engagement route versus the peer-influence route? Which pathway yields a greater effect on profit? Which yields a greater effect on community? We hope that these will be fertile areas of future research.

Our study has important managerial implications for online platforms that employ the freemium model. The idea that a small subset of a site's users converting to premium has ecosystem-level effects on value has not been recognized thus far and is a novel contribution to the IS literature. By quantifying the significant dimensions of social-engagement-related value, we show that premium users have a customer lifetime value beyond the net present value of their individual monetary contributions. Thus, our work will be useful for developing better monetization strategies for freemium models that take into account the additional components of an individual's behavior that contributes to the site's value. In the absence of our work, freemium sites would be underestimating the customer lifetime value of premium users, who they can now recognize, reward, and strive even harder to retain.

Establishing the linkage between payment and social engagement is not straightforward for a variety of issues related to endogeneity. Specifically, observed differences in behavior between free and premium users cannot be easily attributed to the effect of premium subscription itself because of the potential self-selection of premium users. In this study, we propose a novel propensity score matching technique, LA-PSM, which we use in order to carefully disentangle self-selection from causality in observational data. Unlike the traditional PSM, our technique allows us to match the treatment group based not just on the observed attributes, but also on the unobserved time-invariant attributes by matching the treatment group of existing adopters to the control group of future adopters, who currently have not adopted.

This research has opportunities to be extended. Future research, for example, could quantify the strength of the multiplier effect over time and could examine other social engagement activities available on different, specific realizations of the broader freemium model. Additional work also needs to be done to tease out the specific mechanisms that link premium adoption to social engagement. For instance, there could be multiple possible mechanisms for the increase in social engagement that we find. Take the case of the increase in new friends made by premium users. This could happen as users may want to make friends with more engaged users, or users could select to friend premium users because of their status, or premium users themselves could be more proactive about making friends. Teasing out these mechanisms are interesting directions for future research.

In addition, we believe that the LA-PSM methodology would be useful in establishing causality in situations where researchers have access to longitudinal panels of data and where the outcomes of interest are very rare.

## References


Adams, J. S. 1965. "Inequity in Social Exchange," *Advances in Experimental Social Psychology* (2), pp. 267-299.

Anderson, C. 2009. *Free: The Future of a Radical Price*, New York: Random House.

Andreoni, J. 1989. "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence," *Journal of Political Economy* (97:6), pp. 1447-1458.

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

Aral, S. and Walker, D. 2011. "Creating Social Contagion through Viral Product Design: A Randomized Trial of Peer Influence in Networks," *Management Science* (57:9), pp. 1623-1639.

Bapna, R., Ramaprasad, J., Shmueli, G., and Umyarov, A. 2016. "One-Way Mirrors in Online Dating: A Randomized Field Experiment," *Management Science* (62:11), pp. 3100-3122.

Bapna, R., and Umyarov, A. 2015. "Do Your Online Friends Make You Pay? A Randomized Field Experiment on Peer Influence in Online Social Networks," *Management Science* (61:8), pp. 1902-1920.

Bateman, P. J., Gray, P. H., and Butler, B. S. 2011. "Research Note—The Impact of Community Commitment on Participation in Online Communities," *Information Systems Research* (22:4), pp. 841-854.

Beatty, S. E., and Kahle, L. R. 1988. "Alternative Hierarchies of the Attitude-behavior Relationship: The Impact of Brand Commitment and Habit," *Journal of the Academy of Marketing Science* (16:2), pp. 1-10.

Bhargava, H. K., and Choudhary, V. 2001. "Information Goods and Vertical Differentiation," *Journal of Management Information Systems* (18:2), pp. 89-106.

Burke, M., Marlow, C., and Lento, T. 2009. "Feed Me: Motivating Newcomer Contribution in Social Network Sites," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York: ACM, pp. 945-954.

Chellappa, R. K., and Shivendu, S. 2005. "Managing Piracy: Pricing and Sampling Strategies for Digital Experience Goods in Vertically Segmented Markets," *Information Systems Research* (16:4), pp. 400-417.

d'Agostino, R. B. 1998. "Tutorial in Biostatistics: Propensity Score Methods for Bias Reduction in the Comparison of a Treatment to a Non-randomized Control Group," *Statistics in Medicine* (17:19), pp. 2265-2281.

Festinger, L. 1957. *Cognitive Dissonance Theory*, Newbury Park, CA: Sage Publications.

Garg, R., Smith, M. D., and Telang, R. 2011. "Measuring Information Diffusion in an Online Community," *Journal of Management Information Systems* (28:2), pp. 11-38.

