Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Understanding the Social Learning Effect in Contagious Switching Behavior

Mandy Mantian Hu, Sha Yang, Daniel Yi Xu

informs
http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Vol. 65, No. 10, October 2019, pp. 4771–4794
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Understanding the Social Learning Effect in Contagious Switching Behavior

Mandy Mantian Hu,[a] Sha Yang,[b] Daniel Yi Xu[c]

[a] Chinese University of Hong Kong Business School, Shatin, Hong Kong; [b] Marshall School of Business, University of Southern California, Los Angeles, California 90089; [c] Department of Economics, Duke University, Durham, North Carolina 27708

Contact: mandyhu@baf.cuhk.edu.hk, http://orcid.org/0000-0001-6168-4600 (MMH); shayang@marshall.usc.edu, http://orcid.org/0000-0003-4190-5722 (SY); daniel.xu@duke.edu (DYX)

Received: August 26, 2015
Revised: April 15, 2017; April 9, 2018; July 4, 2018
Accepted: July 5, 2018
Published Online in Articles in Advance: July 19, 2019

https://doi.org/10.1287/mnsc.2018.3173

Abstract. We study the contagious switching behavior related to a consumer's choice of wireless carriers, that is, that a consumer is more likely to switch wireless carriers if more of their contacts from the same carrier have switched. Contagious switching (or a positive network effect) can be driven by information-based social learning, as well as other mechanisms related to network size. Although previous marketing literature has documented the social-learning effect, most of the applications studied involve products in which consumers usually do not enjoy any direct benefits from a large network other than from information-based social learning. We explore the importance of the social-learning effect relative to other mechanisms that may also lead to the network effect. We propose a dynamic structural model with interpersonal interactions. To model the social-learning effect, a consumer uses feedback from his or her contacts who have switched from a focal carrier to update his or her quality expectations of alternative carriers. Our model further accounts for two unique aspects of consumer strategic learning: (i) the individual's perception on the signal of alternative carriers from contacts who switch is systematically different according to whether the signal comes from a loyal contact; and (ii) that the perceived noisiness of the signal on alternative carriers from a contact who has switched depends on the strength of the relationship between the individual and the contact. The remaining network effect not captured through social learning is modeled as a function of the size of the network. We solve the model with a two-step dynamic programming algorithm, with the assumption that a consumer is forward-looking and decides whether to stay with the same service carrier in each period by maximizing the total utility received from that day onward. We apply the proposed model to the data set of a mobile network operator in a European country. We find that churning/switching behavior is contagious in the network context and that one-third of general network effects can be attributed to social learning. We also detect strategic learning by consumers from their contacts in two ways: the experience signal on alternative carriers from a more loyal contact who has switched from the focal carrier is perceived to be more positive than that from a less loyal contact; and the social-learning effect is stronger from an individual's closest contacts. The simulation analysis demonstrates the value of our model in helping a company prioritize its customer relationship management effort.

History: Accepted by Matthew Shum, marketing.
Funding: The project was supported by Hong Kong General Research Fund [Grant 14521716].

Keywords: social learning • network effect • contagious switching behavior • customer relationship management

1. Introduction

When people are connected by a network, they inevitably influence each other. The network effect, referring to particular forms of externalities such that actions of peers affect an individual's preference (Katz and Shapiro 1985), has attracted great interest both in practice and research (e.g., Tucker 2008, Goldenberg et al. 2010, Ryan and Tucker 2012). Social learning is one of the mechanisms by which network effect influences consumer behavior and is defined as the process by which individuals use feedback from their peers to update their own expectations of product quality (Moretti 2011). Social learning encompasses a wide range of activities,[1] including "word-of-mouth" learning (Narayan et al. 2011, Chandrasekhar et al. 2012, Ma et al. 2014), influence from expert opinions in movie reviews, magazine articles, and news coverage (Erdem et al. 2005, 2008; Chintagunta et al. 2009), reading online reviews from other consumers (Zhao et al. 2013), observational learning (Cai et al. 2009, Zhang 2010), and peer-based learning (Chan et al. 2014a, b). In addition to social learning, the network effect also manifests through other mechanisms (e.g., Goolsbee and Klenow 2002, Moretti 2011). For example, it could be due to "conformism," because people tend to behave like the peers in their group, to earn a potential social benefit.

Unlike typical consumer products, many high-tech products (such as mobile service) enjoy a strong positive network effect; that is, the utility from the product or service increases with the number of others in a network also using it. The network externality in such contexts leads to one important behavioral pattern: contagious switching behavior (i.e., contacts in one's network that leave will positively affect an individual customer's switching probability). Contagious switching behavior related to consumer choice of wireless carriers has attracted a great deal of attention in the telecommunications industry in recent years (e.g., Richter et al. 2010). According to a comprehensive 2007 survey conducted by ComScore Networks on the behavior and attitudes of U.S. wireless phone subscribers, 13% of respondents considered the influence of friends and family to be the primary reason for switching carriers, which was nearly as important as the second-most-important reason: lower prices (14%). The literature in multiple fields provides many examples that the behavior of other consumers influences an individual's decisions on that same behavior (e.g., Christakis and Fowler 2007, Dasgupta et al. 2008, Aral and Walker 2014). In marketing, such contagious behavior has been well-documented in the context of new product and technology adoption (e.g., Manchanda et al. 2008; Iyengar et al. 2011, 2015).

Although previous marketing literature has documented the social-learning effect, most applications studied involve consumer product goods, movies, and prescribed medicines, for which consumers do not usually enjoy any direct benefits from a large network (other than from information-based social learning). The importance of the social-learning effect relative to other mechanisms that may also lead to the general network effect has not been explored. The main research objective of this study is to propose a modeling framework to understand the social-learning effect within a network and separate it out from the general network effect.

Quantifying and understanding the social-learning effect within a network is important from a marketing perspective. Unlike the network-size effect, which is agnostic about the particular characteristics of network influence, social learning is shaped by characteristics of switching contacts and the strength of the relationship between the two involved parties. As such, insights from the social-learning effect are more actionable to marketers. For example, given the same network size, a focal carrier can predict which customers are more prone to churning due to social learning about alternative carriers a customer gains from their network contacts that have already switched. Therefore, priority of their customer relationship management (CRM) efforts can be given to those customers who have a higher number of closely connected switching contacts. From a competitor's perspective, marketing efforts should be targeted toward the influential customers whose experience signals are viewed to be more credible. A competitor can offer referral programs to incentivize information sharing and acquire more customers.

Our empirical context pertains to an individual's stay/switch decision with one focal wireless carrier. Individual consumers in our model are assumed to be forward-looking and to make the stay/switch decision by maximizing their expected future utility from alternative decision outcomes. After controlling for plan details and individual characteristics, we focus our attention on two factors that explain the stay/switch decision of individual consumers. First is the learning-from-self mechanism. The service quality of the current carrier directly affects the consumer's utility of staying (Iyengar et al. 2007). Because of service variability, a consumer faces uncertainty and learns about the mean service quality of the carrier to whom he or she subscribes on the basis of his or her direct usage experience.

The second is the general network effect, which represents the unique characteristic of mobile service. We deconstruct the network-related utility into two components. The first component corresponds to the utility from service quality, which is affected by an individual's network (i.e., the switching-out contacts) via the social-learning mechanism outlined above. In this paper, we posit that consumers learn about the quality of alternative carriers from those contacts who have switched from the focal carrier, because obtaining direct experience by subscribing to alternative carriers themselves is costly. Their information comes from direct communication with those in their contact network who have already switched. The quality expectation of an alternative carrier will affect a customer's switching probability. More importantly, our model accounts for two unique aspects of consumers' strategic learning from others: (i) the signal on an alternative carrier from a contact who has switched is perceived to be systematically different by an individual according to whether the signal comes from a loyal contact; and (ii) the noisiness of the signal on an alternative carrier from a contact who has switched is dependent on the strength of the relationship between the individual and the contact.

The second component corresponds to the remaining network effect *not* captured through social learning and is modeled as a function of the network-size variable Net. This remaining network effect captures benefits other than information-based social learning from more contacts using the same carrier. It may arise from customer cost considerations. For example, some wireless carriers offer plans with cheaper calling rates to people on the same network. If a consumer's friends switch to other networks, higher costs may be incurred if the focal customer does not follow. It could also be due to "conformism." The strength of this effect is usually modeled as a function of the number of people who make the same choice within a group (i.e., network size). Compared with previous marketing literature on social learning that focuses on consumer product goods or prescribed medicine (e.g., Erdem et al. 2008, Chan et al. 2013), this second component captures a unique feature of high-tech products or services: the more people that use a product, the higher the direct utility from using it. Adding this second component of network-related utility also complicates the modeling and estimation.

The proprietary data set we use provides the opportunity to separate out the social-learning effect of the first component from the general network effect. We rely on the time variation in number of switching contacts, loyalty of each switching contact, and relationship strength between the focal individual and each switching contact to capture the social-learning effects. Chan et al. (2014a) used variation in a pool of peers to study peer-based learning for cosmetics salespeople. Chan et al. (2014b) used coworker variation to identify the short-term impact of peers on individual productivity under different compensation systems. Snijders (2001) developed the SEINA program to use panel information on networks and behavior to differentiate social influence and nonlatent homophily. This approach has been applied to different settings in sociology studies (e.g., Lewis et al. 2012, Brenton 2016) and psychology research (e.g., van Zalk and Denissen 2015). To identify the network-size effect of the second component, we rely on the cross-sectional variation in the initial network size across individuals. In other words, the effect of the size of the network focuses on the aggregate-level effect from size rather than the specific nature of interaction between a focal individual and their network.

Our data came from a European third-party research firm that specializes in the telecommunications industry. The data period spanned from June 2008 to February 2009, and we used individuals' calling behavior to form networks. Eagle et al. (2009) compared mobile network data and self-reported survey data and found that these two sets of data may be distinct from each other but that the data collected from a mobile network can be used to accurately predict cognitive constructs such as friendship. This finding suggests that mobile network information will be useful in studying the influence of others on an individual's stay/switch decision with respect to subscription to a wireless carrier. We used snowball sampling to extract individual networks, and 2,077 individuals from 198 individual networks were used for this study.

Our empirical findings show that (i) churning/switching behavior is contagious in the network context; (ii) one-third of the network effect is attributed to social learning, and while the general effect of the switch of an additional contact is decreasing, the effect of social learning from an additional switcher will not diminish quickly; and (iii) more importantly, the magnitude of social learning is affected by the characteristics and strength of the relationship of the parties involved. Friends are not all equal. Using simple measures of contact frequency as the proxy for the closeness of friends, we are able to identify whose decision might weigh more for an existing customer. These findings lead to important managerial implications on CRM.

We make several contributions to the existing research. First, we separate out the social-learning effect from the general network effect. The social-learning effect represents the information exchange argument of contagious switching behavior and is difficult to quantify in a network setting. We take advantage of several unique features of the data set and fill a gap in the literature by modeling social learning as well as the remaining network effect. We demonstrate that one-third of the general network effect can be attributed to social learning.

Second, we propose a dynamic structural model with strategic interpersonal interactions. Our proposed modeling framework is a complex setting that involves dynamic choice decisions with consumer learning and strategic interactions between multiple players (i.e., an individual consumer and his or her contacts). We extend the learning model by accounting for strategic learning behavior as reflected in the way that an individual interprets signals from others depending on the others' loyalty level and strength of relationship. This extension also accounts for the individual heterogeneity of contacts and the dyadic heterogeneity between an individual and his or her contacts.

Third, we estimate the dynamic forward-looking model using a modified two-step procedure. In this revised two-step procedure, we first parametrically estimate the network's change. This step separates our dynamic model from the static model of contagious behavior, because the variable of network size does not enter our model directly. This alleviates the homophily concern that usually arises in such context. In the second step, we take the estimated network size as a given and estimate the structural parameters using a variation of the technique used by Keane and Wolpin (1997) (Crawford and Shum 2005). The current approach is a variation of the two-step algorithm (BBL) from Bajari et al. (2007). We cannot directly implement the BBL algorithm because the state variables of service quality are not observed, which prevents us from estimating the transition probabilities and policy functions for each individual. To accommodate this complexity, instead of estimating *who* switches in the first step, we estimate *how many* switch. A comparison between the empirical transition matrix of the network from the data and that based on the simulation using model estimates demonstrates the validity of such an approach in our context.

The remainder of the paper is organized as follows. Section 2 describes the data and survey results as the motivation for the model. Section 3 develops a dynamic model with a forward-looking agent that incorporates learning and network effects. Section 4 discusses the identification. Section 5 presents the estimation procedure and the results, followed by robustness checks in Section 6. The policy simulations are presented in Section 7. Section 8 concludes the paper.

