Are Paid Subscriptions on Music Social Networks Contagious? A Randomized Field Experiment

Ravi Bapna, Akhmed Umyarov[1]

December 2, 2011

Demonstrating compelling causal evidence of the existence and strength of peer to peer influence in online social networks has become the holy grail of the modern research in online social networks. While it has been consistently demonstrated that user characteristics and behavior tend to cluster inside social networks both in space and in time, there are several different mechanisms that can cause this observed clustering. Among the most frequently cited explanations for this effect are two rival mechanisms: peer influence and homophily. While both peer influence and homophily mechanisms can each lead to very similar observational data, the policy implications of each of these mechanisms are significantly different: under peer influence an effective policy might be to identify "influential" people and induce the desired behavior among them, while under homophily mechanism this policy may have no effect and a different type of action is needed. Considering the massiveness and continuing growth of online social networks, it is therefore of critical importance both for research and for businesses to reliably identify presence of each of these mechanisms in the general population of online social networks and quantify the strength of them. Traditionally, the econometric identification of peer influence from purely observational data has proven to be a hard challenge. In this paper, we present a novel randomized experiment that tests the existence of causal peer influence in the general population of a particular large-scale online social network. We present our findings starting with insights learned from observational data, followed by a quasi-experiment based on observational data and concluding with the randomized field trial. Both quasi-experiment and randomized experiment demonstrated that new adoptions were significantly higher in the treatment group vs. control group. The quasi-experiment allowed us to determine the adequate sample size for our randomized trial as well as provided an interesting baseline to compare our results against. Both simple t-test and logistic regression indicate that user's adoption of a product causes her online friends to pay for it and adopt it as well. Our point estimates show that, for a median social network user, the odds of adopting the paid subscription increase by 116% due to peer influence when her friend adopts it. In addition, we find that peer influence is significantly stronger for users with smaller number of friends as compared to the ones with large number of friends. Finally, we find that the quasi-experiment tends to produce the results similar to randomized trial, somewhat over-estimating the effect on users with larger number of friends and under-estimating it for the users with smaller number of friends, thus providing the first insights about the nature of bias in estimating peer-effects by the models with self-selected populations.

Keywords: Peer-effects, randomized experiment, social contagion, matching models, music subscription, online social networks, last.fm

[1] Carlson School of Management, University of Minnesota (authors are listed in alphabetical order)

1. Introduction and Background

The general challenge of demonstrating causal inference from observational data has been immortalized in (Manski 1995) reference to the simultaneous movements of a man and his image in the mirror. He asks, "Does the mirror cause the man's movement or reflect them?" and concludes that without understanding optics and human behavior we cannot really tell. Interestingly, this quote from pre-Facebook era is nothing less than extremely relevant to the causality questions that arise in today's digital age featuring massive[2] online social networks such as Facebook and Twitter as well as niche networks such as Last.fm, Spotify, LinkedIn and others. Nowadays, these online social networks are credited with playing roles that range from inspiring political action to driving viral and word-of-mouth spread of products and services (Aral and Walker 2011, Hill et al 2006, Iyengar et al 2011, Manchanda et al 2008, Mayzlin 2006), and as such, represent a vast reservoir of social and economic influence. Central to the tapping into this reservoir is the understanding of causal relationships that drive the spread of products, services and information over these social networks, the focus of this paper.

It has been consistently demonstrated in the literature that in online social networks user characteristics and behavior tend to cluster both in space and in time with users generally being similar to their online friends and acting similar to their online friends (Aral and Walker 2011). Surprisingly, there are several different underlying causal mechanisms that can lead to this observed clustering with the most frequently cited ones being *peer influence* and *homophily*. Under the mechanism of pure *peer influence,* an individual causes her online friends to undertake a certain action, which in turn, leads to the observed clustering of the behavior. On the other hand, under the mechanism of pure *homophily,* an

[2] Online social networks such as Facebook, with 800 M users, and Twitter, with 100 M users, are increasingly consuming a significant and growing portion of our time and attention. A 2010 Nielsen study[2] estimated that the amount of time the average user spent on Facebook was about seven hours per month, and more importantly, was growing at the rate of 10% per month.

individual tends to be online friends with peers that are similar to her on observed and unobserved characteristics and possibly the environment that they face. In this case, it is not a surprise that behavior of that individual is correlated with the behavior of her friends: they may not influence each other at all, but the observed correlation of their actions comes from their intrinsic similarity that forces them to make similar choices independently over and over again.

The importance of disentangling peer influence and homophily mechanisms stems from the fact that despite leading to very similar observational data, the policy implications of each of these mechanisms are vastly different. Under peer influence an effective policy may be to identify “influential” people and induce the desired behavior among them so that it would propagate through the social contagion, while under homophily mechanism this policy may have little effect (Aral 2010). Instead, under homophily, a careful segmentation based targeting strategy would be preferred. Moreover, the mechanisms of peer influence and homophily are not necessarily mutually exclusive and may complement each other, therefore social contagion processes in real online networks may contain a mixture of peer-influence and homophily.

Importantly, peer effects have the added bonus of bringing with them the *social multiplier effect*. (Manski 1995) provides an intuitive example of that effect describing a potential positive feedback loop of peer influence for the context of academic performance of high school students. (Manski 1995) reports that if an increase in individual student’s academic performance causes the increase in the performance of the reference group of her peers, then this reference group may in turn increase the performance of that individual even further, and so on, leading to a positive self-reinforcing feedback loop with the social multiplier effect. On the other hand, homophily-based mechanisms that arise out of similarity of individual characteristics or contextual information do not typically exhibit this multiplier effect, perhaps explaining the fascination amongst researchers and practitioners about viral marketing of product and services. These factors make it critical, both for theory and practice, to reliably identify presence of each of these mechanisms in the context of large-scale online social networks. It is fair to say that causal identification and measurement of peer influence in the general population of online social networks, or put simply, existence and strength of social contagion, has become the holy grail of modern research in online social networks (Aral 2011).

In this paper, we present a novel randomized experiment that identifies the existence of peer influence in the general population of users of online social networks. Our work is inspired by (Aral and Walker 2011) that demonstrates that significant social contagion can be created by embedding viral features into product design and showcases the potential of using randomized experiments to study peer-effects in online social networks. In our study, we examine the presence and strength of peer influence on premium subscription payments in the general population of Last.fm, a multi-million user music listening social network, with our novel non-intrusive treatment that allows us to create a random assignment of subjects and thus avoids voluntary subject recruitment procedures. Observational data that we collected from Last.fm website clearly indicates that paid subscribers are significantly more likely to be connected to paid subscribers even controlling for the number of friends and other known covariates. However, as explained by (Manski 1995), inferring the presence of peer-influence at play from this is not judicious. More specifically, (Aral 2011) identifies several sources of bias in making such an inference: these include simultaneity (Godes and Mayzlin 2004), unobserved heterogeneity (Van den Bulte and Lilien 2001), homophily (Aral et al 2009), and correlated effects (Manski 1993). While multiple attempts have been made at identifying peer effects using network structure based instrument variables (Bramoullé et al. 2009, Oestreicher-Singer and Sundararajan 2010), natural experiments (Tucker 2008) and matched sample counterfactuals (Aral, Muchnik and Sundararajan 2009, Oh, Susarla and Tan 2011, Oestreicher-Singer and Zalmanson 2011), each method has its limitations (Manski 1993, Aral 2011) and the best we have in the absence of randomized or controlled exogenous variation are upper bounds of peer influence (Aral et al 2009).

Interestingly, (Manski 1995) touches upon the possible reasons behind the lack of randomized trials involving general populations of different real-world networks. He reminds the reader that it is particularly harder to draw inference about general population from a self-selected sample of recruited subjects. Also, even if a self-selection bias is not an issue, generalizable analysis is limited to the observations that are made without undue *intrusion*, since people behavior may change when they know they are being observed.

Our study attempts to close this gap in the literature by introducing a randomized field experiment that eliminates any voluntary subject recruitment procedure, thus mitigating a potential self-selection bias. In addition to that, 1) the manipulation is non-intrusive and subjects are watched silently, thus observer bias is not applicable; 2) the manipulation cannot be escaped and subjects cannot withdraw from the study, thus subject mortality bias is not applicable; 3) peer influence has straightforward monetary measurement in this setting since the observed outcome for each subject is a payment transaction (purchase of subscription) and, unlike adoption of free products, subjects must actually pay real money to adopt the subscription; 4) subjects were selected uniformly randomly from the general population of all network users, thus our results provide the inference about the general population of a social network. Also, our overall experimental design provides insights into the nature and extent of the bias that self-selected samples may inflict when analyzed using quasi-experimental techniques such as propensity score matching (Rosenbaum and Rubin 1983) matching only based on the observables.