Goh, K. Y., Heng, C. S., and Lin, Z. 2013. "Social Media Brand Community and Consumer Behavior: Quantifying the Relative Impact of User- and Marketer-Generated Content," *Information Systems Research* (24:1), pp. 88-107.

Greenwood, B. N., and Agarwal, R. 2015. "Matching Platforms and Hiv Incidence: An Empirical Investigation of Race, Gender, and Socioeconomic Status. *Management Science* (62:8), pp. 2281-2303.

Joyce, E., and Kraut, R. E. 2006. "Predicting Continued Participation in Newsgroups," *Journal of Computer Mediated Communication* (11:3), pp. 723-747.

Lampe, C., and Johnston, E. 2005. "Follow the (Slash) Dot: Effects of Feedback on New Members in an Online Community," in *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work*, New York: ACM, pp. 11-20.

Li, C., and Bernoff, J. 2011. *Groundswell: Winning in a World Transformed by Social Technologies*, Boston: Harvard Business Press.

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

Molitor, D., Reichhart, P., Spann, M., and Ghose, A. 2016. "Measuring the Effectiveness of Location-Based Advertising: A Randomized Field Experiment" (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2645281).

Niculescu, M. F., and Wu, D. J. 2010. "When Should Software Firms Commercialize New Products via Freemium Business Models," paper presented at the Workshop on Information Systems and Economics, St. Louis, MO, December.

Oestreicher-Singer, G., and Zalmanson, L. 2013. Content or Community? A Digital Business Strategy for Content Providers in the Social Age," *MIS Quarterly* (37:2), pp. 591-616.

Parameswaran, M., and Whinston, A. B. 2007. "Research Issues in Social Computing," *Journal of the Association for Information Systems* (8:6), pp. 336-350.

Parker, G. G., and Van Alstyne, M. W. 2005. "Two-Sided Network Effects: A Theory of Information Product Design," *Management Science* (51:10), pp. 1494-1504.

Pearl, J. 2009. *Causality*, New York: Cambridge University Press.

Rabin, M. 1993. "Incorporating Fairness into Game Theory and Economics," *The American Economic Review* (83:5), pp. 1281-1302.

Raghunathan, S. 2000. "Software Editions: An Application of Segmentation Theory to the Packaged Software Market," *Journal of Management Information Systems* (17:1), pp. 87-113.

Rishika, R., Kumar, A., Janakiraman, R., and Bezawada, R. 2013. "The Effect of Customers' Social Media Participation on Customer Visit Frequency and Profitability: An Empirical Investigation," *Information Systems Research* (24:1), pp. 108-127.

Rosenbaum, P. R., and Rubin, D. B. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika* (70:1), pp. 41-55.

Shadish, W. R., Cook, T. D., and Campbell, D. T. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Boston: Houghton, Mifflin and Company.

Shiv, B., Carmon, Z., and Ariely, D. 2005. "Placebo Effects of Marketing Actions: Consumers May Get What They Pay For," *Journal of Marketing Research* (42:4), pp. 383-393.

Shmargad, Y. 2016. "The Monetization of Information Broadcasts: A Natural Experiment on an Online Social Network," in *AAAI Spring Symposium*, Association for the Advancement of Artificial Intelligence, Palo Alto, CA.

Vigna, S. D., and Malmendier, U. 2006. "Paying Not to Go to the Gym," *The American Economic Review* (96:3), pp. 694-719.

Xu, K., Chan, J., Ghose, A., and Han, S. P. 2016. "Battle of the Channels: The Impact of Tablets on Digital Commerce," *Management Science* (63:5), pp. 1469-1492.

## About the Authors

**Ravi Bapna** is the Curtis L. Carlson Chair in Business Analytics and Information Systems at the Carlson School of Management, University of Minnesota. His research and scholarship views the emerging digital landscape as a giant global laboratory, a sandbox to gain a deeper causal understanding of how consumers, firms, industries, and societies are being reshaped by the big data and business analytics revolution. His professional interests have resulted in research, consulting and executive education engagements with a variety of leading U.S. and Indian companies.

**Jui Ramaprasad** is an associate professor of Information Systems at the Desautels Faculty of Management at McGill University. She received her doctorate in information systems from the Paul Merage School of Business at the University of California, Irvine, and holds a B.S. from the University of Southern California. Her research interests lie in understanding role and impact of social technologies and IT-enabled features in the digital transformation of business models, studied in the context online music communities and online matching platforms. Her work contributes new understanding of how information technology can influence human behavior, shape social interactions, and impact organizations that have been digitally disrupted.