2. Empirical Context

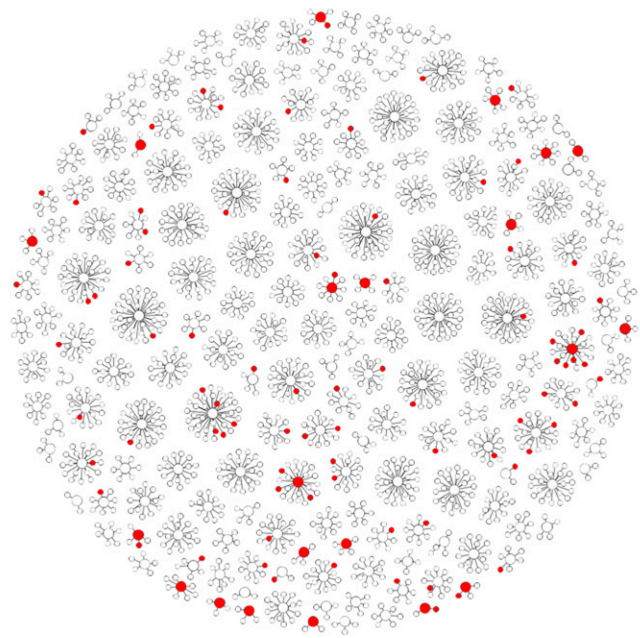
2.1. Data

Data were obtained from a third-party European research firm that specializes in the multinational telecommunication industry. We have two specific datasets. The first one includes calling record data and includes the aggregated information of each individual's mobile service usage (including the number of phone calls and text messages sent and received and their contacts' IDs) over a short period of time at the beginning of the sample period. We use this data set to identify each customer's initial network. The second data set contains the monthly information from June 2008 to February 2009 for each individual subscriber, related to demographics, plan details (postpaid/prepaid, on-network/off-network cost ratio, minutes included, and family plan ID), and churn dates.

Snowball sampling was used to generate a sample of data for this study. The sample was constructed as follows: 20 individuals ("seeds") were randomly selected from the entire mobile network. On the basis of these seeds, we identified everyone they contacted by either call or text message over a short period of time at the beginning of the sample period. We considered those direct contacts to be the neighbors or "friends" of the seeds. Together, they formed a sphere of a network group. In total, we have 198 networks that include 2,077 individuals. We plot out the data in Figure 1.

The network data in our sample have two features. First, our network data are ego-centric. This means that we have information regarding the focal person's friends and the friends' characteristics; however, we do not have information about the focal person's friends' network. As such, only the leave/stay decisions of focal customers (i.e., the one in the center of each small network in the plot) are studied. Using surveys to collect self-reported network information typically results in ego-centric networks, and this type of network has been widely used to study behaviors within a network (e.g., Nair et al. 2010; Iyengar et al. 2011, 2015; Lewis et al. 2012). Second, networks are dynamic, that is, we observe how the networks change over time.

Figure 1. (Color online) Plot of the Ego-Centric Data



Notes. Larger dots represent the focal customers who we have network information for; smaller dots represent contact friends of each focal customer. Black (red online) dots represent the customers who have switched carriers during the sample period. We model each focal customer's stay/leave decision with a mobile service provider in each time period.

The social-learning effect can therefore be inferred from the dynamic change in the focal person's switching decision, as a function of the variation in the switched contact friends. This variation comes from the number of switched friends, and their characteristics (i.e., intensity and loyalty) that we observe in the data provide additional variation. Chan et al. (2014a) used variation in a pool of peers to study peer-based learning for cosmetics salespeople. Chan et al. (2014b) used coworker variation to identify the short-term impact of peers on individual productivity under different compensation systems. Snijders (2001) developed the SEINA program to use panel information on networks and behavior to differentiate social influence and nonlatent homophily. This approach has been applied to different settings in sociology studies (e.g., Lewis et al. 2012, Brenton 2016) and psychology research (e.g., van Zalk and Denissen 2015).

Table 1 provides the summary statistics. "Minutes" indicates the minutes included in the plan. "Relative price" is the ratio between the charges to call someone within the focal carrier's network and someone outside. The value equals 1 if there is no difference, and the smaller the number the larger the difference. "Family plan" is a binary variable in which 1 indicates that the customer is subscribed to a family plan and 0 otherwise. On average, each individual had a network size of 10 people. "Average usage" is the average mobile

Table 1. Summary Statistics

Variable	Mean	Standard deviation	Minimum	Maximum
Male	0.36	0.48	0	1
Age	34.50	7.51	20	57
Minutes	460.05	194.24	15	1,200
Relative price	0.89	0.18	0.23	1.00
Family plan	0.46	0.50	0	1
Network size	10.49	7.55	2	35
Average usage	26.32	28.10	1.67	211.86
Churn	0.11	0.31	0	1
Network male percent	0.32	0.22	0.00	1.00
Network age	34.18	3.17	26.67	45.00
Stay periods	7.49	1.66	0	8
Intensity	28.60	60.53	1	1,214
Loyalty	0.17	0.37	0	1

usage per contacts for each individual in the initial period. We use it to capture the individual tendency of using mobile service. “Intensity” captures the number of minutes of phone calls made by the focal and contacts in the initial period; and “loyalty” is an indicator of whether the contact has renewed his or her contract with focal carrier before or not. “Churn” is an indication of whether the focal customer switched carrier in the sample period or not. During the eight-month sample period, 11% of customers left their focal carrier.

2.2. Simple Data Analysis

To gain some initial understanding, we have taken advantage of the panel nature of our data and estimated a model with an individual fixed effect to control for individual heterogeneity (Chamberlain 1980). We use the fixed effect model to demonstrate that both network and social-learning effects exist and are separable from individual heterogeneity. The fixed effect model allows the focal individual’s staying/churning decision in each period to depend on his or her own experience, network size, cumulative number of friends who churned leading up to the period (a proxy for social learning), observed network-specific characteristics (such as average age, gender ratio, etc.), and an individual-specific fixed effect to control for individual heterogeneity. The results from this fixed-effect model are reported in Table 2. In the fixed effect model, the individual characteristics were dropped owing to the lack of variation over time. Here is the summary of our findings. The larger the network the more likely the person will stay. The effect of cumulative number of switched friends (as proxy for social learning) is significant and negative after controlling for the individual-specific fixed effect. This demonstrates that the social learning exerts a different effect from the network or the unobserved heterogeneity.

We have also done a correlation analysis. We find that the correlation between churn and the total number of switched contacts is 0.092 (p-value = 0.0022), providing a model-free evidence for social learning. We also find that the correlation between churn and initial network size is −0.200 (p-value = 0.0047), providing model-free evidence for the network size effect. We next show two three-dimensional figures based on our data to further illustrate the empirical identification of the two effects. In Figure 2, the x-axis measures the initial network sizes, the y-axis represents the total number of switched contacts over the sampling period, and the z-axis is the switching probability, conditioning on the initial network size and total switched contacts. Figure 2(a) indicates that given initial network size, the more contacts switch over time, the more likely the focal customer will switch carrier. For example, with initial network size of five, if one of the contacts switch, the switching probability of the focal customer is 16.7%. And if no one switches, the switching probability of the focal customer is 11.1%.

Figure 2(b) indicates that given the total number of contacts switch over time, the probability that the focal customer switches is negatively correlated with the initial network size. For example, for those people who have one contact switch in the sampling period, if the initial network size is four, the switching probability of the focal customer is 1.8%; and if the initial network size is two, the switching probability becomes 3.6%.

Figure 3 plots the average loyalty, interaction intensity, and number of switches per period. We find that they all show different variations. In addition, the correlation between intensity and the number of switches is 0.22 (p-value = 0.61), and the correlation between interaction intensity and loyalty is 0.37 (p-value = 0.36). This further suggests that the three pieces of additional data all bring unique information and thus help us identify the social-learning from the network size effect. This is important because, whereas network size effect is identified through the variation in the initial network size, the social-learning effect is identified through the number of switches as well as the specific nature of

Table 2. Estimates of the Fixed-Effect Model

Variable	Coefficient	Standard error
Network size	**0.000**	0.000
Cumulative contact churn	**−0.029**	0.012
Cumulative stay period	−0.002	0.002
Network age	**−0.081**	0.026
Network male percent	**−0.431**	0.134
Network price	0.515	0.407
Network minutes	−0.000	0.000
Constant	**0.946**	0.284
Observations	1,584	—
Groups	198	—
Pseudo R^2	0.702	—

Note. Bold values are significant at the 5% level.

interaction between the focal individual and his or her network.

2.3. Survey on Consumer Learning on Wireless Service Carriers

To motivate model assumptions on consumer learning, we collected additional survey data from m-Turk. We obtained 101 respondents. We asked six questions (Q1–Q6) regarding the learning behavior. We include the exact wording of the six questions in the appendix. The survey includes the following questions: Q1 asks how people learn about other wireless carriers. Q2 asks whether, when they know their friends switch carriers, they would make quality inference of the new carrier. This tests the assumption about what people learn from switching contacts. Q3 asks whether, when they know their friends switch carriers, they will discuss the new carrier with their friends. This tests the assumption of word of mouth. Q4 asks whether, when they know their friends switch carriers, they would make a positive quality inference of the new carrier. This tests the assumption of the normal distribution of the experience signal. Q5 asks how people learn about the carrier they are using. This tests the assumption about learning from own experience. Q6 asks about the learning about current carrier, and whether people would make quality inference based on the staying contacts.

We report below our key findings from the survey. First, among the different sources that an individual learns about the other wireless carriers than the one (s)he is using, 75% of respondents choose friend as one of the sources, and 57% choose family members. Advertisement accounts for 70%, and online discussion board accounts for 33%. This indicates that social learning is an important venue for people to learn about carriers that they are not using, according to responses on Q1.

Second, when we ask why they think their friends switch carriers in Q2, we have respondents give a rating ranging from 1 to 5, where 1 means "because they think the new carrier has better data and call quality and not because of other reasons" and 5 means "because of many reasons not just about quality." The sample average is significantly above the neutral point (mean = 3.94, t-test = 4.162, p-value < 0.001). This indicates that people do not just make quality inferences when their friends switch carriers. They understand that the switching may be caused by many reasons.

Third, Q3 asks respondents how they would respond to friends switching carriers. We have respondents give a rating ranging from 1 to 5, where 1 means "they will not discuss the new carrier with the friends, but based on observation they infer that the alternative carrier has better quality" and 5 means "they would not infer anything based on observation only, but discuss the new carrier with them." The mean is also significantly above the neutral point (mean = 3.65, t-test = 5.973, p < 0.001). This indicates that, when people observe carrier switching behavior, they may not simply infer the positive quality from the observation. People talk and obtain information from the conversations. Thus, word-of-mouth or information-based learning is a plausible assumption.

Fourth, Q4 asks respondents when their friends switch carrier, whether they think their friends "must be satisfied with the new carrier" or it could be either way. The results is also significantly above the neutral point (mean = 3.52, t-test = 3.977, p < 0.001). That indicates the symmetry of the distribution of the signals coming from their switched contacts regarding the quality of alternative carriers. Thus, symmetric assumption on the experience signals from switched contacts on alternative carriers is reasonable in our context.

Fifth, regarding the learning of the carrier people are using, we ask respondents to rate from 1 ("mainly based on my friends' evaluations") to 5 ("mainly based on my own usage experience") in Q5. The sample mean on Q5 is substantially above neutral (mean = 4.52, t-test = 20.263, p < 0.001). This justifies the assumption of learning about the quality of the focal carrier from individuals' own experience.

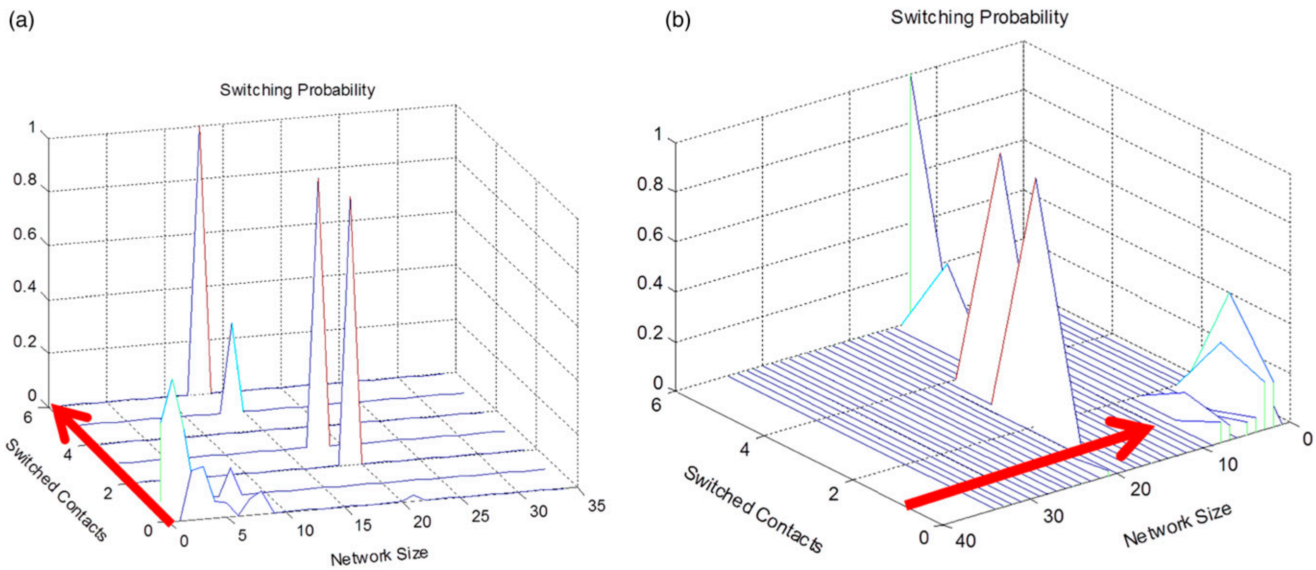
Sixth, regarding the learning from the friends who stay with their wireless carrier, we ask people to rate from 1 ("I usually infer that the carrier has good data and call quality") to 5 ("I will not infer anything about the carrier based on the observation only"). We find that the sample mean is not significantly different from the neutral point, suggesting that people may not necessarily infer quality information from their friends' staying with a carrier (mean = 2.90, t-test = −0.784, p = 0.44). This result is in line with the finding from Q5.

