Improving Diffusion Forecasts Using Social Interactions Data

Olivier Toubia
Columbia Business School
ot2107@columbia.edu

Jacob Goldenberg
Hebrew University

Rosanna Garcia
Northeastern University
Abstract

We propose an approach for using data on social interactions (e.g., number of recommendations received by consumers, number of recommendations given by adopters, number of social ties) in order to improve the forecasts made by extant diffusion models. We extend major extant diffusion models to capture explicitly the generation of social interactions and their impact on adoption. In particular, we extend the discrete-time versions of the Mixed Influence Model (Bass model), the Asymmetric Influence Model, and the Karmeshu-Goswami Model. The extended models may be calibrated using a combination of social interactions data and penetration data. A field study conducted in collaboration with a Consumer Packaged Goods company suggests that the incorporation of social interactions data results in improved diffusion forecasts. The field study also suggests that the benefit of using social interactions data comes in great part from an improved ability to select, based on in-sample fit, the model that will produce the best forecasts.

Keywords: diffusion models, forecasting, innovations, measurement, social networks.
1. Introduction

In the past few years, data on social interactions between consumers, including word-of-mouth (WOM) recommendations and other product-related interactions, have become increasingly available. For example, the Keller Fay group’s TalkTrak® system (www.kellerfay.com) interviews a large, representative sample of the US population on a daily basis, asking participants to record all of their product-related conversations (both offline and online) for 24 hours. Word of mouth (WOM) marketing companies like BzzAgent™ (www.bzzagent.com), shespeaks™ (www.shespeaks.com) and Vocalpoint (www.vocalpoint.com), have assembled large panels of consumers who spread word of mouth about new products and report their interactions with other consumers to the company (Godes and Mayzlin 2009). Market research firms like Nielsen Online (www.nielsen-online.com) mine millions of online communities, discussion boards, blogs and social networks to quantify online conversations about brands and products. Of course, these companies are not the only source of social interactions data, and such data may also be collected directly using traditional methods.

While social interactions and social influence have been studied extensively in the marketing literature, past research has focused primarily on analyzing and quantifying the impact of social interactions on sales and diffusion, as opposed to leveraging social interactions data to improve diffusion forecasts. For example, Godes and Mayzlin (2009) study how WOM created in a viral marketing campaign influences sales, and how characteristics of the transmitter and of the recipient of WOM influence the effectiveness of these social interactions. Chevalier and Mayzlin (2006) study the impact of online book reviews on sales. East, Hammond and Lomax (2006) measure the impact of positive and negative word of mouth on purchase probability. Godes and Mayzlin (2004) study how the quantity and the dispersion of online discussions about
new TV shows are linked to the shows’ ratings. Trusov, Bucklin and Pauwels (2009) study how
electronic invitations sent out by the members of an online social networking site impact the
number of new users, relative to traditional marketing efforts. Several papers have also studied
the relation between online WOM and movie box office performance (see for example
Dellarocas, Zhang and Awad 2008; Duan, Gu and Whinston 2008; and Liu 2006).

The present paper focuses on leveraging social interactions data to improve the diffusion
forecasts made by extant diffusion models. Because social influence is at the heart of diffusion
models, it seems natural to assume that incorporating social interactions data into the calibration
of extant diffusion models may give rise to improved diffusion forecasts.

Many diffusion models used in marketing may be traced back to the Bass model (Bass
1969), also referred to as the Mixed Influence Model (Mahajan and Peterson, 1985), and its
antecedents (e.g., Mansfield 1961). This model has been extended for example to account for
asymmetric influence between different segments of potential adopters (Lehmann and Esteban-
Bravo 2006; Muller and Yogev 2006; Van den Bulte and Joshi 2007), and heterogeneity across
potential adopters (Karmeshu and Goswami 2001).

One of the most managerially relevant applications of diffusion models is forecasting
future diffusion based on past data. However, at early stages of the diffusion process extant
diffusion models are not very useful to forecast future diffusion based on aggregate penetration
data only (see for example Hauser, Tellis, and Griffin, 2006; Mahajan, Muller, and Bass, 1990;
Van den Bulte and Lilien, 1997). The common solution to this problem is to complement
aggregate penetration data with other data. For example, the diffusion rate of past analogous
innovations has been shown to be a useful source of complementary information (Bass et al.
2001; Hahn et al. 1994; Lenk and Rao 1990; Roberts, Nelson and Morrisson 2005; Sultan, Farley,
and Lehmann 1990; Sood, James and Tellis 2009; Talukdar, Sudhir and Ainslie 2002). Data on
the timing of adoption of consumers in a sample may also be used to inform parameter estimates
(Schmittlein and Mahajan 1982; Sinha and Chandrashekaran 1992). In both cases, additional
penetration data is used to complement aggregate penetration data. The framework proposed in
this paper is not incompatible with the use of such ancillary data: it allows using data on social
interactions, in addition to these other sources of data.

One challenge with incorporating social interactions data into diffusion forecasts is that it
is not obvious how social interactions data may be incorporated into the calibration of extant
diffusion models. Consider for example the hazard rate of the Mixed Influence Model,

\[ h(t) = p + qF(t), \]

where \( p \) and \( q \) are the coefficients of external influence and the coefficient of
internal influence respectively, and \( F(t) \) is the cumulative penetration at time \( t \) (proportion of
ultimate adopters that have already adopted). Suppose that data were available on the number of
social ties of a group of consumers, the number of recommendations received by these
consumers, as well as data on which of these consumers have adopted the innovation and how
many recommendations these adopters gave in turn to other consumers. A likelihood function for
these data may not be derived readily from the Mixed Influence Model. This is because the
Mixed Influence Model, like most extant diffusion models, explicitly captures neither the
probability of adopting conditional on a number of recommendations nor the generation of
recommendations.

We address this challenge by extending major extant diffusion models so that they may
be calibrated with a combination of penetration data and social interactions data. More, precisely,
we develop models that nest the discrete-time versions of the Mixed Influence Model (MIM), the
Asymmetric Influence Model (AIM), and the Karmeshu-Goswami Model (KG). The extended
models introduced here capture explicitly the generation of social interactions as well as their resulting impact on adoption. Using a field data set collected in collaboration with a Consumer Packaged Goods company, we illustrate how social interactions data may be combined with penetration data to calibrate the extended models, resulting in improved diffusion forecasts.¹ Our study suggests that the benefit of using social interactions data comes in great part from an improved ability to select, based on in-sample fit, the model that will produce the best forecasts.

