Social interactions in customer churn decisions: The impact of relationship directionality ☆

Michael Haenlein *

Marketing Group at ESCP Europe, 79 Avenue de la République, F-75011 Paris, France

A B S T R A C T

The impact of social factors on individual-level decision making has been a subject of general interest within the marketing field. However, studies analyzing social interactions and social contagion have, to a great extent, focused on the importance of social interactions in the customer acquisition process and have relied on the use of undirected networks. Our study contributes to the literature stream by focusing on two elements that have been analyzed less frequently. Specifically, we focus on the importance of social interactions in the customer retention process within a directed social network. Using the customer base of a mobile phone provider, we rely on call detail records to investigate the churn behavior of 3431 focal actors. We provide evidence for social interactions in customer churn decisions and show that, at any given point in time, a focal actor is significantly and substantially more likely to defect from a provider if other individuals to whom that actor is socially connected have previously defected from the provider. However, this effect is limited to social contacts with whom the focal actor has *outgoing* calling relationships and who have churned relatively recently (in our sample, less than 5 weeks prior to the point in time that is under examination). We therefore provide empirical evidence demonstrating that social effects do play a role in customer retention but only when tie directionality and churn recency are taken into account.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Over the past decades, it has been widely accepted within the marketing field that consumer behavior cannot be fully understood when individual actors are studied in isolation and that social factors play an important role in shaping individual-level decision making. This perspective is particularly prevalent in the analysis of customer acquisition strategies, specifically in the area of new product adoption. Studies in those contexts have argued that socially influenced decisions occur in situations in which two or more actors make a purchase or consumption decision jointly (Hartmann, 2010) or as a result of social contagion, i.e., following exposure to other consumers' knowledge, attitudes or behaviors concerning a product (Iyengar, Van den Bulte, & Valente, 2011). Variables that measure the aspects of social exposure have therefore gained importance in models of innovation diffusion (Libai, Muller, & Peres, 2013; Nair, Manchanda, & Bhatia, 2010).

The vast majority of this research has focused on the importance of social interactions in the customer *acquisition* process and has relied on the use of *undirected* networks, i.e., networks that consider the connections between individuals without taking into account who initiates each connection. Our study contributes to the literature by looking at

two elements that have been analyzed less frequently. Specifically, we investigate the importance of social interactions in the customer *retention* process within a *directed* social network, i.e., a network in which the initiator of each connection is identified. Our analysis uses the call history database of a mobile phone operator to analyze the social interactions in customer churn decisions. Specifically, we look at situations in which two or more people who are socially connected decide to leave a company, either simultaneously or in close temporal proximity to one another. Our approach is consistent with a larger body of research that relies on electronically recorded communication patterns to approximate social relationships within a larger network (e.g., Goldenberg, Han, Lehmann, & Hong, 2009; Hill, Provost, & Volinsky, 2006; Nitzan & Libai, 2011; Trusov, Bucklin, & Pauwels, 2009). For our analysis, we used a directed network that allowed us to differentiate between the impact of churn among *outgoing* calling relationships (i.e., a focal actor calls another actor in the network or an *alter*) and among *incoming* calling relationships (i.e., an alter calls a focal actor).

Our results show that, at any given point in time, a focal actor is significantly and substantially more likely to defect from a provider if other individuals to whom that actor is socially connected have previously defected from the provider. However, this effect is limited to social contacts with whom the focal actor has *outgoing* calling relationships and who churned relatively recently (in our sample, less than 5 weeks prior to the point in time that is under examination). We also show that churn among a focal actor's social contacts has a nonlinear effect on the hazard that the focal actor will defect as well;

specifically, a focal actor's hazard of defection increases quadratically with the number of recent churn events among his or her outgoing calling relationships.

2. Contribution compared to prior research

By focusing on the importance of social interactions in the customer *retention* process within a *directed* social network, our work provides the following two contributions compared to previously published work: First, prior marketing literature on the importance of social influence has focused extensively on customer acquisition, specifically on new product adoption (e.g., Rogers, 2003; Van den Bulte & Wuyts, 2007; Wuyts, Dekimpe, Gijsbrechts, & Pieters, 2010). Relatively few studies have looked at the impact of social networks on consumer behavior once customers have been acquired, i.e., behavior that is contingent on an initial product or brand choice decision. Some notable exceptions include the work of Haenlein (2011), who looks into the social network effects in post-acquisition product usage and customer-level revenue and Nitzan and Libai (2011), who analyze social contagion in the customer retention process.

The lack of research regarding social influence on customer retention is surprising, because customer acquisition and relationship initiation are only one part of a firm's CRM process (Reinartz, Krafft, & Hoyer, 2004), and the analysis of factors influencing retention, relationship duration, and customer churn is an active area of research (e.g., Bolton, 1998; Lemmens & Croux, 2006; Neslin, Gupta, Kamakura, Lu, & Mason, 2006; Schweidel, Fader, & Bradlow, 2008). In addition, studies that have been conducted outside of the marketing field provide clear indications that social effects play an important role not only in initiation but also in termination decisions. Medical research, for example, has highlighted the importance of social effects in the cessation of smoking (Christakis & Fowler, 2008) and of heroin and cocaine consumption (Buchanana & Latkina, 2008). Likewise, studies on human resource management have stressed the importance of communication networks in the context of collective turnover decisions (Bartunek, Huang, & Walsh, 2008; Castilla, 2005; Krackhardt & Porter, 1986). Similar effects have been found in studies of decisions regarding deserting the army (De Paula, 2009) and marriage dissolution (McDermott, Fowler, & Christakis, 2009). In light of these findings, it seems likely that social effects should also influence a customer's decision to remain loyal to a certain company, and, indeed, Nitzan and Libai (2011) have provided empirical evidence that this is the case. Our study continues this stream of research by providing further insight into the dynamics that underlie social influence in the customer retention process.

Second, we contribute to the prior literature by analyzing social influence within a directed (versus an undirected) network. Across a large range of disciplines, including organizational science, finance, psychology and marketing, studies that analyze social networks have focused nearly exclusively on undirected networks. In marketing specifically, this is true for analytical (Zubcsek & Sarvary, 2011) and empirical research (Trusov, Bodapati, & Bucklin, 2010) and applies to social interactions in processes of customer acquisition (Nair et al., 2010) as well as retention (Nitzan & Libai, 2011). In fact, an analysis of the journals that are covered in the EBSCO database reveals only a handful of articles that rely on directed social networks and that specifically compare the difference in influence between the number of incoming connections (indegree) and the number of outgoing connections (outdegree). See Table 1 for an overview of these studies.

Interestingly, the studies that have focused on directed networks reveal that while the impact of incoming and outgoing relationships is comparable in some settings, it is not the case in all situations. This empirical finding can be explained by the fact that, depending on the situation, incoming and outgoing relationships can have fundamentally different meanings. For example, *considering* another person as a friend is not necessarily identical to *being considered* as a friend, and analyzing both types of relationships in the same manner carries the risk

of confounding different effects. Empirical evidence for this finding can be found in Christakis and Fowler (2007), who analyze the spread of obesity within a social network and show that results differ depending on whether A perceives B as a friend, B perceives A as a friend, or both A and B perceive each other as friends.

An additional potential benefit of analyses that rely on directed networks is that they entail fewer identification issues and reflection biases compared with analyses of undirected networks, and therefore they can provide stronger evidence for social contagion. Within an undirected network, it is sometimes difficult to establish whether correlation in behavior is due to social contagion or to other factors, such as endogenous group formation, correlated unobservable variables or simultaneity (Manski, 1993; Moffitt, 2001). Within a directed network, it is possible to compare the impact of incoming versus outgoing relationships. If the two types of relationships differ, for example, in their influence on the time of adoption or time to churn, it can be seen as stronger evidence for social effects (Bramoulle, Djebbari, & Fortin, 2009). Christakis and Fowler (2007), for example, used this approach to rule out the possibility that the spread of obesity is driven by unobserved environmental factors rather than by social contagion. Our study builds on this line of thinking by comparing the differential effects of incoming versus outgoing relationships on the retention process.

3. Relationship directionality and social interactions in customer churn decisions

The occurrence of social influence in customer churn decisions may stem from a variety of factors: On the one hand, the decision of a customer to leave her current supplier in response to the departure of one of her social contacts might be driven by economic or utilitarian considerations, often resulting from naturally occurring or artificially induced network externality effects. In the context of the cellular phone industry, for example, which corresponds to our empirical setting, companies frequently charge different prices for in-network versus out-of-network traffic, and it has been shown that such differential pricing influences customers' choices of supplier as well as their subsequent usage behavior (Gerpott, 2008; Hoernig, 2007; Staahl Gabrielsen & Vagstad, 2008). On the other hand, social influence in customer churn decisions might result from social factors. The churn of a friend can, for example, lead to the loss of certain social benefits that arise from patronizing the same company, which is consistent with what has been discussed from the perspective of social enrichment (Fernandez, Castilla, & Moore, 2000; Schmitt, Skiera, & Van den Bulte, 2011) and balance theory (Krackhardt & Porter, 1985; Nitzan & Libai, 2011). In most real-life situations, economic and/or utilitarian factors and social factors can be expected to influence decision making simultaneously.

The use of a directed network to analyze social interactions in customer churn decisions can provide valuable insights regarding the drivers of such decisions. It seems likely that the impact of churn among incoming and outgoing calling relationships is not identical and that relationship directionality plays an important moderating role. Two arguments can be used to justify such thinking: First, churn among incoming versus outgoing calling relationships can be associated with different sets of economic and/or utilitarian and social consequences for the focal actor. Looking at the economic and/or utilitarian reasons, a churn event among outgoing calling relationships can represent a financial burden for a focal actor, which may increase the hazard that the focal actor will defect. This is the case when the cellular service provider follows the "calling party pays" principle, according to which the mobile phone subscriber does not pay for incoming calls but only pays for outgoing calls (Valletti & Houpis, 2005) and when out-of-network calls are priced higher than in-network calls. In such a situation, the churn of a party from whom a focal actor receives only incoming calls has no financial consequences for the focal actor, as receiving calls does not generate any cost, regardless of whether those calls stem from within

Table 1
Selective review of papers investigating influence within directed networks.