3. Model

3.1. Overview of the Model

We propose a dynamic forward-looking model that captures both the social-learning effect and the network

Figure 2. (Color online) Separating out the Social Learning Effect from the General Network Effect



Notes. (a) For the same network size (network effect), the more contacts who switch over time (more social learning), the greater the likelihood that the customer will switch. (b) For the same number of switched contacts (social learning), the larger the network (larger network effect) the lower the likelihood that the customer will switch.

effect in the context of an individual's decision regarding whether to stay or switch with respect to a wireless carrier. The model consists of three key elements. First, each consumer learns about the quality of the current wireless carrier through his or her own experience. Mobile service is an experience good; customers are usually uncertain about its quality before they actually use it, and it is not enough for them to use it just once. They form a general perception about the quality through repeated usage over time. The perceived quality of the same wireless carrier can vary across customers, potentially owing to network coverage or the efficiency of the devices. Customers have a prior knowledge regarding the quality of the carrier, which is assumed to follow a normal distribution, and we allow customers to update this knowledge in a Bayesian fashion (Erdem and Keane 1996).

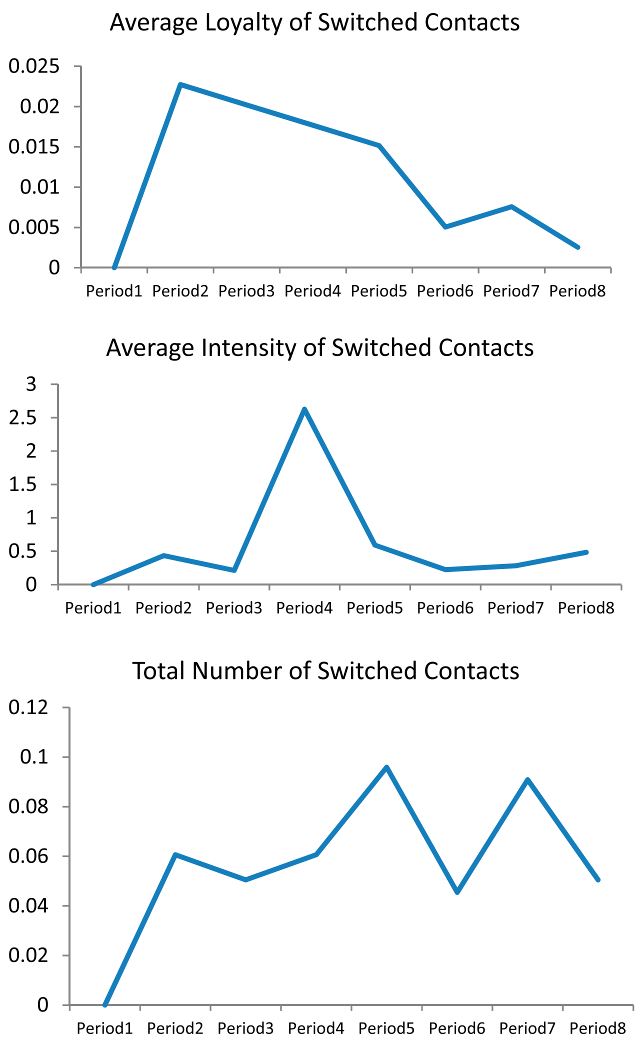
Second, because learning from others about alternative wireless carriers is much less costly than sampling or obtaining direct experience, we assume that people use feedback from their peers to update their own expectations of product quality for these alternative wireless carriers. We infer the occurrence of this social learning when a person's mobile contacts change carriers. Their prior beliefs about the alternatives' service quality vary depending on the customer's idiosyncratic beliefs. As we found in the survey, when contacts switch, word of mouth communication regarding the alternative carriers occurs. We assume that individuals use the new information from the switched contacts to update their beliefs about the alternative carrier in a Bayesian fashion. To capture consumers' strategic learning from others, we allow the signals from others on alternative carriers to be perceived to be systematically different depending on the individual consumer's loyalty status. We also allow the credibility of the signals from others to be dependent on the closeness of the relationship between the individual and the contact who recently switched from the focal carrier.

Third, we further assume that the utility from using the service depends directly on the number of contacts remaining with the network. It is the remaining network effect that is not captured through the social learning and is modeled as a function of the network size variable *Net*. Adding this second component of network-related utility also complicates the modeling and estimation, because it needs to consider the focal customer's expectation on the switching decisions of his or her contacts.

We assume that consumers are forward-looking. Using a questionnaire and experiment, Lemon et al. (2002) showed that "when deciding whether or not to continue a service relationship, consumers not only consider current and post evaluations of the firm's performance, but also incorporate future considerations regarding the services" (p. 1). Because the network size directly affects customer utility, it results in an intertemporal link when customers make their decisions. Consumers who decide to switch today will take into consideration the extra cost that they may incur tomorrow because their network is different from that of their friends. In addition, uncertainty about quality induces a tradeoff between switching today and bearing all of the risks versus waiting and accumulating more information.

Hu, Yang, and Xu: *Social Learning and Network Effect*
Management Science, 2019, vol. 65, no. 10, pp. 4771–4794, © 2019 INFORMS

Figure 3. (Color online) Per-Period Averages of the Switchers' Characteristics and the Switching Numbers



3.2. The Model

Following the dynamic structural literature (e.g., Crawford and Shum 2005, Pakes et al. 2007), we summarize the conditions at time t using a vector of state variables $I_t \in I$ for each individual. Let $z_t \in Z$ be a vector of commonly observed state variables. In our context, z_t includes the individual characteristics and mobile plan information (X_t) and network size (Net_t). Given the state I_t, consumers predict their own network size changes and make switching decisions at the beginning of each period. If they decide to switch, their contract terminates by the end of the period. Let $d_{it} \in \{0, 1\}$ denote individual i's decision at time t. This is a dichotomous indicator that equals 1 if consumer i chooses to stay with the network in period t and 0 otherwise. $\beta \in (0, 1)$ is the common discount factor and is set as 0.95. We further assume that when customers switch they do not return and that there is no new entry into the network because we do not observe a new entry in our sample. This may limit interpretation of our results.

The consumers' single-period utility from staying with their current wireless carrier in time period t, u_{1t} is assumed to be a function of z_t (X_t, Net_t), and S_{ft} is assumed to be the received quality signal of focal network at time t. Following Crawford and Shum (2005), we consider a quasilinear utility specification with a constant absolute risk-aversion specification for the subutility function of S_{ft} and linear terms in z_t,

$$\begin{aligned} u_{1t}(z_t, S_{ft}; \theta) &= -\exp(-rS_{ft}) + \theta X_t + \lambda Net_t + \varepsilon_t(1) \\ &= \bar{u}_{1t} + \varepsilon_t(1), \end{aligned} \tag{1}$$

where r measures the degree of risk aversion and λ captures the network effect.

In our context, we assume that θX_t is a linear utility specification of the plan details and individual characteristics:

$$\begin{aligned} \theta X_t = {} & \theta_1 Price + \theta_2 Min + \theta_3 Fam + \theta_4 Age \\ & + \theta_5 Male + \theta_6 AverageUsage, \end{aligned} \tag{2}$$

where *Price* is the relative on-network/off-network call price of the current plan for each individual, which is predetermined and does not vary across time; *Min* is the number of minutes included in the plan; and *Fam* is the dummy for a family plan.

Furthermore, we define $u_{0t}(S_{at})$ as the *abstract* utility function for the alternative wireless carriers:

$$\begin{aligned} u_{0t}(S_{at}; \theta) &= -\exp(-rS_{at}) + \varepsilon_t(0) \\ &= \bar{u}_{0t} + \varepsilon_t(0), \end{aligned} \tag{3}$$

where S_{at} is the received quality signal of the alternative wireless carriers at time period t.

We assume that before taking action, each individual receives a private choice-specific preference shock $\varepsilon_{it}(d_{it})$ that is known to the individual but is unobserved by the researchers. By assumption, it is drawn independently across agents and over time.

Consumers are assumed to switch carriers if the current network utility plus the discounted switching value and the idiosyncratic taste of switching is greater than the current utility plus the discounted continuation value and the idiosyncratic taste of continuing. The Bellman equation for the value of people who are currently on the network at the beginning of each period is taken from Pakes et al. (2007) and Dunne et al. (2013):

$$\begin{aligned} V(I_t; \theta) = \max\big\{ & \bar{u}_{1t}(z_t, S_{ft}; \theta) + \beta VC(I_t; \theta) \\ & + \varepsilon_t(1), \bar{u}_{1t}(z_t, S_{ft}; \theta) + \beta VX(I_t; \theta) + \varepsilon_t(0) \big\}. \end{aligned} \tag{4}$$

$VC(.)$ is the continuation value, which is the expectation of the next period's realization of the value function:

$$VC(I_t;\theta) = \mathbb{E}[\max\{\overline{u}_{1t+1}(z_{t+1}, S_{ft+1};\theta) + \varepsilon_{it+1}(1) + \beta\int VC(I'_{t+1};\theta)dP(I'_{t+1}|I_{t+1}), \overline{u}_{1t+1}(z_{t+1}, S_{ft+1};\theta) + \varepsilon_{t+1}(0) + \beta\int VX(I'_{t+1};\theta)dP(I'_{t+1}|I_{t+1})\}|I_t]. \quad (5)$$

$VX(.)$ is the expected utility from the alternative wireless carriers, which represents the switching value,

$$VX(I_t;\theta) = \mathbb{E}[\overline{u}_{0t+1}(S_{at+1};\theta) + \varepsilon_{t+1}(0)\,|\,I_t]. \quad (6)$$

3.3. Learning from Own Experience

Although they are uncertain about the quality of experience goods, consumers are assumed to have prior information. Each consumer's prior beliefs about the true quality, Q_{if}, is summarized by a normal distribution:

$$Q_{if} \sim N(\mu_f, \sigma^2_{\mu f}). \quad (7)$$

In this expression, μ_f denotes the prior mean quality and the standard deviation, $\sigma^2_{\mu f}$ measures the precision of the prior means. Consumer i does not know Q_{if} but receives quality signals that allow him or her to update his or her perceptions about the true quality. We assume that the signals are independently and normally distributed across periods around the true service quality (Q_{if}):

$$S_{ift} \sim N(Q_{if}, \sigma^2_f). \quad (8)$$

S_{ift} is observed by the consumer but is unknown to the researchers, whereas Q_{if} is unknown to either. The usage experience and subsequent observed signal S_{ift} do not exactly reveal Q_{if} but do provide information about it. This distribution assumption, along with an initial conjugate prior on Q_{if}, generates a Bayesian learning process in which the consumer's posterior beliefs about Q_{if} are given recursively by the sequence of normal distributions with mean π_{ift} and variance V_{ift} (DeGroot 2005):

$$\pi_{ift+1} = \begin{cases} \dfrac{\dfrac{\pi_{ift}}{V_{ift}} + \dfrac{S_{ift+1}}{\sigma^2_f}}{\dfrac{1}{V_{ift}} + \dfrac{1}{\sigma^2_f}} = \dfrac{\sigma^2_f}{V_{ift} + \sigma^2_f}\pi_{ift} + \dfrac{V_{ift}}{V_{ift} + \sigma^2_f} S_{ift+1} & \text{if } d_{it} = 1, \\ \pi_{ift} & \text{otherwise,} \end{cases} \quad (9)$$

$$V_{ift+1} = \begin{cases} \dfrac{1}{\dfrac{1}{\sigma^2_{\mu f}} + \dfrac{n_{it+1}}{\sigma^2_f}} & \text{if } d_{it} = 1, \\ V_{ift} & \text{otherwise,} \end{cases} \quad (10)$$

where n_{it+1} is the total number of periods that consumer i stays with the current carrier until period t+1.

We assume that consumers have rational expectations so that their prior beliefs equal the actual distribution of the "true" quality Q_{if} across individuals. Thus, μ_f is the population mean quality of the current wireless carrier, and $\sigma^2_{\mu f}$ is the heterogeneity or dispersion parameter. The values for μ_f, $\sigma^2_{\mu f}$, and σ^2_f are to be estimated.

3.4. Modeling Social Learning

People learn from others. It is less costly and more efficient than sampling or obtaining direct experience because subscribing to and learning about a new wireless carrier is often costly. When a contact on the network switches, it not only affects the size of the network but also helps the individual update the quality information about alternative wireless carriers. Q_{ia} denotes the true quality of alternative wireless carriers for individual i. We normalize the prior mean for alternative wireless carriers to 0 and assume the equality of initial variances, for the purposes of empirical identification:

$$Q_{ia} \sim N(0, \sigma^2_{\mu a}) \quad (11)$$

and

$$\sigma^2_{\mu f} = \sigma^2_{\mu a}.$$

Social learning occurs only when a friend on the consumer's network switches, so there may be no learning at all for some periods or multiple learning occurrences for a particular period. Furthermore, we assume that the social-learning process is strategic in two ways. First, the signals on alternative carriers from contacts who have switched may be perceived differently by the individual, depending on the type of contact (e.g., loyalty). Intuitively, if the contact who switched had been a highly loyal customer with the focal carrier, the individual may interpret the signal from this contact to be more positive than the mean. This is because alternative carriers often need to stimulate a better-than-average experience for a switcher who was loyal to a competitor's service. To reflect this logic, we specify the mean of the signals on alternative carriers to equal $Q_{ia} + \delta z_j$, where z_j takes the value of 1 if contact j had been a loyal customer of the focal carrier (i.e., they renewed their contract) and 0 otherwise. Following our rationale, the value of δ would be positive.