Previous attempts to model social interactions directly in diffusion models include Van den Bulte and Lilien (2001) who model the diffusion of the drug Tetracycline across a community of 121 physicians by capturing the structure of the physicians’ social network, and modeling the effect on physician i of the adoption of another physician j to which i is connected.² However this type of approach (see also for example Iyengar, Van den Bulte and Valente 2010; and Strang 1991) requires mapping the complete social network of the entire potential market (or at least a large proportion thereof). Such data are difficult to obtain for most consumer products. In contrast, our proposed approach allows calibrating diffusion models using social interactions data from a sample of consumers and does not require mapping the social network. Mahajan, Muller and Kerin (1984) build a Markov Model of the adoption process in which consumers transition between various adoption states coupled with negative versus positive attitudes towards the innovation (unaware, positive potential customer, negative potential customer, positive current customer, negative current customer) and in which customers with positive (negative) attitudes towards the product have a positive (negative) influence on the diffusion

¹ In this paper we focus on modeling trial as opposed to repeat sales or total sales, and on producing post-launch as opposed to pre-launch diffusion forecasts. We leave extensions to repeat sales and to pre-launch forecasts to future research.
process. The calibration of such a model requires panel data on consumers’ adoption states and attitudes towards an innovation. In contrast, in this paper we focus on cross-sectional data on social interactions between customers. Dellarocas, Zhang and Awad (2008) develop a modified version of the Mixed Influence Model tailored to the entertainment industry. They estimate the parameters of this extended model for a set of movies and link the diffusion parameters to a set of covariates that describe each movie, including measures related to online WOM. This link between diffusion parameters and movie covariates enables them to produce diffusion forecasts for any new movie characterized by a set of covariates. However this approach relies on analogies between movies and therefore requires access to a fairly large dataset of past movies, including longitudinal box office and social interactions data for each movie. Finally, the other papers mentioned above (e.g., Chevalier and Mayzlin 2006, Godes and Mayzlin 2004) focus on reaching substantive insights on the impact of online social interactions on sales, as opposed to using social interactions data to forecast future diffusion. Consequently, the models used by these researchers typically do not produce out-of-sample diffusion forecasts, which is the intent of the present paper.

This paper is organized as follows. In Section 2, we describe the extension of three major extant diffusion models (Mixed Influence Model, Asymmetric Influence Model, Karmeshu-Goswami Model). In Section 3, we illustrate how data on social interactions may be combined with penetration data to calibrate the extended models and improve diffusion forecasts. We conclude in Section 4 and offer directions for future research.

2. Model Development

In order to predict future diffusion when data on past social interactions are available, it is useful to model explicitly how social interactions are generated and how they influence adoption, as
opposed to entering social interactions data as mere covariates in a diffusion model. Indeed, entering social interactions data as covariates into a diffusion model (e.g., by regressing diffusion in each period on the amount of social interactions in that period) usually does not allow predicting future diffusion because the value of these covariates in the future is unknown (the independent variable needs to be known in order to predict the dependent variable). In contrast, if the generation of social interactions and their link to adoption are explicitly captured in the model, future social interactions and future diffusion may both be predicted by the model.

Therefore, our approach is to extend existing diffusion models in a way that explicitly captures the generation of social interactions and their impact on adoption. We focus on extending models that have been studied and validated by many researchers over a long period of time, rather than attempt to develop new, fundamentally different diffusion models. Our intent is not to replace extant models but to demonstrate how social interactions data may be used to supplement these long proven models. This ‘extension’ approach drives our modeling assumptions. In particular, we make assumptions that allow nesting extant diffusion models while deviating as little as possible from them. Many of these assumptions may easily be relaxed, as illustrated in Appendix 1.

For simplicity, in the remainder of the paper we focus on recommendations as the primary source of social interactions. We define a recommendation as an event in which a consumer who has adopted the innovation recommends it to another consumer (that other consumer may or may not have adopted already – if that other consumer has already adopted

_____________________

3One exception is when social interactions and diffusion data are available from a large set of past innovations. For example, Dellarocas, Zhang and Awad (2008) regress diffusion parameters on social interactions covariates across a large set of movies and forecast the diffusion parameters for a new movie based on early social interactions data for that movie. Such approach requires diffusion and social interactions data across a large set of analogous innovations. In contrast, our approach may be used with data from a single innovation.
then the recommendation will have no effect on adoption). In our field study, we measured recommendations with questions such as “To how many people have you recommended the product so far?” and “How many consumers have recommended the product to you?”. Our model may be applied to other forms of social interactions as well, such as observing other consumers using the innovation, etc. When appropriate, we note how such other forms of social interactions may be captured by modifying the definition of the parameters of the models.

For ease of exposition we start with the extension of the most popular diffusion model in marketing, the Mixed Influence Model (MIM) (Bass 1969). We next turn to the more recent Asymmetric Influence Model (AIM) (Van den Bulte and Joshi 2007) and Karmeshu-Goswami Model (KG) (Karmeshu and Goswami 2001). We summarize the parameters of the extended models in Table 1.

[INSERT TABLE 1 ABOUT HERE]

a. Extending the Discretized Mixed-Influence Model (Bass Model)

In order to incorporate social interactions data into diffusion forecasts, we model both adoption conditional on the number of recommendations received as well as the generation of recommendations. We do so in such a way that the resulting aggregate diffusion process is given in closed form, and that the discretized MIM is nested. We first describe the modeling of the probability of adoption conditional on the number of recommendations received, and of the generation of recommendations. Next, we provide closed-form expressions for the resulting aggregate diffusion process, and show formally that the discretized MIM is nested within the extended model.
Adoption conditional on number of recommendations received

Let $i$ index consumers in the potential market and $t$ index (discrete) time periods. Let $r_{it}$ be the number of recommendations received by consumer $i$ in period $t$. The probability of adopting in period $t$ conditional on receiving $r_{it}$ recommendations in period $t$ and on not having adopted in periods 1 to $t-1$, may be modeled with a discrete-time conditional hazard rate $h(r_{it})$. We adopt a specification for this hazard rate that is comparable to the hazard rates assumed by agent-based models (Garber et al. 2004; Goldenberg et al. 2002). We will show later how the following specification allows nesting the MIM: \(^4\)

$$h(r_{it}) = 1 - (1 - p)(1 - q)^r_{it}$$

(1)

The parameters $p$ and $q$ capture similar forces as the parameters of the MIM, with $p$ capturing external effects and $q$ capturing internal effects. However, the parameter $q$ is defined here as the probability that a potential adopter would adopt based on one recommendation, in the absence of external effects. Note that the hazard rate in Equation (1) does not assume that it only takes one recommendation for adoption to take place. Instead, each recommendation has a probability $q$ of triggering adoption. The conditional hazard rate is equal to one minus the probability of “resisting” the innovation, which is equal to the probability of “resisting” the external forces and “resisting” the influence of $r_{it}$ recommendations.

Generation of recommendations

We now model the generation of recommendations, in a way that allows nesting the MIM.