Source	Network type	Indegree	Outdegree	Dependent variable	Key finding
Haunschild and Beckman (1998)	4960 medium- and large-sized firms in four industry sectors	Number of outside directors who sit on the focal firm's board	Number of directors from the focal firm who sit on an outside board	Number of acquisitions completed by the focal firm	Sig. positive impact of outdegree; Impact of indegree not sig.
Hansen (1999, 2002)	41 divisions of a multinational electronics and computer company	Number of divisions that nominate focal division as a source of advice	Number of other divisions that the focal division nominates as a source of advice	Project completion time	Hansen (1999): Impact of indegree and outdegree not sig. Hansen (2002): Sig. negative impact of indegree and outdegree
Hochberg, Ljungqvist, and Lu (2007)	3469 venture capital (VC) funds headquartered in the United States	Number of times a VC firm is invited to co-invest in other VCs' deals	Number of other VCs a VC firm has invited into its own syndicates	Fund performance and Exit rate	Fund performance: Sig. positive impact of indegree and outdegree; Exit rate: Sig. positive impact of indegree
Mercer and Derosier (2010)	1016 third grade students in 11 schools	Number of received friendship nominations in fall (October)	Number of sent friendship nominations in fall (October)	Number of new friendship nominations received in spring (April)	Sig. negative impact of outdegree; sig. positive impact of indegree
Iyengar et al. (2011)	193 physicians in three US cities	Number of other physicians who nominate a particular physician for referral or discussion	Number of nominations given by a particular physician to other physicians for referral or discussion	Adoption and post-adoption usage of newly launched prescription drug	Sig. positive impact of indegree; Impact of outdegree not sig.

Note: Based on all of the articles that were covered in the EBSCO database (Academic Search Premier, Alt Health Watch, Business Source, Business Source Premier, Communication & Mass Media Complete, EconLit, Health Source: Nursing/Academic Edition, MEDLINE, PsycARTICLES, Psychology and Behavioral Sciences Collection, PsycINFO, Sociological Collection) where the full text includes the expressions "social network," "indegree" and "outdegree." Excludes conceptual articles, descriptive analyses, manuscripts that do not report separate results for indegree and outdegree, conference proceedings and papers not written in English language.

the provider's network or from a competing network. The situation is different for outgoing calls, as a focal actor incurs higher costs following the churn of a person whom he or she calls, as communication is subsequently priced according to out-of-network fees rather than in-network fees, which are generally cheaper.[1]

Regarding the social factors, outgoing calling relationships are purposefully initiated by a focal actor. While people cannot necessarily control who calls them, they have full control over the calls that they place themselves. In placing a call to another person, a focal actor essentially signals that the person is "worth the effort" of a call (versus waiting until the other person calls of his or her own accord) and therefore might be more important to the focal actor compared with the social contacts from whom the focal actor only receives calls. Outgoing calling relationships therefore have a different meaning attached to them in comparison to incoming ones and represent links to others who might have a more persuasive influence on the focal person (Hall & Valente, 2007). If this is indeed the case, a focal actor should be more likely to follow suit after the defection of an outgoing calling relationship as opposed to an incoming one.

Further, one should note that a directed network constitutes a more accurate representation of reality compared with an undirected network. Many of the processes that underlie the formation of social networks are directional in nature. Networks of referral (Reingen & Kernan, 1986) or advice (Lazega & Duijn, 1997), for example, build on notions of status and prestige. These concepts are related to the number of incoming relationships or indegree, i.e., the number of people who seek information from A, but not the number of outgoing relationships or outdegree, i.e., the number of people from whom A seeks information (Iyengar et al., 2011; Van den Bulte & Wuyts, 2007). Models that are designed to describe the formation of such networks, such as exponential random graph (p*) models (Holland & Leinhardt, 1981) or

the Jackson–Rogers approach (Jackson & Rogers, 2007) therefore take explicit account of the direction of a relationship. Transforming an inherently directed network into an undirected one involves the addition of certain links that have were not initially present (e.g., A → B becomes A–B, which in turn is equivalent to A → B and B → A). This transformation represents a distortion of the underlying social network structure, which could bias subsequent analyses that are based on an undirected network.

4. Overview of the main constructs and variable definitions

Our study investigates social interactions in customer churn decisions in the context of the cellular phone industry and specifically the different affects of churn among incoming versus outgoing calling relationships. For the purpose of our analysis, we define *churn* as the termination of an existing service agreement between a customer and his or her cellular phone company. Our analysis focuses only on customers who maintain a contractual service agreement with their cellular phone companies and, therefore, excludes non-contractual relationships (prepaid mobile phones). We differentiate between two types of churn – *voluntary churn* (i.e., cases in which the decision to terminate the service agreement is made by the customer) and *involuntary churn* (i.e., cases in which the decision is made by the cellular phone provider) – because different causes of churn have been shown to evolve differently over time (Braun & Schweidel, 2011).

We focus our analysis on situations in which two or more customers between whom there is a calling relationship – referred to herein as *social contacts* – decide to leave a company either simultaneously or in close temporal proximity to one another. We define three types of calling relationships: focal customer A *receives* calls from customer B (an *incoming calling relationship* for person A); focal customer A *calls* customer B (an *outgoing calling relationship* for person A); or customers A and B call each other (a *mutual calling relationship* for person A). The relationship types are not mutually exclusive; e.g., two customers who have a mutual calling relationship with each other also have incoming and outgoing calling relationships with each other. We limit our analysis to first-order effects, i.e., relationships between a focal actor and the other actors who either call or are called by the focal actor directly. Potential higher-order effects (e.g., the impact of

---

[1] To test the difference in pricing between in-network and out-of-network calls in our sample, we performed a regression analysis in which we explained the amount of money that was paid by a customer through the number of outgoing minutes that were spent in-network and out-of-network. Our results indicate that after taking account of a series of dummy variables that indicated service plan choice, out-of-network calls are 2 to 2.5 times more expensive than in-network calls.

churn of actor C in a calling relationship in which person A calls person B and person B calls person C) are beyond the scope of our analysis.

To operationalize the extent of churn among incoming and outgoing calling relationships, we used a two-step approach: First, we collected data on all of the incoming and outgoing calls that were made by a sample of focal actors over a two-month period (time period 1). Based on this information and following the work of Onnela et al. (2007) and Nitzan and Libai (2011), we define two types of *tie strength*, measuring the strength of the incoming or outgoing calling relationship between focal actor i and alter j, who is also a customer of the provider, as follows:

$$tie\_incoming_{i,j} = \frac{Incoming\ communication\ volume\ between\ i\ and\ j}{Total\ incoming\ communication\ volume\ for\ i} \quad (1)$$

$$tie\_outgoing_{i,j} = \frac{Outgoing\ communication\ volume\ between\ i\ and\ j}{Total\ outgoing\ communication\ volume\ for\ i}. \quad (2)$$

The *incoming* (or *outgoing*) *communication volume* between focal actor i and alter j represents the total number of minutes that were spent by focal actor i on incoming calls from customer j (or on outgoing calls to customer j) within time period 1. The total incoming and outgoing communication volume for focal actor i is the total number of minutes of all of the calls that were placed or received by i within time period 1, including calls to and from the customers of the focal provider, customers of other cellular providers, and landlines. In the case of mutual calling relationships, we determine tie strength separately for the incoming and outgoing calling relationship components. The separate determination of tie strength allows us to disentangle the impact of the strength of the incoming relationship between focal actor A and customer B (i.e., the share of person B among all of the calls that were *received* by actor A) from that of the strength of the outgoing relationship between the two parties (i.e., the share of person B among all of the calls that were *made* by actor A).

We subsequently collected data on the churn behavior of all of the focal actors as well as on the other customers with whom they maintained incoming and/or outgoing calling relationships over a second period of six months (time period 2). Based on this information, we determined two time-dependent variables for each focal actor that measure the cumulative impact of churn events that focal actor i was exposed to among his or her incoming and outgoing calling relationships, between the start of time period 2 (t_0) and time t. These variables are defined in the following way:

$$s_{in,i}(t) = \sum_j \sum_{t_0}^t c_j w\left(t - t\_churn_j\right) tie\_incoming_{i,j} \quad (3)$$

$$s_{out,i}(t) = \sum_j \sum_{t_0}^t c_j w\left(t - t\_churn_j\right) tie\_outgoing_{i,j} \quad (4)$$

where c_j is an indicator variable that equals one if customer j churned prior to or at t and zero otherwise; t_churn_j is the time at which customer j churned; and $w()$ ($0 \leq w() \leq 1$) is a weighting function that weights churn events according to the difference between t and t_churn_j.

Our definition of $s_{in,i}(t)$ and $s_{out,i}(t)$ is based on two assumptions: First, we expect the impact of churn among incoming and outgoing calling relationships on focal actor churn to be stronger with increasing levels of tie strength. This assumption is consistent with the observation that individuals are more affected by others with whom they have closer relationships (Brown & Reingen, 1987). It is also consistent with the research on referral programs, which has shown that referral behavior among social contacts is influenced by tie strength (Ryu & Feick, 2007). We account for this effect in the definition of $s_{in,i}(t)$ and $s_{out,i}(t)$ by summing over the product of the churn indicator (c_j) and the tie strength instead of simply summing over c_j, which would be equivalent to weighting all of the churn events identically.