We present our findings starting with the insights gained from observational data, followed by a quasi-experiment matching-based analysis and concluding with the randomized trial. As mentioned later in this paper, our quasi-experiment is made feasible due to the sophistication of our data-crawler that is able to obtain dynamic snapshots of the Last.fm social network. The dynamic nature of the data allows us to simulate the "treatment" – albeit in a confounded (peer influence plus homophily) way – by tracking the natural adoption of the subscription by users and comparing it to a control group. Given the rarity of paid subscriptions (we were looking for the proverbial needle in a haystack), the simulated experiment was instrumental in helping us calibrate the sample size of the actual experiment, as well as in informing us about the nature and extent of bias in non-experimental data. Both, the quasi-experiment and randomized trial demonstrated that new adoptions were significantly higher in the treatment group vs. control group. Moreover, our logistic regression estimates indicate that, for the median social network user, the odds of adopting the paid subscription increase by 116% due to peer influence when her friend adopts it, indicating significant causal peer-effects in the monetization of music listening social networks. In addition, we find that the strength of peer influence can be significantly weakened by the size of the influenced user's friendship circle. Finally, we find that the quasi-experiment tends to produce results similar to randomized trial, somewhat over-estimating the effect on users with larger number of friends and under-estimating it for the users with smaller number of friends, providing the first insights about the nature of bias in estimating peer-effects by the models with self-selected populations.

The remaining sections are structured as follows. Section 2 describes the institutional details of our experimental context. Section 3 formally poses the research question. Section 4 describes the design of our experiment. In Section 5, we describe the data collection process, review the data, and provide summary statistics. Section 6 presents our analysis and results: first for the quasi-experiment based only on observational data, then the results of the randomized experiment. Section 7 draws the conclusion of our results and outlines prospects for the future work.

2. Institutional Details

The music industry today serves as a canonical example of how a long-established, growing and profitable industry can be disrupted and subsequently re-invented by the social machinery of Internet.

One of the important emerging models of today's music consumption in the Internet is a *freemium* social community (Anderson 2009), as exemplified by sites such as Last.fm, Pandora, Spotify and many others. *Freemium* social communities typically operate based on a two-tiered business model that offers free access to the basic set of features and content while charging a fee for more advanced premium features. For example, free subscribers of Last.fm[3] website can listen to the online radio interrupted by commercials, while paid subscribers of Last.fm website enjoy 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 etc.

Freemium communities often employ numerous social computing features (Parameswaran and Whinston 2007), such as, for example, *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 among themselves and exhibit certain *peer influence* on each other. On Last.fm website, for instance, online friends can affect each other's music choices while sharing their own music listening experiences, they can listen to friend's "recommended radio", can review friend's "Loved songs" and so on. Appendix A provides a snapshot of a typical Last.fm user's page. More specifically, (Oestreicher-Singer and Zalmanson 2011) provide a nice overview of the institutional details of Last.fm website as freemium social community. Among the findings of their study is the fact that the music listening on Last.fm is socially driven which means it is based on what your friends are listening, and that a paid subscription appears as a distinct (ostensibly status) symbol visible to your friends. Also, as discussed in these studies of Freemium communities (Oestreicher-Singer and Zalmanson 2011, Pauwels and Weiss 2008), a singular challenge for their long-term economic viability is

[3] http://virtualmusic.tv/2011/02/2010-music-website-heat-map/ indicates that Last.fm, with reportedly 30 million subscribers, received 9.8 million hits per month in 2010.

discerning pathways and strategies for moving users *from-free-to-fee,* that is converting users from the large pool of free users to the elite set of premium paid users.

In this paper, we present a randomized field experiment on Last.fm website providing the evidence that making one person a paid subscriber on Last.fm can cause her online friends to pay for subscription and become subscribers as well. Our experimental design relies on the unique social feature of Last.fm that allows gifting any random user in Last.fm social network with a paid subscription. While this feature of Last.fm website has not yet been studied extensively in the social networks literature, it offers a unique opportunity to create a "gold standard" randomized trial on an online social network. From an experimental design perspective *anyone in Last.fm social network has an equal chance of receiving a gift from us. Last.fm users cannot decline the gift or hide their subscription status from others. They cannot transfer the gift to anyone else, or postpone using it, or share it with someone else, or refund it.* This makes the gifting social feature particularly valuable in an experimental context, a fact this research is the first to bring forth.

3. Research Questions

The main research questions of this study are formulated as the following hypotheses:

<u>Hypothesis 1:</u> In an online social network there exists peer influence such that an individual's adoption *causes* the adoption by her online friends.

<u>Hypothesis 2:</u> The effect of peer influence is moderated by and is decreasing in the number of friends the influenced individual has.

While the first hypothesis is the focal point of this paper and its rationale has been articulated at length already, it is worth dwelling a bit on the basis for the second hypotheses. (Iyengar et al 2011) make a compelling case for looking at moderating factors that may shape the nature and extent of social contagion at work. While it could be argued that, for instance, heavy users are more likely to exert a greater influence on others, (Godes and Mayzlin 2009) note that heavy users may tend to be connected mostly to people already predisposed to be early adopters. While the focus of that prior research is on the influencer side of the equation, such as whether better connected adopters exert more influence than do less connected ones, we position ourselves on the susceptibility to influence side of that equation, since it is natural that impact of peer influence will also depend on the susceptibility of the individual being influenced. A user who has many thousands of friends on Last.fm may be much less susceptible to the influence of the marginal peer's adoption decision, as opposed to those social network users who are more selective in befriending others. Similar distinctions between selective and non-selective tie forming behaviors in the context of trust have been observed in other online social networks such as Facebook (Bapna et al 2011).

In order to address our research questions we first need to establish a causal link between person's B decision to subscribe due to the influence from B's friend - person A. In this paper, our conceptualization of *peer influence* is due to (Aral 2011). This conceptualization is rooted in utility theory in that the actions of one's peers change the utility one expects to receive from engaging in a certain behavior and thus the likelihood that one will engage in that behavior (Aral 2011). Such a conceptualization is flexible and encompassing with respect to the myriad influence mechanisms that could lead to social contagion. In other words, in order to demonstrate the presence of *peer influence* we do not seek to explain which influence mechanism from person A *causes* person B to subscribe: be it awareness raising, explicit or tacit persuasion, observational or social learning, imitation or any other mechanism. It is only required to demonstrate that person A causes person B to subscribe. It is important to note that in this study, we do not raise the question of disentangling the general peer influence into the exact types of peer influence mechanisms as above. This disentanglement would require collecting very different type of data and possibly a different experimental design.

For illustrative purposes, we present the following intuitive analogy before we describe the actual experimental setup using strict formalism. Assume that the paid subscription is like a disease caused by virus, albeit a benevolent one. We call this the U1B1-B (for Umyarov-1-Bapna-1 Benevolent[4]) virus. Our data shows that people sick with this virus tend to be friends with other sick people, but this alone is not evidence that the "disease" is contagious. This clustering could easily be explained by the fact that people tend to befriend people who are of similar "age" and in a similar "health" condition and therefore belong to the same risk group and are equally likely to catch the U1B1-B virus from the environment (rather than from a peer), causing the observed clustering. Therefore, the question of our experiment would be: is the U1B1-B "subscription disease" contagious or is it just caught from the "environment" by people in "poor health"? For the experiment, we will select the *manipulated group M* of 1000 Last.fm users who will be randomly chosen to receive the subscription gifts, which is akin to getting randomly infected by the U1B1-B virus, over which they have no control, ruling out any self-selection, and individual characteristics or contextual (observed or unobserved) homophily-based decisions that confound the analysis of observational data. We will also select the *not-manipulated group NM* of 1000 random Last.fm users who do not get "infected" by us. After a period of time, we compare the occurrence of the "disease" among the friends of M and friends of NM groups. Given the initial uniform randomization of groups M and NM, both observed and unobserved statistical properties of M and NM are expected to be statistically identical before the manipulation. Therefore, if any statistical difference is observed in the outcomes among friends of M and friends of NM groups, this difference should be considered as being caused by our manipulation.

Our works relates to and builds upon the propensity score (Rosenbaum and Rubin 1983) matching based approaches of (Aral, Muchnik and Sundararajan 2009), (Oh, Susarla and Tan 2011) as well as (Oestreicher-Singer and Zalmanson 2011). A key advancement of our work is that while propensity score matching accounts for observable user characteristics in crafting usable control groups, it is widely recognized (Aral, Muchnik and Sundararajan 2009, Oestreicher-Singer and Zalmanson 2011) that other

[4] We hope to use this methodology for other settings in the future, hence the numeric indexing.

unobservable user characteristics (say amount of free-time an individual has, income level, sensitivity to commercials etc) or contextual affects such as marketing (van den Bulte and Stremersch 2004) could as well be influencing the propensity to be treated and be linked to homophily. This limitation is overcome in our study through randomization such that there is no reason to believe the treatment group and the control group (described in the next section) should have any systematic difference in observable and latent/unobservable characteristics. In the absence of randomization, the best we can get are upper bounds of the true estimate of contagion (Aral and Walker 2011).