**Akhmed Umyarov** is an assistant professor at the Carlson School of Management, University of Minnesota. He received his Ph.D. from New York University's Stern School of Business. His research focuses on causal identification of human behavior phenomena in online settings. He designed and conducted a number of large-scale field experiments in online platforms for research questions ranging from measuring peer influence in online social networks to exploring subtle signals of interest in online dating to examining the effects of reputation in online labor markets.

MIS Quarterly

RESEARCH ARTICLE

# MONETIZING FREEMIUM COMMUNITIES: DOES PAYING FOR PREMIUM INCREASE SOCIAL ENGAGEMENT?

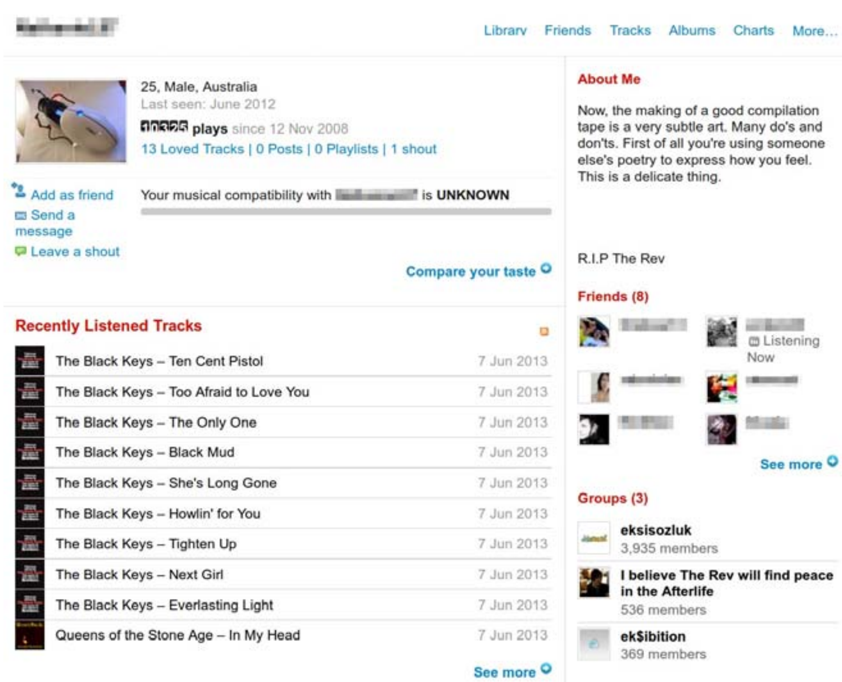
**Ravi Bapna**
Carlson School of Management, University of Minnesota,
Minneapolis, MN 55455 U.S.A. {rbapna@umn.edu}

**Jui Ramaprasad**
Desautels Faculty of Management, McGill University,
Montréal, QC H4A 1G5 CANADA {jui.ramaprasad@mcgill.ca}

**Akhmed Umyarov**
Carlson School of Management, University of Minnesota,
Minneapolis, MN 55455 U.S.A. {akhmed.umyarov@gmail.com}

# Appendix A

## Profile Page of a Random Last.fm User




**Note:** The user profile page lists user attributes such as age, gender, country of origin, number of friends, and others.

# Appendix B

## Falsification Tests for Regular Methods

**Table B1. Falsification Test for Regular PSM**

| | Dependent Variables | | | |
|---|---|---|---|---|
| | **Posts (1)** | **Playlists (2)** | **SongsListened (3)** | **FriendCount (4)** |
| Constant | 0.0079** | 0.0004 | 1.8039*** | 0.1295*** |
| | (0.0037) | (0.0010) | (0.0593) | (0.0088) |
| Tenure | 0.00001*** | 0.000001*** | 0.0002*** | 0.00004*** |
| | (0.000001) | (0.000000) | (0.00001) | (0.000002) |
| Age | 0.0002*** | 0.0001** | 0.0530*** | 0.0026*** |
| | (0.0001) | (0.00002) | (0.0012) | (0.0002) |
| Gender (Male = 1) | 0.0042*** | 0.0005 | 0.5386*** | 0.0114*** |
| | (0.0012) | (0.0003) | (0.0195) | (0.0029) |
| LastFMCountry | 0.0015 | 0.0001 | 0.3174*** | 0.0079*** |
| | (0.0011) | (0.0003) | (0.0185) | (0.0027) |
| Adopter | 0.0162*** | 0.0012*** | 2.2330*** | 0.1051*** |
| | (0.0015) | (0.0004) | (0.0242) | (0.0036) |
| After | 0.0004 | 0.0002 | 0.0136 | 0.0001 |
| | (0.0015) | (0.0004) | (0.0241) | (0.0036) |
| Adopter * After | 0.0013 | 0.0002 | 0.0919*** | 0.0065 |
| | (0.0021) | (0.0006) | (0.0338) | (0.0050) |
| Observations | 105,417 | 105,417 | 105,367 | 105,432 |
| $R^2$ | 0.0036 | 0.0008 | 0.1702 | 0.0271 |