Second, we expect that churn events among incoming or outgoing calling relationships that happened a longer time ago should have a lesser impact on focal actor churn rates compared with churn events that occurred relatively recently. Such an assumption is consistent with the observation that the effect of word-of-mouth decays exponentially over time for both customer acquisition (Strang & Tuma, 1993; Trusov et al., 2009) and customer retention (Nitzan & Libai, 2011). This is reflected in Eqs. (3) and (4) by requiring that the weighting function $w()$ decreases with an increase in the difference between t and t_churn_j. Combined with the fact that $w()$ should be bound between zero and one, this condition implies that $w(0) = 1$ and $\lim_{t \to \infty} w(t - t\_churn_j) = 0$. For our analysis, we choose a weighting function that is based on the Gompertz function and that can accommodate a large variety of different shapes, ranging from the equal weighting of all churn events at one end to exponential decay at the other (see Appendix A for a more elaborate explanation):

$$w(x) = \frac{1}{1 - e^{\alpha}}\left(1 - e^{\alpha e^{\beta x}}\right) with\ \alpha, \beta < 0. \quad (5)$$

Using such a weighting function results in the following operationalization of $s_{in,i}(t)$ and $s_{out,i}(t)$:

$$s_{in,i}(t) = \sum_j \sum_{t_0}^t c_j \frac{1}{1 - e^{\alpha}}\left[1 - e^{\alpha e^{\beta\left(t - t_{churnj}\right)}}\right] tie\_incoming_{i,j} \quad (6)$$

$$s_{out,i}(t) = \sum_j \sum_{t_0}^t c_j \frac{1}{1 - e^{\alpha}}\left[1 - e^{\alpha e^{\beta\left(t - t_{churnj}\right)}}\right] tie\_outgoing_{i,j}. \quad (7)$$

See Appendix B for a numerical example that illustrates the calculation of $s_{in,i}(t)$ and $s_{out,i}(t)$.

5. Research model

To test the extent to which churn among incoming and outgoing calling relationships influences focal actor churn, we estimated a stratified Cox Proportional Hazards model in which we included the two previously defined social impact measures, $s_{in,i}(t)$ and $s_{out,i}(t)$, as time-dependent covariates.

Hazard rate models are the method of choice when dealing with right-censored duration times (Helsen & Schmittlein, 1993). Taking account of potential censoring is particularly important when analyzing social influence because truncation can trick researchers into believing that they have found evidence of social contagion when there is actually none (Van den Bulte & Iyengar, 2011). The dependent variable in a hazard model is the hazard rate $h(t)$, which describes the conditional probability that an event (e.g., churn) will occur at time t, given that it did not occur prior to t:

$$h(t) = \frac{f(t)}{1 - F(t)}. \quad (8)$$

where $f(t)$ and $F(t)$ describe, respectively, the probability density function and the cumulative density function of the duration time process.

A (semi-parametric) Cox Proportional Hazards model (Cox, 1972) assumes that $h(t|x_{it})$ – the hazard rate of an event at time t for actor i, who is characterized by a vector of covariates x_{it} – can be expressed as the product of an unspecified baseline hazard function $h_0(t)$ at time t, which is identical for all individuals, and the relative risk or hazard ratio $\exp(\gamma x_{it})$. We use a variation on the traditional Cox Proportional Hazards model, the stratified Cox Proportional Hazards model (Lunn & McNeil, 1995; Wei, Lin, & Weissfeld, 1989), which allows for the incorporation of multiple unspecified baseline hazard functions, which are associated with different types of events. This model enables us to

analyze the two competing risks of voluntary and involuntary churn separately, using the following functional form:

$$h(t|x_{it}) = [j\, h_1(t) + (1-j)h_2(t)]e^{\gamma x_{it}}, \qquad (9)$$

where j is an indicator variable that equals one when churn is voluntary and zero otherwise, and $h_1(t)$ and $h_2(t)$ are the baseline hazard functions for voluntary churn and involuntary churn, respectively.

As highlighted above, working with a directed network protects us to some extent from identification issues and reflection bias, which are a common problem in the analysis of social influence (Manski, 1993; Moffitt, 2001). The inclusion of a non-parametric baseline hazard in our model also protects against identification issues (Lin & Wei, 1989).[2] To minimize such issues even further, we include three types of control variables in our model: First, we account for the fact that people prefer to interact with others who are similar to themselves (i.e., homophily, see McPherson, Smith-Lovin, & Cook, 2001; Reagans, 2005), suggesting that people who share similar characteristics – and in particular similar consumption patterns – are more likely to be socially connected to one another. For this factor, we obtained information on actors' observable characteristics, including basic demographics (i.e., age, gender, marital status) and socio-demographics (i.e., membership in one of 13 socio-demographic groups), and included them as time-independent control variables in our model.

Second, we controlled for correlated unobservable variables that could cause similarities in behavior among social contacts. Several types of correlated unobservable variables have been discussed in the social network literature, including sales force activities (Nair et al., 2010), advertising and detailing (Van den Bulte & Lilien, 2001) and (geographical) neighborhood effects (Moffitt, 2001). To control for bias that arises from correlated unobservable variables, we obtained information on the geographic location (ZIP-code) of each actor within our sample. Given that our sample included a substantial number of different ZIP-codes, we grouped adjacent codes into larger areas. We subsequently included a full set of area-specific fixed effects in our model to take account of the similarities in exposure to marketing efforts, network quality, preferences, and other effects that result from geographical proximity.

Third, we take account of other factors that have traditionally been shown to influence individual-level churn behavior, specifically, acquisition channel and contract characteristics (i.e., service plan and contract duration[3]). Prior research has shown the influence of acquisition channel on loyalty and cross-buying (Verhoef & Donkers, 2005), the interrelationship between current and future usage levels (Bolton & Lemon, 1999), and the duration-dependence of churn and retention rates (e.g., Fader & Hardie, 2007; Schweidel et al., 2008). To account for the potential influence of these variables, we included an additional series of time-independent dummy variables that indicate acquisition channel and service plan choice. Contract duration was included as a time-dependent covariate and operationalized by the number of months that elapsed since the start of the contract.

The considerations that are outlined above lead to the following specification of the covariate vector x_{it} that is included in Eq. (9):

$$x_{it} = \gamma_1 z_i + \gamma_2 d_i(t) + \gamma_3 s_{in,i}(t) + \gamma_4 s_{out,i}(t) + \gamma_5 s_{in,i}(t)^2 + \gamma_6 s_{out,i}(t)^2, \qquad (10)$$

where z_i is a vector of time-independent control variables (comprising age, gender, marital status, socio-demographic group membership, geographical area, acquisition channel and service plan); $d_i(t)$ is a time-dependent covariate that measures contract duration at time t; and $s_{in,i}(t)$ and $s_{out,i}(t)$ are the social impact measures that are defined above for churn among the incoming and outgoing calling relationships.

Note that Eq. (9) allows for a quadratic influence of the two social impact measures $s_{in,i}(t)$ and $s_{out,i}(t)$, which enables us to take account of potential nonlinearities in returns to scale of social influence. For example, a focal actor who experiences the churn of multiple social contacts might show super-additive effects (i.e., a combined effect that is larger than the sum of the individual effects) if one loss reinforces and legitimizes the other, or sub-additive effects (i.e., a combined effect that is smaller than the sum of the individual effects) if each additional effect provides less incremental information than the previous one. Quadratic effects allow us to account for such mechanisms to a certain extent.

6. Data collection

Our study is integrated into a larger body of research that uses electronically recorded communication patterns to approximate social relationships within a larger network. The substantial effort that has traditionally been associated with collecting social network data (Reingen & Kernan, 1986), combined with the inherent measurement bias of self-reported information on social relationships (Feld & Carter, 2002), has motivated researchers to identify alternative data collection techniques for social network analysis. Recently, the use of large-scale databases has gained increasing popularity. Examples of applications in the marketing field include the work of Hill et al. (2006) and Goldenberg et al. (2009) in the area of new product adoption, as well as that of Trusov et al. (2009), who compared the effectiveness of word-of-mouth and traditional advertising. Our analysis follows this stream of research and, similarly to Nitzan and Libai (2011) and Haenlein (2011), uses the call history database of a cellular service provider to reconstruct the social network of a set of customers.

Specifically, our dataset stems from a mobile phone operator in a major European country and was collected between August 2006 and April 2007. In that period, there were five major mobile phone operators in this market, four (including our company) with market shares between 22% and 26% and one with a market share of 4%. The market is among the top ten markets for mobile phone services in Europe and can be considered to be very mature, with a penetration rate in excess of 100%. As is the case for most mobile phone operators in Europe, the billing scheme that is applied by our company follows the "calling party pays" principle. This implies that the mobile phone subscriber does not pay for incoming calls and only pays for outgoing calls (Valletti & Houpis, 2005).

Our main data source was the customer database of the mobile phone operator, which includes basic information on each customer (e.g., acquisition date) and information on all of the incoming and outgoing calls that were made (i.e., calling party, calling time and call duration).[4] Our first step in collecting the data was to select a random sample of customers from the database. Because every customer

---

[2] As has been discussed by Manski (1993) and Moffitt (2001), among others, the analysis of social contagion is subject to a series of potential biases that are introduced by simultaneity (i.e., person A's actions affect person B's actions and vice versa), correlated unobservable variables (i.e., omitted variables that affect the actions of A and B) and endogenous tie formation (i.e., the formation of ties that are based on common behavior). As shown by Lin and Wei (1989), Cox Proportional Hazards models with a non-parametric baseline hazard provide robust parameter estimates, even if the model is misspecified. This finding implies that the inclusion of a non-parametric baseline hazard prevents these problems to some extent, including, specifically, the bias that is introduced due to omitted correlated unobservable variables.

[3] Mobile phone providers (including the one we collaborated with) usually sell a wide range of different service plans that represent contractual relationships of a fixed duration. Once a contract expires, the user has the option to subscribe to a new contract. Contract duration, which is the duration of the current contract, therefore differs from the total duration of the customer–company relationship (i.e., the time elapsed since acquisition).