4. Experimental Design

4.1 Methodology

In this section, we describe our experimental design beginning with our subject selection procedure. At the onset of the study we developed a custom multi-threaded, cloud-based, web crawler that was used to identify the largest closed connected component of the Last.fm social network which consists of roughly 3.8 million users. Because there are a considerable number of inactive accounts in the network, we decided to direct our attention only to the active users for receiving subscription gifts, where a user is considered *active* if she listened to a song in the last 30 days before the manipulation. It turned out there are roughly 1.26 million active listeners in the Last.fm connected component. Let's call this list L. We form the group G as a random sample of 2000 users drawn uniformly randomly from L with no replacement. Group G contains the 2000 users who will then be randomly split into manipulated and non-manipulated groups. Consequently, we form the manipulated group M as a random sample of 1000 users drawn uniformly randomly from G with no replacement. Finally, we form the non-manipulated group NM as NM = G \ M, that is the rest 1000 users that were left in G after we picked M group.

We define our *treatment group* T as all immediate friends of M who are not themselves in M and who are not friends of someone in NM. Symmetrically, we define our *control group* C as all immediate friends of NM who are not themselves in NM and are not friends of someone in M.

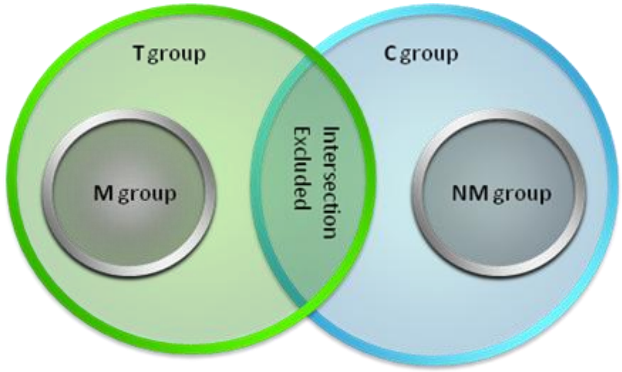


Figure 1 – Venn Diagram Showing Treatment, Control, Manipulated and Non-Manipulated Groups

Figure 1 presents an intuitive Venn diagram for these sets of users. Note that since group T and group C contain all friends, they may contain both active and inactive users.

Our experiment procedure is composed of four stages. In the first stage we randomly assign users to groups M and NM, crawl their current friend network, and thus calculate groups T and C. In the second stage, we deploy a pre-treatment check and crawl the current status of M, NM, T and C groups immediately before the treatment. In the third stage we deploy the U1B1-B virus by giving 1000 gifts to group M using our Paypal account bot. Finally, we crawl the current status of M and NM groups immediately after manipulation to make sure our manipulation worked. Given (Manski 1995) concern about subjects' behavior changing when they know they are being observed, we directed users to our Last.fm page (see Appendix B) where took great care[5] in "explaining" to the users that these were left-over funds from another project that we were simply giving away, and expected nothing in return. The messaging worked, as can be gleaned by their comments left on our wall.

4.2 Strengths of the Experiment in Mitigating Threats to Validity

Our design has several intuitive benefits that help us overcome the myriad challenges (van den Bulte and Stremersch 2004, Aral 2011) in making causal detection of social contagion from observational data, separating out homophily from peer influence. As mentioned above, one of the ways in which

[5] Needless to say our protocol was approved by our IRB.

homophily manifests itself in observational data is through self-selection bias, when treatments are not randomly assigned, which is not the case in our study. Also, in contrast to other experimental studies involving voluntary subjects, we have no attrition or mortality bias. This is because users are selected randomly and they cannot escape, decline or withdraw from the manipulation. It is also important to mention that each person's network will be collected immediately before the manipulation, immediately after the manipulation and with different levels of delay after the manipulation. Only "immediately before the manipulation" friend network is used to determine treatment group T and control group C. Clearly, if a person starts self-selecting subscriber friends after the manipulation has occurred, it will not have any effect on the experiment. Further, because the subscriptions themselves are not transferrable and not refundable, we can rule out any direct treatment diffusion effect, suggesting that any effect that is observed must be through some peer influence other than just direct transfer of the subscription. It is however possible given the social network setting that the control group C may be influenced by manipulated group M through 2nd degree friendship connections, i.e. there may be the possibility of an indirect treatment diffusion effect. This however would likely lead to an underestimation of the observed difference, not overestimation. The effect can also be mitigated by post-experimental controls on shortest distances between the control group C and the treatment group T. Finally we can rule out any compensatory rivalry/resentful demoralization or experimenter bias, since neither treatment group T nor control group C know that they are being treated and watched. Only manipulated group M receives a gift from us. However, manipulated group M is told that the gift is given out of the expiring left-over funds from a prior survey and that gift receiver is not required to do anything, thus M group itself is also not aware of being manipulated.

5. Data Description

5.1 Snapshot Data

Our dataset consists of panel data on 3,880,144 users that make up the largest connected component of Last.fm network forming over 23 million friendship pairs. These users have been tracked consistently as a panel since May 2011 with updates roughly every 2 weeks. These dynamic updates provided us with fresh snapshots of the entire social network containing the renewed list of friends and subscription status for every user. In addition to this information, we have been tracking 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:

- $Age_{i,t}$. Self-reported age of user i. Age distribution was truncated to the interval between 8 and 79 in order to eliminate outlier data points that are likely fake.
- $Gender_{i,t}$. Self-reported gender of user i. Dummy variable.
- $FriendCnt_{i,t}$. Total count of number of friends of user i.
- $SubscriberFriendCnt_{i,t}$. Total count of number of subscriber friends of user i.
- $SongsListened_{i,t}$. Total count of all songs ever listened and reported to Last.fm by user i.
- $Playlists_{i,t}$. Total count of playlists ever made by user i on Last.fm.
- $Posts_{i,t}$. Total count of forum posts ever made by user i.
- $Shouts_{i,t}$. Total count of shouts[6] ever received user i.
- $LovedTracks_{i,t}$. Total count of all tracks that were "loved" by user i.
- $RegDate_i$. User i original registration date on the website measured as number of days since January 1, 1960 (standard date representation of SAS statistical package).

[6] Shout is a Last.fm slang for a wall-post on the user's "wall".

- *LastfmCountry$_{i,t}$*. Dummy 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 are slightly different[7] in the official Last.fm countries ("USA", "Germany", "UK") versus the rest of the world.
- *Subscriber$_{i,t}$*. Dummy variable indicating whether user *i* is currently a premium subscriber.

The descriptive summary statistics for 1,251,464 active[8] Last.fm users are displayed in Table 1 below. This table provides a breakdown of statistics for active subscribers and active non-subscribers for one particular snapshot of data collected around September 8, 2011 before our manipulation.

From this data, we find that subscribers are consistently different from non-subscribers in a variety of metrics: they are older, tend to have more friends and disproportionally more subscriber-friends, more playlists, loved tracks and registered earlier than non-subscribers. These observations confirm the observed clustering of subscription behavior indicating the underlying homophily or peer influence. Our summary data are remarkably in line with 2009 Last.fm data reported by (Oestreicher-Singer and Zalmanson 2011).

The histogram[9] on Figure 2 demonstrates the distribution of the number of friends for all users in Last.fm social network. As we can see from the histogram more than 60% of the users have less than 20 friends. This observation will prove important when we talk about the increased strength of marginal effects of peer influence on users who have small number of friends.

---

[7] Even though the premium subscription costs the same amount for every country, the subscription is more valuable for people outside USA, Germany and UK. Several Last.FM services that are normally free for US/Germany/UK users require premium subscription for the rest of the world because of music licensing contracts.

[8] Active user means a user who listened to at least 1 song within 30 days prior to the collection of that particular snapshot of data.

[9] Note that the histogram is "censored" at 200, because the non-informative "long-tail" goes up to tens of thousands. In this histogram, the last bin should be interpreted as "everyone who has 200 friends or more"

subscriber	N Obs	Label	Mean	Std Dev	Missing	Median	Minimum	Maximum
0	1214303	Age	23.21	6.18	385200	22.00	8.00	79.00
		Gender (Male=1)	0.66	0.48	234278	1.00	0.00	1.00
		FriendCnt	24.18	70.65	0	10.00	1.00	11780.00
		SubscriberFriendCnt	0.65	2.85	0	0.00	0.00	541.00
		SongsListened	24913.30	32365.72	1	15022.00	0.00	1000472.00
		Playlists	0.53	3.32	0	0.00	0.00	2291.00
		Posts	7.67	141.70	0	0.00	0.00	64108.00
		Shouts	42.19	271.02	27717	5.00	0.00	131765.00
		LovedTracks	128.15	406.44	0	35.00	0.00	99109.00
		RegDate	17838.23	636.71	584	17902.00	15642.00	18877.00
		LastfmCountry	0.30	0.46	0	0.00	0.00	1.00
1	37161	Age	30.26	9.25	14165	28.00	8.00	78.00
		Gender (Male=1)	0.76	0.43	8449	1.00	0.00	1.00
		FriendCnt	33.73	116.62	0	10.00	1.00	9788.00
		SubscriberFriendCnt	2.85	10.35	0	1.00	0.00	709.00
		SongsListened	31996.64	43938.95	0	18139.00	0.00	1000070.00
		Playlists	1.44	5.38	0	1.00	0.00	496.00
		Posts	27.74	465.16	0	0.00	0.00	50740.00
		Shouts	85.31	531.56	1275	5.00	0.00	36508.00
		LovedTracks	370.05	1104.95	0	149.00	0.00	63595.00
		RegDate	17678.54	628.82	1	17735.00	15642.00	18868.00
		LastfmCountry	0.28	0.45	0	0.00	0.00	1.00

