MIT Sloan School of Management

MIT Sloan School Working Paper 5177-16

Peer Effect of iPhone Adoptions on Social Networks

T. Tony Ke and Clair Zhuqing Yang

Last Revised: January 2017

Peer Effect of iPhone Adoptions on Social Networks

T. Tony Ke* Clair Zhuqing Yang†

January 2, 2017

Abstract

This paper studies the peer effect of iPhone adoptions in China. We use a unique data set of iPhone adoption records from a provincial capital city in China, in a span of over four years starting from iPhone's first introduction to mainland China. We construct a social network using six month's call transactions among iPhone adopters and all other users on the carrier's network, with strength of social ties measured by duration of calls. Based on the network structure, We test whether an individual's adoption decision is influenced by his friends' adoptions. A fixed-effect model shows that, on average, a friend's adoption increases an individual's adoption probability in next month by 0.89%, and the marginal effect decreases in the size of his current neighboring adopter-base. To further control for potential time-varying correlated unobservables, we instrument adoptions of one's friends by their birthdays and the IV estimation shows a similar peer effect at 0.94%. We also investigate how network structures modulate the magnitude of peer influence. Our results show that peer effect is stronger when the influencer has a stronger relationship with the influencee or when he has less friends.

Keywords: peer effects, identification, social networks, product diffusion
JEL Classification: L14, O33, M31, C26

*Marketing Department, MIT Sloan School of Management, Massachusetts Institute of Technology. Email: kete@mit.edu

†Department of Global Economics and Management, MIT Sloan School of Management, Massachusetts Institute of Technology. Email: clazyang@mit.edu

1 Introduction

Peer effects occur when the action of one agent directly affects others' choices, usually those that are socially close to him, as opposed to via the intermediation of market. Peer effects are ubiquitous in life. Teenagers look to their peers when deciding what songs to listen and what movies to watch, consumers consult their friends before shopping for cars or houses, and even firms and organizations try to learn from others before deciding whether or not to adopt new technologies.

These peer effects are of primary importance to corporate managers as well as policy-makers since they allow a stimulus to one individual to be multiplied through the network. Management has long been aware of the importance of peer effect in launching a successful new product. Firms frequently give out free samples to selected customers, and consciously design effective marketing campaigns to leverage peer effects on social medias (Aral and Walker 2011). Policy interventions, such as school desegregation and busing, have used social interactions as the major goal to alleviate stratification by income, education, race, and to improve social equality (Moffitt et al. 2001). Quantifying the magnitude of peer effect therefore is critical to constructing sound network interventions in both the public and private sectors.

In this paper, we study the peer effect in adoption of a new consumer technology—iPhones—using individual-level iPhone adoption data from a provincial capital city Xining in northwestern China. Our sample spans a period of over four years starting from iPhone's first introduction to China market. During this period, the mobile phone carrier China Unicom has the exclusive right to sell iPhones in mainland China, hence our data include almost all iPhone users adopted during the time period[1]. We construct a social network using half a year's call transactions between iPhone adopters and all other users on a carrier's network. Based on the network structure, we test whether or not an individual's iPhone adoption decision was affected by his/her friends' decisions. We quantify the peer effect using both fixed effect and instrumental variable approaches, and investigate how network structures modulate the magnitude of peer influence.

Xining, where our data come from, is the largest city on Tibetan Plateau, with population of around two million and GDP per capita of around $7,000$ US dollar in 2013. iPhones were considered a very novel technology when first introduced to Xining, and we expect peer influence plays an important role in people's purchase decisions. On a fundamental level, iPhone adoption could be subject to peer effect due to informational, behavioral, social preference, or network externality reasons. Friends' recommendations and user reviews give a consumer much information about how good an iPhone is and whether it fits their need—this is the informational channel, so-called word-of-mouth. It is also possible that a consumer observes other people's purchase and usage decisions and infer iPhone is a good product—this is the behav-

[1]Except a very few people who bought iPhones from overseas and brought it back to China to use.

ior channel, or behavior learning. Alternatively, iPhone could also be viewed as a fashionable product. Using an iPhone speaks something about one's personality and taste; and observing other people using it changes directly the utility one can get from using the same product, either positively or negatively—this is the social perference channel, e.g., bandwagon or snob effect. Lastly, iPhone is a communication device with many add-on applications. Having one more person in the network makes it more attractive for others to join since they will have more people to communicate with[2]. In this paper, we do not attempt to distinguish these channels. Instead, we focus on quantifying the peer effect, which could be an aggregate of all the channels.

Identification of peer effect has long been a challenge to social scientists. Peer effect implies that the behavior of connected agents on a network tends to be correlated. However, correlation in the behaviors *per se* does not necessarily imply that one agent's action has a causal effect on that of others. Other factors besides peer effect could also give rise to such behavioral correlations. From a policy and strategy point of view, only causal peer effects are of primary interest because it impacts the outcome of individual-level policy interventions. Moffitt et al. (2001) summarized that the primary factors confounding the identification of peer effects are: simultaneity, endogenous group formation and correlated unobservables. Simultaneity problem arises if one person's action influences the others, and *vice versa* (Manski 1993). Fortunately, our setting does not suffer from this simultaneity problem, because of the natural sequence of individual adoptions across time in the panel data. Endogenous group formation problem arises when the outcome variable also affects the likelihood of two agents being connected, which in our case, means that two strangers start calling each other because both of them use iPhone. Arguably, it is true that this might happen theoretically; yet, we believe it would be too subtle an effect to intervene with the identification, especially since using a smart phone or not is irrelevant to the quality and cost of phone call service in the market where our data is collected. To further support this exogenous network formation argument, we also provide evidence in section 5.4 showing that the network structure has been stable throughout my sample period and adopting an iPhone does not change agents' phone call patterns.

Therefore, the only serious potential confounding factor remained in our setting is correlated unobservables. Adoption decisions of one's peers can be endogenous for his adoption decision, because people who know each other tend to face similar unobserved environment to adopt the technology. To address this issue, we come up with two approaches. To control for time-invariant correlated unobservables, we apply an individual fixed-effect model. To further control for time-varying correlated unobservables, we come up with an instrumental variable approach. We use an individual's friends' birthdays as an instrumental variable for his friends' adoption

[2] Strictly speaking, the network externality effect operates via add-on applications but not phone calls. This is because, during the time when we collected our data, using a smart phone or not is irrelevant to the quality and cost of phone call services. Hence, we do not expect using iPhone to change one's preference of phone calls.

decisions, and see how these birthday-induced adoptions by his friends affects his own adoption decision.

Our fixed effect (FE) and instrumental variable (IV) models show results similar in magnitudes. A friend's adoption decision indeed has a positive impact on one's own. According to the FE model, having one more friend adopting iPhone increases an individual's probability of adoption by 0.89% in the next month. The IV regression yields a similar estimate of 0.94%, after clearing away effects of potential time-varying correlated unobservables. We also show that this peer effect decreases in the number of current adopters. In other words, as more friends have already adopted, the marginal impact of an additional friend becomes smaller. Using our estimates, a firm looking to promote iPhone sales in a setting similar to ours would be able to compute the average external peer effect of a new user[3]. And these numbers would be of great interests for managers when designing optimal promotion schemes.

We also investigate how heterogeneity in network structure impacts the magnitude of peer influence. It is one of the frontier questions to study the impact of individuals' network attributes on social influence (e.g. Tucker 2008, Banerjee et al. 2013, Aral and Walker 2014). Our results show both an "intimacy" effect and a "popularity" effect. Our results show that the more time a pair of friends spent on talking to each other during the six months, the greater the peer effect is between them. Whereas a person with more friends has a smaller peer effect on each of his friends.

The paper unfolds itself as the following. Section 2 summarizes relevant literature and our connections to previous studies. Section 3 gives the background and basic patterns of our data sample. Major empirical results are provided in section 4. Section 5 explores the impact of network heterogeneity on peer effect and various robustness checks. And section 6 concludes.

2 Literature Review

Individuals make decisions in almost every social aspect under the influence of friends, neighbors, or professional peers: from education (Sacerdote 2001, Epple and Romano 2011), criminal activities (Glaeser et al. 1996, Bayer et al. 2009), welfare program participation (Bertrand et al. 2000, Duflo and Saez 2003), to physicians' prescriptions (Manchanda et al. 2008, Nair et al. 2010, Iyengar et al. 2011), etc. In product market especially, there is widely recorded phenomenon of peer influence on purchasing behaviors: from computers (Goolsbee and Klenow 2002), online groceries (Bell and Song 2007), automobiles (Narayanan and Nair 2013), TV service (Nam et al. 2010), in-flight purchase (Gardete 2015), insurance plans (Cai et al. 2015), to solar panels (Bollinger and Gillingham 2012), etc.