Consider consumer $i$ who has adopted the innovation on or before period $t-1$. We model the

\(^4\) Goldenberg, Lowengart and Shapira (2009) recently showed that a continuous-time version of this hazard rate converges to the hazard rate of the Mixed Influence Model if each adopter has the same influence on each potential adopter. Our nesting result does not rely on such an assumption.
number of recommendations given by consumer $i$ in period $t$, $g_{it}$, using a binomial distribution. The number of trials equals the number of social ties this individual has in the social network that is relevant to the diffusion of the innovation under study. This number may be measured for example using sociometric surveys (see for example Coleman, Katz and Menzel 1966; Iyengar, Van den Bulte and Valente 2010). The success probability of this binomial process equals the probability that an adopter would recommend the product to each of his or her ties in each period. Formally:

$$ g_{it} \sim \text{Binomial}(\text{ties}_i, a) \Rightarrow P(g_{it} \mid \text{ties}_i) = \binom{\text{ties}_i}{g_{it}} a^{g_{it}} (1 - a)^{\text{ties}_i} $$

(2)

where $P(g_{it} \mid \text{ties}_i)$ is the probability mass function of the variable $g_{it}$ conditional on the number of social ties $\text{ties}_i$, the parameter $a$ captures the probability that a consumer recommends the innovation to each of his or her ties in each period conditional on having adopted the innovation, and the parameter $\text{ties}_i$ captures the number of social ties of consumer $i$.

Note that other forms of social interactions, different from recommendations, may be captured as well by modifying the definition of the parameter $a$. For example, if social influence works through potential adopters observing other consumers using the innovation, the parameter $a$ may be defined as the probability that an adopter will be using the innovation while interacting with each of his or her ties.

We also model the number of recommendations received by a potential adopter $i$ in period $t$, $r_{it}$, using a binomial distribution. The number of trials equals the number of social ties of consumer $i$ and the success probability equals the probability that each of these ties would recommend the product to $i$ in period $t$. This latter probability is expressed as the probability that
a given tie would recommend the product to consumer $i$ conditional on the tie having adopted (captured by the parameter $a$ introduced above), multiplied by the probability that the tie has adopted on or before period $t-1$. Formally:

$$r_{it} \sim \text{Binomial}(\text{ties}_i, aF_{t-1}) \Rightarrow P(r_{it} | \text{ties}_i) = \binom{\text{ties}_i}{r_{it}} (aF_{t-1})^{r_{it}} (1 - aF_{t-1})^{\text{ties}_i - r_{it}}$$

(3)

where $P(r_{it} | \text{ties}_i)$ is the probability mass function of the variable $r_{it}$ conditional on the number of social ties $\text{ties}_i$, $a$ is as defined in Equation (2), and $F_{t-1}$ is the cumulative penetration in period $t-1$, which equals the probability that a randomly selected consumer in the potential market has adopted the innovation by period $t-1$. Note that Equations (2) and (3) do not assume that all adopters will recommend the product, but rather that each adopter has some probability of recommending the product to each of his or her ties in each period.

We note that our specification distinguishes between ties and recommendations. Ties reflect the social network of consumers and describe relatively stable dyadic relationships. Recommendations describe events that occur between consumers linked in that social network. Social ties and recommendations are different constructs which may both be measured. Capturing them separately allows incorporating data on both of these constructs into diffusion forecasts.

**Aggregate diffusion process**

We now show how the individual-level processes captured in equations (1) to (3) may be aggregated to obtain closed-form expressions of the aggregate diffusion process. We drop the subscript $i$ when integrating over the distribution of consumers in the population. Let $P(\text{ties})$ denote the probability mass function (i.e., distribution across consumers) of the number of social ties. The marginal penetration in period $t, f_t$, is equal to the proportion of non-adopters in the
target market before period $t$, $1 - F_{t-1}$, multiplied by the expected value of the hazard rate in period $t$, where the expected value is taken over the distribution of ties and over $r_t$, the number of recommendations received. We have the following:

$$f_t = (1 - F_{t-1}) \ E [h(r_t)]$$

$$(4)$$

$$= (1 - F_{t-1}) \sum_{ties} \sum_{r_t=0}^{ties} (1 - (1 - p)(1 - q)^{r_t}) P(r_t \mid ties) P(ties)$$

$$= (1 - F_{t-1}) \sum_{ties} \sum_{r_t=0}^{ties} (1 - (1 - p)(1 - q)^{r_t}) \left( \frac{ties}{r_t} \right) (aF_{t-1})^{r_t} (1 - aF_{t-1})^{(ties-r_t)} P(ties)$$

Given a number of social ties $ties$, the number of recommendations received $r_t$ may vary between 0 and $ties$, which explains the summation from 0 to $ties$ in the above equation. The hazard rate corresponding to each possible value of $r_t$, $(1 - (1 - p)(1 - q)^{r_t})$, is weighted by the probability of that value of $r_t$ occurring, $P(r_t \mid ties) = \left( \frac{ties}{r_t} \right) (aF_{t-1})^{r_t} (1 - aF_{t-1})^{(ties-r_t)}$.

This equation provides a closed-form expression of the marginal penetration in period $t$ given the cumulative penetration in the previous period. Marginal penetration in any period unconditional on past penetration is obtained recursively, without using any simulation or numerical approximation.

**Relation to Mixed Influence Model (MIM)**

Finally, we show how the discretized MIM may be obtained as a special case, in which the number of social ties is assumed to be homogeneously equal to 1. Under the assumption that $ties=1$ for all consumers, the number of recommendations received by a potential adopter in period $t$, $r_t$, is 1 with probability $aF_{t-1}$ and 0 with probability $(1-aF_{t-1})$. The expected value of the
hazard rate over \( r_t \) becomes equal to the hazard rate of the discretized MIM, with \( p^{\text{MIM}} = p \) and 
\[ q^{\text{MIM}} = q(1-p)a. \]

\[
E [h(r_t | \text{ties})] = \sum_{r_t = 0}^{\infty} (1-(1-p)(1-q)^r)P(r_t | \text{ties} = 1) \\
= pP(r_t = 0 | \text{ties} = 1) + (1-(1-p)(1-q))P(r_t = 1 | \text{ties} = 1) \\
= p(1-aF_{r-1}) + (p + q(1-p))aF_{r-1} = p + q(1-p)aF_{r-1}
\]

(5)

Note that this special case is presented here only in order to establish that the model described in equations (1) to (4), which we will refer to as the extended MIM, nests the original (discretized) MIM. We will not set the parameter \( \text{ties} \) to 1 in our field application.