Table 1 – Summary Statistics of Historical Data

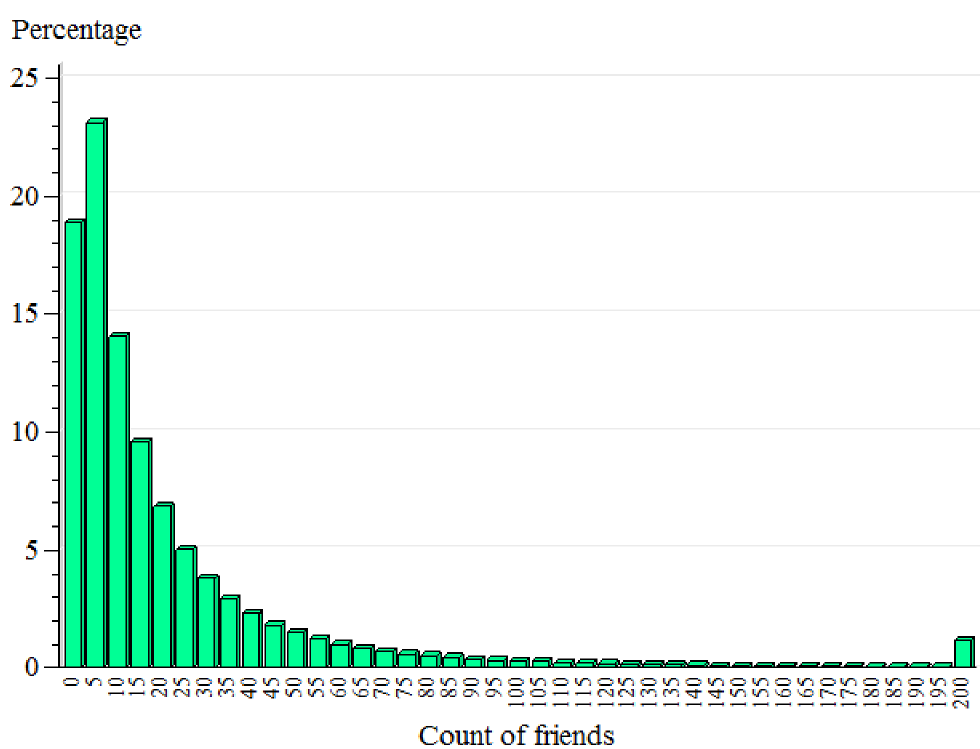


Figure 2 – Distribution of Number of Friends of Last.fm Users

5.2 Dynamic Data

The collection of snapshots allows us to look into the dynamics of user characteristics in the social network as well as the dynamics of the social network itself. The following network dynamic variable is the variable of interest in this particular study:

- $Adopter_{i,[t,t+1]}$. Dummy variable indicating whether user *i* who had not been a paid subscriber at time *t* adopted subscription and became a paid subscriber at time *t+1*

Since the minimum possible unit of a premium subscription available is 1 month and we collected our data with the intervals of at most 2-3 weeks, our data collection process has not missed any single subscription event for any user in the network beginning from May 2011 and till the moment this paper is being read by our readers. Therefore, $Adopter_{i,[t,t+1]}$ variable is a guaranteed indicator of adoption or non-adoption in time period [t,t+1]. The table below displays the summary statistics for *the dynamic data* of adopters vs. non-adopters. Please note that there is a subtle, but very important difference between the types of information displayed by Table 1 and Table 2. Table 1 displays the information on "mature subscribers" - people who may have subscribed long time ago and stayed with the subscription until now, while Table 2 displays the characteristics of current new adopters. Nevertheless, we observe that there is a similarity between Tables 1 and 2 suggesting a remarkable consistency in the data generation process over the years: new adopters tend to resemble the large mass of existing subscribers based on the observed characteristics.

We used this dynamic data to simulate and calibrate our experiment before we actually ran it. In particular, because paid subscriptions are a rare event in our network, a key experimental challenge for us was to decide on the sample size for the manipulation so as to be able to pick up statistically any peer effect that may be there. More details on this in Section 6.2.

Adopter	N Obs	Label	Mean	Std Dev	Missing	Median	Minimum	Maximum
0	1211366	Age	23.20	6.18	384294	22.00	8.00	79.00
		Gender (Male=1)	0.66	0.48	233726	1.00	0.00	1.00
		FriendCnt	24.16	70.43	0	10.00	1.00	11780.00
		SubscriberFriendCnt	0.65	2.80	0	0.00	0.00	465.00
		SongsListened	24912.04	32363.37	1	15024.00	0.00	1000472.00
		Playlists	0.53	3.32	0	0.00	0.00	2291.00
		Posts	7.67	141.83	0	0.00	0.00	64108.00
		Shouts	42.14	271.01	27602	5.00	0.00	131765.00
		LovedTracks	127.97	406.32	0	35.00	0.00	99109.00
		RegDate	17838.13	636.65	584	17902.00	15642.00	18877.00
		LastfmCountry	0.30	0.46	0	0.00	0.00	1.00
1	1099	Age	26.31	7.13	346	25.00	11.00	74.00
		Gender (Male=1)	0.70	0.46	204	1.00	0.00	1.00
		FriendCnt	42.70	196.79	0	14.00	1.00	4730.00
		SubscriberFriendCnt	2.76	17.58	0	1.00	0.00	541.00
		SongsListened	31984.12	38619.43	0	18991.00	0.00	423529.00
		Playlists	1.05	1.98	0	1.00	0.00	27.00
		Posts	13.08	96.25	0	0.00	0.00	2266.00
		Shouts	93.17	381.14	43	7.00	0.00	6247.00
		LovedTracks	310.65	542.01	0	133.00	0.00	6143.00
		RegDate	17712.48	651.39	0	17734.00	15642.00	18877.00
		LastfmCountry	0.24	0.43	0	0.00	0.00	1.00

Table 2 – Summary Statistics of Dynamic (2-3 weeks) Data

6. Analysis and Results

6.1 Model Specification

Since our outcome variable *Adopter*$_{i,[t,t+1]}$ is a binary choice variable we decided to use logistic regression as the apparatus to control for the observed covariates and determine causality in our scenario. The formula below depicts our logistic regression model, treatment variable and controls:

$$\begin{aligned}
&\text{logit}\ \{\Pr(\text{Adopter}=1)\} = \alpha + \beta_1 \cdot \text{OurTreatment} + \beta_2 \cdot \text{OtherTreatment} + \\
&+\beta_3 \log(\text{FriendCnt}) + \beta_4 \log(\text{FriendCnt}) \cdot \text{OurTreatment} + \beta_5 \log(\text{FriendCnt}) \cdot \text{OtherTreatment} + \\
&+\beta_6 \log(\text{SubscriberFriendCnt}) + \beta_7 \cdot \text{Age} + \beta_8 \cdot \text{AgeMissing} + \\
&+\beta_9 \cdot \text{LastfmCountry} + \beta_{10} \cdot \text{CountryMissing} + \beta_{11} \cdot \text{RegDate} + \beta_{12} \log(\text{SongsListened}) + \\
&+\beta_{13} \log(\text{Posts}) + \beta_{14} \log(\text{Playlists}) + \beta_{15} \log(\text{Shouts}) + \beta_{16} \log(\text{LovedTracks})
\end{aligned}$$

The following variables are used as manipulation variables in this particular study:

- *OurTreatment*$_{i,t}$. This manipulation variable represents the count of how many friends of user *i* were manipulated by us at time *t.* Since this variable was manipulated by us independently of the user group assignment, its coefficient represents *pure* peer influence.
- *OtherTreatment*$_{i,[t-1,t]}$. This variable represents the count of how many friends of user *i* adopted the subscription on their own in the time interval *[t-1,t]* independently from our manipulation. While technically not being a treatment, this variable controls for other "treatment" that the influenced user *i* receives from the network besides ours. Unlike *OurTreatment,* the coefficient of *OtherTreatment* represents both peer influence and homophily combined: this "other" friend who adopted subscription on her own may exhibit peer influence on user *i* or she may serve as an indicator that user *i* belongs to a "risk group of likely adopters" or both.

All other controls variables were explained in the previous section.

6.2 Quasi-Experiment Description

Before conducting the actual randomized experiment, we first constructed a quasi-experiment that simulates our randomized experiment using only observational data. The quasi-experiment study was conducted in order to determine the appropriate sample size, check if the effect can be observed in observational-only data as well as to compare the ultimate result of the quasi-experiment against the future randomized experiment. While we acknowledge that the influence of unobserved characteristics cannot be ruled out by the quasi-experiment, we could still control for the observed characteristics of users as a "first-order approximation." Methodologically, this approach is similar to the matching based quasi-causal techniques seen in (Aral, Muchnik and Sundararajan 2009, Oh, Susarla and Tan 2011) as well as (Oestreicher-Singer and Zalmanson 2011).