[3]For example, at the initial stage of iPhone diffusion when an average individual have about one user friend, the direct impact of a new iPhone user on his peers would be about 1.01%, compared to a much smaller effect of 0.65% at a later stage when an average individual have about 100 adopted friends.

It is worth mentioning that there is a subtle difference between peer effect and network effect (for the latter, see a survey by Birke 2009). Network effect relies on network externality, which captures the phenomena that having more people in a certain group makes the utility of later joiners even higher. Typical examples include adoption of industry standards (e.g. David 1985, Augereau et al. 2006), choice of business platforms (Brown and Morgan, 2009, Hendel et al., 2009, Cantillon and Yin, 2008), and membership of social media websites such as Facebook and LinkedIn. Peer effect, on the other hand, encompasses a much broader meaning. In addition to being triggered by network externality, peer effect could also be due to informational, behaviorial or social reasons: consumers could learn about a product from others' comments and choices, or simply find it fashionable to go for what is "hot". Regardless of the underlying mechanism, peer effect manifests itself as a causal influence of one's action upon his peers' choices.

The peer influence that we study in this paper can be seen as a special case of general social interaction effects (Manski 1993, Moffitt et al. 2001). Social interaction effects usually include both contextual effects—the direct influence of others' characteristics on one's choice—and peer effects—the influence by others' actions[4]. Many attempts have been made to demonstrate and quantify the peer effects. The early literature on aggregate diffusion has been trying to quantify "peer effects" by treating the entire population of past adopters as the reference group (Bass 1969, Mahajan et al. 1990). With access to more micro-level data, recent studies have taken on a more subtle view of reference groups, emphasizing the role of social structures in channeling peer effects based on geographic locations (e.g. Bell and Song 2007, Bollinger and Gillingham 2012, Narayanan and Nair 2013), ethnic or culture proximity (e.g. Bandiera and Rasul 2006), friend or family relationships (e.g. Conley and Udry 2010), or some combination of these factors.

However, a closer look at these heterogenous peer effects poses an identification challenge aforementioned. Some of the studies have tried controlling for detailed individual-level information to alleviate the correlated unobservable problem. Several studies also use specific identification strategies to study peer effect. For example, Bollinger and Gillingham (2012) identified the peer effects in adoption of solar photovoltaic panels, by leveraging the time delay of installations after the initial request. Their analysis is based on zip code level data without network structures; also the validity of their empirical strategies hinges on the assumption that there is no covariate that influences two adoption decisions that are separated by the installation delay or longer. Nam et al. (2010) studies the adoption of a video-on-demand service, where

[4]A literature somewhat relates to ours are those that use identification strategies to study the influence of an individual's social activities or characteristics on his *own* behavior or social network structure. For example, Shriver et al. (2013) studies whether online users' (surfing-related) content-generation activity affects their social ties and vice versa, by exploiting changes to wind speeds at various surfing locations. Our question here is different from and conceptually harder than theirs, because peer effect captures the spread of the same behavior among individuals and it is usually harder to find exogenous shocks only to some people but not to their friends for the same behavior.

random fluctuation in the signal quality adds exogenous shocks to the content of message communicated from friends to friends, but not the initial adoption decisions. Also, they are based on geographical locations without network details. Godinho de Matos et al. (2014) also studied peer effects of iPhone adoptions on social networks. They combined a network-structure-based community identification algorithm with instrumental variables to identify the peer effect. The instrument they used for one's friends' adoptions is the friends' friends' adoptions, which however, may still be subject to potential endogeneity issues such as homophily. Also, their data do not cover all iPhone users in the local market.

More recent studies have been utilizing randomized field experiments to get clean identifications of peer effects (e.g., Sacerdote 2001, Duflo and Saez 2003, Bursztyn et al. 2014, Cai et al. 2015, among others). But oftentimes the population sample under study is limited due to feasibility constraints of experiments, and the social network structure is either unavailable, or measured by subjective surveys. In contrast, our sample includes all adopters and their peers in a metropolis, and we construct social network from objective measurements, using duration of phone calls among consumers to approximate social tie strength. Such a communication network would be able to preserve more detailed and nuanced information than location proximities or networks obtained by surveys. And we also use an instrumental variable approach to get a clean identification of peer effect.

The research that is the closest to ours is Tucker (2008). In her paper, she identified the network externality in adoption of a video-messaging technology, by utilizing a stand-alone use of the technology (watching local TV programs) as an instrument. Methodologically, her approach is close to ours. Nonetheless, her study is on a very specific setting: adoptions of a technology occur in a corporation instead of the marketplace, and individuals do not incur any pecuniary cost to adopt the technology.

To summarize, our paper studies peer effect of a mainstream consumer product (iPhone) on a social network, which is constructed from objective phone calls among consumers. In contrast to the large literature on prediction of product diffusion using network structures (e.g., Hill et al. 2006, Katona et al. 2011, among others), the main goal of this paper is not to make predictions of individual adoptions. However, based on a consistent estimate of peer effects, our findings could indeed help predict future sales of similar products.

3 Background and Data Description

3.1 Data Background

IPhone was first introduced in China on October 30, 2009. For a rather long time, iPhone was offered to subscribers of China Unicom exclusively, until January 17, 2014, when China Mobile started to offer iPhone on its network. There are three players in the mobile phone telecommunication market in China: China Unicom, China Mobile, and China Telecom; and all of which are state-owned public companies. Currently, China Mobile owns roughly 70% market share of mobile telecommunications in China, whereas China Unicom about 20% and China Telecom the rest 10%.

Our data set includes monthly call transactions and iPhone adoption records of all iPhone adopters in China Unicom network who adopted before the end of October 2013 in capital city Xining of Qinghai province in northwestern China. The data on call transaction is available from May 2013 to October 2013. Whereas the dataset of iPhone adoption is complete: running from the first adoption in November 2009 to October 2013.

We also get complimentary data containing individual information of the adopters, such as cell phone monthly usage charge, service plan subscription, and most importantly, individual's birthday, which we will use as an instrumental variable for adoption time. Data on individual information is available for a subgroup (72.3% of the population of all adopters).

3.2 Adoption Pattern

There are in total $82,471$ adoption instances from November 2009 to October 2013. Some adopters later stopped using iPhone (by either dropping out of the carrier's network or by replacing it with a phone of other brands), and there are $47,727$ (57.9%) active iPhone users by the end of October 2013.[5] The monthly adoption and usage trend is shown in Figure 1. We can see that the adoption rate grows exponentially during the sample period. Therefore, in the framework of a S-curve diffusion (Griliches 1957, Rogers 1962), our dataset covers an early stage of iPhone adoptions.

In the data, we observe dives and surges of new adoptions in some months. This is mainly due to three reasons: consumers' strategic waiting before launching of new models, occasional limited supply capacity in certain months, and seasonalities, such as the National Day Sales in October, and Spring Festival around the beginning of each year. We include monthly fixed effect in all our regressions to control for these time trends.

3.3 Call Pattern

Our call transaction data set includes all people who have adopted iPhone from November 2009 to October 2013 and remained active during May to October 2013, as well as all of their friends, who have either made or answered a call with an iPhone

[5]The four-year cumulative attrition rate of over 40% seems relatively high, which we think is an idiosyncratic feature of China's market. As iPhone's exclusive carrier in the sample period, China Unicom owns only 20% share of the telecommunication market, compared with China Mobile's 70% market share. We expect a significant amount of people stops using iPhone because they switches from China Unicom to Chine Mobile, which in general has better network signal coverage.