Second, we anticipate that consumers may perceive signals from other people to carry different levels of credibility. Intuitively, signals from closer contacts are less noisy than signals from more distant contacts. To reflect this logic, we allow the variability of the experience signal to be a function of the distance between the individual consumer and the contact who switched, $\exp\left(\alpha_0 + \frac{\alpha_1}{D_{ij}}\right)$. D_{ij} measures the contacting frequency between the contact j and the individual i (i.e., $1/D_{ij}$ measures the distance between the two). Therefore, the value of α_1 is expected to be positive. Econometrically, the use of those two aspects allows us to account for the individual heterogeneity of the contacts and the dyadic heterogeneity between the individual and his or her contacts.

To reflect such a strategic social-learning behavior, we specify quality signals for individual i at time t from contact j on alternative carriers as follows:

$$S_{iatj} \sim N\left(Q_{ia} + \delta z_j, \exp\left(\alpha_0 + \frac{\alpha_1}{D_{ij}}\right)\right). \tag{12}$$

The updating process is as follows. For contacts $1, 2 \ldots, j, \ldots J$,

$$\pi'_{iat} = \begin{cases} \dfrac{\dfrac{\pi'_{iat}}{V'_{iat}} + \dfrac{S_{iat+1j} - \delta z_j}{\exp\left(\alpha_0 + \frac{\alpha_1}{D_{ij}}\right)}}{\dfrac{1}{V'_{iat}} + \dfrac{1}{\exp\left(\alpha_0 + \frac{\alpha_1}{D_{ij}}\right)}} & \text{if } d_{itj} = 1, \\ \pi'_{iat} & \end{cases} \tag{13}$$

$$V'_{iat} = \begin{cases} \dfrac{1}{\dfrac{1}{\sigma^2_{\mu a}} + \dfrac{m_{it}+1}{\exp\left(\alpha_0 + \frac{\alpha_1}{D_{ij}}\right)}} & \text{if } d_{itj} = 1, \\ V'_{iat} & \end{cases} \tag{14}$$

$$m_{it} = \begin{cases} m_{it} + 1 \text{ if } d_{itj} = 1 \\ m_{it}. \end{cases} \tag{15}$$

After updating through all of the contacts,

$$\begin{aligned} \pi_{iat+1} &= \pi'_{iat}, \\ V_{iat+1} &= V'_{iat}, \end{aligned} \tag{16}$$

where m_{it} denotes the accumulated number of consumer i's contacts who switch before the time period t+1. A rational-expectation assumption is also made here. Notice that with the prior mean of alternative wireless carrier equal to 0, μ_f as defined in the previous section actually indicates the relative quality of the current wireless carrier rather than the absolute quality.

3.5. Modeling the Network Effect

Following Pakes et al. (2007), we let Net_{it} be the size of individual i's network at the beginning of each period and let NSW_{it} be the number of contacts who switch during time period t (which is unknown at the time of the consumer's decisions). To calculate $VC(.)$ in Equation (5), we need to form the consumer's perceptions about the likely numbers of contacts who switch conditional on the current state variables. These perceptions generate the probability distribution $P(\text{NSW}_{\text{it}} \,|\text{Net}_{\text{it}}, X_{it})$. It is required that $P(\text{NSW}_{\text{it}} \,|\text{Net}_{\text{it}}, X_{it})$ be consistent with the contacts' behavior so as to generate equilibrium—in other words, consumers have rational expectations. We let

$$\begin{aligned} Net^e_{it} &= Net_{it} - NSW^e_{it} \\ &= Net_{it} - \int NSW_{it} dP(\text{NSW}_{\text{it}} \,|\text{Net}_{\text{it}}, X_{it}), \end{aligned} \tag{17}$$

where Net^e_{it} and NSW^e_{it} denote the expectations conditional on the current state.[2]

3.6. State Variables and Transition Probabilities

Our dynamic learning model has four types of state variables. The first includes individual i's posterior mean values for the quality of the current service provider and the alternative providers π_{ift} and π_{iat}. The second includes counts of the number of periods that consumer i stays in the current network (i.e., n_{it}, and the number of contacts who switch, m_{it}). The third category consists of individual and group characteristics X_{it}, z_j, and D_{ij} and the network size Net_{it}. Finally, there are idiosyncratic errors $\varepsilon_{it}(1)$ and $\varepsilon_{it}(0)$. We use I_t to denote the vector of state variables. (π_{ift}, π_{iat}, n_{it}, m_{it}, X_{it}, Net_{it}, z_j, D_{ij}, $\varepsilon_{it}(1)$, $\varepsilon_{it}(0)$) for period t.

The transition rules for Net_{it}, π_{ift} and π_{iat} were given before, and X_{it} is predetermined. Similarly, z_j and D_{ij} are predetermined characteristics of contacts. As for n_{it} and m_{it}:

$$n_{it+1} = \begin{cases} n_{it} + 1 & \text{if stay in period } t+1 \\ n_{it} & \text{if churns in period } t+1, \end{cases} \tag{18}$$

$$m_{it+1} = \begin{cases} m_{it} + \#\text{contacts churns in period } t+1 \\ m_{it} \quad \text{no one churns in period } t+1. \end{cases} \tag{19}$$

3.7. Rational Expectations in Equilibrium and Dynamic Decision Making

The model exhibits an analogue to a dynamic simultaneous-move game with incomplete information. We focus on pure strategy Markov perfect equilibria (Pakes and McGuire 1994) such that each individual's behavior depends only on the current state and the current private shock.

We assume that the consumers make their switching decisions simultaneously at the beginning of each period. The dependence of utility on the network size makes this decision endogenous. That is, the choice probability for each individual is a nonlinear mapping of the choice probabilities of all other members on the network, $\text{Prob}(d_{it} \,|\, I_{it}) = \text{H}(\text{Prob}(d_{-it} \,|\, I_{-it}))$.

The choice-specific value function of staying for individual i is

$$\begin{aligned} V_1(I_t) &= \overline{u}_{i1t}(S_{ift}, X_{it}, Net_{it}) + \beta VC(I_t) + \varepsilon_{it}(1) \\ &= \mathbb{E}[\overline{u}_{i1t}(S_{ift}, X_{it}, Net_{it}) + \beta VC(I_t) + \varepsilon_{it}(1)|S_{ift}, S_{iatj}] \\ &= -\exp\left(-r\pi_f{}^t + \frac{1}{2}r^2\left(v_f{}^t + \sigma_f^2\right)\right) + \theta X_{it} \\ &\quad + \lambda Net^e_{it}(X_{it}, Net_{it}, \gamma) + \varepsilon_{it}(1) \\ &\quad + \beta\mathbb{E}[\mathbb{E}[V(I_{t+1})|S_{ift}, S_{iatj}]|I_t]. \end{aligned} \tag{20}$$

The choice-specific value function of switching is

$$\begin{aligned} V_0(I_t) &= \overline{u}_{i1t}(S_{ift}, X_{it}, Net_{it}) + \beta VX(I_t) + \varepsilon_{it}(0) \\ &= \mathbb{E}[\overline{u}_{i1t}(S_{ift}, X_{it}, Net_{it}) + \beta VX(I_t) + \varepsilon_{it}(0)|S_{ift}, S_{iatj}] \\ &= -\exp\left(-r\pi_f{}^t + \frac{1}{2}r^2\left(v_f{}^t + \sigma_f^2\right)\right) + \beta X_{it} \\ &\quad + \lambda Net^e_{it}(X_{it}, Net_{it}, \gamma) + \varepsilon_{it}(0) \\ &\quad + \beta\mathbb{E}\left[-\exp\left(-r\pi_a{}^t + \frac{1}{2}r^2\left(v_a{}^t + \exp\left(\alpha_0 + \frac{\alpha_1}{\mathrm{Dij}}\right)\right)\right)|I_t\right]. \end{aligned} \tag{21}$$

We assume that the choice-specific preference shocks $\varepsilon_{it}(d_{it})$ have an extreme value distribution; the conditional choice probability then has a very convenient expression:

$$\mathrm{Prob}(d_{it} = 1|I_t) = \frac{\exp(V_1(I_t))}{\exp(V_1(I_t)) + \exp(V_0(I_t))} \tag{22}$$

$$\mathrm{Prob}(d_{it} = 0|I_t) = \frac{\exp(V_0(I_t))}{\exp(V_1(I_t)) + \exp(V_0(I_t))}. \tag{23}$$

Finally, the integrated value function is

$$\begin{aligned} V(I_t) &= \int \max\{V_1(I_t), V_0(I_t)\}\mathrm{dP}(\varepsilon_{\mathrm{it}}) \\ &= \log(\exp(V_1(I_t)) + \exp(V_0(I_t)). \end{aligned} \tag{24}$$

4. Identification

We will discuss the identification of our model along the research on social interactions. Social interaction effect has been widely studied in economics, especially after the seminal paper by Manski (1993), which discussed the identification of endogenous social effects as a reflection problem. Social interaction effect is defined by Manski as "the propensity of an individual to behave in some way varies with the prevalence of that behavior in some reference group containing the individual," (p. 531).

Identification of social interaction effect often arises with cross-sectional data. As Manski (1993) classified, similar behaviors from individuals in the same group can be driven by three effects: (i) endogenous effect or the true social interaction effect; (ii) exogenous or contextual effects; and (iii) correlated effects. This is likely the general background, on the basis of which our review team raised their identification concerns on our study.

For now, let us focus on the endogeneity issue on identification of social interaction effect. We use the following linear model as a simple illustration, following the literature. Let y denote an outcome (such as grade point average [GPA]), x denote exogenous covariates, i and j index people, ε be the error term, and N be the total number of people in the analysis.

$$y_i = \theta\sum_{j=1}^{N} w_{ij}y_j + \beta x_i + \varepsilon_i, \tag{25}$$

where $w_{ij} = 1$ if i and j are in the same group, = 0 otherwise.

Endogeneity arises in two cases. In the first case, grouping is entirely random. Because of the simultaneity, y_j will be a function of y_i and thus a function of ε_i for those i and j belonging to the same group. This leads to the classic endogeneity problem due to the correlation between y_j and the error term ε_i in Equation (25) In the second case, grouping is not random, and based on y. Endogeneity arises not only owing to simultaneity (as shown in the first case) but also owing to endogenous group formation. Imagine if i and j belong to the same group because they have high academic performance (i.e., correlated with the outcome y); we can easily show that w_{ij} will be correlated with ε_i in Equation (25). In sum, identification of social interaction effects involves endogeneity driven by simultaneity and nonrandom group formation.

Modeling social-learning effect when taking network effect into consideration in our setting is just as in modeling social interaction effect, but often in a panel setting, and we try to understand how an individual's behavior at t is affected by behavior of connected others in $t-1$. Equation (25) now becomes

$$y_{it} = \theta\sum_{j=1}^{N} w_{ij,t-1}y_{j,t-1} + \beta x_{it} + \varepsilon_{it}; \tag{26}$$

$w_{ijt} = 1$ if i and j are in the same group at t, = 0 otherwise.

The dynamic setup eliminates the endogeneity issue due to simultaneity, to a large degree. It is for this benefit, most research related to influence from others in marketing has adopted this specification. Our model resembles the above general structure, but with differences. In our context, y can be interpreted as the preference for alternative carriers. x includes the size of person i's network who stays with the focal carrier and other covariates. We next address potential identification issues in our empirical context.

4.1. About Homophily

The identification is achieved from both functional form and data variation. In our empirical context, the endogeneity issue driven by nonrandom group formation is less likely for two reasons. First, our social-learning mechanism is specified on consumer preference for an alternative carrier. In this case, $w_{ij,t-1}$ (where i is the focal individual, j stands for the one in i's network who switched to an alternative carrier) is not very likely to be determined by ε_{it}. ε_{it} is the unrealized shock on individual i's preference for an alternative carrier at t. $w_{ij,t-1}$ basically identifies the person in i's network who switched in $t-1$. i and j are connected because they once use the same carrier, but whether j switched out in $t-1$ is not likely to be determined by some unrealized shock in i's preference on the alternative carrier.

Second, the initial network in our data was determined before our study period. In particular, the data provider identified the network on the basis of individuals' contact information over a time window before our study period. In such case, the initial network structure is unlikely to be affected by consumers' future preference on an alternative carrier (i.e., switching out).

A common underlying trait shared by friends may also affect individual behavior. However, different from the static model, the variable of network size does not enter our model directly. We have a dynamic forward-looking model, and we use two-stage dynamic programming to solve the model. In the first stage of the estimation, we model the expected number of friends who may switch (for details refer to Section 5). This step essentially models the network change. We include individual- and network-level characteristics as well as month dummies to control for the time effect in this step. The assumption here is that the individual forms rational expectations and that those expectations will be realized. In the second step we model how the switching decision is affected by the expected direct network effect and social learning and control for individual and network specific characteristics, as well as the learning from own usage. Therefore, the separation of social learning and correlated unobservables only requires the residual of anticipated network change in the first stage to be uncorrelated with the residual of the proposed model and unexplained by the factors of the current period in the second stage.

Following Nair et al. (2010) and Iyengar et al. (2011, 2015), we have included each focal customer's friends' average characteristics as covariates of the model, and we modeled the unobserved individual specific effect through the prior specification, to control for homophily.