We show in Appendix 2 that when \( \text{ties} \) follows any general distribution, the first-order linear approximation in \( q \) of the expected value of the hazard rate in Equation (4) is equal to the hazard rate of the discretized MIM, with \( p^{\text{Bass}} = p \) and \( q^{\text{Bass}} = q(1-p)a \cdot E(\text{ties}) \), where \( E(\text{ties}) \) is the expected value of \( \text{ties} \) across consumers.

**Identification**

We now discuss identification issues. As mentioned above, the number of social ties of a set of consumers, \( \{\text{ties}_i\} \), may be measured directly for example using sociometric surveys. The parameters \( p, q, \) and \( a \) are identified when the following additional individual-level data are available from the sample of consumers: adoption status (i.e., whether each consumer has adopted the innovation), the number of recommendations received by each consumer, and the number of recommendations given by those who have adopted. The number of recommendations given (conditional on adoption) does not depend on \( p \) or \( q \), which allows identifying the probability of recommending the innovation, parameter \( a \), from \( p \) and \( q \) (given \( \text{ties}_i \)). Similarly,

---

\(^5\) While the hazard rate in the continuous-time MIM is a function of \( F_n, F_{r,i} \) is used in the discrete-time version.
adoption conditional on a number of recommendations received does not depend on \( a \), which allows identifying \( p \) and \( q \) from \( a \). The parameters \( p \) and \( q \) are identified from each other because the number of recommendations received influences the hazard rate only through \( q \), and not \( p \).

**Relaxing some of the assumptions**

Finally, we highlight a set of assumptions made only for ease of exposition and in order to nest extant models in a parsimonious fashion. A list of these assumptions is provided in Table 2. These assumptions may be relaxed. In Appendix 1 we introduce a more general diffusion model (i) in which all the assumptions listed in Table 2 are relaxed, (ii) that accepts the models presented in this paper as special cases, (iii) for which closed-form expressions or the aggregate diffusion process are still available. First, the models presented in this paper assume heterogeneity in the parameter \( ties \), but homogeneity in \( p \) and \( q \) within each segment (the extended AIM and KG models presented next assume the existence of multiple segments). The model in Appendix 1 assumes instead that these parameters are distributed across consumers according to any discrete probability distribution. Second, the models in this paper assume that the probability that an adopter will recommend the product to each of his or her ties is constant over time. The model in Appendix 1 captures non-uniform influence (Easingwood, Mahajan and Muller 1983) by making the parameter \( a \) a function of the period at which the adoption occurred. Third, the conditional hazard rate in Equation (1) assumes that the number of recommendations relevant to the adoption decision of consumer \( i \) in period \( t \) is the number of recommendations received in the same period by this consumer, \( r_{it} \). While this assumption seems reasonable in situations in which the time window is large enough, the model in Appendix 1 assumes instead that adoption in period \( t \) is influenced by any linear combination of the number of recommendations received by \( i \) in each period 1 to \( t \) (e.g., number of recommendations received
in period $t-1$, cumulative number of recommendations received, higher weight on more recent recommendations, etc.). Fourth, while social ties are assumed to be symmetric ($A$ connects to $B$ implies that $B$ connects to $A$) in the models presented in this paper, the model in Appendix 1 allows for asymmetries in social ties.

[INSERT TABLE 2 ABOUT HERE]

As a step towards using social interactions data to improve diffusion forecasts, we have extended the discretized MIM to capture explicitly the generation of social interactions and their impact on adoption. We now use a similar approach to extend the AIM and the KG Model. The same assumptions made in the extended MIM and listed in Table 2 are made in the extended AIM and the extended KG model for ease of exposition and in order to nest extant models in the most parsimonious fashion. These assumptions may still be relaxed and the general model presented in Appendix 1 accepts the extended AIM and KG models as special cases.

**b. Extending the Discretized Asymmetric Influence Model (AIM)**

The Mixed Influence Model (Bass 1969) is probably the best-known diffusion model in marketing, and it has been used in a large number of applications. Since it was introduced, many theoretical developments have been published. One of the latest models proposed in the literature, which is viewed by many as the state of the art in marketing diffusion models, is the Asymmetric Influence Model of Van den Bulte and Joshi (2007). This model assumes the existence of two segments with asymmetric influence on one another (see also Lehmann and Esteban-Bravo 2006, and Muller and Yogev 2006). These two segments, labeled as “innovators” and “imitators,” are such that innovators are only influenced by other innovators, while imitators are influenced both

---

6 We refer the readers to Van den Bulte and Joshi (2007) for a review of the literature supporting the existence of two such segments.
by innovators and by imitators. Following Van den Bulte and Joshi (2007), we refer to the innovators segment as segment 1 and to the imitators segment as segment 2. Like in the original model, we assume that the proportion of innovators in the potential market is given by $\theta$. We extend this model using a similar approach as we used to extend the MIM. We first describe the modeling of adoption conditional on the number of recommendations, followed by the modeling of the generation of recommendations. Next, we provide closed form expressions for the aggregate diffusion process and verify that the AIM is indeed nested within the extended model.

**Adoption conditional on number of recommendations received**

Adoption conditional on the number of recommendations received in the innovators segment is modeled similarly to Equation 1:

$$h^1(r^1_{it}) = 1 - (1 - p_1)(1 - q_1)^{r^1_{il}}$$

(6)

where $r^1_{il}$ refers to the number of recommendations received by innovator $i$ from other innovators (innovators are assumed not to be influenced by imitators).

The probability of adoption conditional on the number of recommendations received in the imitators segment is modeled similarly, with the exception that imitators receive recommendations from both innovators and other imitators:

$$h^2(r^1_{it}, r^2_{it}) = 1 - (1 - p_2)(1 - q_2)^{r^1_{il} + r^2_{il}}$$

(7)

where $r^1_{il}$ refers to recommendations received by innovator $i$ from innovators and $r^2_{il}$ refers to recommendations received by imitator $i$ from other imitators. The parameters $p_1$ and $p_2$ capture external effects in segments 1 and 2 respectively. The parameter $q_1$ in Equation (6) (respectively $q_2$ in Equation (7)) is the probability that an innovator (respectively imitator) would adopt based on one recommendation, in the absence of external effects.
Generation of recommendations

The number of recommendations given in period \(t\) by a consumer \(i\) who has adopted the innovation on or before period \(t-1\) is modeled as follows if \(i\) is an innovator:

\[
g_{it}^{1\rightarrow 1} \sim \text{Binomial}(\text{ties}_{i}^{1\rightarrow 1}, a^{1\rightarrow 1})
\]

\[
g_{it}^{1\rightarrow 2} \sim \text{Binomial}(\text{ties}_{i}^{1\rightarrow 2}, a^{1\rightarrow 2})
\]

where \(g_{it}^{1\rightarrow 1}\) is the number of recommendations made by \(i\) to other innovators and \(g_{it}^{1\rightarrow 2}\) is the number of recommendations made by \(i\) to imitators; \(\text{ties}_{i}^{1\rightarrow 1}\) (respectively \(\text{ties}_{i}^{1\rightarrow 2}\)) is the number of social ties between consumer \(i\) and other innovators (respectively other imitators); \(a^{1\rightarrow 1}\) (respectively \(a^{1\rightarrow 2}\)) captures the probability that an innovator would recommend the innovation to an innovator (respectively imitator) to which he or she is connected in each period following adoption. If \(i\) is an imitator we have the following instead:

\[
g_{it}^{2\rightarrow 2} \sim \text{Binomial}(\text{ties}_{i}^{2\rightarrow 2}, a^{2\rightarrow 2})
\]

where \(g_{it}^{2\rightarrow 2}\), \(\text{ties}_{i}^{2\rightarrow 2}\) and \(a^{2\rightarrow 2}\) are defined as above.