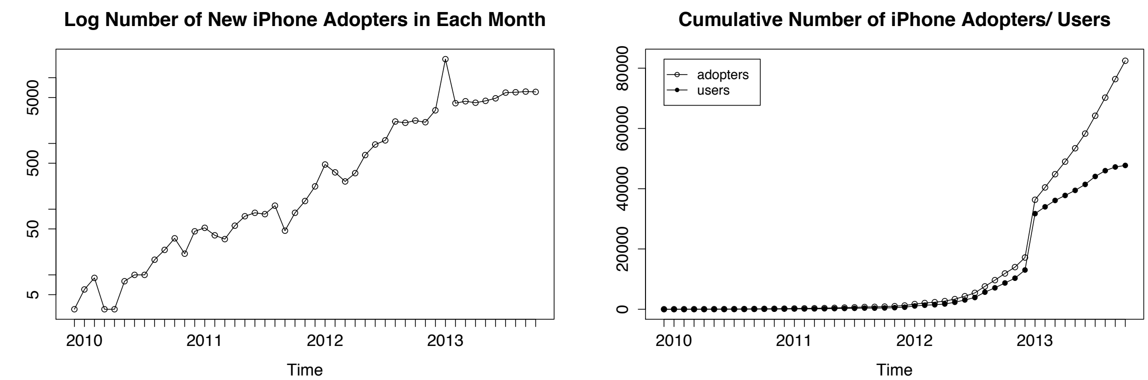

Figure 1: Adoption and usage trend of iPhone after introduction in Nov-2009.

adopter between May 2013 and October 2013. IPhone users must be China Unicom subscribers, while their friends are not necessarily China Unicom subscribers. Out of 82, 471 iPhone adopters, 74, 967 (90.9%) remained active in the carrier's network in the period from May to October 2013.[6] Our following analysis will be based on the sample of 74, 967 active adopters. There are in total, 4, 030, 156 friends of all active iPhone adopters. Call transactions are aggregated by month. Each transaction consists of the following information: anonymized phone number of the caller, anonymized phone number of the receiver, and their monthly call duration. Between two users on the network, if they did not make a single call during the sample period, their (null) transaction is not included in the dataset.

There are 10, 762, 428 call transaction records in total between May 2013 and October 2013. We use these call transactions to construct social network for iPhone adopters and their friends. Therefore our social network embeds 90.9% of the entire sample population who have ever adopted an iPhone between Novermber 2009 and October 2013 in Xining. The total number of people who made calls is 82, 420, and the total number of people who received calls is 4, 328, 013 during this period.[7] Combining both callers and receivers (iPhone adopters and their friends), there are 4, 105, 123 individuals on the phone-call network. The average monthly call duration for each pair of contacts, who at least made one call in that month, is roughly 11 minutes.

The following Table 1 provides the summary statistics for our sample. As we can see, the phone call network proves to be quite stable over the sample period, a point

[6] By "being active", we mean a consumer made or received at least one call in the period. In China, when a consumer switches mobile phone carrier, he has to change his phone number. Our record of a consumer discontinued when he left China Unicom.

[7] The huge difference between numbers of callers and receivers comes from calls from outside the carrier's network. We do not have information on incoming calls from outside network, and can only observe outgoing calls to outside network.

which will be further confirmed by more rigorous network analysis in section 5.4.

Table 1: Summary Statistics for the Data Sample

	# Observations	Mean	Std Dev	p25	p50	p75
Panel A: Adoption (November 2009 to October 2013)						
adoption	74,967					
individual birthday	54,171					
Panel B: Call Transaction (May 2013 to October 2013)						
duration-05	2,911,112	11.49	52.65	1.00	2.57	7.57
duration-06	2,943,157	11.29	51.29	1.00	2.60	7.60
duration-07	3,201,634	10.98	49.22	0.98	2.55	7.45
duration-08	3,290,485	10.97	50.03	1.00	2.57	7.50
duration-09	3,232,633	11.21	51.47	1.00	2.55	7.47
duration-10	3,238,310	11.25	52.50	0.98	2.52	7.40
total duration	10,762,428	19.57	130.05	.0.97	2.65	8.97

Note: The panel consists of all iPhone adopters in China Unicome network in a provincial capital city Xining in northwestern China by the end of October 2013. Monthly (total) call duration includes aggregated call transactions of all pairs with non-zero call duration in that month.

4 The Empirical Model

Our empirical model follows the linear-in-mean model of social interactions (Manski 1993), which we interpret as a reduced form of the behavioral process generating adoption decision across the population network.

Strictly speaking, there are two variations of the linear-in-mean model, one of which has the absolute number of adopters as the explanatory variable whereas the other has the fraction of adopters out of all friends. Both specifications are theoretically justifiable, and researchers usually make their own choices as which one to use. In this paper, we use absolute number of adopters as the explanatory variable for both the main models and the extension on network heterogeneity. Similar results using the fraction of adopters as explanatory variable are discussed in the robustness check section 5.3.

4.1 The Network Structure

We index consumers (network nodes) by i. We first construct a social network by aggregating all call transactions from May 2013 to October 2013. If there is a call from Alice to Bob in the six months, we establish a directed link from Alice to Bob. In this way we get a directed social network, which has 443 weakly connected maximal components in the network, among which the largest one consists of 4,102,936 individuals (99.95% of the whole population).

Given the network, we define inward neighbors for consumer i as all other consumers that have a directed link to i; and similarly we define the outward neighbors for i as all others that have a directed link from i. Then, we construct the panel dataset with the following variables

$$(i, t, ADOPT_{it}, INSTALLBASE\_IN_{it}, INSTALLBASE\_OUT_{it})$$

where $ADOPT_{it}$ is the adoption indicator of individual i in month t: $ADOPT_{it} = 1$ if i first adopted iPhone in month t, and $ADOPT_{it} = 0$ otherwise.[8] $INSTALLBASE\_IN_{it}$ is the cumulative number of i's inward neighbors who have adopted iPhone by month t; and similarly $INSTALLBASE\_OUT_{it}$ is the cumulative number of i's outward neighbors who have adopted iPhone by month t. By combining our adoption dataset and the social network constructed from transaction data set, we get a panel dataset of 3,100,442 individual-year observations. The summary statistics for the panel is provided in the appendix.

In this section, we use absolute number of (inward/outward) neighbouring iPhone users as a measure of peer effects on the network, by assuming homogeneous peer effect. In section 5.2, we discuss various extension of the basic model and try to incorporate more network structure into the measures.

4.2 Fixed Effect Model

In this paper, we define a focal consumer's *peers* as those who are directly linked to him on the phone call network. In other words, we are estimating a *local* network effect. Arguably, a consumer's adoption decision could also be affected by macro-level network features, such as iPhone diffusion on the overall network. We would not be able to test such macro-level features in this paper since we only have data on one city (and hence, one network). But we do control for these variables by including time trends in the regression.

The two-way fixed effect (FE) regression of peer effect on adoption is

$$ADOPT_{it} = \beta_1 INSTALLBASE_{it-1} + \alpha_i + \gamma_t + \varepsilon_{it}, \tag{1}$$

where ε_{it} is assumed i.i.d. across individuals i and time t, and α_i, γ_t are individual and time fixed effect respectively. The variables $INSTALLBASE\_IN_{it}$ and $INSTALLBASE\_OUT_{it}$ are highly correlated, with a correlation coefficient at 0.998, hence we include only one of them in the regression equation as $INSTALLBASE_{it-1}$ to avoid collinearity.

[8]Aforementioned, we consider adoption instead of usage decisions, therefore, with $ADOPT_{it} = 1$, all records of individual i after t are dropped from our panel.

In this paper, we choose a linear probability model over a logit for its simplicity in incorporating instrumental variable and a flexible fixed effect structure with a panel structure[9]. Another reason that makes it specifically difficult to implement a logit model in our setting is quasi-separation issue of our data, which leads to a non-convergence and potential bias of logit maximum likelihood estimator (Albert and Anderson, 1984)[10]. Hence, in this paper, we use a linear probability model to estimate the peer effect on iPhone adoption averaged across the population.

By including only past values of *INSTALLBASE* in the regression, our model specification features a one-direction influence (usually termed as *passive* social interaction in the literature, see Hartmann et al., 2008) instead of a feedback loop. Theoretically, a setting would best fit a one-direction framework if the action of one focal agent affects his neighbours but not the other way around, or, more commonly, if the agent is concerned only about the *realized* actions of his neighbors and is myopic enough to not foresee how his current decision will impact his future self by influencing others around him. In this paper, we implicitly make the latter assumption that an agent is concerned only about the past actions of his neighbors and is myopic. Hence, we quantify peer effect via equation (1).

Table 2 summarizes the estimation results. As we can see that having one more friend adopting iPhone, on average, increases an individual's probability of adoption by 0.89% in the next month. This holds true for either inward- or outward-phone-call definition of friend. A quadratic model shows that this peer effect decreases in the number of current adopters. In other words, as more friends have already adopted, the marginal impact of an additional friend becomes smaller.

Figure 2 plots the coefficients of peer effect by month from January 2010 to October 2013. Due to very few adoptions in the earlier stage and small variation in the explanatory variable, the estimates of peer effect coefficient appear insignificant with very wide confidence intervals before 2012. However, iPhones underwent an acceleration of diffusion towards the end of 2012. And a positive peer effect started to manifest itself. From then on, the estimate of peer effect remained significant and mostly stable throughout the sample period.

[9] As pointed out by Narayanan and Nair (2013), having a flexible fixed effect structure is vital to getting a consistent estimate when individual characteristics correlate with install bases, while choosing a linear probability model over a non-linear one does not compromise the uncovering of the true value even when underlying process is a non-linear one.