4.2. About the Correlated Unobservables

The identification of the correlated unobservables due to strategic behaviors of networked customers is achieved from our empirical setup. Because our network data are ego-centric, we only model focal customers' leave/stay decision. These focal customers are not in each other's direct networks, which could reduce the concern of correlated unobservables due to strategic behaviors of networked customers.

To control for the correlated unobservable in preferences across people, we have developed a correlated learning model as robustness check. The identification is achieved from functional form and data variation. The model follows the spirit of Erdem (1998) and Yang and Allenby (2003). More specifically, we allow an individual's prior on the focal carrier's quality (relative to an alternative carrier) to be correlated according to a spatial autoregressive (SAR) form. The SAR model has been widely adopted in economics, statistics, and marketing to control for correlated unobservables (e.g., Yang and Allenby 2003, Lee 2007). For model details and results please refer to Section 6 Robustness check.

4.3. Identify Social Learning Effect from the Network Size Effect

The identification is achieved from both the data variation and the functional form. Whereas social learning is identified according to temporal variation, identification of the direct network effect relies on cross-sectional differences. The cross-sectional information of the network sizes in the first period helps us identify the effect of network size in the model because the social-learning effect has not arisen then. Figure 4 plots the distribution of the network size in the first period. The variation of the initial network size is substantial across consumers, which helps identify the network size effects. Imagine two individuals with the same number and type of switching contacts (i.e., the same social-learning experience) but different number of staying contacts in their network; then their different churning probabilities can be attributed to the differences in their network size, which helps identify the network size effect.

On the other hand, the functional form of social learning then quantifies the effect from switching contacts under such circumstance. We have panel data of network changes. As such, variation on the switching contacts across time captures this social-learning effect on the alternative carrier (Chan et al. 2014a, b).

Furthermore, we use extra data to identify the social-learning effect. As shown in Equation (12) of signals from an alternative carrier, the received signal from each switched contact is affected by (i) the loyalty of the switcher to the focal carrier z_j before switching; and (ii) the interaction level between the switcher and the focal person D_{ij}. These two covariates will only affect the nature of social learning but not the network size. This further suggests that the three pieces of additional data all bring unique information and thus help us identify the social-learning from the network size effect.

4.4. Identify the Network Size Effect from Consumer Heterogeneity

We now clarify how we can ensure that the network size effect is not confounded with consumer heterogeneity. The identification is achieved from both data variation and functional form. In our proposed model, to capture observed customer heterogeneity, we use individual demographic information, plan characteristics, and average usage per contact in the initial period. Individuals who have high preference for chatting with friends are likely to have a large network and less likely to switch because of high preference. To control for this possibility, we have now incorporated the average usage per contacts in the initial period as an individual-specific covariate in the per-period utility function of the full model. For unobserved heterogeneity, we use the assumption of the distribution of priors as a control. Because we have panel data, we controlled for the unobserved heterogeneity on people's prior preference for the focal carrier (the prior preference is drawn from a normal distribution). As such, the network size effect can be separately estimated and is not confounded with the consumer heterogeneity.

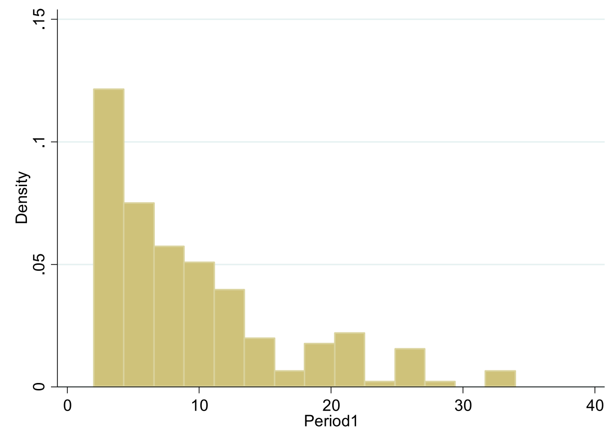
We have estimated the fixed-effect model. Results are reported in Table 2. As a result, the larger the network, the more likely the person will stay. The effect of cumulative switched friends (as proxy for social learning) is significant and negative after controlling for the individual-specific fixed effect. This demonstrates that the social-learning effect can be separated from the network effect and the unobserved heterogeneity. The fixed effect model and the proposed model are not directly comparable. The fixed effect model demonstrates that the social-learning and network effects can be separated from unobserved heterogeneity, and the proposed structural model explains how.

5. Estimation and Results

Direct implementation of the method used by Bajari et al. (2007) in our context is not feasible because the service qualities are unobserved, which prevents us from estimating the transition probabilities and the policy functions for each individual. Therefore, our approach uses the idea of two-step models (Aguirregabiria and Mira 2007, Bajari et al. 2007, Pakes et al. 2007) but embeds the first step in a nested fixed-point algorithm. In other words, instead of estimating who switches in the first step, we estimate *how many* switch.

In equilibrium, the perceptions of network size change must be consistent with actual observations. Our proposed estimation consists of two steps. In the first, we estimate NSW_{it}^e for each individual. Figure 5(a) shows the empirical transition probabilities of the network size observed from the data. The x-axis indicates the network size at time t, and the y-axis indicates the network size at time $t+1$. The blocks are the percentage

Figure 4. (Color online) Distribution of the Network Size in the Initial Time Period



(probability) of a network size change from time t to time $t+1$ according to the data. Darker shades indicate a higher probability. We notice that for each period the maximum number of switching neighbors is two. We therefore assume that NSW_{it}^e can be approximated by a multinomial ($NSW_{it}^e = 0, 1, 2$) logit model of Net_{it} and X_{it}. Because the consumers' expectations must be consistent with the average realizations, these estimations will converge to the true expected number of switching neighbors:

$$\text{Prob}(NSW_{it}^e = j) = \frac{\exp(\gamma_{j1} Net_{it} + \gamma_{j2} X_{it})}{1 + \sum_{j=1}^{2} (\gamma_{j1} Net_{it} + \gamma_{j2} X_{it})} \quad j = 1, 2; \tag{27}$$

$$\text{Prob}(NSW_{it}^e = 0) = \frac{1}{1 + \sum_{j=1}^{2} (\gamma_{j1} Net_{it} + \gamma_{j2} X_{it})}. \tag{28}$$

The second step treats the estimate of NSW_{it}^e as known and calculates the network size by subtracting this number from the network size at the beginning of the period. We then estimate the structural parameters using an adaption of the approximation method used by Keane and Wolpin (1997) (Crawford and Shum 2005).[3] We use simulation to integrate the two sets of signals, S_{ift} and S_{iatj}, because they are unobservable to researchers. Thus, the simulated likelihood is as follows,

$$\prod_{i=1}^{I} \left(\frac{1}{l} \sum_{l=1}^{L} \prod_{t} \left(\frac{\exp(V_1^l)}{\exp(V_1^l) + \exp(V_0^l)} \right)^{d_{it}} \cdot \left(\frac{\exp(V_0^l)}{\exp(V_1^l) + \exp(V_0^l)} \right)^{(1-d_{it})} \right). \tag{29}$$

We use $L = 10{,}000$ draws for the simulation and the Berndt-Hall-Hall-Hausman method to approximate the Hessian matrix.

A comparison between the empirical transition matrix of the network from the data and that based on the simulation using the model estimates demonstrates the validity of such an approach in the current context.

To do so, we generate simulated decision sequences using our estimated parameters and compare the summary statistics from these simulated data with the analogous statistics for the actual estimation data.[4] In particular, we focus on the network transition probability matrix, not only because it is a good summary of the network changes but also because it provides a test of robustness on the model assumptions about the expected number of neighbor terminations, NSW^{e}_{it}. In Figure 5(b) we present the fitted matrix and compare it with Figure 5(a) (i.e., that from the data). To measure the similarity between the two matrices, we stack the columns of each matrix and calculate the correlation between the two vectors. The correlation is 0.976 (p-value $<$ 0.000). Similar to the data, the simulated network size is highly likely to remain the same over time. The size changes are small and rare. The results demonstrate the validity of the model.

The last column in Table 3 report the estimates of the parameters of the proposed dynamic structural model. The point estimates for the parameters of the true quality distribution indicate substantial heterogeneity in service quality across customers. Notice that μ_f actually captures the average quality difference between the current wireless carrier and the alternatives. The relative mean service quality estimate (2.191; standard error [s.e.] 0.157) indicates that the focal wireless carrier has a higher average service quality than the alternative carriers. The estimate of the standard deviation of the true service quality, $\sigma_{\mu f}$ (1.2; s.e. 0.102), is large, which suggests a rather large variation in service quality. The standard deviations of the quality signals, σ_f (1.192; s.e. 0.104), measures how quickly consumer uncertainties about true service quality decrease owing to learning from their own usage experience.

The estimate of δ is significantly positive (1.038; s.e. 0.100), indicating that the signals on alternative carriers from a more loyal contact who has switched are interpreted in a more positive manner than are those from a less loyal customer of the focal carrier. The significant and positive signs of α_1 (0.770; s.e. 0.081) demonstrate that a closer relationship between the switched contact and the individual leads to the reception of a less noisy signal. The strategic learning has never been detected before. The results are intuitive. The individual's perception of the information from her contacts depends on the characteristics of the contact and the relationship with the contact. If the contact used to be loyal to the focal carrier but eventually switched, the information regarding the alternative carrier from this contact would be perceived as more positive. The consumer may suspect that the contact must have very positive knowledge about the alternative carrier that she does not have, so her contact decides to switch; and it is also quite natural that people will trust the information from her close contacts more than otherwise. So the learning process differs depending on the information source, and our results demonstrate that.

According to the estimation results, the inclusion of more minutes and family plans are both effective strategies to retain customers. The estimate of network effects, λ (0.081; s.e. 0.017), is positive and significant, which indicates that in addition to social learning, the actions of friends directly affect a customer's probability of staying. Average usage is significant and positive (0.078; s.e. 0.014), indicating that there is heterogeneity in the tendency of using the mobile service.

We also present the estimation results of a duration model in the second column of Table 3. A duration model is commonly used by practitioners to predict customer switching decisions. We are not able to obtain a significant estimate of the network size, but it is confounded with many interpretations. As a comparison, in column 3 of Table 3 we report the estimation results using a model similar to that which we propose but without taking into consideration the social-learning details. We obtained similar results for the parameters, but our model fits the data better according to the Bayesian information criterion.

6. Robustness Check

6.1. Correlated Unobservables

We have developed a correlated learning model to test the issue of unobserved correlation. In the typical learning model it is assumed that the priors of consumer quality perceptions come from normal distribution with mean and variance as parameters to be estimated and the priors are independent from each other. Erdem (1998) proposed a correlated learning model and allowed the priors of different *products* to be correlated. Here, we account for the potential correlated errors by allowing the priors of different *people* to be correlated with a SAR form. This follows the spirit of the model proposed by Yang and Allenby (2003) and Lee (2007). The SAR model has been widely used in economics, statistics, and marketing to control for correlated unobservables. The details are as follows.

Let Q_{if} denote the true quality of the focal wireless carrier relative to the alternative carrier, μ_f as the prior mean, and η_i as the error term, for individual i:

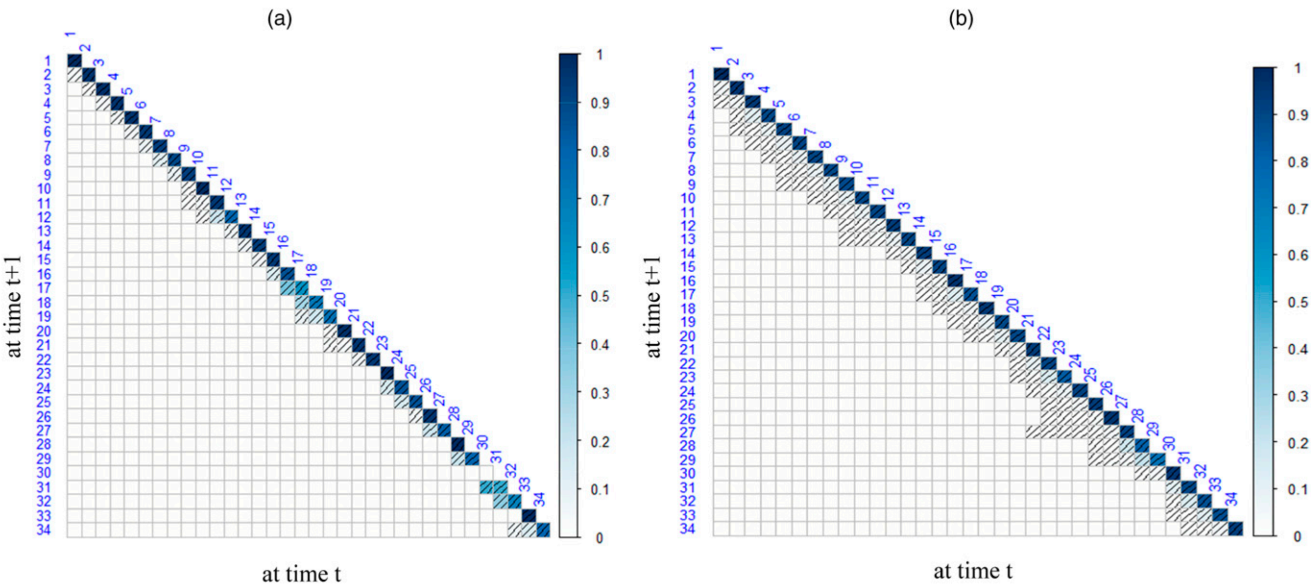
$$Q_{if} = \mu_f + \eta_i, \tag{30}$$

$$\eta = \rho W \eta + u, \tag{31}$$

$$u \sim N(0, \sigma^2_{\mu f} I), \tag{32}$$

$$Q_f \sim MVN\left(\begin{pmatrix}\mu_f \\ \vdots \\ \mu_f\end{pmatrix}, \sigma^2_{\mu f}(I-\rho W)^{-1}(I-\rho W')^{-1}\right) = MVN(\overline{\mu_f}, \sigma^2_{\mu f}\Sigma). \tag{33}$$

Figure 5. (Color online) Network Transition Probabilities (Data vs. Simulation)



If ρ is significantly different from zero, then Equation (31) implies that consumer prior beliefs are correlated owing to unobservable. Equation (33) spells out the spatial autoregressive relationship that potentially captures the unobserved correlation among individuals through the adjacency matrix W. The adjacency matrix W is constructed by following Yang and Allenby (2003):

$$w^1_{i,j} = \frac{1}{abs\,|\,age_i - age_j\,|},$$

$w^2_{i,j} = 1$ if person i and person j belong to the same gender; 0 otherwise,

$w^3_{i,j} = 1$ if person i and person j use the same brand phone; 0 otherwise,

$w^4_{i,j} = 1$ if person i and person j share a family plan; 0 otherwise,

$$W = \sum_{g=1}^{4} \phi_g W^g, \tag{34}$$

$$\phi_g = \frac{\exp(\zeta_j)}{\sum_{g=1}^{4} \exp(\zeta_g)}. \tag{35}$$

We normalize ζ_4 to be 1.