In order to introduce the design of the quasi-experiment, consider 3 consecutive times in the evolution of our data: t-1, t and t+1 each separated by at least 2 weeks. If we look into our data across at least 2 week period [t-1,t], we will typically see that thousands of users suddenly became subscribers in that time period [t-1,t]. We will refer to them as "0->1" users. Let us randomly select 1000 of these 0->1 users into a group "M" [10]. It is also very typical that in the same time period [t-1,t], we will likely see more than 1mln active users who remained non-subscribers. We will refer to them as "0->0" users. For every user in group "M", we would like to find her alter-ego, that is a person who has certain properties identical to the user but happened to remain a "0->0" in the same time frame [t-1,t]. We match every 0->1 user from group "M" with a random 0->0 alter-ago based on the exact matching of the observed count of friends and subscriber friends[11] and thus form a group "NM" of 1000 alter-egos. Figure 3 below depicts the nature of the quasi-experiment, with "manipulated" users actually representing natural adopters.

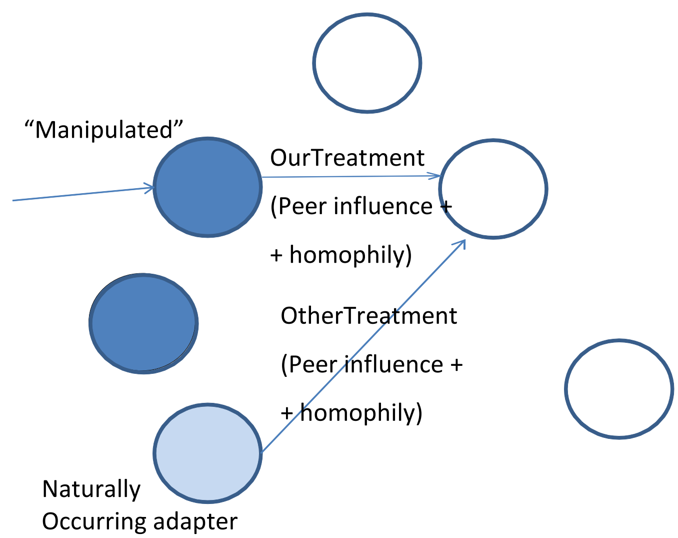


Figure 3 – Dynamic Data and Quasi-Experiment Used to Calibrate Experiment

[10] We denote quasi- groups with quotes: like group "M" as opposed to group M so as to clearly separate places where we talk about the quasi-experiment from places where we talk about the randomized experiment

[11] This is just one particular type of matching, and a different type of matching (like propensity score matching) could have definitely been done here. We chose our particular matching in order to make sure that the social network characteristics of "M" and "NM" groups are as similar as possible since this is the subject of our study. Fortunately, the sheer size of our dataset allowed us a plenty of precise matches for every 0->1 user.

Similarly to our experimental setup, we define our *quasi-treatment group* "T" as all immediate friends of "M" who are not themselves in "M" and who are not friends of someone in "NM". Symmetrically, we define our *quasi-control group "*C" as all immediate friends of "NM" who are not themselves in "NM" and are not friends of someone in "M". Clearly, because of the matching, groups "M" and "NM" are identical in terms of the matched characteristics at time t-1. By comparing the subscription changes in groups "T" and "C" during the subsequent time period [t,t+1] and controlling for all known observed characteristics of each user, we are able to tell whether being a friend of "M" has any effect on the subscription behavior as compared to being a friend of "NM".

6.3 Results of Quasi-Experiment

Our logistic regression analysis of adoption among friends of groups "M" and "NM" is presented in Table 3 below**.** This analysis demonstrates that *OurTreatment* variable is statistically significant in explaining the decision to subscribe after controlling for variety of demographic, network and social activity variables. Similar results occur if we select different time periods, different seeds for random samples and with other robustness checks.

Variable	**Estimate**	**Std Err**	**Wald χ^2**	**Pr > χ^2**
Intercept	-10.2650	2.3595	18.9265	<.0001
OurTreatment	**1.2182**	**0.3459**	**12.4058**	**0.0004**
OtherTreatment	**1.1634**	**0.4200**	**7.6737**	**0.0056**
log(FriendCnt)	-0.4467	0.1075	17.2750	<.0001
OurTreatment * log(FriendCnt)	**-0.1778**	**0.0758**	**5.5041**	**0.0190**
OtherTreatment * log(FriendCnt)	**-0.2055**	**0.0804**	**6.5291**	**0.0106**
log(SubscriberFriendCnt)	**0.8729**	**0.1015**	**73.9233**	**<.0001**
Age	0.00726	0.00895	0.6566	0.4178
AgeMissing	0.0376	0.2840	0.0176	0.8945
LastfmCountry	-0.6470	0.1482	19.0610	<.0001
CountryMissing	0.3154	0.1779	3.1449	0.0762
RegDate	0.000119	0.000119	1.0014	0.3170

Variable	Estimate	Std Err	Wald χ^2	Pr > χ^2
log(SongsListened)	0.2129	0.0515	17.0557	<.0001
log(Posts)	0.0199	0.0407	0.2387	0.6252
log(Playlists)	0.0510	0.0993	0.2633	0.6078
log(Shouts)	0.0715	0.0504	2.0121	0.1561
log(LovedTracks)	0.2382	0.0370	41.3896	<.0001

Table 3 – Quasi-experiment Results Based on Matching

Note also that all other "peer influence"-related variables like *OtherTreatment, SubscriberFriendCnt* as well as their interactions are also statistically significant. The results provide a strong suggestion for the possibility to observe the desired causal effect in the randomized experiment. However, as we have already mentioned, the quasi-experiment only takes observed variables into account. The randomized experiment explained in the next section would be able to account for all the unobserved variables in this study.

6.4 Randomized Experiment Description

Given the results of the quasi-experiment mentioned in the previous section, we determined that the sample size of 1000 is adequate to determine the causal peer influence effect if it exists. Therefore, we sampled a group M of 1000 users uniformly randomly from the full lists of all active users who were not subscribers at the moment. Each person in this group M subsequently received a 1 month subscription gift from us, with the gifts being distributed over the period of several hours by a GreaseMonkey script. From the list of the remaining active users we sampled a group NM of 1000 users uniformly randomly. The users from group NM did not receive any gift or any other communication from us and were only being silently tracked. Figure 4 below can be contrasted with Figure 3 above. Table 4 below presents the descriptive summary statistics of groups M and NM that were achieved after the randomization.

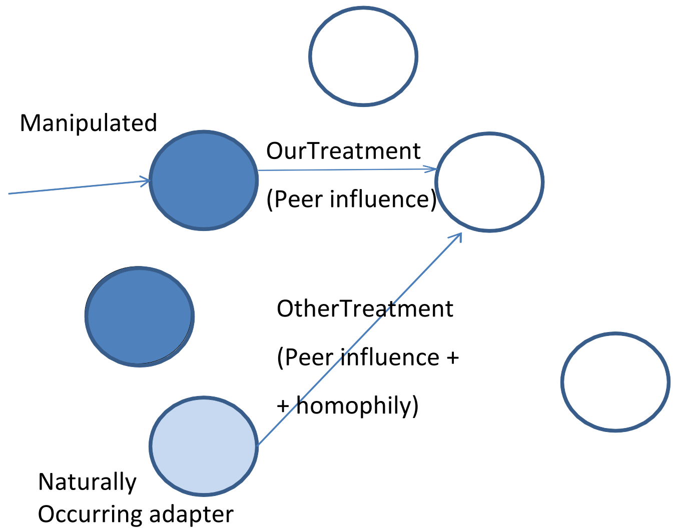


Figure 4 – Randomized Experiment Design;
Naturally Occurring "OtherTreatment" Offers an Interesting Contrast to OurTreatment

As the corresponding t-tests demonstrate, our randomization technique worked correctly and the group means did not differ statistically on any observed attribute prior to manipulation. Also, a manipulation check was done immediately after distributing gifts and web pages of all 1000 users demonstrated that they received a gift and became premium subscribers immediately.