[10] Heinze and Schemper (2002) proposed using a penalized maximum likelihood estimation originally developed by Firth (1993) to solve the separation problem. However, such a method would pose much complexity to a panel discrete choice model, which is already subject to the incidental parameters problem due to adding fixed effect to a non-linear logistic or probit model. Hence, in this paper, we try to shield from these technical difficulties by using a linear probability model.

Table 2: Estimation of Peer Effect with FE Model

VARIABLES	(1) $ADOPT_{it}$	(2) $ADOPT_{it}$	(3) $ADOPT_{it}$	(4) $ADOPT_{it}$	(5) $ADOPT_{it}$
INSTALLBASE_IN	0.00802***	0.00890***		0.0114***	
	(0.000269)	(0.000636)		(0.000353)	
INSTALLBASE_OUT			0.00886***		0.0114***
			(0.000629)		(0.000354)
INSTALLBASE_IN^2				-4.30e-05***	
				(6.79e-06)	
INSTALLBASE_OUT^2					-4.29e-05***
					(6.80e-06)
Constant	-3.03e-11	-0.00232***	-0.00232***	-0.00229***	-0.00229***
	(0.0004836)	(8.54e-05)	(8.54e-05)	(8.57e-05)	(8.57e-05)
Individual FE	N	Y	Y	Y	Y
Monthly FE	Y	Y	Y	Y	Y
Observations	3100442	3,100,442	3,100,442	3,100,442	3,100,442
R-squared	0.2568	0.275	0.275	0.277	0.277
Number of i		74,967	74,967	74,967	74,967

Note: * denotes significance at 10% level, ** at 5% level, and *** at 1% level. All estimations above use robust standard error to control for heteroscedasticity in linear probability models.

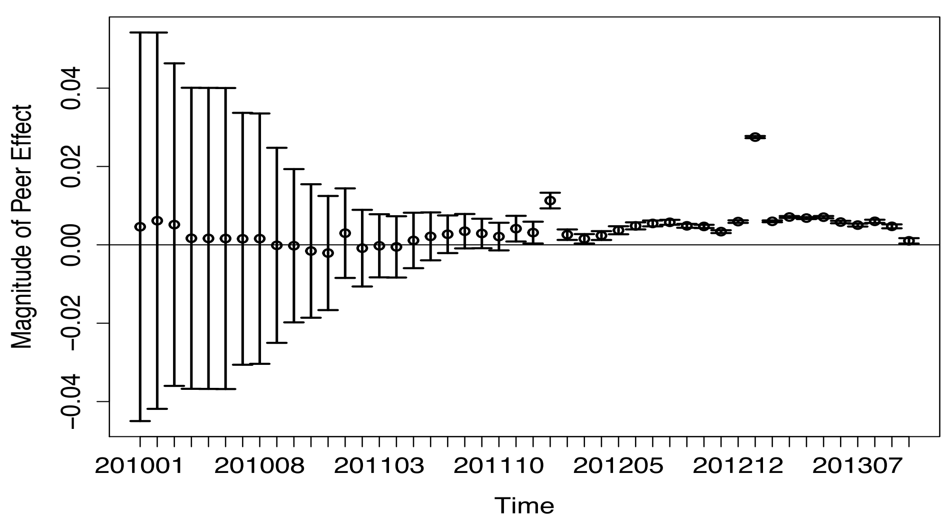

Figure 2: Peer Effect on IPhone Adoption by Month

4.3 Instrumental Variable Model

As discussed earlier, the only potential confounding factor remained in our setting is correlated unobservables. Our fixed effect model in the previous section controls for time-invariant unobservables, but the concern for time-variant correlated unobservables remains. In other words, if there exist omitted variables that both correlate with the network structure and are time-varying in nature, our identification with a FE model could be compromised. As an example, let us consider the following situation with only two types of people in the population, fashion-followers, who are more likely to purchase a product when the overall sales are high (compare to peer effect that depends on adopters in the local friends network), and non-followers. Fashion-followers are naturally more likely to adopt in the later stage of product diffusion, hence cannot be controlled by individual FE. Moreover, if fashion-followers are also more likely to become friends, then it cannot be controlled by monthly FE either and will induce a spurious peer effects among friends. Another example will be targeting marketing campaigns, which are usually time-varying and offered to people with similar characteristics. Since people with similar characteristics are more likely to be friends (known as homophily), targeting marketing campaigns can be a time-varying unobservable. In this section, we use an instrumental variable (IV) approach to further control for such confounding effects.

We use an individual's birthday as an IV for his adoption decision, and test how this affects his neighbors' subsequent adoptions. The basic idea is that it is more likely for people to adopt iPhones on their birthdays, because either they are more likely to reward themselves with a long fancied product, such as an iPhone, or they are more likely to receive one as a gift on their birthdays. So we expect to see a higher probability of iPhone adoption around an individual's birthday. Being totally random, birthdays would satisfy the exclusion restriction automatically[11]. All we need to check is the inclusion requirement to make it a legitimate instrumental variable, which we test with weak instrument tests below.

For individual i, we define $BDAY_{it}$ as a dummy variable which equals to one if i's birthday is in month t and zero otherwise. By running an OLS regression of $ADOPT_{it}$ on $BDAY_{it}$, we find that the coefficient estimate as 0.088, which is statistically significant with p value less than 1%. This means that on average an individual is 8.8% more likely to adopt an iPhone during his birth month. We denote individual i's adoption time as $\tau(i)$. Individual i's installed base can be constructed by aggregating adoptions among his neighbours $\mathcal{N}(i)$ up till time $t-1$:

$$INSTALLBASE_{it-1} = \sum_{j \in \mathcal{N}(i)} \sum_{s=1}^{\min\{\tau(j), t-1\}} ADOPT_{js} \tag{2}$$

In other words, variable $INSTALLBASE_{it-1}$ is the number of current iPhone users

[11] To further justify the usage of birthday as an IV, China Unicom did not have advertisements or promotions based on customers' birthdays in the sample period.

among i's friends at time t. Here, we take a first step by using the following IV[12]:

$$IV\_BDAY_{it-1} = \sum_{j \in \mathcal{N}(i)} \sum_{s=1}^{\min\{\tau(j),t-1\}} BDAY_{js}. \tag{3}$$

We use a standard two stage least square estimation (2SLS) procedure. The first stage regression becomes the following:

$$INSTALLBASE_{it-1} = \delta_1 IV\_BDAY_{it-1} + \omega_i + \pi_t + \epsilon_{it}. \tag{4}$$

Here, ω_i and π_t are the individual and time fixed effects. By estimating the above regression equation, we have $\widehat{INSTALLBASE}_{it-1}$ as the independent variable of interest in the second stage regression:

$$ADOPT_{it} = \beta_1 \widehat{INSTALLBASE}_{it-1} + \alpha_i + \gamma_t + \varepsilon_{it} \tag{5}$$

Here, again, ε_{it} is assumed i.i.d. across i and t, and α_i, γ_t are individual and time fixed effect respectively.

Table 3 gives the estimation results for the first stage IV regression. We get an estimate of δ_1 at about 0.099 with a significant p-value as 0.000. Similar results hold for both inward- and outward- definition of linked friends. This shows that friends' adoption decisions are indeed affected by their birthdays[13].

Table 4 shows the IV regression of peer effect on iPhone adoptions. Here we get estimate of peer effects similar in magnitude. According to the IV estimates, having one additional friend adopting iPhone increases an individual's probability of adoption by about 0.94% in the next month (compared to the 0.89% estimate by FE model). Similar to the previous section, our IV estimates also find this marginal impact of newly adopted friends decreasing when the size of already adopted user base gets bigger.

[12]From this equation, we know for each $j \in \mathcal{N}(i)$ and $s = 1, \cdots, \min\{\tau(j), t-1\}$, $BDAY_{js}$ can be used to instrument $INSTALLBASE_{it-1}$. Intuitively, each friend's past adoption induced by his birthday will cause an exogenous shock to one's installed base. Therefore, $INSTALLBASE_{it-1}$ is over-identified, and the most effective IV can be obtained by GMM estimation.

[13]Arguably, there might be some heterogeneity in the effectiveness of the IV. For example, some people, like teenagers, are more likely to get an iPhone as a gift on their birthdays than others. Since we have strong reasons to believe in the monotonicity of such effect, i.e. no person will be *less* likely to adopt iPhone on his birthday, this would not compromise our identification. It might, however, add some subtleties to the interpretation of the results. As in the usual case of heterogeneous treatment effect, our IV regression estimates the average causal effect for those that are affected by the instrument.