As before, signals are coming from normal distributions,

$$S_{ft} \sim MVN(Q_f, \sigma_f^2 I). \tag{36}$$

The learning then follows multivariate Bayesian updating, and the posterior mean and variance-covariance matrix are updated as

$$\mu_N = \sigma_f^2 I(\Sigma_{Nt} + \sigma_f^2 I)^{-1}\pi_{ft} + \Sigma_{Nt}(\Sigma_{Nt} + \sigma_f^2 I)^{-1} S_{ft}, \tag{37a}$$

$$\Sigma_{Nt+1} = ((1/\sigma_{\mu f}^2)\Sigma^{-1} + N_{t+1}(1/\sigma_{\mu f}^2) I^{-1})^{-1}. \tag{37b}$$

To reduce the computation burden, we take advantage of the following approximation to simplify the inversion calculation, when estimating the model

$$(I - \rho W')^{-1} \cong I + \rho W' + \rho W W'.$$

We estimated this correlated learning model and reported the results in the last column of Table 4. As shown, ρ is not significant, indicating that the correlated unobservable is not significant in our empirical context. The other coefficient estimates remain largely consistent with those from our proposed model. Although the correlated unobservable is not found significant, we now include this as a future research area. If richer data are available, researchers can quantify the magnitude of each information source that may lead to similar behavior among individuals.

The correlation is not just defined according to demographic information but also phone model and family plan to capture other common factors. For example, mobile phones are sold by wireless carriers, so a common factor that affects people's choice of phone model may also affect their perception of the carrier. As a conclusion, we did not find evidence of correlated unobservables in our empirical context. Estimation results remain largely unchanged with this new specification.

6.2. Learning from Staying Contacts on the Quality of Focal Carrier

Although we justified our assumption of consumer learning from switching contacts on alternative carriers on the basis of the survey results, we still tested this alternative assumption and report the results in Table 5. As shown, our results remain qualitatively unchanged.

Table 3. Model Comparison

	Duration model	Dynamic model	Proposed model
Constant	−0.457	—	—
	(3.124)	—	—
Relative price	−0.990	−0.126	−0.577
	(1.638)	(0.191)	(0.443)
Plan minute	**0.002**	**0.153**	**0.184**
	(0.001)	(0.037)	(0.041)
Age	0.018	**0.416**	**0.390**
	(0.029)	(0.108)	(0.092)
Male	−0.309	**0.602**	**0.640**
	(0.454)	(0.188)	(0.084)
Family	0.655	0.244	0.006
	(0.512)	(0.147)	(0.171)
Network	0.122	**0.093**	**0.081**
	(0.085)	(0.016)	(0.017)
Network age	0.076	—	—
	(0.069)	—	—
Network male percent	−0.001	—	—
	(0.852)	—	—
Average usage	0.342	**0.300**	**0.078**
	(0.311)	(0.107)	(0.014)
μ_f	—	**1.112**	**2.191**
	—	(0.329)	(0.157)
$\sigma_{\mu f}$	—	**1.118**	**1.200**
	—	(0.103)	(0.102)
σ_f	—	**1.038**	**1.192**
	—	(0.106)	(0.104)
$\alpha 0$ (intercept)	—	—	**−0.339**
	—	—	(0.031)
$\alpha 1$ (intensity)	—	—	**0.770**
	—	—	(0.081)
δ (loyal)	—	—	**1.038**
	—	—	(0.100)
Risk aversion	—	**0.988**	**0.974**
	—	(0.102)	(0.101)
σ_a	—	**0.900**	—
	—	(0.101)	—
LL	−54.237	−255.926	−180.675
BIC	161.357	617.436	497.100

Notes. Bold values are significant at the 5% level. Standard errors are in parentheses. LL, log-likelihood; BIC, Bayesian Information Criterion.

6.3. Distribution Assumption for Social Learning

Following Zhang (2010), we have derived the distribution of the signals using Equation (4). Given the nth draw of the state variables, we want to equate the continuation value *VC* (Equation (5)) and the switching value *VX* (Equation (6)) by solving for the signals of social learning. Because there is no analytical solution for this equation, we solved the signals numerically. To simplify, we did not include the loyalty and intensity of the switchers, because our objective is to determine whether the distribution of the signal of social learning is truncated. We use the Newton-Raphson method and solve for the value of the social-learning signal that equates *VC* and *VX* iteratively. In Figure 6 we plot the signals that are solved from the equating *VC* and *VX*. As shown, the distribution of the signals from social learning is not truncated.

We assume the signals from switched contacts are independent and identically distributed. This assumption is valid in the current context because there is no ordering of contacts in the network setting. For *VC*, the change of network size downward will directly affect the utility, which leads to the decrease of *VC* in the next period; and it will also affect the future value of *VX*. For *VX* of next period, the change of network size will have two effects, adjusting the posterior mean of

Table 4. Correlated Learning Model Comparison

	Dynamic model	Proposed model	Correlated prior
Relative price	**−0.324**	−0.506	−0.533
	(0.146)	(0.295)	(0.477)
Plan minute	**0.093**	**0.164**	**0.161**
	(0.023)	(0.032)	(0.035)
Age	**0.228**	**0.300**	**0.300**
	(0.04)	(0.056)	(0.099)
Male	**0.291**	**0.572**	**0.549**
	(0.084)	(0.158)	(0.110)
Family	0.156	0.118	0.179
	(0.08)	(0.152)	(0.112)
Network	**0.066**	**0.080**	**0.065**
	(0.008)	(0.034)	(0.018)
μ_f	**0.241**	**0.242**	**0.237**
	(0.007)	(0.020)	(0.004)
$\sigma_{\mu f}$	**1.200**	**1.154**	**1.169**
	(0.051)	(0.067)	(0.031)
σ_f	**1.076**	**1.141**	**1.126**
	(0.051)	(0.019)	(0.043)
α0 (intercept)	—	**−1.019**	−0.265
	—	(0.302)	(0.151)
α1 (intensity)	—	**0.365**	**0.800**
	—	(0.115)	(0.137)
δ (loyal)	—	**1.052**	**1.045**
	—	(0.013)	(0.013)
Risk aversion	**0.962**	**0.977**	**0.978**
	(0.018)	(0.032)	(0.021)
σ_a	**0.966**	—	—
	(0.005)	—	—
ζ1	—	—	**1.013**
	—	—	(0.130)
ζ2	—	—	**1.012**
	—	—	(0.133)
ζ3	—	—	**1.011**
	—	—	(0.134)
rho	—	—	0.171
	—	—	(0.743)
LL	−249.064	−223.998	−357.822
BIC	581.087	546.038	881.561

Notes. Bold values are significant at the 5% level. Standard errors are in parentheses. LL, log-likelihood; BIC, Bayesian Information Criterion.

the alternative carrier and sharpening the perception by decreasing the posterior variance. The latter will increase VX because we assume the individual is risk averse. We normalize the prior of the quality to 0. As a result, to equate VC to VX, the signals of social learning can be assumed to come from a normal distribution (i.e., there is no need for a cutoff point). If we draw a positive signal, the posterior mean will adjust upward, leading to the increase of VX in the next period, which will partially cancel out the effect from the decrease of network size on VC. If we draw a negative signal, the posterior mean will adjust downward, however, owing to the positive effect of posterior variance on VX the results becomes uncertain. In either case, we can solve for a point that equates VC and VX.

Table 5. Robustness Check by Assuming Staying Contacts as Signals for the Focal Carrier Quality

	Proposed model	Staying contacts as signals
Relative price	−0.506	**−0.563**
	(0.295)	(0.211)
Plan minute	**0.164**	**0.164**
	(0.032)	(0.054)
Age	**0.300**	**0.303**
	(0.056)	(0.143)
Male	**0.572**	**0.588**
	(0.158)	(0.252)
Family	0.118	0.173
	(0.152)	0.190
Network	**0.080**	**0.063**
	(0.034)	(0.023)
μ_f	**0.242**	**2.391**
	(0.020)	(0.261)
$\sigma_{\mu f}$	**1.154**	**1.177**
	(0.067)	(0.103)
σ_f	**1.141**	**1.239**
	(0.019)	(0.616)
δ (loyal)	**−1.019**	−0.284
	(0.302)	(0.343)
α0 (intercept)	**0.365**	0.753
	(0.115)	(0.484)
α1 (intensity)	**1.052**	**1.078**
	(0.013)	(0.105)
Risk aversion	**0.977**	**0.975**
	(0.032)	0.103
LL	−223.998	−390.231
BIC	546.038	878.504

Notes. Bold values are significant at the 5% level. Standard errors are in parentheses. LL, log-likelihood; BIC, Bayesian Information Criterion.

6.4. Different Discount Rates

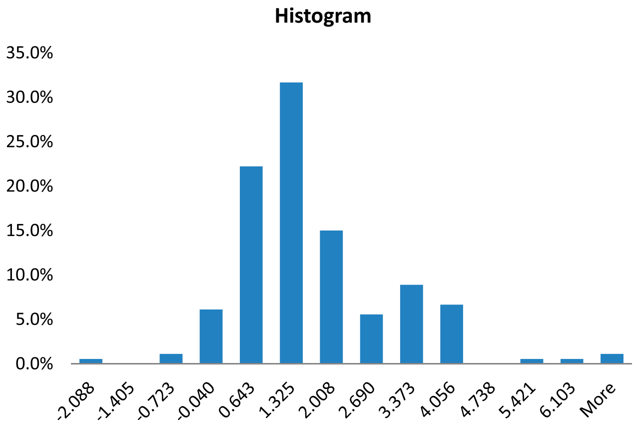
We have estimated our proposed model under different assumptions of discount rate (d = 0.9, d = 0.95, and d = 0.99), and the results are available upon request. Results are largely consistent. In the paper, we use d = 0.95.

6.5. Established Identification of Three Different Mechanisms via a Simulation Exercise

To further establish identification, we also simulate consumer churning patterns under different model specifications: (a) social-learning effect; (b) direct network effect; and (c) homophily. The simulation results show that the three mechanisms generate different churning patterns, providing evidence that social-learning effect can be separated from direct network effect or homophily.

6.5.1. The Data Simulation Process. In the first step, on the basis of the estimation results of the proposed model, we generate the outcome data (i.e., leave or stay decision) in each time period for each individual. This data-generation process incorporates social-learning

Figure 6. (Color online) Distribution of Social-Learning Signals Solved by Equating the Value of Continuation to the Value of Switching



effect, direct network effect, learning from own usage, and individual- and network-specific characteristics, following the proposed model.

In the second step, on the basis of the simulated data from the full specification as laid out above, we run estimations using three model specifications: (a) the social-learning effect only; (b) the direct network effect only; and (c) homophily only.[5] For the models under (a) and (b), we simply shut down the related network effect part and the related social-learning part from our proposed full model. To estimate the model under (c), we use a proportional hazard model with Weibull distribution. Homophily implies that there is some unobserved factor that makes individuals become friends and also affects their switching behavior. For example, an individual's attitude toward novelty can affect his or her choice of friends and behavior. This means that for each time period, whether the contact friends switch, the unobserved factor will affect the focal person's switching probability. In other words, the focal customer has a constant switching probability over time due to the unobserved factor. No matter how many friends have switched in one period, the switching probability of the focal person in each period due to the unobserved factor should not vary. This switching probability may vary by individual. The proportional hazard model is a good candidate to capture this, because it assumes a constant switching probability for each individual in each period and accounts for variation across consumers given the individual specific and initial network characteristics.

In the last step we then simulate the churning patterns according to the estimation results of the three models respectively. More specifically,

- We simulate the binary outcome 1,000 times for each individual in each time period, according to the estimates of the three model specifications.
- For each time period in each of the 1,000 simulations, we sum up the predicted staying probabilities across individuals to calculate the total number of people who stay with the current carrier.
- We then calculate the average and the confidence intervals across the 1,000 simulations for each period to get the churning patterns under each model specification.