The number of recommendations received by a potential adopter \(i\) in period \(t\) is as follows if \(i\) is an innovator:

\[
r_{it}^{1\rightarrow 1} \sim \text{Binomial}(\text{ties}_{i}^{1\rightarrow 1}, a^{1\rightarrow 1} F_{t-1}^{1})
\]

where \(r_{it}^{1\rightarrow 1}\) is the number of recommendations received by \(i\) from other innovators and \(F_{t-1}^{1}\) captures penetration among innovators. If \(i\) is an imitator we have the following:

\[
r_{it}^{1\rightarrow 2} \sim \text{Binomial}(\text{ties}_{i}^{1\rightarrow 2}, a^{1\rightarrow 2} F_{t-1}^{1})
\]

\[
r_{it}^{2\rightarrow 2} \sim \text{Binomial}(\text{ties}_{i}^{2\rightarrow 2}, a^{2\rightarrow 2} F_{t-1}^{2})
\]
Aggregate diffusion process

Closed-form expressions of the aggregate diffusion process are obtained as follows, where $P(ties^{k\rightarrow j})$ refers to the probability mass distribution of $ties^{k\rightarrow j}$:

$$f^1_t = (1 - F^{1}_{t-1}) \sum_{\nu^{i\rightarrow j}} [h^1(r^{1\rightarrow i}_{t})]$$

$$= (1 - F^{1}_{t-1}) \sum_{tie^{i\rightarrow j}} \sum_{\nu^{i\rightarrow j}} (1 - (1 - p)(1 - q))^{\nu^{i\rightarrow j}} (a^{1\rightarrow i} F^{1}_{t-1})^{\nu^{i\rightarrow j}} (1 - a^{1\rightarrow i} F^{1}_{t-1})^{(n\nu^{i\rightarrow j} - \nu^{i\rightarrow j})} P(ties^{1\rightarrow i})$$

(14)

$$f^2_t = (1 - F^{2}_{t-1}) \sum_{\nu^{i\rightarrow j}} [h^2(r^{1\rightarrow i}_{t}, r^{2\rightarrow j}_{t})]$$

$$= (1 - F^{2}_{t-1}) \sum_{tie^{i\rightarrow j}} \sum_{\nu^{i\rightarrow j}} \sum_{\nu^{2\rightarrow j}} (1 - (1 - p)(1 - q))^{\nu^{2\rightarrow j}} \left(a^{2\rightarrow j} F^{2}_{t-1}ight)^{\nu^{2\rightarrow j}} (1 - a^{2\rightarrow j} F^{2}_{t-1})^{(n\nu^{2\rightarrow j} - \nu^{2\rightarrow j})} P(ties^{2\rightarrow j}, ties^{2\rightarrow j})$$

Relation to Asymmetric Influence Model (AIM)

Consider the special case in which $ties^{l\rightarrow l}=1$ for all consumers and $\{ties^{l\rightarrow 2}, ties^{2\rightarrow 2}\} = \{1,0\}$ for a proportion $w$ of consumers and $\{0,1\}$ for the remaining proportion (1-$w$) of consumers. This corresponds to a situation in which each innovator has one social tie with another innovator, and each imitator has one social tie, which is with an innovator with probability $w$ or with an imitator with probability (1-$w$). In that case, the expected hazard rates are as follows:

$$h^1_t = \sum_{\nu^{1\rightarrow i}} [h^1(r^{1\rightarrow i}_{t})] = p_1(1 - a^{1\rightarrow i} F^{1}_{t-1}) + (1 - (1 - p)(1 - q))(a^{1\rightarrow i} F^{1}_{t-1}) = p_1 + q_1(1 - p)(1 - q)a^{1\rightarrow i} F^{1}_{t-1}$$

(15)

$$h^2_t = \sum_{\nu^{2\rightarrow j}} [h^2(r^{1\rightarrow i}_{t}, r^{2\rightarrow j}_{t})] = w[p_2 + q_2(1 - p)]a^{1\rightarrow i} F^{1}_{t-1} + (1 - w)[p_2 + q_2(1 - p)]a^{2\rightarrow j} F^{2}_{t-1}$$

(16)

$$= p_2 + q_2(1 - p)[wa^{1\rightarrow i} F^{1}_{t-1} + (1 - w)a^{2\rightarrow j} F^{2}_{t-1}]$$

These hazard rates are identical to those of the discretized AIM, with $p_{1,AIM}=p_1$, $q_{1,AIM}=q_1(1-p_1)a^{1\rightarrow i}$, $p_{2,AIM}=p_2$, $q_{2,AIM}=q_2(1-p_2)[wa^{1\rightarrow i} + (1 - w)a^{2\rightarrow j}]$. 
\( w^{\text{AIM}} = \frac{w a^{1 \rightarrow 2}}{w a^{1 \rightarrow 2} + (1-w)a^{2 \rightarrow 2}} \). (The readers are referred to Van den Bulte and Joshi 2007, page 402, for the notations of the AIM.) Just like with the extended MIM, this special case is presented here only for theoretical purposes. Moreover, we also show in Appendix 2 that when \( \{ \text{ties}^{1 \rightarrow 1}, \text{ties}^{1 \rightarrow 2}, \text{ties}^{2 \rightarrow 2} \} \) follow any discrete distribution, the first-order linear approximation in \( q_1 \) and \( q_2 \) of the expected value of the hazard rates in Equation (14) are equal to the hazard rates of the discretized AIM, with

\[
\begin{align*}
    w_1^{\text{AIM}} &= p_1, \\
    q_1^{\text{AIM}} &= q_1(1-p_1) a^{1 \rightarrow 1} E(\text{ties}^{1 \rightarrow 1}), \\
    p_2^{\text{AIM}} &= p_2, \\
    q_2^{\text{AIM}} &= q_2(1-p_2) [a^{1 \rightarrow 2} E(\text{ties}^{1 \rightarrow 2}) + a^{2 \rightarrow 2} E(\text{ties}^{2 \rightarrow 2})], \\
    w^{\text{AIM}} &= \frac{a^{1 \rightarrow 2} E(\text{ties}^{1 \rightarrow 2})}{a^{1 \rightarrow 2} E(\text{ties}^{1 \rightarrow 2}) + a^{2 \rightarrow 2} E(\text{ties}^{2 \rightarrow 2})},
\end{align*}
\]

where \( E(\text{ties}^{k \rightarrow j}) \) is the expected value of \( \text{ties}^{k \rightarrow j} \).