[4] Note that while we know when each customer has been acquired (i.e., how long she or he has been a customer of the mobile phone operator) and when the current contract the customer is subject to was established, we do not know how much time is left on the current contract. We return to this caveat when we discuss the limitations of our analysis.

can be uniquely identified by his or her mobile phone number, we generated a list of 1.25 million random numbers and matched them to the customer database. This process resulted in a random sample of 4163 customers (100%). We then removed 424 customers from our dataset who had joined the cellular service after August 1, 2006 (the beginning of the data collection timeframe) and for whom, therefore, only partial information was available, which resulted in 3739 customers (90%). Out of these 3739 customers, 3450 customers (83%) could be matched to a second database containing churn and revenue data, and of these, 3431 customers (82%) could be matched to a third database that contains demographic data. For these 3431 customers, we downloaded information about all of the incoming calls that were received and outgoing calls that were made (phone number and duration of call) over a two-month time period (August 1, 2006 to September 30, 2006, time period 1). This information was subsequently used to define the incoming and outgoing calling relationships and to determine tie strength according to Eqs. (1) and (2).

Our network, which represents the egocentric networks of the final sample of 3431 focal actors, includes 19,668 nodes that are linked through 25,799 calling relationships, or ties. On average, over the course of time period 1, each focal actor maintained 1.07 exclusively incoming calling relationships (corresponding to alters who call the actor but whose calls are not reciprocated), 1.01 exclusively outgoing calling relationships (corresponding to alters who are called by an actor but who do not reciprocate the call) and 2.72 mutual calling relationships (see Table 2 for a full distribution of the number of calling relationships per focal actor).[5] The total number of calling relationships that were maintained by each focal actor is hence 4.8 (1.07 + 1.01 + 2.72). While this finding might appear to be relatively low, it should be noted that our analysis only covers calling relationships in which *both* partners are customers of the focal cellular service provider (a necessary condition that ensures our ability to evaluate the churn behavior of both parties). Tie strength ranges from a minimum of 0% to a maximum of 100% for both incoming and outgoing relationships. The average tie strength in our sample was 7.1% for incoming relationships and 4.2% for outgoing relationships.

A total of 547 focal actors churned between November 1, 2006, and April 30, 2007 (time period 2), which corresponds to a churn rate of 15.9% (the overall churn rate of the company during this period was approximately 15.5%). Approximately 80% (447) of these churn events were voluntary, and 20% (100) were involuntary; the majority of the latter (95) were terminations that were due to bad debt (i.e., non-payment of outstanding bills). Among the 3431 focal actors, 1213 (35.4%) had at least one social contact who churned within the observation period. Out of those, 155 focal customers subsequently churned themselves, which represents a total of 28.3% of all of the 547 focal actors who churned. In terms of descriptive statistics, the average age was 36.2 years for voluntary churners, 33.5 years for involuntary churners, 38.9 years for non-churners and 38.4 years overall. The average contract duration was 1.4 years (504 days) for voluntary churners, 0.9 years (326 days) for involuntary churners, 1.4 years (500 days) for non-churners and 1.4 years (495 days) overall. The average duration of staying with the company (i.e., time elapsed since acquisition) was 2.3 years (858 days) for voluntary churners, 1.8 years (651 days) for involuntary churners, 4.2 years (1535 days) for non-churners and 3.9 years (1421 days) overall. A full list of descriptive sample statistics can be found in the Web Appendix.

Table 2
Distribution of number of calling relationships.

	Number of actors			
	Voluntary churn	Involuntary churn	No churn	Total
Number of incoming calling relationships				
0	214	44	1277	1535
1	137	30	910	1077
2	51	11	377	439
3	25	10	153	188
4	7	4	71	82
5	3	1	32	36
6	3	0	22	25
7	1	0	13	14
8	2	0	9	11
9+	4	0	20	24
Total	447	100	2884	3431
Number of outgoing calling relationships				
0	231	40	1341	1612
1	141	28	873	1042
2	38	13	348	399
3	21	7	164	192
4	7	6	70	83
5	4	0	41	45
6	1	3	20	24
7	2	2	9	13
8	1	0	7	8
9+	1	1	11	13
Total	447	100	2884	3431
Number of mutual calling relationships				
0	68	21	354	443
1	135	24	794	953
2	98	22	568	688
3	57	6	399	462
4	39	9	250	298
5	17	5	175	197
6	12	3	103	118
7	9	3	63	75
8	2	2	45	49
9+	10	5	133	148
Total	447	100	2884	3431

7. Estimation results

We estimated the model that is specified in Eqs. (1)–(2), (6)–(7) and (9)–(10) using the survival package (Version 2.35-4) within the R computing environment (Version 2.12.1). Specifically, we encapsulated the likelihood of the stratified Cox Proportional Hazards model in a constrained optimization algorithm to identify optimal values for α and β ($\alpha, \beta < 0$) that result in maximum likelihood, given an estimation of γ contingent on α and β. The optimization algorithm that was used is a limited-memory modification of the BFGS quasi-Newton method.

7.1. Model comparison

We specified and compared five different models (elaborated in Table 3). *Model 1* (*Model 2*) incorporates no-churn recency and assumes that the two social impact measures influence the hazard of churn in a linear (quadratic) way; *Model 3* is a baseline model that includes only the covariates without any consideration of social impact measures; *Model 4* and *Model 5* mirror *Model 1* and *Model 2* but allow for churn recency effects. Table 3 shows the fit of our final model (*Model 5*), as well as of the rival model specifications. Model fit was measured by the Bayesian information criterion (BIC) and by the Akaike information criterion (AIC). Two findings are of particular interest.

---

[5] In total, we observed 3674 exclusively incoming calling relationships, 3475 exclusively outgoing calling relationships and 9325 mutual calling relationships. The total number of ties (25,799) was equal to the total number of incoming ties (3674 + 9325 = 12,999) plus the total number of outgoing ties (3475 + 9325 = 12,800).

Table 3
Model fit.

ID	Model		Log-likelihood	BIC	AIC
Model 1	No churn recency	Linear social impact	−2372.09	5973.55	5134.19
Model 2	No churn recency	Quadratic social impact	−2364.89	5971.76	5123.79
Model 3	No churn recency	Covariates only	−2374.11	5964.97	5134.21
Model 4	Churn recency	Linear social impact	−2360.48	5950.32	5110.95
Model 5	Churn recency	Quadratic social impact	−2350.48	5942.93	5094.95

Note: *Models 1–4* can be considered as constrained versions of the more general *Model 5*, which is specified in Eqs. (1)–(2), (6)–(7) and (9)–(10). Specifically, *Models 1–3* ("No churn recency") assume $\beta = 0$ in Eqs. (6) and (7); *Model 3* ("Covariates only") assumes $\gamma_3 = \gamma_4 = \gamma_5 = \gamma_6 = 0$ in Eq. (10); and *Model 1 and Model 4* ("Linear social impact") assume $\gamma_5 = \gamma_6 = 0$ in Eq. (10).

First, models that take churn recency into account (*Model 4 and Model 5*), i.e., which weight churn events differently depending on how long ago they occurred, show better fit than models that do not (*Models 1–2*). Interestingly, when churn recency is not explicitly accounted for, a model without any social impact measures (*Model 3*) fits the data better than one that takes churn into account among incoming and outgoing calling relationships (*Models 1–2*). This surprising result can be explained as follows: As we will show later, only churn events that occur within a relatively small window of time are significantly related to focal actor churn. Considering all churn events to be independent of their recency therefore introduces a substantial amount of noise into the calculation of the social impact measures. Our results show that the inclusion of these noisy variables into the model deteriorates model fit instead of improving it.

Second, once churn recency is taken into account, the inclusion of churn among incoming and outgoing relationships does improve model fit. This finding implies that, in line with our basic hypothesis, social interactions affect customer churn decisions. The best-fitting model is *Model 5*, which includes the two social impact measures in quadratic form. This finding implies that the influence of social impact measures on focal actor churn is nonlinear. *Model 5* has the lowest BIC and AIC of all of the models and, as indicated by an analysis of deviance, it represents a significant improvement compared to both *Model 3* and *Model 4*.[6]

7.2. Churn recency

The optimal values for the weighting function that result in maximum likelihood for the stratified Cox Proportional Hazards model are $\alpha = -28.00$ and $\beta = -1.20$. Fig. 1 illustrates the shape of the weighting function that corresponds to these two values. The churn events that occurred less than two weeks prior to t are accounted for nearly to their full extent in the calculation of $s_{in,\,i}(t)$ and $s_{out,\,i}(t)$ ($w > 0.90$), while the churn events that happened more than five weeks prior to t have essentially no impact ($w < 0.10$). Between these two extreme values, the weighting function follows a sigmoid shape, which is characteristic of the Gompertz function. From a managerial perspective, this finding implies that once a churn event has occurred, a company has a very small window in which it can attempt to prevent churn among the defecting customer's social contacts. We will come back to this point in our discussion section.

---

[6] The χ^2 difference ($-2 \times$ Difference in LL) between *Model 3* and *Model 4* is 27.26, and the χ^2 difference between *Model 4* and *Model 5* is 20.00.

7.3. Parameter estimates—Control variables

Table 4 provides the detailed results for *Model 5*, i.e., parameter estimates, standard errors and *p*-values, as well as the parameter estimates for a baseline model without social effects (*Model 3*) for comparison. Looking at the demographic control variables, we see, on the one hand, that age is significantly negatively related to churn (−0.02). This implies that customers become *less* likely to churn with increasing age, which is consistent with prior literature on the impact of age on repeat buying (Lambert-Pandraud & Laurent, 2010; Lambert-Pandraud, Laurent, & Lapersonne, 2005). On the other hand we observe that widowed customers are *more* likely to churn (3.41). Because these customers tend to be older (the mean age among widowed customers is 63.25 years, ranging between 51 and 83), this might be seen as a contradiction to the previous finding. However, because prior research has highlighted that elderly people are more vulnerable to fraudulent and unethical marketing practices (Yoon et al., 2005), it is conceivable that the relationship between age and loyalty might not be linear but an inverse U-shaped one. Our empirical findings support this notion.

With respect to socio-demographics, we see that, with one exception (Group 07), all of the variables that indicate socio-demographic group membership are significantly positively correlated with churn propensity. Group 12 has the highest coefficient (2.86) and corresponds to customers for whom socio-demographic group information is not available ("unclassified"). The second-highest coefficient (2.08) corresponds to Group 08, which represents pensioners with an active lifestyle and an above-average income. Again, this finding shows that customers with a relatively high age tend to be more likely to churn than other types of clients, which is consistent with our previous results.