Variable	Group	Not Missing	Mean	Std Err	Minimum	Maximum	t-test Method	t Value	Pr > |t|
Age	NM	697	23.3529	0.2425	11.00	69.00	Pooled	0.16	0.8698
	M	709	23.2990	0.2225	14.00	66.00	Satterthwaite	0.16	0.8699
Gender (Male=1)	NM	819	0.6606	0.0166	0	1.00	Pooled	0.98	0.3256
	M	819	0.6374	0.0168	0	1.00	Satterthwaite	0.98	0.3256
FriendCnt	NM	1000	26.4890	1.9297	1.00	918.0	Pooled	-0.24	0.8089
	M	1000	27.3050	2.7683	1.00	2248.0	Satterthwaite	-0.24	0.8090
SubscriberFriendCnt	NM	1000	0.7860	0.1169	0	100.0	Pooled	0.65	0.5138
	M	1000	0.6930	0.0813	0	40.00	Satterthwaite	0.65	0.5138
SongsListened	NM	1000	28260.9	1286.6	35.00	536568	Pooled	0.93	0.3539
	M	1000	26723.8	1045.5	32.00	365165	Satterthwaite	0.93	0.3539
Playlists	NM	1000	0.5560	0.0275	0	11.00	Pooled	1.31	0.1901
	M	1000	0.4990	0.0337	0	22.00	Satterthwaite	1.31	0.1901
Posts	NM	1000	6.7290	1.4856	0	846.0	Pooled	-0.10	0.9210

Variable	Group	Not Missing	Mean	Std Err	Minimum	Maximum	t-test Method	t Value	Pr > |t|
	M	1000	6.9410	1.5355	0	696.0	Satterthwaite	-0.10	0.9210
Shouts	NM	970	39.1918	4.4609	0	2530.0	Pooled	0.05	0.9595
	M	975	38.9036	3.5029	0	1528.0	Satterthwaite	0.05	0.9595
LovedTracks	NM	1000	144.5	13.8871	0	8214.0	Pooled	0.11	0.9090
	M	1000	142.5	11.5661	0	6396.0	Satterthwaite	0.11	0.9090
RegDate	NM	1000	17763.2	19.3645	15778.0	18753.0	Pooled	0.64	0.5221
	M	1000	17745.6	19.3291	15815.0	18760.0	Satterthwaite	0.64	0.5221

Table 4 – Groups M and NM have Similar Statistical Properties

In one month after the manipulation was done, we collected new snapshot of the social network and compared adoption behavior among all friends of group M versus all friends of group NM as described in the experiment design.

The interesting aspect here is that even while our treatment is in progress, users in groups T and C still have naturally occurring adopter friends from the other parts of the network. These “natural treatments” are captured by *OtherTreatment* control variable. Unlike *OurTreatment* however, this variable contains both peer influence and homophily signal giving us the option to directly contrast the estimated effect of *OurTreatment* vs. *OtherTreatment*. We believe this is a unique feature of our design that gives us an insight on the extent of the homophily strength that has been so hard to quantify earlier.

6.5 Randomized Experiment Results

Before presenting the results of our logistic regression model, we would like to start with the results of a simple t-test on the means of distributions of adopters in groups T and C:

Treatment	N	Mean	Std Dev	Std Err	Method	Variances	t Value	Pr > |t|
0	21284	0.00197	0.0444	0.000304	**Pooled**	Equal	-2.06	0.0394
1	21981	0.00296	0.0543	0.000366	**Satterthwaite**	Unequal	-2.07	0.0388

As demonstrated by the results of the t-test, treatment group experienced significantly more adopters as compared to the control group. It is important to note that given randomization of our treatment, even a simple t-test can serve as a statistical evidence for supporting our hypothesis ignoring all the controls.

Given that many control variables are useful in explaining the adoption decision, we are interested in utilizing them in a more efficient model that takes them into account. For that, we use the same logistic regression model that was presented in Section 6.1 and Table 5 below demonstrates the results of the estimation on the data gathered from the randomized experiment:

Label	Estimate	Std Err	Wald χ^2	Pr > χ^2
Intercept: adopter=0	-4.8567	3.3430	2.1106	0.1463
OurTreatment	**1.3570**	**0.5860**	**5.3629**	**0.0206**
OtherTreatment	**1.5822**	**0.4787**	**10.9245**	**0.0009**
log(FriendCnt)	0.0167	0.1529	0.0119	0.9130
OurTreatment * log(FriendCnt)	**-0.2547**	**0.1235**	**4.2549**	**0.0391**
OtherTreatment * log(FriendCnt)	**-0.2631**	**0.0830**	**10.0523**	**0.0015**
log(SubscriberFriendCnt)	**0.4779**	**0.1514**	**9.9686**	**0.0016**
Age	0.0322	0.0119	7.3139	0.0068
AgeMissing	0.6690	0.3909	2.9284	0.0870
LastfmCountry	-0.5555	0.2199	6.3780	0.0116
CountryMissing	0.1688	0.3044	0.3076	0.5791
RegDate	-0.00023	0.000172	1.8028	0.1794
log(SongsListened)	0.0464	0.0696	0.4442	0.5051
log(Posts)	0.0991	0.0560	3.1276	0.0770
log(Playlists)	0.3543	0.1415	6.2654	0.0123
log(Shouts)	0.0259	0.0684	0.1432	0.7051
log(LovedTracks)	0.2119	0.0548	14.9566	0.0001

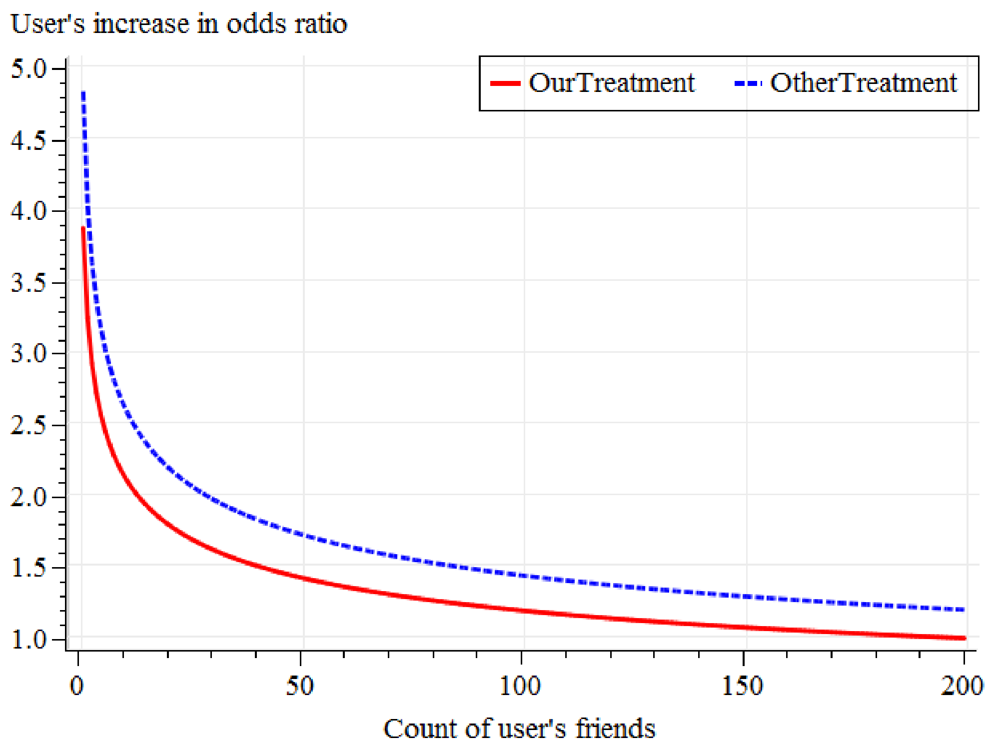
Table 5 – Results of Randomized Trial Indicate Significant Peer Effects which are Moderated by Influencees' Susceptibility

As evident from these results, *OurTreatment* variable is statistically significant. Moreover, since *OurTreatment* was assigned independently of whether an individual is a friend of group M or group NM, this coefficient has causal interpretation: *OurTreatment causes* the adoption of subscription, thus providing a positive answer to Hypothesis 1 of our research question. It is important to note that the estimated coefficients of *OtherTreatment* and log(*SubscriberFriendCnt*) are also statistically significant and positively associated with the likelihood of adoption of subscription: the effect that Hypothesis 1 would imply.

In addition to that our results demonstrate that the interaction of *OurTreatment* and log(*FriendCnt*) is also statistically significant with the negative sign implying that people with large number of friends are less susceptible to *OurTreatment* other things being equal. Once again, given the independence of the assignment of the treatment, the result has causal interpretation and confirms Hypothesis 2 of our research question. Interestingly, the interaction term of observed *OtherTreatment* and log(*FriendCnt*) is also significant with negative sign: the effect that Hypothesis 2 would imply.

Since both *OurTreatment* and *OtherTreatment* enter the model with interactions terms, the marginal effect of either of these variables is not constant and depends on the exact characteristics of the influenced user. In order to provide a point estimate of the marginal effect of each of these variables, we take a median Last.fm user from the social network and apply *OurTreatment* to that user varying only her number of friends and holding all other variables constant (assuming also *OtherTreatment* = 0). We repeat the same procedure for *OtherTreatment* by varying only the median user's number of friends and holding all other variables constant (assuming also *OurTreatment* = 0).

The two resulting curves are displayed in the picture below with the Red line for the marginal effect of *OurTreatment* depending on the number of friends and the Blue line for the marginal effect of *OtherTreatment* respectively: the horizontal axis represents the number of friends the influenced user has, while the vertical axis represents the increase in the user's odds ratio to adopt the subscription. For example, for a median user with 10 friends a unit of *OurTreatment* would increase the user's odds of buying subscription 2.16 times (that is, 116% increase in odds).



As the picture above demonstrates users with small number of friends are the ones who are the most susceptible to peer influence demonstrating hundreds of percents increases in odds ratios. Moreover, given that users with small number of friends represent the vast majority of the social network, this finding shows not only the statistical significance of our result, but also its economic significance.