**Notes**: *p < 0.1; **p < 0.05; ***p < 0.01
Robust heteroscedasticity-consistent standard errors in parentheses

**Table B2. Falsification Test for CEM**

| | Dependent Variables | | | |
|---|---|---|---|---|
| | **Posts (1)** | **Playlists (2)** | **SongsListened (3)** | **FriendCount (4)** |
| Constant | 0.0098*** | 0.0005 | 1.6451*** | 0.1365*** |
| | (0.0035) | (0.0010) | (0.0577) | (0.0089) |
| Tenure | 0.00001*** | 0.000001*** | 0.0001*** | 0.0001*** |
| | (0.000001) | (0.000000) | (0.00001) | (0.000002) |
| Age | 0.0003*** | 0.0001*** | 0.0487*** | 0.0026*** |
| | (0.0001) | (0.00002) | (0.0011) | (0.0002) |
| Gender (Male = 1) | 0.0036*** | 0.0004 | 0.5626*** | 0.0156*** |
| | (0.0012) | (0.0003) | (0.0190) | (0.0029) |
| LastFMCountry | 0.0014 | 0.0001 | 0.3093*** | 0.0083*** |
| | (0.0011) | (0.0003) | (0.0181) | (0.0028) |
| Adopter | 0.0178*** | 0.0015*** | 2.2725*** | 0.1139*** |
| | (0.0015) | (0.0004) | (0.0237) | (0.0037) |
| After | 0.0003 | 0.0003 | 0.0075 | 0.0039 |
| | (0.0015) | (0.0004) | (0.0237) | (0.0037) |
| Adopter * After | 0.0018 | 0.0007 | 0.0689** | 0.0047 |
| | (0.0020) | (0.0006) | (0.0332) | (0.0051) |
| Observations | 110,785 | 110,785 | 110,731 | 110,795 |
| $R^2$ | 0.0047 | 0.0009 | 0.1723 | 0.0304 |

**Notes**: *p < 0.1; **p < 0.05; ***p < 0.01
Robust heteroscedasticity-consistent standard errors in parentheses

# Appendix C

## Dynamic LA-PSM Falsification Test

**Table C1. Falsification Test for Dynamic LA-PSM**

| | Dependent Variable | | | |
|---|---|---|---|---|
| | **Posts (1)** | **Playlist (2)** | **SongsListened (3)** | **FriendCount (4)** |
| Constant | 0.0260*** | 0.0013 | 4.0010*** | 0.3104*** |
| | (0.0054) | (0.0016) | (0.0718) | (0.0133) |
| Tenure | 0.00001*** | 0.000002*** | 0.0002*** | 0.0001*** |
| | (0.000001) | (0.000000) | (0.00002) | (0.000003) |
| Age | 0.0004*** | 0.0001** | 0.0708*** | 0.0048*** |
| | (0.0001) | (0.00003) | (0.0015) | (0.0003) |
| Gender (Male = 1) | 0.0069*** | 0.0009 | 0.4449*** | 0.0321*** |
| | (0.0018) | (0.0005) | (0.0237) | (0.0044) |
| LastFMCountry | 0.0016 | 0.0002 | 0.5662*** | 0.0009 |
| | (0.0018) | (0.0005) | (0.0234) | (0.0043) |
| Adopter | 0.0022 | 0.0014** | 0.0403 | 0.0029 |
| | (0.0022) | (0.0006) | (0.0295) | (0.0054) |
| After | 0.0006 | 0.0006 | 0.0258 | 0.0031 |
| | (0.0023) | (0.0007) | (0.0302) | (0.0056) |
| Adopter * After | 0.0014 | 0.0014 | 0.0163 | 0.0125 |
| | (0.0032) | (0.0009) | (0.0416) | (0.0077) |
| Observations | 70,930 | 70,930 | 70,925 | 70,931 |
| $R^2$ | 0.0011 | 0.0007 | 0.0562 | 0.0233 |

**Notes**: *p < 0.1; **p < 0.05; ***p < 0.01
Robust heteroscedasticity-consistent standard errors in parentheses