Finally, regarding acquisition process characteristics, one can see that two out of twelve dummy variables for the acquisition channel are significant. Channel 07 corresponds to retail outlets that are directly maintained and owned by the cellular service provider itself. The negative coefficient (−0.65) implies that customers that are acquired directly have a lower-than-average likelihood of churning. Channel 11 represents an internal channel that handles customers who have already threatened to leave a company in the past. The significant positive coefficient that is associated with this channel (1.75) shows that customers who have been retained using special means tend to be less loyal in the long run than other types of clients.

7.4. Parameter estimates—Social impact measures

Looking at the main variables of interest for our study, the impact of churn among incoming and outgoing calling relationships, two points are particularly noteworthy: First, while the two impact measures that are related to churn among *outgoing* calling relationships influence focal actor churn significantly (p-value ≤ 0.00005), the two impact measures that are related to churn among *incoming* calling relationships do not (p-value ≥ 0.16). This finding implies that, consistent with our theoretical expectations, the direction of the calling relationship matters because only churn among outgoing calling relationships influences the hazard of churn for the focal actor. The fact that we observed different results depending on relationship directionality supports the assumption that the significance of social impact measures is unlikely to be caused by confounds such as endogenous group formation, correlated unobservable variables and simultaneity (Manski, 1993; Moffitt, 2001) and provides an empirical indication for the presence of social contagion effects.

Second, churn among outgoing calling relationships influences the hazard of churn for the focal actor in a nonlinear way. For low levels of $s_{out,\,i}(t)$, a focal customer's hazard of defection *increases* with increasing numbers of defections among the social contacts with whom he or she has outgoing calling relationships. For example, a focal actor who
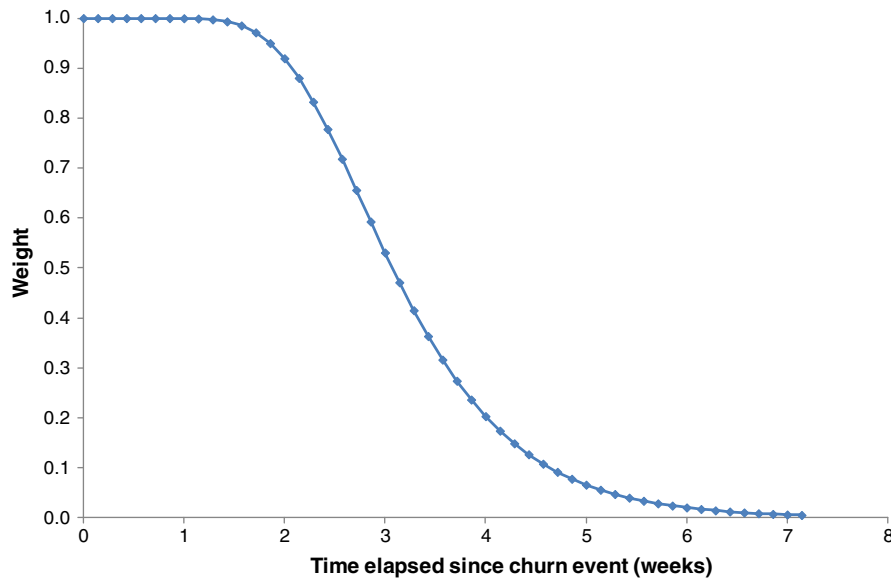
Fig. 1. Weighting function for *Model 5* ($\alpha = -28.00$; $\beta = -1.20$).

experiences the churn of an outgoing calling relationship with a recency-weighted tie strength of 1% (or, alternatively, of multiple outgoing calling relationships with a total tie strength that is equal to 1%) has a hazard rate that is approximately 1.6 times larger than that of a customer who was not exposed to churn among outgoing calling relationships. If, however, churn among outgoing calling relationships increases from 1% to 2%, this ratio increases from 1.6 to 2.5.[7]

To better interpret these results, it is interesting to contrast them to a model without social influence. The parameter estimates for such a model (*Model 3*) are shown in the first column of Table 4. As observed, a model without social effects results in very similar parameter estimates for all of the other variables than one that includes social effects. All of the coefficients have the same sign (with the exception of acquisition channel 3, a variable that is not significant in both models) and are of comparable orders of magnitude. Looking at the parameter estimates of *Model 3* also helps to interpret the change in the hazard ratio that is associated with a certain social effect. For example, we see that the churn of an outgoing calling relationship with a recency-weighted tie strength of 2% (which is equal to 2.5, as discussed above) corresponds to a change in age of approximately 40 years.[8] This finding shows that, at least in our sample, social effects have a considerable impact on the hazard of churn.

8. Robustness checks

Given the significant difference we observed in the impact of churn among incoming versus outgoing calling relationships, it is conceivable that other factors in addition to relationship directionality influence the

importance of certain churn events. To test for such effects, we estimated three variations of *Model 5* in which we classified incoming and outgoing calling relationships based on whether the two relationship partners (a) share the same service plan, (b) share the same acquisition channel or (c) share the same ZIP code. An analysis of deviance for the three alternative models on four degrees of freedom (df; each model includes four additional parameters) results in χ^2 values of 9.8 for service plan, 2.3 for acquisition channel and 6.5 for ZIP code.[9] This shows that two of these three alternative models (those that classify relationships according to same or different acquisition channels or ZIP code) yield no significant improvement over *Model 5*. The improvement of the fit of the third model (classifying relationships according to same or different service plans) is significant with a p-value of 4.42%.

Table 5 shows the estimation results for the eight social impact measures (s_{in} and s_{out} for same and different service plans in both linear and quadratic forms) for a model in which relationships are distinguished according to service plan (*Model 5.1*). As observed, our substantive conclusions also hold in this model: churn among incoming calling relationships has no significant impact on focal actor churn, while churn among outgoing calling relationships impacts focal actor churn significantly.

We also tested a variation of *Model 5*, in which we differentiated social impact measures that were based on alter churn type (voluntary versus involuntary), which allowed us to test for differences regarding these two types of churn in the parameter estimates in addition to the unspecified baseline hazard function. This alternative specification did not improve model fit significantly (χ^2 difference of 5.8 on 4 df).

It is interesting to note that the time-dependent covariate "contract duration" is not significantly related to the hazard of churn within *Model 5* (p-value: 0.5238).[10] This finding might seem surprising and could be considered as an indication of a confounding effect between contract duration and the social impact measures: Assume a focal customer who, once acquired, does not subscribe to a new contract and who experiences churn within her social network early in the company relationship. For this client, contract duration would *increase* over

---

[7] The change in the hazard rate that is associated with a certain value of recency-weighted $s_{out, i}$ can be determined as: exp $(0.4715 \times s_{out, i} - 0.0087 \times s_{out, i}^2)$. It should be noted that the increase in the hazard of defection with increasing levels of churn among outgoing calling relationships is only true for low levels of recency-weighted $s_{out, i}$ up to a maximum value of 27.0%. Beyond that point, the hazard of defection decreases with increasing levels of churn among outgoing calling relationships. That said, recency-weighted churn among outgoing calling relationships ranges from a minimum of 0% to a maximum of 100% in our analysis sample, with a mean of 1.05% over all observations and a mean of 3.91% for all cases in which recency-weighted $s_{out, i}$ is larger than zero. The results for high levels of recency-weighted churn among outgoing calling relationships therefore need to be interpreted with caution.

[8] Calculated as ln (2.5) / 0.0223.

[9] The critical χ^2 value at 4 df is 9.49.

[10] In fact, contract duration is not a significant predictor in any of the models that are shown in Table 3 (p-values equal to or above 0.30).

Table 4
Model estimation results for *Model 5* ($\alpha = -28.00$; $\beta = -1.20$).

Variable	Baseline	Estimate	SE	Pr(>|z|)
s_{in}		−0.040	0.046	0.383
s_{out}		0.472	0.104	0.000
s_{in}^2		0.001	0.001	0.161
s_{out}^2		−0.009	0.002	0.000
Age	−0.022	−0.023	0.007	0.002
Contract duration	−0.0003	−0.0002	0.000	0.524
Gender (Female)	0.097	0.039	0.134	0.771
Marital status—Married	0.300	0.280	0.167	0.093
Marital status—Divorced	0.072	0.090	0.371	0.809
Marital status—Other	0.475	0.539	0.406	0.185
Marital status—Separated	−0.101	−0.060	0.832	0.943
Marital status—Widowed	2.227	3.406	0.920	0.000
Marital status—Refused info	−1.322	−1.224	1.660	0.461
Socio-demographic group 02	0.913	0.849	0.230	0.000
Socio-demographic group 03	1.397	1.418	0.228	0.000
Socio-demographic group 04	1.475	1.469	0.228	0.000
Socio-demographic group 05	1.372	1.349	0.256	0.000
Socio-demographic group 06	1.502	1.468	0.269	0.000
Socio-demographic group 07	1.020	0.705	0.471	0.135
Socio-demographic group 08	2.036	2.078	0.385	0.000
Socio-demographic group 09	0.926	0.862	0.315	0.006
Socio-demographic group 10	0.856	0.674	0.331	0.041
Socio-demographic group 11	1.158	1.051	0.415	0.011
Socio-demographic group 12	2.702	2.862	0.745	0.000
Socio-demographic group 13	1.351	1.538	0.717	0.032
Acquisition channel 02	−0.257	−0.161	0.212	0.447
Acquisition channel 03	−0.025	0.069	0.234	0.767
Acquisition channel 04	−0.307	−0.251	0.219	0.250
Acquisition channel 05	0.106	0.169	0.292	0.562
Acquisition channel 06	−0.478	−0.462	0.267	0.084
Acquisition channel 07	−0.810	−0.655	0.292	0.025
Acquisition channel 08	−0.655	−0.867	0.747	0.246
Acquisition channel 09	0.235	0.894	0.558	0.109
Acquisition channel 10	−1.930	−1.695	1.231	0.169
Acquisition channel 11	1.622	1.749	0.773	0.024
Acquisition channel 12	−0.973	−0.876	1.066	0.411
Acquisition channel 13	0.158	0.088	1.106	0.937

Table 5
Estimation results for different specifications of *Model 5* (social impact measures only).