Table 3: First Stage With Birthday IV

VARIABLES	(1) *INSTALLBASE_IN*	(2) *INSTALLBASE_OUT*	(3) *INSTALLBASE_IN*	(4) *INSTALLBASE_OUT*
IV_BIRTHDAY_IN	0.0590***		0.0992***	
	(5.34e-05)		(9.27e-05)	
IV_BIRTHDAY_OUT		0.0591***		0.0994***
		(5.35e-05)		(9.29e-05)
Constant	-0.112***	-0.113***	0.0218***	0.0218***
	(0.000964)	(0.000966)	(0.00470)	(0.00471)
Individual FE			Y	Y
Monthly FE			Y	Y
Observations	3,100,442	3,100,442	3,100,442	3,100,442
R-squared	0.283	0.283	0.400	0.399
Number of i			74,967	74,967

Table 4: Second Stage With Birthday IV

VARIABLES	(1) $ADOPT_{it}$	(2) $ADOPT_{it}$	(3) $ADOPT_{it}$	(4) $ADOPT_{it}$	(5) $ADOPT_{it}$	(6) $ADOPT_{it}$
INSTALLBASE_IN	0.00936***		0.00952***		0.0126***	
	(9.62e-05)		(0.00308)		(0.000121)	
INSTALLBASE_OUT		0.00931***		0.00946***		0.0126***
		(9.58e-05)		(0.00306)		(0.000121)
$INSTALLBASE\_IN^2$					-7.59e-05***	
					(1.19e-06)	
$INSTALLBASE\_OUT^2$						-7.59e-05***
						(1.19e-06)
Individual FE	Y	Y	Y	Y	Y	Y
Monthly FE	Y	Y	Y	Y	Y	Y
Error			HAC/Clu(i,t)	HAC/Clu(i,t)		
Estimator	IV-2SLS	IV-2SLS	IV-GMM	IV-GMM	IV-2SLS	IV-2SLS
Wald F-stat	1.20E+06	1.20E+06	1.20E+06	1.20E+06	4.40E+05	4.40E+05
Observations	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442
R-squared	0.275	0.275	0.275	0.275	0.276	0.276
Number of i	74,967	74,967	74,967	74,967	74,967	74,967

Note: HAC stands for heteroskedasticity-autocorrelation (HAC) robust, and Clu(i,t) stands for 2-way clustered standard error (Cameron et al. 2006, Thompson 2009) that are robust to arbitrary heteroskedasticity and intra-group correlation with respect to both time and individual dimensions.

5 Extensions and Robustness Checks

5.1 Alternative Measure of Birthday IV

In our main model, both the installed base $INSTALLBASE_{it-1}$ and the birthday IV IV_BDAY_{it-1} have increasing time trends. Even after controlling for time fixed effect in the first stage of the 2SLS regression, there could still be spurious significance due to heterogeneous time trends across individuals. In this section, we try to provide further evidence of our result by considering an alternative measure of birthday IV using all friends' birthdays over the last three months. To be more precise, we have

$$IV\_BDAY3_{it-1} \equiv \sum_{j \in \mathcal{N}(i)} \sum_{s=\max\{1,t-3\}}^{\min\{\tau(j),t-1\}} BDAY_{js}. \tag{6}$$

which is the total number of friends' birthdays over the past three months.

The following table (5) and (6) shows the first and second stage results using the alternative birthday IV, with table 7 tests a slightly modified version of birthday IV over six months. All these modifications yield similar results.

Table 5: First Stage With 3-Month-Total Birthday IV

VARIABLES	(1) *INSTALLBASE_IN*	(2) *INSTALLBASE_OUT*	(3) *INSTALLBASE_IN*	(4) *INSTALLBASE_OUT*
IV_BDAY_IN3	0.0984***		0.0984***	
	(0.000710)		(0.00659)	
IV_BDAY_OUT3		0.0991***		0.0991***
		(0.000712)		(0.00663)
Constant	-0.00529	-0.00526	-0.00529**	-0.00526**
	(0.00550)	(0.00552)	(0.00208)	(0.00208)
Individual FE	Y	Y	Y	Y
Monthly FE	Y	Y	Y	Y
Error	OLS	OLS	Robust	Robust
Observations	3,100,442	3,100,442	3,100,442	3,100,442
R-squared	0.178	0.177	0.178	0.177
Number of i	74,967	74,967	74,967	74,967

5.2 Network Structure and Heterogenous Peer Effect

Heterogeneous social interactions have important implications for policy design and for firms' allocation of marketing efforts. Peer effects, as one of many influences that channeled through social interactions, could be very sensitive to the social and structural conditions under which the interaction happens. Studies, especially in sociology, have long recognized that early adopters with different "status" or level of "popularity" would have varied influence over later new comers, and hence different

Table 6: Second Stage With 3-Month Birthday IV

VARIABLES	(1) $ADOPT_{it}$	(2) $ADOPT_{it}$	(3) $ADOPT_{it}$	(4) $ADOPT_{it}$	(5) $ADOPT_{it}$	(6) $ADOPT_{it}$
INSTALLBASE_IN	0.0118***		0.0121***		0.0173***	
	(0.000635)		(0.00462)		(0.000974)	
INSTALLBASE_OUT		0.0118***		0.0122***		0.0175***
		(0.000630)		(0.00460)		(0.000976)
INSTALLBASE_IN^2					-0.000126***	
					(9.70e-06)	
INSTALLBASE_OUT^2						-0.000129***
						(9.64e-06)
Individual FE	Y	Y	Y	Y	Y	Y
Monthly FE	Y	Y	Y	Y	Y	Y
Error			HAC/Clu(i,t)	HAC/Clu(i,t)		
Estimator	IV-2SLS	IV-2SLS	IV-GMM	IV-GMM	IV-2SLS	IV-2SLS
Wald F-stat	19E+04	19E+04	19E+04	19E+04	4589.04	4603.06
Observations	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442
R-squared	0.275	0.275	0.274	0.274	0.272	0.271
Number of i	74,967	74,967	74,967	74,967	74,967	74,967

Note: HAC stands for heteroskedasticity-autocorrelation (HAC) robust, and Clu(i,t) stands for 2-way clustered standard error (Cameron et al. 2006, Thompson 2009) that are robust to arbitrary heteroskedasticity and intra-group correlation with respect to both time and individual dimensions.

Table 7: Second Stage With 6-Month Birthday IV

VARIABLES	(1) $ADOPT_{it}$	(2) $ADOPT_{it}$	(3) $ADOPT_{it}$	(4) $ADOPT_{it}$	(5) $ADOPT_{it}$	(6) $ADOPT_{it}$
INSTALLBASE_IN6	0.00970***		0.0101***		0.0131***	
	(0.000418)		(0.00338)		(0.000540)	
INSTALLBASE_OUT6		0.00967***		0.0101***		0.0131***
		(0.000416)		(0.00337)		(0.000539)
INSTALLBASE_IN6^2					-7.13e-05***	
					(3.98e-06)	
INSTALLBASE_OUT6^2						-7.17e-05***
						(3.97e-06)
Month FE	Y	Y	Y	Y	Y	Y
Indiv FE	Y	Y	Y	Y	Y	Y
Error	OLS	OLS	HAC/Clu(i.t)	HAC/Clu(i.t)	OLS	OLS
Estimator	IV-2SLS	IV-2SLS	IV-GMM	IV-GMM	IV-2SLS	IV-2SLS
Observations	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442
R-squared	0.275	0.275	0.275	0.275	0.276	0.276
Number of i	74,967	74,967	74,967	74,967	74,967	74,967

Note: The above table tests a modification of equation 6. The regressions are similar to those in table 6, except that the birthday IV used here is a summation over an individual's friends' birthday over the past six months.

impact on the speed and pattern of diffusion (e.g. Nair et al. 2010, Banerjee et al. 2013). In this section, we investigate whether these more "important" agents would have greater impacts on the adoption decisions of their neighbours.

We use standard network indices, including inward and outward degrees and tie strength to measure importance of a friend to the focal agent (Banerjee et al. 2013, Tucker 2008). Degree index approximates the popularity of an individual; while tie strength, measured by total duration of phone calls, evaluates the extent of the friendship between a pair of contacts[14].

Figure 3 gives a histogram of the logarithmic degree of iPhone adopters on our network. It is worth noting that our data includes all iPhone users and their friends, but not necessarily their friends' friends. In other words, it enables us to calculate degree of iPhone adopters precisely but not that of the non-users, the latter of which, luckily, is not among what we need for the empirical purpose of this paper[15]. Figure 4 shows the histogram of logarithmic total call duration over the six months among all pairs of contacts in the network. Again, Figure 4 leaves out calls among non-users, which we do not need.