The results from the simulation exercise outlined above are presented in Figure 7. The third line from the top represents the "true" data, indicating the number of people who stay with the focal carrier in each time period as predicted by our proposed full model. The second line, fourth line, and first line represent the same metric under social-learning effect, direct network effect, and homophily, respectively. The dotted lines represent the 95% confidence intervals for the simulated outcomes.

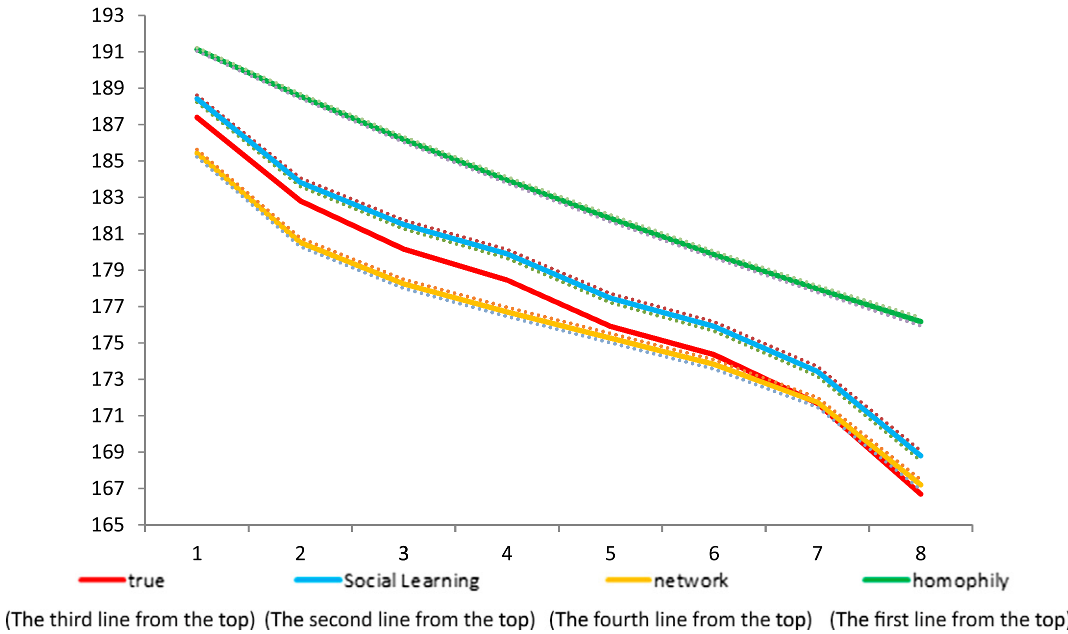
6.5.2. Findings from the Simulation Exercise. The churning pattern from the proposed full model is not entirely driven by social-learning effect, direct network effect, or homophily. In other words, none of the three mechanisms can fully explain the "true" data (the third line) generated from the proposed full model. So the simulation results provide some evidence that the three mechanisms are separable from each other.

The social-learning model (the second line) underestimates the average churning probability in the sample, because it assumes away the direct network effects (by boosting the utility of staying when friends switch) and assumes that friends' switching could have either positive or negative effects on the focal individual's churning probability. Thus, the effect is not monotonic with the number of friend switches. Overall, the churning frequency obtained from the social-learning model is lower than that from the "true" data. However, because the social-learning model considers the effect of friends' churning behavior on the focal person, the churning pattern predicted by the social-learning model is closer to that from the "true" data than what is obtained from the homophily model, which assumes a constant switching probability over time.

The model under the direct network assumption (the fourth line) overestimates the churning probability in the sample, because it assumes that the number of friends who stay has a monotonic positive effect on the focal individual's probability of staying. The more friends who switch, the more likely that the focal individual will churn. The comparison between the "true" data and the data generated from the direct network effect model shows that there exist some other mechanisms (such as social learning) that mitigate the direct effect from the network.

The hazard model (the first line), which assumes a constant probability of churning, severely underestimates the churning probability in the sample. We

Figure 7. (Color online) Simulation of Churning Patterns Under Different Mechanisms



Note. The dotted lines represent the 95% confidence intervals for the simulated outcomes.

assume that some unobserved factor directly influences the behavior choice of the focal individual at each time t by the same magnitude but varies by the individual- and network-specific characteristics (e.g., gender, age, plan choice, phone usage, initial network size, initial network gender ratio, etc.) at the initial time t_0 across individuals. The model also assumes away the influence from the contact friends. The results indicate that homophily cannot fully replicate the data pattern. We want to consider individual interactions and their influences within a network.

The churning pattern under social learning is closer to that from the "true" data in the beginning. However, toward the end, the churning pattern under the direct network effect is closer to that from the true data. This result is consistent with our findings regarding the elasticity of social learning and the direct network effect in the counterfactual analysis in Section 7. From the study of "elasticity," we learned that the network effect has the largest effect from initial switchers compared with later switchers on the focal customer's switching probability. However, for social learning, even later switchers have an impact on the focal person's probability of switching. This finding indicates that, for social learning, the effect of an additional switcher will not diminish quickly. So, in the model that only considers the direct network effect, the churning pattern should be quite different from the data generated by both the direct network effect and the social-learning effect in the beginning, when the network effect has a larger impact and severely overestimates the churning probability. As time passes, the network effect diminishes when an additional friend switches, mitigating the overestimation. In contrast, the social-learning effect kicks in later, so its churning pattern deviates more from the generated data in later periods.

In conclusion, the churning patterns are very different under the three assumptions (social learning, direct network effects, or homophily). This suggests that the three mechanisms have different marginal effects on the focal individual's churning behavior. Such a finding demonstrates that the social-learning effect can be separable from the other two. In our study, with a parameterized social-learning model (with assumptions justified by survey results), our goal is to capture the social-learning effect in a network setting and separate it from the general network effect, so that we can better understand the underlying mechanisms leading to consumer churning behavior.

7. Policy Simulations

7.1. Social-Learning Effect Within a Network

When contact neighbors leave a wireless carrier, the associated network effect decreases as the quality uncertainty of alternative wireless carriers (learning effect). Both effects lead to a higher probability for customers to switch. In the current context, one empirical question would concern which effect is more prominent. We simulate customer decision sequences and network evolvements using our model and model estimates to calculate the network elasticity, which is the percentage change of the customer retention probability over the percentage change of network size, $elasticity = \frac{\Delta\%\,\text{Prob(staying)}}{\Delta\%\,\text{network size}}$. The elasticity estimate is 0.127 (s.d. = 0.174), which means that a 1% increase in the network size makes customers 12.7% more likely to stay. This elasticity can be decomposed into two parts: the elasticity from the social-learning effect and that from the network effect. A larger effect leads to a larger magnitude of elasticity.

We then shut down the network effect (by setting the coefficient λ to 0) and simulate the decision sequences and network evolution again. This time, the network elasticity drops to 0.038 (s.d. = 0.840). By not allowing a network effect, the change in the network size affects the retention probability only through the learning-from-others effect. A 1% change in the network size results in an increase of approximately 0.038% in the probability of staying. If we shut down the social-learning effect, we obtain the elasticity as 0.085 (s.d. = 0.156).[6] This result shows that approximately one-third of the network elasticity can be attributed to the social-learning effect. Intuitively, this result is not surprising because social-learning effect works through noisy signals received from contacts who leave the wireless carrier. Israel (2005) similarly found that a direct reduction of the lock-in utility would increase welfare more dramatically than faster learning.

We have also plotted the elasticities per each period for the total network effect, social learning, and the effect of network size (along with the 95% confidence interval bounded by dotted lines) in Figure 8. Figure 8(a) indicates the total effect of network, which is declining over time. This means as time passes, the effect of the switch of an additional contact is decreasing. Figure 8(b) indicates the effect of social learning, and Figure 8(c) indicates the effect of network size. Network size has the largest effects from initial switchers compared with later switchers on focal people's switching probability. However, for social learning, even later switchers have an impact on the focal person's switching probability. That means the effect of an additional switcher for social learning will not diminish quickly.

7.2. Targeting Based on Social Learning

Companies want to identify the potential switchers and take action to prevent it from happening. Often with limited resource, it becomes especially important to correctly rank customers according to their switching probability and target those who are mostly likely to switch. Imagine the company wants to target the top 10% ranked customers in terms of the switching probability. We compare the following three strategies and calculate the hit rate under each targeting strategy with our data. The hit rate in this case is calculated as the number of switched customers in the real data that have been picked up according to the strategy over the total number of switched customers in the data.

We use three strategies: in the first strategy, we rank the customers according to the total number of switched customers and pick the top 1% on the list as the potential switchers. In the second strategy, we rank the customers according to the ratio of total number of switched customers to the initial network size. The last strategy is to rank the customers according to the social-learning effect according to the estimation

This simulation analysis is reported in Table 6. It indicates that switching is contagious, such that the more of the contacts switch the more likely the focal customer will switch, and it also demonstrates the value of our model in helping prioritize a company's CRM effort. Using the ranking approach based on the calculated social-learning effects, which takes into consideration the strategic interactions among the focal customer and his contacts, we correctly identify more than 40% of the potential switchers.

Figure 8. (Color online) Network Elasticities over Time

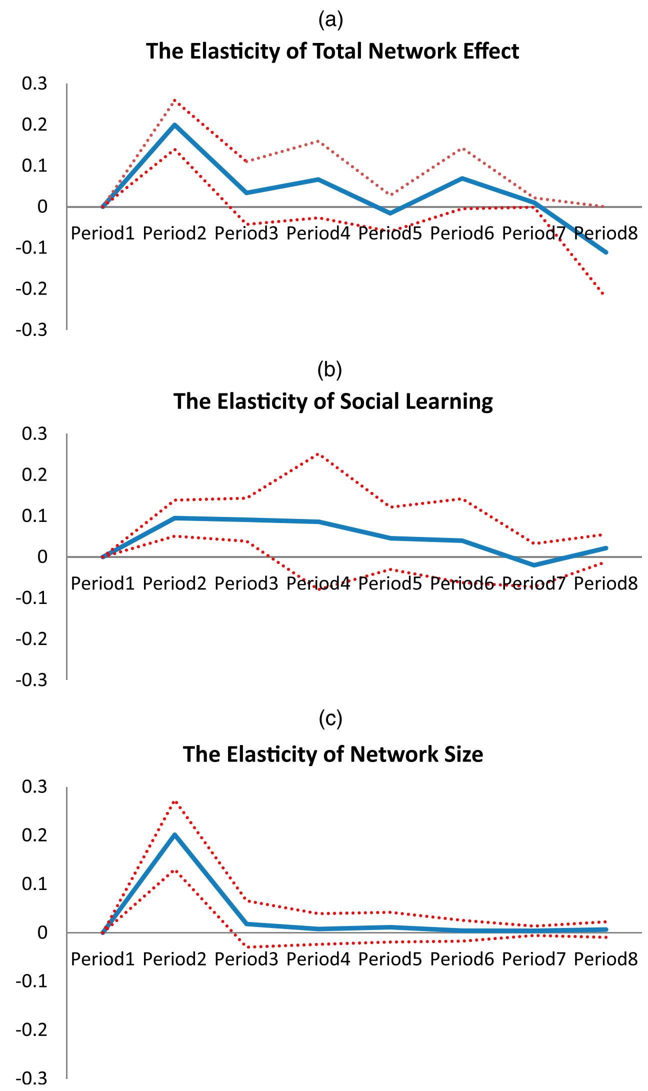


Table 6. Policy Simulation Results

	Hit rate (%)
Strategy 1	33.3
Strategy 2	23.8
Strategy 3	42.9

7.3. Closeness of Friends

The model allows us to take into consideration the heterogeneity of the relationship among different friends. We assume that different contact frequency will affect the weight attached to the signals received from different friends. We therefore conduct the following simulation. We rank the friends according to their closeness to the focal customer, as measured by contact frequency. We assume that each time a certain number of friends decide to churn, they are the highest on the ranking among the available peers on the network. Our objective is to compare the number of people who stay after certain periods.

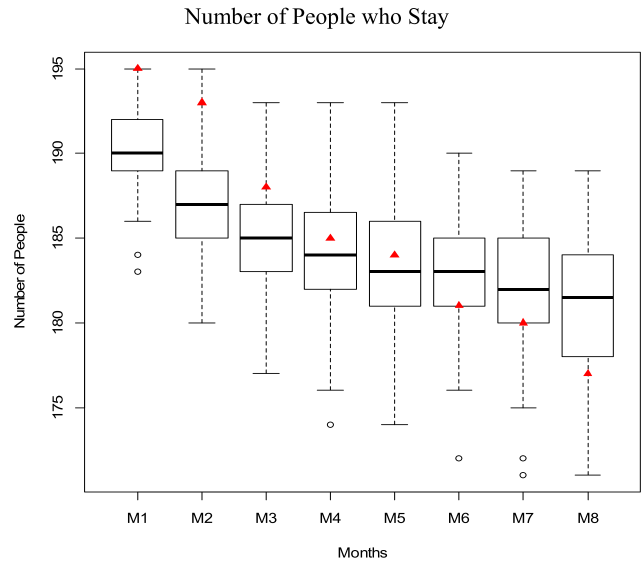
The results are shown in Figure 9. The boxplot depicts the distribution of the number of people who stay with the carrier at the end of each period under the new condition. The triangle indicates the numbers from the real data. There are significant changes in the beginning, and we observe more people churn. However, when the closeness of the friends decreases, the speed of the change decreases. The results indicate that friends are not all created equal. Using simple measures of contact frequency as the proxy for the closeness of friends, we are able to identify whose decision might weigh more for the existing customers.