Model	Variable	Estimate	SE	Pr(>|z|)
Model 5 (baseline)	s_{in}	−0.040	0.046	0.383
	s_{out}	0.472	0.104	0.000
	s_{in}^2	0.001	0.001	0.161
	s_{out}^2	−0.009	0.002	0.000
Model 5.1 (differentiation by service plan)	$s_{in,\ same\ service\ plan}$	0.028	0.115	0.810
	$s_{in,\ different\ service\ plans}$	−0.004	0.051	0.940
	$s_{out,\ same\ service\ plan}$	0.912	0.338	0.007
	$s_{out,\ different\ service\ plans}$	0.505	0.173	0.004
	$s_{in,\ same\ service\ plan}^2$	−0.001	0.002	0.650
	$s_{in,\ different\ service\ plans}^2$	0.001	0.001	0.170
	$s_{out,\ same\ service\ plan}^2$	−0.054	0.022	0.014
	$s_{out,\ different\ service\ plans}^2$	−0.010	0.003	0.001
Model 5.2 (threshold for contract duration)	s_{in}	−0.041	0.047	0.382
	s_{out}	0.467	0.106	0.000
	s_{in}^2	0.001	0.001	0.138
	s_{out}^2	−0.009	0.002	0.000
Model 5.3 ("contract duration" as total duration of client relationship)	s_{in}	−0.042	0.046	0.369
	s_{out}	0.475	0.105	0.000
	s_{in}^2	0.001	0.001	0.147
	s_{out}^2	−0.009	0.002	0.000

time (because the contract is not renewed, contract duration essentially measures the duration of the company–client relationship). At the same time, the two social impact measures would *decrease* over time due to decreasing churn recency over time. This finding makes it conceivable that the effects of the two variables on the hazard rate cancel each other out, leading the contract duration variable to appear insignificant in our preferred model.[11]

To test for such a confounding effect, we conducted a series of additional robustness checks: First we tested whether contract duration shows threshold effects in the sense that its impact on the hazard of churn is different when contract duration is above or below a certain threshold. In total we tested 100 different thresholds that ranged from a minimum of 5 to a maximum of 500 days. The specification that improved fit most substantially was a model with a threshold at 205 days (χ^2 difference compared to *Model 5* of 9.46 on 1 df, p-value: 0.0021, hereafter referred to as *Model 5.2*). In this model, contract duration influences the hazard of churn positively (coefficient: 0.0062) when it is below the threshold of 205 days (p-value: 0.0027) and is otherwise insignificant (p-value: 0.7627). As shown in Table 5, this specification does not affect our substantive conclusions regarding the two social impact measures, s_{in} and s_{out}.

Second, we performed the same threshold test in a model without social impact measures. The results were very similar to those of the first threshold test: The specification that yielded the highest

improvement in model fit was a model with a threshold at 205 days (χ^2 difference compared to *Model 3* of 10.15 on 1 df, p-value: 0.0014). Within this model, contract duration influences the hazard of churn positively (coefficient: 0.0062) when it was below the threshold of 205 days (p-value: 0.0023) and was not significant otherwise. The fact that the result regarding contract duration was identical in models with and without social impact measures makes the presence of any confounding effects unlikely.

Finally, we estimated a variation of *Model 5* using a different specification of the contract duration variable. In the original specification of *Model 5*, we operationalized contract duration as the duration of the *current* contract. This measure was potentially subject to substantial, discrete shifts because contract duration increases steadily until a new contract is signed, and the measure falls back to zero. To test the extent to which these discrete shifts create problems in the parameter estimation process, we estimated an alternative model (*Model 5.3*) in which we defined contract duration as the total duration of the customer–company relationship (i.e., time elapsed since acquisition). As observed in Table 5, this again did not affect our substantive conclusions regarding s_{in} and s_{out}.

9. Discussion

9.1. Theoretical implications

In terms of contribution to theory, our results provide three important theoretical implications: First, consistent with the work of Nitzan and Libai (2011), we find an empirical indication for social interactions in customer churn decisions. Churn of customers with whom a focal actor maintains outgoing calling relationships has a significant and substantial impact on the hazard of churn of the focal actor. The fact that we used a model that takes account of network directionality and includes controls for homophily, correlated unobservable variables, and the characteristics of the acquisition process and contract reduce the likelihood that our findings are caused by spurious effects and provides an indication for the presence of social contagion rather than a simple correlation in observed behavior. Our results therefore confirm that social network effects, which have previously mainly been discussed in the context of customer acquisition, are also present in the case of customer retention and hence remain important even after customers have been acquired. Specifically, they imply that

---

[11] We thank an anonymous reviewer for having pointed this out.

individual-level retention rates and customer lifetime have a certain degree of positive network autocorrelation.

Second, our work provides insight into the dynamics that underlie social influence in customer retention by showing that a focal actor's likelihood of churn is differentially influenced by churn among his or her social contacts and that this influence is dependent on relationship directionality and churn recency. Specifically, we show that only churn events among *outgoing* calling relationships are related to focal actor churn and that these events affect an actor's likelihood of churning at a given time only if they occurred within a short period (in our sample, less than five weeks) prior to that time. If churn recency is not explicitly accounted for, then churn among a focal actor's social relationships does not have a significant impact on the likelihood that the focal actor will churn. Furthermore, the influence of churn among social relationships on the hazard of churn for a focal actor is nonlinear in the sense that the hazard of defection increases quadratically with increasing numbers of recent churn events among outgoing calling relationships.

Combined, these results allow us to speculate about the drivers behind social influence processes, which are very often difficult or even impossible to distinguish empirically (DiMaggio & Powell, 1983; Van den Bulte & Lilien, 2001). For example, one possible interpretation of our findings is that customers simply try to minimize cost and therefore base their retention decisions on churn among their social relationships. Because within our example only outgoing calling relationships represent a cost for a focal actor (due to the "calling party pays" principle), high levels of churn among these relationships provide an incentive for defection to minimize the share of expensive out-of-network calls.[12] However, while such an interpretation can explain the importance of churn among outgoing versus incoming calling relationships, it is not consistent with the importance of churn recency that we observed in our sample. Given that the accumulated cost for out-of-network calls increases with the amount of time that elapsed since the churn of a social contact, we would expect that churn events among outgoing calling relationships that happened longer ago should have the same, if not a higher, impact on focal actor churn than recent events. Instead, we observed the opposite result in our sample. This decay of social influence is consistent with the previous research on the exponential decay of word-of-mouth effects over time (Strang & Tuma, 1993; Trusov et al., 2009). This finding implies that our results are more consistent with an interpretation of social contagion than of simple cost minimization.

Third, our finding that relationship directionality plays an important role in determining the likelihood that a focal actor will churn following the churn of his or her peers also has implications for other research contexts that make use of electronically recorded data on social relationships. Researchers increasingly use data from email networks (Kossinets & Watts, 2006), internet communities (Grabowski, 2007), social networking sites (Trusov et al., 2009) and call detail records (Eagle, Pentland, & Lazer, 2009; Onnela et al., 2007; Palla, Barabasi, & Vicsek, 2007) to reconstruct social networks. Some of these studies have resulted in surprising and sometimes counterintuitive findings, for example on the importance of strong versus weak ties (Onnela et al., 2007) or the degree of overlap between observational data from mobile phones and self-reported survey data (Eagle et al., 2009). The same is true for certain simulation studies that have relied on undirected networks and provided results that contradict commonly held beliefs (Watts & Dodds, 2007). Given that ties of different directions might have different importance for social processes, it is conceivable that these counterintuitive results may in part be driven by the use of undirected rather than directed networks. We therefore encourage researchers to take tie directionality into account in both empirical and simulation works, especially when relying on

electronic tie data. If possible, researchers should prefer directed data over undirected data.

9.2. Managerial implications

From a managerial viewpoint, our analysis indicates that monitoring churn within the social network of a customer and specifically among outgoing relationships can provide useful information in the context of churn prediction and proactive churn management. As shown by Neslin et al. (2006), even a small improvement in the accuracy of churn prediction models can be associated with substantial financial gains for companies. Especially in the telecommunications industry, where customer churn has been shown to be a major cost item and where information on social relationships can be obtained easily through the analysis of call detail records, calculating the impact measures of social network churn is likely to be particularly beneficial.

That said, the high importance of churn recency that we observed in our sample implies that, following a customer's churn, companies only have a very small window in which they can attempt to prevent subsequent churn among the defecting customer's social contacts. Therefore, in many situations it might not be wise to wait until churn among social contacts actually occurs because at that point proactive churn prevention techniques might no longer be feasible. Instead, it seems more advisable to include the positive autocorrelation among individual-level lifetimes and churn propensities that we observe in our analysis within the churn prediction model itself to allow for more accurate predictions. Information on the estimated churn propensity of one customer could constitute additional input towards estimating the churn propensity of another client. Linear network autocorrelation models have previously been used in social network analysis to account for and measure social contagion effects (e.g., Leenders, 2002). Recent advances in spatial modeling now also allow the inclusion of network autocorrelation in logit models, which are frequently used to predict churn propensities (Carl & Kuehn, 2007; Dow, 2008).

9.3. Methodological contribution

In addition to these theoretical and managerial implications, our work also provides a methodological contribution by proposing an approach to incorporate continuous recency effects into Cox Proportional Hazards models. The usual approach to account for recency in nonparametric survival models has been to work with lagged variables as was done, for example, by Nitzan and Libai (2011). This technique, however, only tests for differences at *discrete* time points and does not allow a *continuous* recency effect to be included in the model estimation. Our approach, which uses a continuous weight function to weight churn events according to recency, overcomes this limitation.