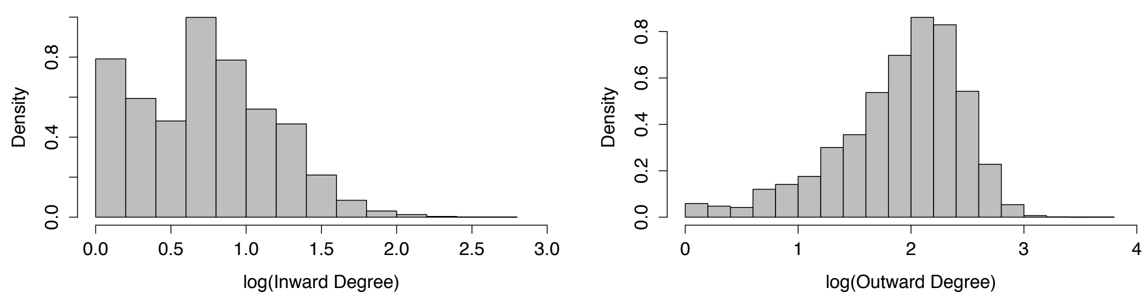

Figure 3: Histogram of degrees of iPhone adopters on the social network.

To identify the heterogenous peer effects, we partition each individual's friends into groups according to their quantiles ranking in the distribution of individuals' degree and distribution of call durations. As shown by Table 8, both distributions of inward degrees and call durations have fat right tail, so we partition individuals into groups of 0-25%, 25-50%, 50-75%, 75-90%, 90-95%, 95-99%, and 99-100%, so as to get a more evenly spaced distribution of degrees and tie strength. Then we regress

[14] Ideally, we would like to explore more network-based indices, such as betweenness, closeness, and centrality, among others (Banerjee et al. 2013; Tucker 2008). However, our data set includes only iPhone users and their friends, but not their friends' friends. In other words, our network is not complete in a way that would make the other network indices precisely calculable. Hence, we can only include degree and tie strength in the current paper.

[15] This is because the non-adopters have $ADOPT_{it}$ always equals to zero. Hence, leaving out the degree of non-adopters does not impact our empirical analysis.

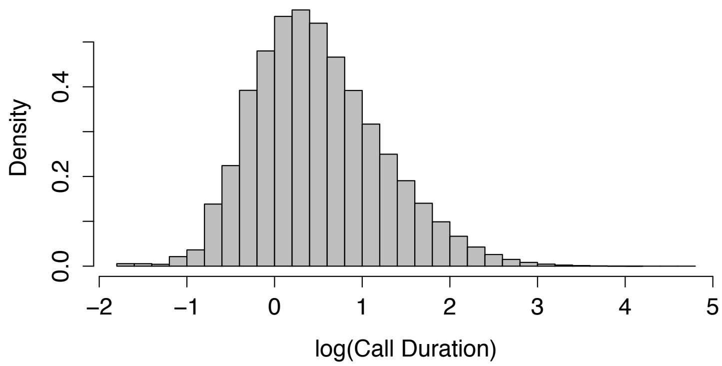

Figure 4: Histogram of six months' total call duration for pairs of contacts.

an individual's number of friend adopters from each of the quantile groups on his adoption decision separately, and the results are given in table 9.

Table 8: Quantiles of Inward Degrees and Call Durations

Probability	0	0.25	0.5	0.75	0.9	0.95	0.99	1
Inward Degrees	0	1	4	9	18	27	56	454
Call Durations	0.02	0.95	2.63	8.94	31.12	67.12	282.72	48303.92

Panel A investigates the "intimacy effect", i.e. whether friends with a stronger tie with the focal individual tend to have a bigger or smaller peer effect. Theories have not yet agreed on the relative magnitudes of peer effect from a good friend compared to that of a casual acquaintance. Some believe that individuals are more readily to be influenced by their close friends, while others argue that information from a "weak tie" contact might prove to be more useful (Granovetter, 1973). In our setting of iPhone diffusion, individual adoption decisions are more affected by strong-tie contacts (friends that they communicated more with). As we can see, having a friend adopter from the group of weakest tie (0-25%) increases an individuals subsequent adoption probability by 1.9%, whereas having a friend adopter from the group of strongest tie (99-100%) increases adoption probability by 9.9%, as estimated by an FE model; and 0.9% compared to 12% as estimated by an IV model.

Panel B investigates the "popularity effect". Interestingly, contrary to what usually found in the literature, a friend with more friends of his own tend to have a smaller peer effect in our setting, though not quite significant once he reaches a certain degree of moderate popularity. This might due to the fact that intimacy effect dominates in our setting; and popular individuals usually spend less time with each one of their friends.

5.3 Absolute Number of Adopters vs. Fraction of Adopters

In previous sections, we use absolute number of adopters as the explanatory variable. As mentioned earlier, there are two variations of the linear-in-mean model, and in this section, we present results as a robustness check using the ratio of adopters as the explanatory variable.

Variable *FRACTION_IN* is defined as total number of adopters among one's friends *INSTALLBASE_IN* divided by his total number of friends *DEGREE_IN*. The subfix *IN* indicates that these variables are defined over the inward-phone-call network, and similar definition holds for variables with subfix *OUT*. Using the adopter ratio *FRACTION* instead of adopter number *INSTALLBASE* as the explanatory variable, column (1) and (2) in table 10 give the results for the FE model as in equation (1), while column (3) shows that for IV model as in equation (5). As we can see, the IV model gives an estimate of peer effect at about 52.8% (for inward friends), meaning that having an extra one tenth of one's friends adopt iPhone would increase his probability of adoption by about 5.3%.

Overall, however, the results using *FRACTION* are much less robust than that using *INSTALLBASE*. This is mainly due to the small fraction of adopters compare to the large number of contacts on the network. This is especially true for outward friends, since our population of phone call receivers includes all land-lines and users of other mobile phone carriers that China Unicom users ever called. Those users could never adopt iPhone (unless change carrier to China Unicom first); and their existence in the network greatly decreases the identifying variation in the explanatory variable *FRACTION*, a point that is shown clearly in the table of panel data summary statistics A.1.

5.4 Exogenous Network Formation

In this section, we provide evidence showing that the network structure has been stable over the sample period.

We investigate the distribution of standard network structure indices, such as degree and tie strength (measured by monthly call duration), and see whether these indices trend with the phase of iPhone diffusion. Box plots of the monthly distributions of four individual-level network indices: in-degree, out-degree, in-call-duration, and out-call-duration, are provided in figure 5. In-degree and in-call-duration are for callers (iPhone adopters), while out-degree and out-call-duration includes anyone that has ever received phone calls on our network, including landlines and non-China Unicom users. As we can see, the distributions of the four indices remained stable throughout our sample period. The diffusion of iPhone did not bring any detectable change to our phone call network, which seems to confirm our exogeneity assumption

Table 9: Peer Effect with Heterogeneity in Degrees and Tie Strength

	Fixed Effect Model							*Instrumental Variable Model*						
VAR	(1) $ADOPT_{it}$	(2) $ADOPT_{it}$	(3) $ADOPT_{it}$	(4) $ADOPT_{it}$	(5) $ADOPT_{it}$	(6) $ADOPT_{it}$	(7) $ADOPT_{it}$	(8) $ADOPT_{it}$	(9) $ADOPT_{it}$	(10) $ADOPT_{it}$	(11) $ADOPT_{it}$	(12) $ADOPT_{it}$	(13) $ADOPT_{it}$	(14) $ADOPT_{it}$
Panel A: "Intimacy Effect" Heterogeneity in the Tie Strength, measured by duration of call														
INST_TIE	0.019***							0.009***						
(0-25%)	(0.00218)							(0.00046)						
INST_TIE		0.021***							0.010***					
(25-50%)		(0.00278)							(0.00045)					
INST_TIE			0.024***							0.015***				
(50-75%)			(0.00131)							(0.00045)				
INST_TIE				0.035***							0.029***			
(75-90%)				(0.00083)							(0.00074)			
INST_TIE					0.055***							0.095***		
(90-95%)					(0.0014)							(0.0023)		
INST_TIE						0.059***							0.12***	
(95-99%)						(0.0015)							(0.0028)	
INST_TIE							0.099***							2.77***
(99-100%)							(0.0074)							(0.0636)
Constant	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***							
	(8.6e-05)	(8.6e-05)	(8.6e-05)	(8.6e-05)	(8.6e-05)	(8.6e-05)	(8.7e-05)							
Panel B: "Popularity Effect" Heterogeneity in the Friends' Degree														
INST_DEG	0.056***							0.099***						
(0-25%)	(0.0036)							(0.0069)						
INST_DEG		0.044***							0.082***					
(25-50%)		(0.0016)							(0.0018)					
INST_DEG			0.034***							0.051***				
(50-75%)			(0.00256)							(0.00069)				
INST_DEG				0.024***							0.035***			
(75-90%)				(0.0028)							(0.00039)			
INST_DEG					0.031***							0.047***		
(90-95%)					(0.0017)							(0.00054)		
INST_DEG						0.023***							0.029***	
(95-99%)						(0.00073)							(0.00035)	
INST_DEG							0.029***							0.035***
(99-100%)							(0.0023)							(0.0011)
Constant	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***							
	(8.6e-05)	(8.6e-05)	(8.6e-05)	(8.6e-05)	(8.6e-05)	(8.6e-05)	(8.7e-05)							
R^2	0.268	0.270	0.272	0.273	0.272	0.272	0.269	0.268	0.268	0.271	0.272	0.271	0.272	0.269
Month FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Indiv FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Estimation	OLS	OLS	OLS	OLS	OLS	OLS	OLS	IV	IV	IV	IV	IV	IV	IV
Obs	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442	3,100,442
R^2	0.271	0.272	0.273	0.273	0.271	0.272	0.268	0.270	0.271	0.272	0.273	0.270	0.268	-0.010
Num Indiv	74,967	74,967	74,967	74,967	74,967	74,967	74,967	74,967	74,967	74,967	74,967	74,967	74,967	74,967