**c. Extending the Discretized Karmeshu-Goswami Model (KG)**

Karmeshu and Goswami (2001) proposed an extension of the Mixed Influence Model that assumes heterogeneity in \( p \) and \( q \) where \( p \) may take the value \( p_1 \) or \( p_2 \) and \( q \) may take the value \( q_1 \) or \( q_2 \). Karmeshu and Goswami further assume that these two parameters are uncorrelated and that diffusion within a segment of the population defined by a specific pair of values \((p,q)\) is independent from the other segments.\(^7\) In other words, diffusion in each segment is captured by the MIM with the values of \( p \) and \( q \) corresponding to that segment. The model from Section 2.a may be extended in a similar fashion. The resulting diffusion process is as follows, where \( \theta_p \) refers to the proportion of consumers with \( p=p_1 \) and \( \theta_q \) refers to the proportion of consumers

\(^7\) Karmeshu and Goswami (2001) also proposed a different version of their model in which \( p \) and \( q \) are correlated. This version may also be nested using the proposed approach.
with $q=q_1$, and diffusion in a segment characterized by parameters $(p_i,q_j), f^{i,j}$, is given by Equation (4):\(^8\)

$$f_i = \theta_p \theta_q f_i^{1,1} + (1 - \theta_p) \theta_q f_i^{2,1} + \theta_p (1 - \theta_q) f_i^{1,2} + (1 - \theta_p)(1 - \theta_q) f_i^{2,2}$$ (17)

As before, the discretized KG model is a special case in which the number of social ties, ties, is set to 1 for all consumers. When ties follow any discrete distribution, the first-order linear approximation in $q_1$ and $q_2$ of the expected value of the hazard rate in each segment is equal to the corresponding hazard rate in the discretized KG model.

### 3. Field Study

In the previous section we have extended the Mixed Influence Model, the Asymmetric Influence Model, and the Karmeshu-Goswami Model to capture explicitly the probability of adopting conditional on the receipt of recommendations and the generation of recommendations. In this section we describe a field study which illustrates how social interactions data may be combined with penetration data to calibrate these extended models in practice.

Our focus is on whether incorporating social interactions data using the proposed approach would allow an analyst to better forecast future out-of-sample penetration based on a calibration dataset. Conceptually, we consider an analyst who would estimate all three original models considered here (MIM, AIM and KG) and their extensions, and then be faced with the decision of which of these models to use for forecasting. Model selection is typically based on in-sample fit. Would our analyst be better off basing his or her forecasts on the original model with the best in-sample fit or on the extended model with the best in-sample fit? (Note that in-sample fit may not be compared between original and extended models because they use

---

\(^8\) For identification purposes and without loss of generality, we impose $\theta_p > 0.5$ and $\theta_q > 0.5$. 
different data – extended models use social interactions data – therefore the choice between original and extended models may not be based on in-sample fit). Before answering this question, we first describe the set up and data of our field study and develop a Bayesian MCMC procedure for calibrating the models developed in the previous section.

**Set up and data**

Our study was conducted in collaboration with a major US-based CPG manufacturer. The company was interested in the penetration of a new oral care product launched by one of its competitors. This product offered a significantly new benefit and represented an innovation in that category. For confidentiality reasons, we will refer to this new product as PROD.

There is evidence that social interactions play a significant role in the diffusion of oral care products, making this setting a reasonable one for applying the models developed in this paper. In particular, Silk (1966) asked a sample of consumers the following question: “Have you recently been asked your advice or opinion about ____?”, for 4 categories of interest: electric toothbrush, mouthwash, toothpaste, and regular toothbrush. A total of 32.77% of the respondents answered “yes” for at least one of these categories. More generally, in their classic study, Katz and Lazarsfeld (1955) found that approximately one third of brand switching for household goods involves personal influences, and Du and Kamakura (2010) found empirical evidence for social contagion across a wide range of CPG categories.

We define one time period as 4 weeks. Our social interactions data came from a survey conducted at \( t=5 \). The survey was administered online through a professional marketing research company. The respondents were 1,239 members of a representative panel of consumers. Our survey started with a set of screening questions designed to ensure the quality of the responses. We presented respondents with a list of oral care brands and asked them to indicate which ones
they were aware of. The list included a set of fictitious brands, and we removed from the analysis all respondents who indicated that they were aware of at least one fictitious brand. After this screening, we were left with $N = 584$ respondents. The very high proportion of respondents screened out, despite the fact that the survey was performed by a professional market research company, suggests that great care should be taken to ensure the quality of online data.

The heart of the survey consisted of the following four questions:

- “How many consumers have recommended PROD to you?” (numerical answer). Let $r_i$ be the number entered by respondent $i$, which in our data varied between 0 and 3.
- “Have you purchased PROD before?” (Yes/No). Let $y_i=1$ if respondent $i$ answered “yes” to that question, and 0 otherwise. A total of 53 respondents answered “yes” to that question.
- If $y_i=1$: “To how many people have you recommended PROD so far?” (numerical answer). Let $g_i$ be the number entered by respondent $i$, which in our data varied between 0 and 8.
- “Please think about your cell phone directory or other personal address book. With how many people in this directory did you have a non-business related conversation in the past seven days (over the phone or in person) or exchanged non-business related emails / text messages?” (numerical answer). Let $ties_i$ be the number entered by respondent $i$, multiplied by 4 (one time period corresponds to 4 weeks in our model). The average value of $ties_i$ was 38.29.