9.4. Future research

With respect to areas of future research, we see three questions of particular importance: First, in some calling relationships, for example among spouses or other family members, the relationship partners might actively try to balance economic costs between each other by offering to terminate a call and to call the other person back. Such actions, if they occur, would represent a distortion of our measure of relationship directionality, especially in the case of mutual calling relationships. One way to assess the extent of bias that this distortion creates would be to test the robustness of our results when excluding calling relationships that belong to the same family or household. Another option would be to consider other measures (e.g., qualitative assessments) that characterize the type of relationship in addition to tie strength and relationship directionality. Additionally, as highlighted above, while we know the acquisition date and contract duration of each customer, we

---

[12] We thank an anonymous reviewer for having provided this explanation.

do not know how much time is left on the current contract. It is therefore conceivable that this missing information could be confounded with the social impact measures that we include in our analysis. If possible, future studies should correct for this potential bias by considering the remaining contract duration explicitly in the model. In sum, these modifications would increase confidence in our results and also allow further insight to be provided into the relative importance of certain types of relationships in driving focal actor churn decisions.

Second, future studies could combine our finding of social network effects in individual-level churn behavior with recent evidence on social network effects in customer-level revenue (Haenlein, 2011), to provide insight into the extent of network autocorrelation in customer lifetime value. Customer lifetime value is a measure of the cumulative discounted profit that a customer generates over his or her lifetime. Given that individual-level churn behavior, customer lifetime and individual-level revenue are all subject to positive network autocorrelation, it can be expected that similar autocorrelation will be observed – possibly to an even greater extent – in customer lifetime value. Assessing the extent of this autocorrelation would provide valuable insights for customer relationship management and for managing the customer base as a whole versus on an individual basis. Traditional approaches to calculating customer lifetime value (e.g., Berger & Nasr, 1998; Fader, Hardie, & Lee, 2005) or agent-based models such as stochastic cellular automata (Goldenberg, Libai, & Muller, 2001) could be applied in this context.

Third, in recent years, research on social influence in customer acquisition has started to move from providing evidence that social contagion exists to giving explanations on why it occurs (Iyengar, Van den Bulte, & Choi, 2012). The literature has discussed several causal mechanisms of social influence, including information transfer, normative pressures, competitive concerns and performance network effects (Van den Bulte & Lilien, 2001). Applied to social interactions in customer churn decisions, this finding implies that different causal mechanisms might operate in parallel and explain the findings we observed. For example, following the churn of a social contact, a focal actor might be more likely to churn because she or he becomes aware of a more attractive offer (information transfer), becomes concerned that remaining with the old provider may result in status disadvantages (normative pressures), or seeks to reduce the costs that are incurred as a result of the differential pricing of in-network and out-of-network calls (performance network effects). A more detailed investigation of the mechanisms through which social contagion occurs will provide insights into the importance of each of these drivers, which could then be used for proactive churn management actions.

Appendix A. Weighting function

The Gompertz function $f(x)$ with parameters a, b and c ($b, c < 0$) is defined as:

$$f(x) = ae^{be^{cx}}. \tag{A1}$$

Based on this definition of $f(x)$, we define a new function $g(x)$ as:

$$g(x) = a - f(x) = a - ae^{be^{cx}} \text{ with } a = \frac{1}{1-e^b}, \tag{A2}$$

which is equivalent to:

$$g(x) = \frac{1}{1-e^b}\left(1 - e^{be^{cx}}\right). \tag{A3}$$

It can easily be observed that $g(x)$ fulfills the following three conditions:

- $g(0) = \frac{1}{1-e^b}\left(1 - e^b\right) = 1$

- $\lim_{x\to\infty}g(x) = \frac{1}{1-e^b} - \frac{1}{1-e^b}e^0 = 0$
- $g(x) \geq g(x + \varepsilon)$ for all x.

This implies that $g(x)$ can be used as a weighting function in defining $s_{in,\,i}(t)$ and $s_{out,\,i}(t)$.

The function $g(x)$ allows for a large variety of different shapes, depending on the choice of b and c. Specifically, we see that:

- $\lim_{c\to 0}g(x) = \frac{1}{1-e^b}\left(1 - e^b\right) = 1$
- $\lim_{b\to 0}g(x) = e^{cx}$.

The limiting cases of $g(x)$ are therefore equal weighting of all churn events (for $c \to 0$) and exponential decay (for $b \to 0$).

Appendix B. Calculation of social impact measures

The use of time-dependent covariates (i.e., contract duration and churn among incoming or outgoing calling relationships) made it necessary to transform the input data before using it for model estimation. Table B.1 illustrates the process of this data transformation. The upper panel shows an input dataset for an actor who maintains four different calling relationships with three alters. The lower panel (the analysis dataset) shows the transformed version of this input data, which serves as the input to estimate the stratified Cox Proportional Hazards model.

The first line in the analysis dataset covers the period between the actor acquisition date (01.01.2005) and the churn date of the first alter (Alter ID 2, 15.11.2006). Because no churn occurred among any of the alters during this period, the impact measures for incoming and outgoing relationships are zero. Given that the current contract of the actor started on 01.07.2006, the contract duration at the churn date of the first alter is 137 days.

The second line covers the period between the churn of Alter ID 2 (15.11.2006) and the churn of the Alter ID 3 (15.12.2006). The focal actor maintained an outgoing calling relationship with Alter ID 2, with a tie strength that was equal to 0.05. The time difference between the churn date of Alter ID 3 (15.12.2006) and the churn date of Alter ID 2 (15.11.2006) is 30 days. Therefore, the outgoing impact measure for that period is $0.05 \times w(30)$, where $w(30)$ is the value of the weighting function for 30 days. The incoming impact measure remains zero. The focal actor's contract duration at the churn date of Alter ID 3 is 167 days.

The third line covers the period between the churn of Alter ID 3 (15.12.2006) and the focal actor's churn date (01.01.2007). The focal actor maintained a mutual calling relationship with Alter ID 3. The tie strength of the incoming calling component of this relationship is equal to 0.10, and the tie strength of the outgoing calling component is equal to 0.15. The time difference between the actor's churn date (01.01.2007) and the churn of Alter ID 2 (15.11.2006) is 47 days, while the time difference between the actor's churn date and the churn date of Alter ID 3 (15.12.2006) is 17 days. The outgoing impact measure for this period is therefore $0.05 \times w(47) + 0.15 \times w(17)$, and the incoming impact measure is $0.10 \times w(17)$. Note that because Alter ID 3 is a mutual calling relationship, both social impact measures s_{in} and s_{out} increase. The increase in each variable depends on the proportion of the total incoming or outgoing communication volume that Alter ID 3 accounts (i.e., tie strength). The contract duration at the actor's churn date is 184 days.

To summarize, the transformed dataset that is used to estimate the stratified Cox Proportional Hazards model includes $n + 1$ data points for each customer, where n is the number of alters who have churned prior to the actor's churn date or the end of the observation period. Note that within our example, the churn of the last alter (Alter ID 4) is irrelevant because it occurred after the actor's churn date.

Table B.1
Exemplary input data and analysis dataset.

Input data

Actor ID	Alter ID	Tie strength	Type of calling relationship	Actor acquisition date	Start of current contract	Actor churn date	Alter churn date
1	2	0.05	Outgoing	01.01.2005	01.07.2006	01.01.2007	15.11.2006
1	3	0.10	Incoming	01.01.2005	01.07.2006	01.01.2007	15.12.2006
1	3	0.15	Outgoing	01.01.2005	01.07.2006	01.01.2007	15.12.2006
1	4	0.20	Incoming	01.01.2005	01.07.2006	01.01.2007	01.02.2007

Analysis dataset

Actor ID	Start	End	Contract duration d	Incoming impact s_{in}	Outgoing impact s_{out}
1	01.01.2005	15.11.2006	137	0.00	0.00
1	15.11.2006	15.12.2006	167	0.00	$0.05\ w(30)$
1	15.12.2006	01.01.2007	184	$0.10\ w(17)$	$0.05\ w(47) + 0.15\ w(17)$

Appendix C. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ijresmar.2013.03.003.

References

Bartunek, J. M., Huang, Z., & Walsh, I. J. (2008). The development of a process model of collective turnover. *Human Relations, 61*(1), 5–38.

Berger, P. D., & Nasr, N. I. (1998). Customer lifetime value: Marketing models and applications. *Journal of Interactive Marketing, 12*(1), 17–30.

Bolton, R. N. (1998). A dynamic model of the duration of the customer's relationship with a continuous service provider: The role of satisfaction. *Marketing Science, 17*(1), 45–65.

Bolton, R. N., & Lemon, K. N. (1999). A dynamic model of customer's usage of services: Usage as an antecedent and consequence of satisfaction. *Journal of Marketing Research, 36*(2), 171–186.

Bramoulle, Y., Djebbari, H., & Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics, 150*(1), 41–55.

Braun, M., & Schweidel, D. A. (2011). Modeling customer lifetimes with multiple causes of churn. *Marketing Science, 30*(5), 881–902.

Brown, J. J., & Reingen, P. H. (1987). Social ties and word-of-mouth referral behavior. *Journal of Consumer Research, 14*(3), 350–362.

Buchanana, A. S., & Latkina, C. A. (2008). Drug use in the social networks of heroin and cocaine users before and after drug cessation. *Drug and Alcohol Dependence, 96*(3), 286–289.

Carl, G., & Kuehn, I. (2007). Analyzing spatial autocorrelation in species distributions using Gaussian and logit models. *Ecological Modelling, 207*(2–4), 159–170.

Castilla, E. J. (2005). Social networks and employee performance in a call center. *The American Journal of Sociology, 110*(5), 1243–1283.

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *The New England Journal of Medicine, 357*(4), 370–379.

Christakis, N., & Fowler, J. H. (2008). The collective dynamics of smoking in a large social network. *The New England Journal of Medicine, 358*(21), 2249–2258.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological), 34*(2), 187–220.

De Paula, A. (2009). Inference in a synchronization game with social interactions. *Journal of Econometrics, 148*(1), 56–71.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review, 48*(2), 147–160.

Dow, M. M. (2008). Network autocorrelation regression with binary and ordinal dependent variables: Galton's problem. *Cross-Cultural Research, 42*(4), 394–419.