Table 10: Peer Effect with Fraction of Adopters

VARIABLES	(1) $ADOPT_{it}$	(2) $ADOPT_{it}$	(3) $ADOPT_{it}$	(4) $ADOPT_{it}$	(5) $ADOPT_{it}$
$FRACTION\_IN$	0.0734***	0.0820***		0.528***	0.250***
	(0.000793)	(0.00155)		(0.102)	(0.0212)
$FRACTION\_IN$			0.121***		
			(0.00454)		
Constant	-0.00240***	1.053***	-0.00243***		
	(0.000540)	(0.00364)	(0.000486)		
Individual FE	Y	Y	Y	Y	Y
Monthly FE	Y	Y	Y	Y	Y
Estimator	FE	FE	FE	IV-2SLS	IV-2SLS
Sample Start	2010	2012	2010	2010	2012
Observations	2,501,265	920,066	3,099,985	2,501,265	919,709
R-squared	0.270	0.279	0.268	0.172	59,931
Number of i	60,933	60,288	74,957	60,933	0.269

Note: Variable $FRACTION$ is the ratio of adopters among one's friends, defined as the total number of adopted friends ($INSTALLBASE$) divided by the number of friends ($DEGREE$). The FE estimation above uses robust standard error.

of the network.

5.5 Non-linear Probability Model

As discussed in previous section, our data set suffers from a quasi separation problem when trying to fit it with a logit model. This is mainly due to the (near perfect) predictability of yearly effect on adoption decisions. In table (11), we provide results estimating logit model using yearly subsamples.

Table 11: Peer Effect with Logit Model By Year

VARIABLES	(1) ADOPT	(2) ADOPT	(3) ADOPT
INSTALLBASE_IN	0.685***	0.110***	0.0522***
	(0.0661)	(0.00214)	(0.00101)
Constant	-7.222***	-4.241***	-1.514***
	(0.0391)	(0.00918)	(0.00520)
Marginal effect	0.0005099	0.0016116	0.0082776
Sample	2011	2012	2013
Observations	895,993	848,833	306,343

Note: Each column in the above table uses adoption observations in the corresponding year only. Hence, it mitigates the quasi-separation problem due to (near perfect) predictability of yearly effect on adoption decisions.

The logit result mainly confirms the linear probability model. The logit peer effect is positive and increasing over time, both in terms of its significance and magnitude. Yet, one should interpret the logit results with caution since the logit model

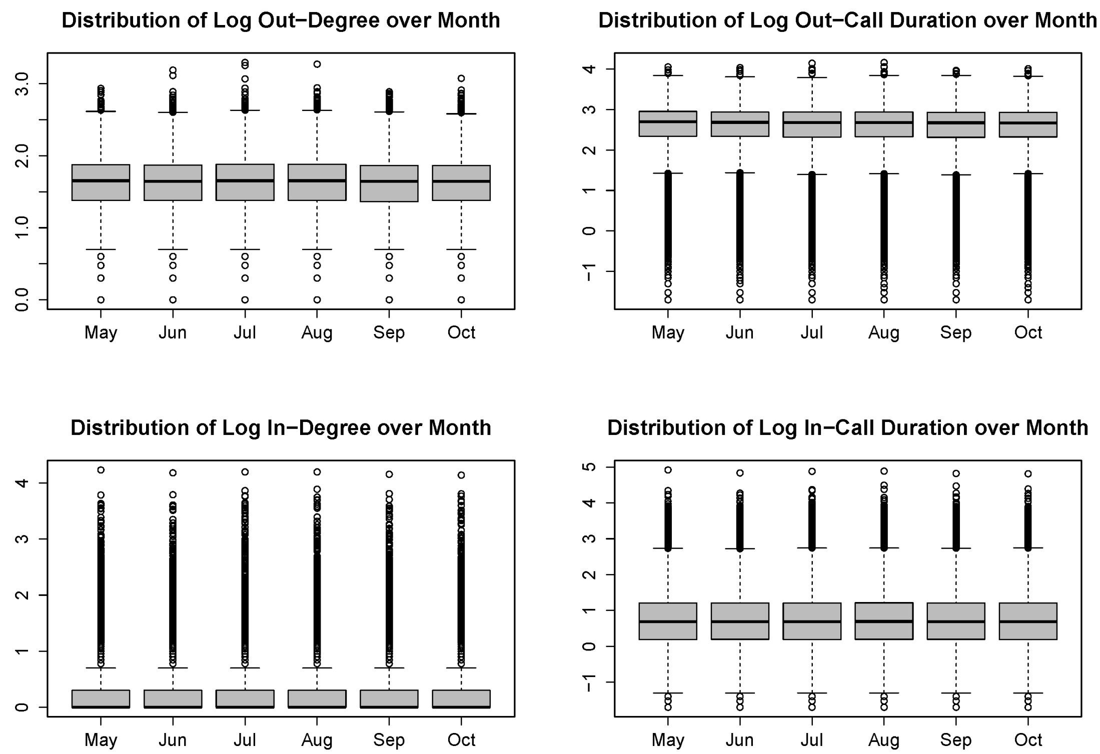

Figure 5: Evolution of Network Structure over Time.

is estimated without controlling for individual fixed effect[16].

6 Conclusion

A Peer effect occurs when the action of one agent directly affects its peers' choices outside the market channel. Understanding peer influence is critical to estimating product demand and diffusion, creating effective viral marketing, and designing "network interventions" to promote positive social changes.

In this paper we study whether the adoption of a consumer technology, in our case an iPhone, is affected causally by his network neighbours' decisions and network characteristics of the other adopters. The empirical setting is to measure the peer effect of iPhone adoption in a provincial capital city in China, during a four-year period starting from the introduction of the first iPhones to Mainland China. We use a unique panel dataset of phone call records by person and by time, that allows us to construct each iPhone adopter's social network, by using half a years call transactions between iPhone adopters and all other users on a carriers network. We measure strength of a social network pairwise tie by the duration of calls. Based on the network structure, We quantify the peer effect of iPhone adoptions, and investigate how the network structure modulates the magnitude of peer influence.

The main specification to identify peer effects is to see how the probability of an individual adopting an iPhone is affected by the measures of networks we create. Of course network size and strength is not randomly assigned. Identification of peer effects, therefore, is a challenge. Peer effect implies that the behavior of connected agents on a network tends to be correlated. However, other factors besides peer effect could also give rise to such behavioral correlation. For example, adoption decisions of ones neighbors can be endogenous for his adoption decision, because people who know each other tend to face similar unobserved environment to adopt the technology.

The identification in our paper has two approaches. To control for time-invariant correlated unobservables, we apply a fixed-effect model, and shows that a friend's adoption increases one's adoption probability in next month by 0.89%. To further control for potentially time-varying unobservables, we instrument adoptions of one's friends by their birthdays, based on the fact that consumers are more likely to adopt iPhones on birthdays. The IV estimation shows a slightly bigger peer effect at 0.94%, after clearing away impacts of potential correlated unobservables. Both models show that the marginal effect of peer influence decreases in the number of current peer

[16]The standard ways to incorporate fixed effect with logit model do not seem to be suitable for our setting. To incorporate FE, there are two standard solutions: random effect, which assumes individual FE exogenous to explanatory variable, and conditional FE, which assumes $\sum_{t=1}^{T} y_{it} = 1$. The first method suffers from endogeneity issue due to homophily while the second does not fit our unbalanced panel.

adopters. In other words, as more friends have already adopted, the marginal impact of an additional friend becomes smaller.