In addition to these data, we received aggregate penetration data about the product from a professional market research company, for periods $t=1$ to $t=11$. Aggregate penetration is measured as the proportion of households in the market who purchased PROD for the first time in each period.
Calibration

We calibrate the three extant diffusion models (MIM, AIM, KG) considered in the previous section and their respective extensions (extended MIM, extended AIM, extended KG). We calibrate the extended models using the aggregate penetration data combined with the full survey data, which consists of $\text{ties}_i$ (number of social ties), $r_i$ (number of recommendations received), $y_i$ (adoption), and for those consumers who adopted, $g_i$ (number of recommendations given). For each extended model, the likelihood for $\{y_i, g_i, r_i\}$ given $\text{ties}_i$ is as follows:

$$
P(y_i, g_i, r_i | \text{ties}_i) = P(g_i | y_i, r_i, \text{ties}_i)P(y_i | r_i)P(r_i | \text{ties}_i)
$$

where each term is based on the equations presented in the previous section. The extended models capture the number of recommendations and adoption in each period. In contrast, our data capture only the cumulative number of recommendations received and made by each consumer in the panel over the first 5 periods ($r_i, g_i$) and we only have information on whether each consumer has adopted in the first 5 periods ($y_i$) but not the period in which the adoption took place. As a result, we treat the periods at which recommendations and adoption took place as latent variables which we integrate out. Another interesting challenge arises in the estimation of the extended AIM and extended KG. Because these models rely on the existence of different segments of consumers, the posterior probability of belonging to each segment must be computed appropriately for each consumer and for each component of the likelihood function. Details are provided in Appendices 3 to 5.

When calibrating the original models, $y_i$ is the only piece of survey data that enters the likelihood function such that: $P(y_i=1)=mF_5$, where $F_5$ is the cumulative aggregate penetration at $t=5$ predicted by the model being estimated and $m$ is the market potential (expressed as a proportion of the full market and estimated by each model).
For each original and extended model, we also specify a likelihood function for the aggregate penetration data. For each model we make the standard assumption (see for example Srinivasan and Mason 1986) that the aggregate penetration in period \( t \), \( S_t \), is equal to the penetration predicted by the model, plus a normal i.i.d. noise which captures the effects of sampling errors, excluded variables, and misspecifications of the density function (see Srinivasan and Mason 1986, p. 170-171):

\[
S_t = m f_t + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2)
\]  

(19)

where \( S_t \) is the measured (incremental) number of adopters in period \( t \), \( m \) is the market potential and \( f_t \) is the marginal penetration in period \( t \) given by the model being estimated.

We estimated each original and each extended model with a similar Bayesian MCMC procedure (see for example Dellarocas et al. 2007 and Lenk and Rao 1990 for other uses of Bayesian MCMC in diffusion research). Details are provided in Appendices 3 to 5.

**Results**

We use the last three periods of aggregate penetration data as holdout periods. We estimate each model using the first 5, 6, 7 and 8 aggregate penetration data points (in addition to the survey adoption data for original models and the full survey data for the extended models). For each model and each number of aggregate penetration calibration data points, we compute three measures of in-sample fit and two measures of holdout predictive ability. We compute the log marginal density of the data under each model (Rossi and Allenby 2003), which captures the likelihood of the data according to the model. We also compute the mean squared error (MSE) and the mean absolute percentage error (MAPE) between the true aggregate penetration (measured as the proportion of households in the market who purchased PROD for the first time in each period) and the point estimates of these quantities suggested by each model. We compute
MSE and MAPE for both the calibration and the holdout aggregate penetration data. The log marginal density is an appropriate criterion for model selection, and the MSE and MAPE on the holdout aggregate penetration data are our metrics for measuring holdout predictive ability.

As mentioned above, our focus in this paper is on whether incorporating social interactions data into extant diffusion models using the proposed approach would allow an analyst to better forecast future penetration based on in-sample data. Conceptually, we consider an analyst whose toolkit consists of the original and extended versions of the MIM, the AIM and the KG model. After calibrating each model under consideration, this analyst would then need to select a model on which to base his or her forecasts, where the selection would typically be based on a measure of in-sample fit such as log marginal density. Our motivating managerial question was: “Would our analyst be better off basing his or her forecasts on the original model with the best in-sample fit or on the extended model with the best in-sample fit?” (Note again that in-sample fit may not be compared between original and extended models because they use different data – extended models use social interactions data – therefore the choice between original and extended models may not be based on in-sample fit). Table 3 shows, for each number of aggregate penetration data points used for calibration (i.e., 5 to 8), the performance of the original and extended models with the highest log marginal density (i.e., the original model with the highest log marginal density among original models and the extended model with the highest log marginal density among extended models). We see that in each case, the extended model outperforms the original model on holdout predictive ability. In other words, an analyst who would consider all three models (MIM, AIM, KG), calibrate each model, pick the best fitting model and use that model to forecast future penetration would be able to improve his/her forecasts by using social interactions data.
Note that the extended models do not necessarily fit the calibration penetration data better, since the likelihood function for the extended models includes not only the calibration penetration data and the survey adoption data, but also the social interactions data. Therefore, less emphasis is put in these models on fitting aggregate penetration data. Figure 1 compares the actual penetration curve with the penetration curves predicted by original and extended models, for each number of aggregate penetration data points used. A vivid illustration of the impact of using social interactions data may be found in the third panel of Figure 1, i.e., in the case in which the first 7 aggregate penetration data points are used for calibration. The original KG model fits these calibration data extremely well, but does a very poor job predicting holdout penetration. In contrast, the extended KG model fits aggregate penetration a little less well but predicts holdout penetration much better. The use of social interactions data reduces the reliance on calibration penetration data, which reduces the fit on these data but also reduces the risk of overfitting and improves holdout predictions. We also report point estimates of the parameters in Table 4.

A deeper understanding of our results may be obtained by considering the in-sample fit and holdout predictive ability for all models and all numbers of calibration periods. We report in Table 5 the in-sample log marginal density and the holdout MSE for all six models and all numbers of aggregate penetration data points used for calibration. Several interesting findings emerge. First, as was already apparent from Table 3, the identity of the best fitting model, both in-sample and out-of-sample, varies based on the number of aggregate penetration data points used for calibration. Therefore, no single “winner” emerges from our study. This reinforces the notion that an analyst should not systematically use the same diffusion model to forecast penetration, but instead test several models and select the most appropriate model in each
situation. More interestingly, the results suggest that the benefit of using social interactions data comes in great part from an improved ability to select, based on in-sample fit, the model that will produce the best holdout penetration forecasts. In particular, in all four cases (i.e., different numbers of calibration periods) the extended model with the best in-sample fit (highest log marginal density) is also the extended model with the lowest holdout MSE. In contrast, the original model with the best in-sample fit is often not the one with the lowest holdout MSE. In fact, in three out of the four cases, there actually exists an original model that achieves a holdout MSE lower than that of any extended model. However this is not relevant managerially because in each case that model does not achieve the best in-sample fit among original models, and therefore an analyst would not be able to identify it based on in-sample data. In other words the identity of the original model with the lowest holdout MSE is neither constant nor revealed by the data that are available at the time the forecasts are produced. The results also suggest that the extended models do not dominate the original models in a Pareto sense. In particular, while the extended KG always predicts better than its original counterpart, this is not the case for the extended AIM and MIM. While counterintuitive at first, this finding is in fact not surprising given that the three extended models considered in our analysis make different assumptions on the diffusion process and on the generation of social interactions and given that we fit the same data with these three models. When we have three models that assume different processes, all three models cannot be correct at the same time. One should not necessarily expect the same data to improve the forecasts made by different models that make different assumptions. Rather, one should expect the additional data to improve the forecasts made by a model only to the extent that this model captures well the underlying data generating process. This further confirms that
the benefit of using social interactions data relies greatly on improved model selection ability and less on a systematic improvement in forecasting accuracy.