Notably, *OtherTreatment* is estimated to be stronger than *OurTreatment* across the entire spectrum of user friend counts. This result is not surprising, however, since as we mentioned earlier *OtherTreatment* contains both homophily signal and peer influence, while *OurTreatment* is pure peer influence. In an attempt to explain the size of the estimated difference between the two and warn against over-generalizing this result, we would hypothesize that homophily in this setting can be a weaker force that acts continuously, while peer influence is a sudden and stronger force, but more short-lived, therefore homophily may not manifest itself enough over the short periods of time such as our experiment, while it can potentially manifest itself considerably over longer periods of time. We also acknowledge that peer influence of a person who received a gift subscription as in *OurTreatment* can potentially be weaker than peer influence of a person who paid for subscription herself as in *OtherTreatment*, however that would only make our estimates more conservative .

In addition to confirming both Hypothesis 1 and Hypothesis 2, we also independently observed several effects that are intuitive and confirmed by already published results:

- The estimation results suggest that older people are more likely to adopt subscription; also subscribers and adopters tend to be older and registered earlier than general population confirming the earlier findings of (Oestreicher-Singer and Zalmanson 2011) that independently collected Last.fm data for a different study.
- We discovered that being in non-Lastfmcountry (that is being outside of the US, UK or Germany) provides a significant increase in the likelihood of adopting: a finding that is consistent with the fact that premium subscription gives much more features to people outside of the US, UK and Germany even though it costs the same amount.
- We observed that even after our "gift" manipulation had expired, some people in group M who received a gift from us decided to renew subscription on their own. The difference between the count of "renew-ers" in group M and the count of "adopters" in group NM was statistically significant despite the fact that these groups were chosen initially at random, thus confirming the well-known effect of free promotions.

While these findings are not the main research question of this study, they serve as an important robustness check and confirm that the effects already known in the literature are also observed in this domain.

6.6 Results of Randomized Experiment compared with Quasi-Experiment

An interesting aspect of our study is that it allows us to directly compare the estimates obtained from a quasi-experiment with the estimates obtained from the randomized experiment in order to get an idea of potential biases that can be introduced by doing quasi-experiments with observational data in the research of online social networks.

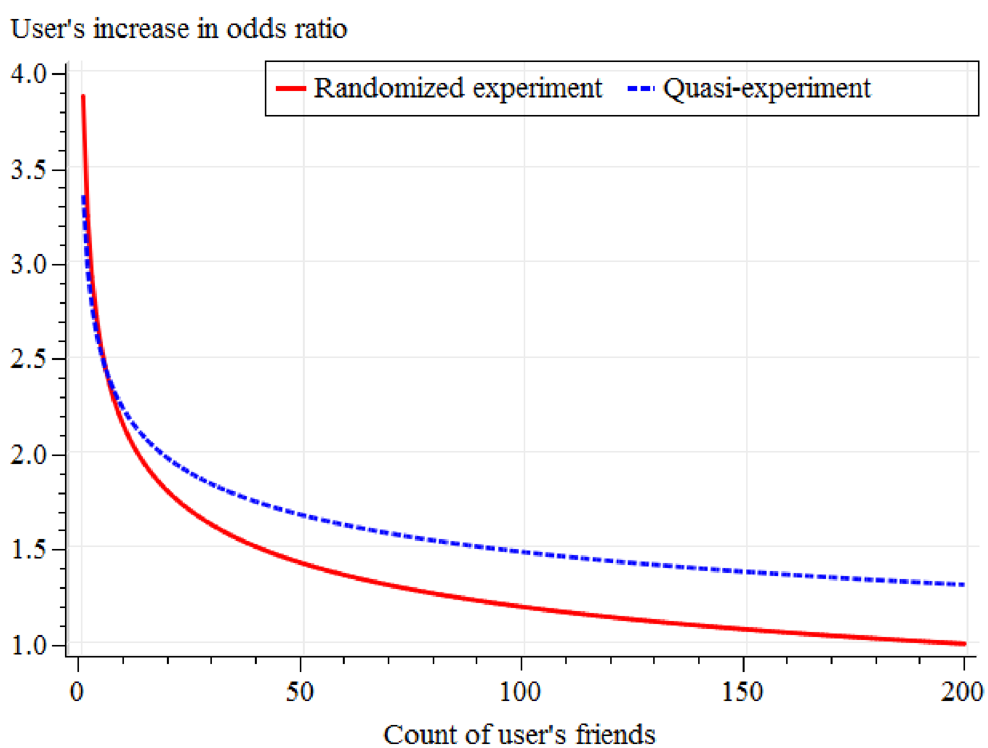
Table 6 below lists the estimates obtained for the same period of time by using randomized experiment and a quasi-experiment respectively.

	Randomized Experiment				Quasi-Experiment			
Variable	**Estimate**	**Std Err**	**Wald χ^2**	**Pr > χ^2**	**Estimate**	**Std Err**	**Wald χ^2**	**Pr > χ^2**
Intercept: Adopter=0	-4.8567	3.3430	2.1106	0.1463	-10.2650	2.3595	18.9265	<.0001
OurTreatment	**1.3570**	**0.5860**	**5.3629**	**0.0206**	**1.2182**	**0.3459**	**12.4058**	**0.0004**
OtherTreatment	**1.5822**	**0.4787**	**10.9245**	**0.0009**	**1.1634**	**0.4200**	**7.6737**	**0.0056**
log(FriendCnt)	0.0167	0.1529	0.0119	0.9130	-0.4467	0.1075	17.2750	<.0001
OurTreatment * log(FriendCnt)	**-0.2547**	**0.1235**	**4.2549**	**0.0391**	**-0.1778**	**0.0758**	**5.5041**	**0.0190**
OtherTreatment * log(FriendCnt)	**-0.2631**	**0.0830**	**10.0523**	**0.0015**	**-0.2055**	**0.0804**	**6.5291**	**0.0106**
log(SubscriberFriendCnt)	**0.4779**	**0.1514**	**9.9686**	**0.0016**	**0.8729**	**0.1015**	**73.9233**	**<.0001**
Age	0.0322	0.0119	7.3139	0.0068	0.00726	0.00895	0.6566	0.4178
AgeMissing	0.6690	0.3909	2.9284	0.0870	0.0376	0.2840	0.0176	0.8945
LastfmCountry	-0.5555	0.2199	6.3780	0.0116	-0.6470	0.1482	19.0610	<.0001
CountryMissing	0.1688	0.3044	0.3076	0.5791	0.3154	0.1779	3.1449	0.0762
RegDate	-0.00023	0.000172	1.8028	0.1794	0.000119	0.000119	1.0014	0.3170
log(SongsListened)	0.0464	0.0696	0.4442	0.5051	0.2129	0.0515	17.0557	<.0001
log(Posts)	0.0991	0.0560	3.1276	0.0770	0.0199	0.0407	0.2387	0.6252
log(Playlists)	0.3543	0.1415	6.2654	0.0123	0.0510	0.0993	0.2633	0.6078
log(Shouts)	0.0259	0.0684	0.1432	0.7051	0.0715	0.0504	2.0121	0.1561
log(LovedTracks)	0.2119	0.0548	14.9566	0.0001	0.2382	0.0370	41.3896	<.0001

Table 6 – Head to Head Comparison of Quasi-Experiment and Randomized Experiment

As evident from Table 6 both quasi-experiment and randomized experiment were able to discover the significance of *OurTreatment* and its interaction terms as well as significance of *OtherTreatment* [12] and log(*SubscriberFriendCnt*).

Since the variable *OurTreatment* features an interaction term in both the randomized experiment and the quasi-experiment, the marginal effects of it are not constant and vary depending on the characteristics of the user. Using an approach similar to the one that we used in the previous section, we plot marginal effects of *OurTreatment* on a median social network user for the randomized experiment and quasi-experiment respectively.



As is demonstrated by the plot above, quasi-experiment tends to over-estimate the effect of the treatment on users with larger number of friends and tends to under-estimate it for the users with

[12] It is interesting to note that, unlike randomized experiment, for quasi-experiment the strength of *OurTreatment* is not expected to be weaker than the strength of *OtherTreatment* since both of them contain homophily and peer influence.

smaller number of friends. Moreover, this pattern of over-estimation and under-estimation is robust across multiple randomization seeds and different runs of the quasi-experiment. To the best of our knowledge this is the first indication of the nature of the quasi-experimental bias in estimating peer-effects in online social networks.

This result is, again, not surprising, since quasi-manipulated group "M" is constructed not from a random sample of general population, but from a random sample of adopters. As we described earlier, adopters have consistently more friends than general population. Because of likely homophily, not only friends of group "M" will have larger number of friends than expected for general population in this scenario, but also the friends of group "M" will be more likely[13] to adopt subscription than general population with that count of friends. Therefore, it is expected that the logistic regression model will overestimate the likelihood of adoption for users with large number of friends given the quasi-experimental data.

Nevertheless, this comparison shows that the quasi-experiment was useful in predicting the basic statistical significance results for the treatment as well as a ball-park estimate of the effect strength. However, the effects of controls were very different in randomized experiment vs. quasi-experiment given that underlying samples were statistically different as well.