8. Conclusions

This paper sets out to understand the contagious switching decisions of customers of a mobile service carrier. We focus on separating out social-learning effects from the general network effects. Our framework merges three research areas: dynamic learning effects, social-learning effects, and network effects. The unique characteristics of mobile service provides us with the opportunity to study the social-learning effect, while at the same time taking into consideration other direct benefits from network size. This can deepen our understanding of contagious switching behavior and help companies prioritize their CRM efforts. However, doing this exerts several challenges in terms of modeling and estimation.

With regard to modeling, in this setting, identification becomes more challenging. We use survey results about social learning in mobile services to motivate the model specifications. We carefully discuss identification issues, such as unobserved heterogeneity and endogenous network formation along the research line of social interaction, in Section 4. To control for unobserved correlation, we build an alternative model of social learning with correlated prior as a robustness check.

Figure 9. (Color online) Heterogeneity of Friends



With regard to estimation, we use a dynamic forward-looking model. As state variable, the expected network size is endogenously embedded in the system. To facilitate the estimation complexity, we propose a two-step dynamic programming algorithm based on the nested fixed-point algorithm proposed by Keane and Wolpin to deal with unobserved state variables, and we take advantage of the two-step models initiated by Hotz and Miller (1993) and Pakes, Ostrovsky, and Berry (2007), as well as the BBL, to solve the dynamic equilibrium problem.

We propose a dynamic structural model with strategic interpersonal interactions. To model the social-learning effect, a consumer will use feedback from his or her contacts in the focal carrier who have switched carriers to update his or her expectations of quality on alternative carriers. Our model further accounts for two unique aspects of consumers' strategic learning. (i) The signals on alternative carriers from the contacts who has switched are perceived to be systematically different by the individual, depending on whether they come from a loyal contact who switched from the focal carrier. (ii) The perceived noisiness of the signal on alternative carriers from a contact who has switched depends on the strength of the relationship between the individual and the contact. To model the remaining network effect, we allow the switching decisions to be affected by the consumer's network size directly.

We present an empirical application of this framework to data from a wireless carrier in a European country in 2008. The results indicate, after controlling for mobile plan details and demographic information, that learning from one's own usage experience, learning from contact neighbors' decisions, and the remaining network effects all have significant effects on a customer's decisions to switch mobile services. Moreover, policy simulations reveal that a 1% change in the number of contact neighbors will lead to a 12.7% change in customer retention rate and that approximately one-third of that change can be attributed to the social-learning effect. This finding suggests that strategies that target consumers on the basis of the social-learning effect would be very effective in retaining customers, because it takes into consideration the heterogeneous characteristics and interactions of contacts.

Our study has several limitations, mainly due to data restrictions. First, as with many previous studies of mobile service usage, we only have data from one carrier. Richer information about consumer switching behavior between different wireless carriers would enable modeling of industry competition in a dynamic social setting. Second, we only incorporate limited characteristics about peers (loyalty) and focal-peer relationship (contact frequency), whereas the social network literature suggests that the structure of a network also affect choice (Watts and Dodds 2007). Thus, another extension would be to incorporate richer information related to network structure into the model. Finally, with richer data, we may find significant effects from correlated learning, and researchers can quantify the magnitude of each information source that leads to similar behavior among individuals. We hope that our study can spawn further interest in understanding contagious behavior by consumers in the social-network context.

Acknowledgments
The authors thank Andrew Ching, Juanjuan Zhang, and session participants at the 2012 Marketing Science conference for helpful comments.

Appendix: Survey Questionnaire and Summary Statistics

Welcome!

Thank you for participating in our survey!

Please enter your worker ID.

____

In this survey, we are interested in your opinions about wireless carriers.

There are no right or wrong answers.

Please proceed to the next pages and share with us your true thoughts.

1. I learn about the other wireless carriers than the one I am using through the following channels (can select multiple choices):
   - Friends
   - My family members
   - Advertisements
   - Specialized websites about wireless carriers
   - Other (please specify)
2. When I know my friends switch wireless carrier, I think the reason is

   simply because they think the new carrier has better data and call quality and not because of other reasons — because reasons not just about quality (such as family plan, phone selection, customer service, retail store locations, and etc.)

   1 ----------------------2--------------------3------------------4--------------------5
3. If my friends switch wireless carrier,

   I will not discuss the new carrier with them; instead, from the behavior only, I will infer that they think the new carrier has better call and data quality — I will not infer anything based on the observation only; instead, I am very likely to discuss the new carrier with them

   1 ----------------------2--------------------3------------------4--------------------5
4. If my friends switch wireless carrier, comparing to the old one,

   I think they must be satisfied with the new one — I think they can either be satisfied or unsatisfied with the new one

   1 ----------------------2--------------------3------------------4--------------------5

5. My evaluation about the wireless carrier that I am using is

mainly based on my friends' evaluations — mainly based on my own usage experience

1 ----------------------2--------------------3-------------------4--------------------5

6. From my friends who stay with their wireless carrier,

I usually infer that the carrier has good data and call quality — I will not infer anything about the carrier based on the observation only

1 ----------------------2--------------------3-------------------4--------------------5

7. What wireless carriers are you currently using?
   - AT&T
   - Verizon
   - Sprint
   - Tmobile
   - Others (please specify)
8. What is your age?
9. What is your gender?
   - Male
   - Female
10. Are you a native speaker of English?
   - Yes
   - No

Endnotes

[1] Please refer to Ching et al. (2013, 2016) for a comprehensive literature review on dynamic learning models.

[2] We will form the expected change of network size in that period for the dynamic model in the first stage of the estimation. So the network size is updated for each period.

[3] Details are available upon request.

[4] Details are available upon request.

[5] The estimation results are available upon request.

[6] This number is similar to the estimate of network effect in the model (0.08).

References

Aguirregabiria V, Mira P (2007) Sequential estimation of dynamic discrete games. *Econometrica* 78(2):1–53.

Aral S, Walker D (2014) Tie strength, embeddedness, and social influence: A large-scale networked experiment. *Management Sci.* 60(6):1352–1370.

Bajari P, Benkard L, Levin J (2007) Estimating dynamic models of imperfect competition. *Econometrica* 75:1331–1370.

Brenton RA (2016) Corporate governance and nested authority: Cohesive network structure, actor-driven mechanisms, and the balance of power in American corporations. *Amer. J. Sociol.* 122:661–713.

Cai H, Chen Y, Fang H (2009) Observational learning: Evidence from a randomized natural field experiment. *Amer. Econom. Rev.* 99(3): 864–882.

Chamberlain G (1980) Analysis of covariance with qualitative data. *Rev. Econom. Stud.* 47(1):225–238.

Chan T, Narasimhan C, Xie Y (2013) Treatment effectiveness and side effects: A model of physician learning. *Management Sci.* 59(6): 1309–1325.

Chan TY, Li J, Pierce L (2014a) Learning from peers: Knowledge transfer and sales force productivity growth. *Marketing Sci.* 33(4): 463–484.

Chan TY, Li J, Pierce L (2014b) Compensation and peer effects in competing sales teams. *Management Sci.* 60(8):1965–1984.

Chandrasekhar A, Larreguy H, Xandri JP (2012), Testing models of social learning on networks: Evidence from a framed field experiment. Working paper, Massachusetts Institute of Technology, Cambridge.

Ching AT, Erdem T, Keane MP (2013) Invited Paper—Learning models: An assessment of progress, challenges, and new developments. *Marketing Sci.* 32(6):913–938.

Ching AT, Erdem T, Keane MP (2016) Empirical models of learning dynamics: A survey of recent developments. Rotman School of Management Working Paper No. 2810950, University of Toronto, Toronto.

Chintagunta PK, Jiang R, Jin GZ (2009) Information, learning, and drug diffusion: The case of Cox-2 inhibitors. *QME* 7(4): 399–443.

Christakis NA, Fowler JH (2007) The spread of obesity in a large social network over 32 years. *New England J. Medicine* 357(4):370–379.

Crawford GS, Shum M (2005) Uncertainty and learning in pharmaceutical demand. *Econometrica* 73:1137–1173.

Dasgupta K, Singh R, Viswanathan B, Chakraborty D, Mukherjea S, Nanavati AA, Joshi A (2008) Social ties and their relevance to churn in mobile telecom networks. *Proc. 11th Internat. Conf. Extending Database Tech.: Adv. Database Tech.* (ACM, New York), 668–677.

DeGroot MH (2005) *Optimal Statistical Decisions*, vol. 82 (John Wiley & Sons, Hoboken, NJ).

Dunne T, Klimek SD, Roberts MJ, Xu DY (2013) Entry, exit, and the determinants of market structure. *RAND J. Econom.* 44(3):462–487.

Eagle N, Pentland AS, Lazer D (2009) Inferring friendship network structure by using mobile phone data. *Proc. Natl. Acad. Sci. USA* 106(36):15274–15278.

Erdem T (1998) An empirical analysis of umbrella branding. *J. Marketing Res.* 35(3):339–351.

Erdem T, Keane MP (1996) Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Sci.* 15:1–20.

Erdem T, Keane MP, Sun B (2008) A dynamic model of brand choice when price and advertising signal product quality. *Marketing Sci.* 27(6):1111–1125.

Erdem T, Keane MP, Öncü TS, Strebel J (2005) Learning about computers: An analysis of information search and technology choice. *Quant. Marketing Econom.* 3(3):207–247.

Goldenberg J, Libai B, Muller E (2010) The chilling effects of network externalities. *Internat. J. Res. Marketing* 27(1):4–15.

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

Hotz VJ, Miller R (1993) Conditional choice probabilities and the estimation of dynamic models. *Rev. Econom. Stud.* 60: 265–289.

Israel M (2005) Services as experience goods: An empirical examination of consumer learning in automobile insurance. *Amer. Econom. Rev.* 95(5):1444–1463.

Iyengar R, Ansari A, Gupta S (2007) A model of consumer learning for service quality and usage. *J. Marketing Res.* 44(4):529–544.

Iyengar R, Van den Bulte C, Lee JY (2015) Social contagion in new product trial and repeat. *Marketing Sci.* 34(3):408–429.

Iyengar R, Van den Bulte C, Valente TW (2011) Opinion leadership and social contagion in new product diffusion. *Marketing Sci.* 30(2):195–212.

Katz M, Shapiro C (1985) Network externalities, competition and compatibility. *Amer. Econom. Rev.* 75(3):424–440.

Keane MP, Wolpin KI (1997) The career decisions of young men. *J. Political Econom.* 105(3):473–522.

Lee L (2007) Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *J. Econometrics* 140(2):333–374.

Lemon K, White T, Winer R (2002) Dynamic customer relationship management: Incorporating future considerations into the service retention decision. *J. Marketing* 66(1):1–14.

Lewis K, Gonzalez M, Kaufman J (2012) Social selection and peer influence in an online social network. *Proc. Natl. Acad. Sci. USA* 109(1):68–72.

Ma L, Krishnan R, Montgomery AL (2014) Latent homophily or social influence? An empirical analysis of purchase within a social network. *Management Sci.* 61(2):454–473.

Manchanda P, Xie Y, Youn N (2008) The role of targeted communication and contagion in product adoption. *Marketing Sci.* 27(6): 961–976.

Manski CF (1993) Identification of endogenous social effects: The reflection problem. *Rev. Econom. Stud.* 60(3):531–542.

Moretti E (2011) Social learning and peer effects in consumption: Evidence from movie sales. *Rev. Econom. Stud.* 78(1):356–393.

Nair HS, Manchanda P, Bhatia T (2010) Asymmetric social interactions in physician prescription behavior: The role of opinion leaders. *J. Marketing Res.* 47(5):883–895.

Narayan V, Rao VR, Saunders C (2011) How peer influence affects attribute preferences: A Bayesian updating mechanism. *Marketing Sci.* 30(2):368–384.

Pakes A, McGuire P (1994) Computing Markov-perfect Nash equilibria: Numerical implications of a dynamic differentiated product model. *RAND J. Econom.* 25:555–589.

Pakes A, Ostrovsky M, Berry S (2007) Simple estimators for the parameters of discrete dynamic games (with entry/exit examples). *RAND J. Econom.* 38(2):373–399.

Richter Y, EladYT, NoamS (2010) Predicting customer churn in mobile networks through analysis of social groups. *Proc. 2018 SIAM Internat. Conf. Data Mining* (SIAM, Philadelphia).

Ryan SP, Tucker C (2012) Heterogeneity and the dynamics of technology adoption. *Quant. Marketing Econom.* 10(1):63–109.

Snijders TA (2001) The statistical evaluation of social network dynamics. *Sociol. Methodology* 31(1):361–395.

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

van Zalk M, Denissen J (2015) Idiosyncratic versus social consensus approaches to personality: Self-view, perceived, and peer-view similarity. *J. Personality Soc. Psych.* 109:121–141.

Watts DJ, Dodds PS (2007) Influentials, networks, and public opinion formation. *J. Consumer Res.* 34(4):441–458.

Yang S, Allenby GM (2003) Modeling interdependent consumer preferences. *J. Marketing Res.* 40(3):282–294.

Zhang J (2010) The sound of silence: Observational learning in the US kidney market. *Marketing Sci.* 29(2):315–335.

Zhao Y, Yang S, Narayan V, Zhao Y (2013) Modeling consumer learning from online product reviews. *Marketing Sci.* 32(1):153–169.