We also investigate how heterogeneity in network structure impacts the magnitude of peer influence. Our results show both a "popularity" and an "intimacy" effect. It is shown that the "popularity" of an individual, as measured by the number of his first-degree contacts, will affect how much influence he can exert on his fellow peers. The higher an individual's degree is, the smaller his peer effect would be on his neighbours. The peer effect is also stronger between "closer" friends. The more time a pair of friends spent on talking to each other during the six months, the greater the peer effect is between them.

Our results could provide useful insights for managers. We studied the diffusion of a mainstream consumer product, the iPhone. The empirical setting is based in a provincial capital city in China, Xining. The city, with a population of 2.3 million and GDP per capita at $6999 in 2013, is a good representation of a typical city in China as well as that of a mid-level developing country. Business practitioners launching promotions for similar products might find our results useful in designing optimal marketing strategies in regions comparable to ours.

References

Adelin Albert and JA Anderson. On the existence of maximum likelihood estimates in logistic regression models. *Biometrika*, 71(1):1–10, 1984.

Sinan Aral and Dylan Walker. Creating social contagion through viral product design: A randomized trial of peer influence in networks. *Management Science*, 57(9):1623–1639, 2011.

Sinan Aral and Dylan Walker. Tie strength, embeddedness, and social influence: A large-scale networked experiment. *Management Science*, 2014.

Angelique Augereau, Shane Greenstein, and Marc Rysman. Coordination versus differentiation in a standards war: 56k modems. *The RAND Journal of Economics*, 37(4):887–909, 2006.

Oriana Bandiera and Imran Rasul. Social networks and technology adoption in northern mozambique. *The Economic Journal*, 116(514):869–902, 2006.

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

F Bass. A new product growth model for consumer durables. *Management Science*, 1969.

Patrick Bayer, Randi Hjalmarsson, and David Pozen. Building criminal capital behind bars: Peer effects in juvenile corrections. *The Quarterly Journal of Economics*, 124 (1):105–147, 2009.

David R Bell and Sangyoung Song. Neighborhood effects and trial on the internet: Evidence from online grocery retailing. *Quantitative Marketing and Economics*, 5 (4):361–400, 2007.

Marianne Bertrand, Erzo FP Luttmer, and Sendhil Mullainathan. Network effects and welfare cultures. *The Quarterly Journal of Economics*, 115(3):1019–1055, 2000.

Daniel Birke. The economics of networks: a survey of the empirical literature. *Journal of Economic Surveys*, 23(4):762–793, 2009.

Bryan Bollinger and Kenneth Gillingham. Peer effects in the diffusion of solar photovoltaic panels. *Marketing Science*, 31(6):900–912, 2012.

Jennifer Brown and John Morgan. How much is a dollar worth? tipping versus equilibrium coexistence on competing online auction sites. *Journal of Political Economy*, 117(4):668–700, 2009.

Leonardo Bursztyn, Florian Ederer, Bruno Ferman, and Noam Yuchtman. Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *Econometrica*, 82(4):1273–1301, 2014.

Jing Cai, Alain De Janvry, and Elisabeth Sadoulet. Social networks and the decision to insure. *American Economic Journal: Applied Economics*, 7(2):81–108, 2015.

Estelle Cantillon and Pai-Ling Yin. Competition between exchanges: Lessons from the battle of the bund. 2008.

Timothy G Conley and Christopher R Udry. Learning about a new technology: Pineapple in ghana. *The American Economic Review*, pages 35–69, 2010.

Paul A David. Clio and the economics of qwerty. *The American economic review*, pages 332–337, 1985.

Esther Duflo and Emmanuel Saez. The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *The Quarterly Journal of Economics*, 118(3):815–842, 2003.

Dennis Epple and Richard Romano. Peer effects in education: A survey of the theory and evidence. *Handbook of social economics*, 1(11):1053–1163, 2011.

David Firth. Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1): 27–38, 1993.

Pedro M Gardete. Social effects in the in-flight marketplace: Characterization and managerial implications. *Journal of Marketing Research*, 52(3):360–374, 2015.

Edward L Glaeser, Bruce Sacerdote, and José A Scheinkman. Crime and social interactions. *The Quarterly Journal of Economics*, 111(2):507–548, 1996.

Miguel Godinho de Matos, Pedro Ferreira, and David Krackhardt. Peer influence in the diffusion of the iphone 3g over a large social network. *Management Information Systems Quarterly (Forthcoming)*, 38(4):1103–1133, 2014.

Austan Goolsbee and Peter J Klenow. Evidence on learning and network externalities in the diffusion of home computers. *Journal of Law and Economics*, 45(2):317–343, 2002.

Mark S Granovetter. The strength of weak ties. *American journal of sociology*, pages 1360–1380, 1973.

Zvi Griliches. Hybrid corn: An exploration in the economics of technological change. *Econometrica, Journal of the Econometric Society*, pages 501–522, 1957.

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

Georg Heinze and Michael Schemper. A solution to the problem of separation in logistic regression. *Statistics in medicine*, 21(16):2409–2419, 2002.

Igal Hendel, Aviv Nevo, and François Ortalo-Magné. The relative performance of real estate marketing platforms: Mls versus fsbomadison. com. *American Economic Review*, 99(5):1878–98, 2009.

Shawndra Hill, Foster Provost, Chris Volinsky, et al. Network-based marketing: Identifying likely adopters via consumer networks. *Statistical Science*, 21(2):256–276, 2006.

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

Zsolt Katona, Peter Pal Zubcsek, and Miklos Sarvary. Network effects and personal influences: The diffusion of an online social network. *Journal of Marketing Research*, 48(3):425–443, 2011.

Vijay Mahajan, Eitan Muller, and Frank M Bass. New product diffusion models in marketing: A review and directions for research. *The Journal of Marketing*, pages 1–26, 1990.

Puneet Manchanda, Ying Xie, and Nara Youn. The role of targeted communication and contagion in product adoption. *Marketing Science*, 27(6):961–976, 2008.

Charles F Manski. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531–542, 1993.

Robert A Moffitt et al. Policy interventions, low-level equilibria, and social interactions. *Social dynamics*, 4:45–82, 2001.

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

Sungjoon Nam, Puneet Manchanda, and Pradeep K Chintagunta. The effect of signal quality and contiguous word of mouth on customer acquisition for a video-on-demand service. *Marketing Science*, 29(4):690–700, 2010.

Sridhar Narayanan and Harikesh S Nair. Estimating causal installed-base effects: A bias-correction approach. *Journal of Marketing Research*, 50(1):70–94, 2013.

Everett M Rogers. *Diffusion of innovations*. Free Press, New York, 1962.

Bruce Sacerdote. Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly Journal of Economics*, 116(2):681–704, 2001.

Scott K Shriver, Harikesh S Nair, and Reto Hofstetter. Social ties and user-generated content: Evidence from an online social network. *Management Science*, 59(6): 1425–1443, 2013.

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

Appendix

Summary Statistics for the Panel

The following table A.1 provides summary statistics for the panel used in section 4 and 5, combining both phone-call network and adoption data set.

Table A.1: Summary Statistics for the Panel (TBD)

	Observation	Mean	Std Dev	p25	p50	p75
i (individual)	74,967					
t (month)	2009/11 – 2013/10					
Panel A: Basic Panel						
ADOPT	3,100,442	0.02	0.15	0	0	0
INSTALLBASE_IN	3,100,442	0.35	1.80	0	0	0
INSTALLBASE_OUT	3,100,442	0.35	1.81	0	0	0
$INSTALLBASE\_IN^2$	3,100,442	3.37	168.39	0	0	0
$INSTALLBASE\_OUT^2$	3,100,442	3.39	168.84	0	0	0
Panel B: IV Panel for Section 4.3						
IV_BDAY_IN	3,100,442	7.91	16.24	0	3	9
IV_BDAY_OUT	3,100,442	7.91	16.25	0	3	9
$IV\_BDAY\_IN^2$	3,100,442	326.17	3,932.44	0	9	81
$IV\_BDAY\_OUT^2$	3,100,442	326.68	3,939.51	0	9	81
Panel C: Adoptor Fraction Panel for Section 5.3						
FRACTION_IN	2,501,265	0.06	0.18	0	0	0
FRACTION_OUT	3,099,985	0.00	0.02	0	0	0

Note: The panel consists of (almost) all iPhone adoptions in the city of Xining from Nov-2009 to Oct-2013. The social network is constructed from call transactions between May-2013 and Nov-2013.