7. Conclusions, Discussion and Future Work

In this study we demonstrated the existence of causal peer influence in the general population of an online social network and quantified the strength of this peer influence. We discovered that the strength of peer influence decreases with the size of the friendship circle of the influenced user. In addition to that we compared the point estimates of the pure peer influence effect vs. peer influence and homophily combined. While these estimates provide a way of quantifying the strength of homophily

[13] This is because they are friends with group "M" and group "M" consists completely of users who have purchased the subscription by themselves.

vs. peer influence in the setting of our experiment, a separate study is needed to identify the longitudinal effects of both of these forces. As a concluding remark, we also compared the results of observational quasi-experiment with the randomized experiment and concluded that quasi-experiments tend to overestimate the strength of peer influence for users with large number of friends, while underestimating it for users with small number of friends.

Our work is silent with respect to the exact peer influence tactics that are at work in the ongoing social contagion process: we do not distinguish between tactics like persuading a friend to subscribe vs. raising awareness about the product vs. imitation of a friend etc, as we combine all of them under the umbrella of *peer influence* mechanism. In this paper, we also do not study whether the influence comes from elite highly influential users or a large number of low influential users: our goal for this paper is to demonstrate that significant social contagion is at work on average in a general population of a social network such as Last.fm.

We expect the following as the important directions for the future work:

- Identification of social network characteristics of influential people and people highly susceptible to peer influence beyond the characteristics discovered in this study by using stratified samples and rare-event detection techniques
- Incorporation of dynamic time-series and survival models that are capable of using multiple snapshots of network and network dynamics into the model in order to better explain the adoption behavior over the long-term
- Study of the longitudinal effects and strength of peer influence vs. homophily

Finally, we believe our experimental methodology is something that can be practically carried out by both researchers and practitioners. A venture capitalist could use our design and approach as a dipstick to examine the nature and strength of social contagion in competing networks. We expect to see more such random acts of kindness to solve interesting problems facing business and society.

References

1. Aral, S., Muchnick, L., and Sundararajan, A. 2009. “Distinguishing Influence-based Contagion from Homophily-driven Diffusion in Dynamic Networks,” *Proceedings of the National Academy of Sciences of the United States of America* (106:51), pp. 21544–21549.
2. Aral, S. 2011, “Identifying Social Influence: A Comment on Opinion Leadership and Social Contagion in New Product Diffusion.” *Marketing Science*., 30, pp. 217-223;
3. Aral, S., Walker, D. 2011 “Creating Social Contagion through Viral Product Design: A Randomized Trial of Peer Influence in Networks. Forthcoming *in Management Science*.
4. Bapna, R., Gupta, A., Rice, S., Sundararajan, A., 2011, “Trust, Reciprocity and Strength of Friendship Ties: Evidence from Facebook,” University of Minnesota working paper.
5. Bramoullé, Y., H. Djebbari, B. Fortin. 2009. “Identification of peer effects through social networks.” *J. Econometrics,* 150(1) 41–55.
6. Godes, D., Mayzlin, D. (2004) “Using Online Conversations to Study Word-of-Mouth Communication.” *Marketing Science* 23(4):545-560.
7. Godes, D., D. Mayzlin. 2009. “Firm-created word-of-mouth communication: Evidence from a field test.” *Marketing Science,* 28(4), 721–739
8. Hill, S., Provost, F., Volinsky, C. (2006) “Network-Based Marketing: Identifying Likely Adopters via Consumer Networks.” *Statistical Science*, 21(2):256-276.
9. Iyengar R., Van den Bulte, C., and Valente, T. W., 2011, “Opinion Leadership and Social Contagion in New Product Diffusion,” *Marketing Science* 30, pp. 195-212.
10. Manchanda P., Xie Y., Youn N. (2008) “The Role of Targeted Communication and Contagion in Product Adoption.” *Marketing Science,* 27(6):961-976.
11. Manski C. F. 1995 ”Identification Problems in the Social Sciences.” Harvard University Press.

12. Manski C. F. 1993 “Identification Problems in the Social Sciences.” *Sociological Methodology* 23:1-56.

13. Mayzlin, D. 2006 “Promotional Chat on the Internet.” *Marketing Science,* 25(2) 155-976.

14. Oestreicher-Singer, G., Sundararajan, A. 2010. “The visible hand of peer networks in electronic markets.” Working paper, Leonard N. Stern School of Business, New York University, New York. http://ssrn.com/abstract=1268516 .

15. Oestreicher-Singer, G., Zalmanson, G., 2011, “Paying for Content or Paying for Community? The Effect of Social Computing Platforms on Willingness to Pay in Content Websites,” Working paper, Tel-Aviv University.

16. Oh, J., Susarla, A., Tan, Y., 2011, “Informational Cascades and Contagion in Online Social Networks,” forthcoming in *Information Systems Research*.

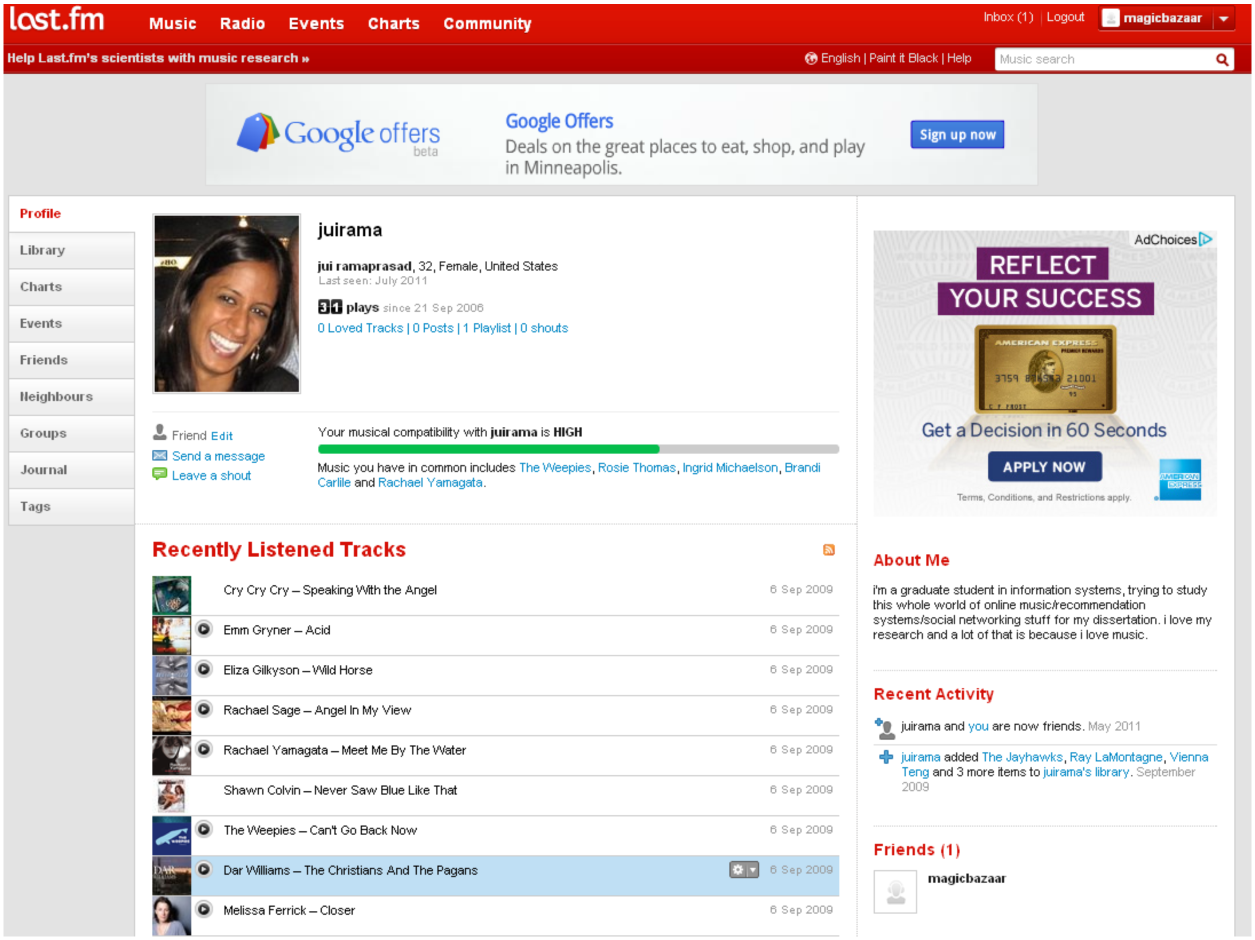
17. Rosenbaum P.R., Rubin D.B. 1983, “The central role of the propensity score in observational studies for causal effects,” *Biometrika,* **70**:41–55.

18. Sacerdote, B. 2001. “Peer effects with random assignment: Results for Dartmouth Roommates.” *Quarterly. J. of Economics*. 116(2) 681–704.

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

20. Van den Bulte C., Lilien GL, 2001, “Medical Innovation Revisited: Social Contagion versus Marketing Effort.” *The American Journal of Sociology* 106(5):1409-1435.

Appendix A – Last.fm User Page



Appendix B – Our Gift Landing Page