Eagle, N., Pentland, A. S., & Lazer, D. (2009). Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences of the United States of America (PNAS), 106*(36), 15274–15278.

Fader, P. S., & Hardie, B. G. S. (2007). How to project customer retention. *Journal of Interactive Marketing, 21*(1), 76–90.

Fader, P. S., Hardie, B. G. S., & Lee, K. L. (2005). RFM and CLV: Using iso-value curves for customer base analysis. *Journal of Marketing Research, 42*(4), 415–430.

Feld, S. L., & Carter, W. C. (2002). Detecting measurement bias in respondent reports of personal networks. *Social Networks, 24*(4), 365–383.

Fernandez, R. M., Castilla, E. J., & Moore, P. (2000). Social capital at work: Networks and employment at a phone center. *The American Journal of Sociology, 105*(5), 1288–1356.

Gerpott, T. J. (2008). Termination-discriminatory pricing in European mobile communications markets. *International Journal of Mobile Communications, 6*(5), 564–586.

Goldenberg, J., Han, S., Lehmann, D. R., & Hong, J. W. (2009). The role of hubs in the adoption process. *Journal of Marketing, 73*(2), 1–13.

Goldenberg, J., Libai, B., & Muller, E. (2001). Talk of the network: A complex systems look at the underlying process of Word-of-Mouth. *Marketing Letters, 12*(3), 211–223.

Grabowski, A. (2007). Interpersonal interactions and human dynamics in a large scale social network. *Physica A, 385*(1), 363–369.

Haenlein, M. (2011). A social network analysis of customer-level revenue distribution. *Marketing Letters, 22*(1), 15–29.

Hall, J. A., & Valente, T. W. (2007). Adolescent smoking networks: The effects of influence and selection on future smoking. *Addictive Behaviors, 32*(12), 3054–3059.

Hansen, M. T. (1999). The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Administrative Science Quarterly, 44*(1), 82–111.

Hansen, M. T. (2002). Knowledge networks: Explaining effective knowledge sharing in multiunit companies. *Organization Science, 13*(3), 232–248.

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

Haunschild, P. R., & Beckman, C. M. (1998). When do interlocks matter? Alternate sources of information and interlock influence. *Administrative Science Quarterly, 43*(4), 815–844.

Helsen, K., & Schmittlein, D. C. (1993). Analyzing duration times in marketing: Evidence for the effectiveness of hazard rate models. *Marketing Science, 11*(4), 395–414.

Hill, S., Provost, F., & Volinsky, C. (2006). Network-based marketing: Identifying likely adopters via consumer networks. *Statistical Science, 21*(2), 256–276.

Hochberg, Y. V., Ljungqvist, A., & Lu, Y. (2007). Whom you know matters: Venture capital networks and investment performance. *Journal of Finance, 62*(1), 251–301.

Hoernig, S. (2007). On-net and off-net pricing on asymmetric telecommunications networks. *Information Economics and Policy, 19*(2), 171–188.

Holland, P. W., & Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *Journal of the American Statistical Association, 76*(373), 33–50.

Iyengar, R., Van den Bulte, C., & Choi, J. (2012). *Distinguishing among mechanisms of social contagion in new product adoption: Framework and illustration.* Working paper: The Wharton School.

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

Jackson, M. O., & Rogers, B. W. (2007). Meeting strangers and friends of friends: How random are social networks? *The American Economic Review, 97*(3), 890–915.

Kossinets, G., & Watts, D. J. (2006). Empirical analysis of an evolving social network. *Science, 311*, 88–90.

Krackhardt, D., & Porter, L. W. (1985). When friends leave: A structural analysis of the relationship between turnover and stayers' attitudes. *Administrative Science Quarterly, 30*(2), 242–261.

Krackhardt, D., & Porter, L. W. (1986). The snowball effect: Turnover embedded in communication networks. *Journal of Applied Psychology, 71*(1), 50–55.

Lambert-Pandraud, R., & Laurent, G. (2010). Why do older consumers buy older brands? The role of attachment and declining innovativeness. *Journal of Marketing, 74*(5), 104–121.

Lambert-Pandraud, R., Laurent, G., & Lapersonne, E. (2005). Repeat purchasing of new automobiles by older consumers: Empirical evidence and interpretations. *Journal of Marketing, 69*(2), 97–113.

Lazega, E., & Duijn, M. v (1997). Position in formal structure, personal characteristics and choixs of advisors in a law firm: A logistic regression model for dyadic network data. *Social Networks, 19*(4), 375–397.

Leenders, R. T. A. J. (2002). Modeling social influence through network autocorrelation: Constructing the weight matrix. *Social Networks, 24*(1), 21–47.

Lemmens, A., & Croux, C. (2006). Bagging and boosting classification trees to predict churn. *Journal of Marketing Research, 43*(2), 276–286.

Libai, B., Muller, E., & Peres, R. (2013). Decomposing the Value of Word-of-Mouth Seeding Programs: Acceleration Versus Expansion. *Journal of marketing research, 50*(2), 161–176.

Lin, D. Y., & Wei, L. J. (1989). The robust inference for the Cox Proportional Hazards Model. *Journal of the American Statistical Association, 84*(408), 1074–1078.

Lunn, M., & McNeil, D. (1995). Applying Cox regression to competing risks. *Biometrics, 51*(2), 524–532.

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

McDermott, R., Fowler, J. H., & Christakis, N. A. (2009). *Breaking up is hard to do, unless everyone else is doing it too: Social network effects on divorce in a longitudinal sample followed for 32 years.* Social Science Research Network (Available at SSRN: http://ssrn.com/abstract=1490708)

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1), 415–444.

Mercer, S. H., & Derosier, M. E. (2010). Selection and socialization of internalizing problems in middle childhood. *Journal of Social and Clinical Psychology*, 29(9), 1031–1056.

Moffitt, R. A. (2001). Policy interventions, low-level equilibria, and social interactions. In S. N. Durlauf, & H. P. Young (Eds.), *Social dynamics*. Cambridge, Massachusetts: MIT Press.

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

Neslin, S. A., Gupta, S., Kamakura, W., Lu, J., & Mason, C. H. (2006). Defection detection: Measuring and understanding the predictive accuracy of customer churn models. *Journal of Marketing Research*, 43(2), 204–211.

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

Onnela, J. P., Saramaeki, J., Hyvoenen, J., Szabo, G., Lazer, D., Kaski, K., et al. (2007). Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 104(18), 7332–7336.

Palla, G., Barabasi, A. -L., & Vicsek, T. (2007). Quantifying social group evolution. *Nature*, 446(5), 664–667.

Reagans, R. (2005). Preferences, identity and competition: Predicting tie strength from demographic data. *Management Science*, 51(9), 1374–1383.

Reinartz, W., Krafft, M., & Hoyer, W. D. (2004). The customer relationship management process: Its measurement and impact on performance. *Journal of Marketing Research*, 41(3), 293–305.

Reingen, P. H., & Kernan, J. B. (1986). Analysis of referral networks in marketing: Methods and illustration. *Journal of Marketing Research*, 23(4), 370–378.

Rogers, E. M. (2003). *Diffusion of innovations.* New York–London–Toronto–Sydney: Free Press.

Ryu, G., & Feick, L. (2007). A penny for your thoughts: Referral reward programs and referral likelihood. *Journal of Marketing*, 71(1), 84–94.

Schmitt, P., Skiera, B., & Van den Bulte, C. (2011). Referral programs and customer value. *Journal of Marketing*, 75(1), 46–59.

Schweidel, D. A., Fader, P. S., & Bradlow, E. T. (2008). Understanding service retention within and across cohorts using limited information. *Journal of Marketing*, 72(1), 82–94.

Staahl Gabrielsen, T., & Vagstad, S. (2008). Why is on-net traffic cheaper than off-net traffic? Access markup as a collusive device. *European Economic Review*, 52(1), 99–115.

Strang, D., & Tuma, N. B. (1993). Spatial and temporal heterogeneity in diffusion. *The American Journal of Sociology*, 99(3), 614–639.

Trusov, M., Bodapati, A. V., & Bucklin, R. E. (2010). Determining influential users in internet social networks. *Journal of Marketing Research*, 47(4), 643–658.

Trusov, M., Bucklin, R. E., & Pauwels, K. (2009). Effects of Word-of-Mouth versus traditional marketing: Findings from an internet social networking site. *Journal of Marketing*, 73(5), 90–102.

Valletti, T. M., & Houpis, G. (2005). Mobile termination: What is the "right" charge? *Journal of Regulatory Economics*, 28(3), 235–258.

Van den Bulte, C., & Iyengar, R. (2011). Tricked by truncation: Spurious duration dependence and social contagion in hazard models. *Marketing Science*, 30(2), 233–248.

Van den Bulte, C., & Lilien, G. L. (2001). Medical innovation revisited: Social contagion versus marketing effort. *The American Journal of Sociology*, 106(5), 1409–1435.

Van den Bulte, C., & Wuyts, S. (2007). *Social networks and marketing.* Cambridge, MA: Marketing Science Institute.

Verhoef, P. C., & Donkers, B. (2005). The effect of acquisition channels on customer loyalty and cross-buying. *Journal of Interactive Marketing*, 19(2), 31–43.

Watts, D. J., & Dodds, P. S. (2007). Influentials, networks, and public opinion formation. *Journal of Consumer Research*, 34(4), 441–458.

Wei, L. J., Lin, D. Y., & Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association*, 84(408), 1065–1073.

Wuyts, S., Dekimpe, M. G., Gijsbrechts, E., & Pieters, R. (2010). *The connected customer: The changing nature of consumer and business markets.* New York, London: Routledge.

Yoon, C., Laurent, G., Fung, H., Gonzalez, R., Gutchess, A., Hedden, T., et al. (2005). Cognition, persuasion and decision making in older consumers. *Marketing Letters*, 16(3/4), 429–441.

Zubcsek, P. P., & Sarvary, M. (2011). Advertising to a social network. *Quantitative Marketing and Economics*, 9(1), 71–107.