[INSERT TABLES 3, 4 and 5 AND FIGURE 1 ABOUT HERE]

4. Conclusions and Directions for Future Research

We have proposed an approach for using social interactions data, which have become increasingly available in the past few years, to improve the diffusion forecasts made by major extant diffusion models. In order to accommodate these data, we have extended major diffusion models to capture explicitly the generation of social interactions and their impact on adoption. Empirically, combining social interactions data with penetration data appears to give rise to improved diffusion forecasts. Moreover, our study suggests that the benefit of using social interactions data comes in great part from an improved ability to select, based on in-sample fit, the model that will produce the best forecasts. The use of social interactions data does not preclude the use of other ancillary data suggested in previous research (e.g., analogous past innovations, timing of adoption).

We believe that several areas for future research may be identified. First, the approach developed here does not readily apply to studying the diffusion of past innovations for which relevant social interactions data were not collected when the innovation was diffusing and for which these data may not be collected retrospectively. Future research may attempt to address this limitation, either by incorporating social interactions data that are already recorded systematically, or by developing efficient tools for systematically collecting relevant social interactions data. Second, the framework itself may be extended for example to capture negative
word-of-mouth (Mahajan, Muller and Kerin, 1984), specific network structures (Barabási and Albert 1999, Shaikh, Rangaswamy, and Balakrishnan, 2007; Watts and Strogatz 1998), different types of ties or relationships (Ansari, Koenigsberg and Stahl 2010; Iyengar, Van den Bulte and Valente 2010), or to include covariates such as marketing mix variables (Bass, Krishnan and Jain, 1994; Horsky and Simon, 1983; Kalish and Sen, 1986; Robinson and Lakhani, 1975). Third, the proposed models use discrete time intervals, making the parameters a function of the data frequency. Future research may explore continuous-time versions. Finally, the propose models may be extended to capture repeat sales, and the estimation approach may be extended to produce pre-launch forecasts.

---

9 Negative recommendations could be modeled for example using a hazard rate such as:

\[ h(r_t^-, r_u^-) = [1 - (1 - p)(1 - q^-)\gamma](1 - q^-)^\gamma, \]

where \( q^- \) and \( r_t^- \) would be as \( q \) and \( r_t \) in the present paper, \( r_u^- \) would be a number of negative recommendations (also coming from a binomial process), and \( q^- \) would be the probability of deciding not to adopt in period \( t \) based on one negative recommendation.
References


Muller, Eitan, and Guy Yogev (2006), “When does the majority become a majority? Empirical analysis of the time at which main market adopters purchase the bulk of our sales,” *Technological Forecasting and Social Change*, 73(9), 1107-1120.


### Tables and Figures

<table>
<thead>
<tr>
<th>Name of variable in extended Mixed Influence Model</th>
<th>Similar variables in other extended models</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$</td>
<td>$p_1, p_2$</td>
<td>Captures the effect of external forces on adoption</td>
</tr>
<tr>
<td>$q$</td>
<td>$q_1, q_2$</td>
<td>Probability of adopting based on one recommendation</td>
</tr>
<tr>
<td>$g_{it}$</td>
<td>$g_{it}^{1\rightarrow 1}, g_{it}^{1\rightarrow 2}, g_{it}^{2\rightarrow 2}$</td>
<td>Number of recommendations given by consumer $i$ in period $t$</td>
</tr>
<tr>
<td>$r_{it}$</td>
<td>$r_{it}^{1\rightarrow 1}, r_{it}^{1\rightarrow 2}, r_{it}^{2\rightarrow 2}$</td>
<td>Number of recommendations received by consumer $i$ in period $t$</td>
</tr>
<tr>
<td>ties$_i$</td>
<td>ties$_i^{1\rightarrow 2}, ties_i^{1\rightarrow 1}, ties_i^{2\rightarrow 2}$</td>
<td>Number of social ties of consumer $i$</td>
</tr>
<tr>
<td>$a$</td>
<td>$a_1^{1\rightarrow 1}, a_1^{1\rightarrow 2}, a_2^{2\rightarrow 2}$</td>
<td>Probability that an adopter recommends the innovation to each of his or her social ties in each period following adoption</td>
</tr>
<tr>
<td>$f_t$</td>
<td>$f_t^{1}, f_t^{2}, f_t^{1,1}, f_t^{1,2}, f_t^{2,1}, f_t^{2,2}$</td>
<td>Marginal penetration in period $t$</td>
</tr>
<tr>
<td>$F_t$</td>
<td>$F_t^{1}, F_t^{2}, F_t^{1,1}, F_t^{1,2}, F_t^{2,1}, F_t^{2,2}$</td>
<td>Cumulative penetration by the end of period $t$</td>
</tr>
</tbody>
</table>

Table 1: list of variables.

<table>
<thead>
<tr>
<th>Assumption</th>
<th>Implication on extended Mixed Influence Model</th>
<th>Possible relaxation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homogenous diffusion parameters within a segment</td>
<td>{$p, q$} are homogenous</td>
<td>Parameters follow any discrete probability distribution across consumers</td>
</tr>
<tr>
<td>Uniform influence over time: the probability that an adopter recommends the innovation is not a function of when adoption took place</td>
<td>$a$ constant over time</td>
<td>Recommendation probability is (any) function of the number of periods since the adoption took place</td>
</tr>
<tr>
<td>Only recommendations from period $t$ influence adoption at $t$</td>
<td>Hazard rate is a function of $r_{it}$</td>
<td>Recommendations from periods 1 to $t$ may have a (different) impact on adoption at time $t$</td>
</tr>
<tr>
<td>Social ties are symmetric</td>
<td>Number of recommendations received and the number of recommendations given are (different) functions of the same parameter ties$_i$.</td>
<td>Social ties are asymmetric</td>
</tr>
</tbody>
</table>

Table 2: Relaxable assumptions.

The above assumptions were made in order to nest the discretized versions of the MIM, AIM and KG models in a parsimonious fashion. However they are not necessary to obtain closed-form expressions of the aggregate diffusion process. Appendix 1 presents a more general model that relaxes all these assumptions.
### Table 3: Holdout predictive ability and in-sample fit.

Using social interactions data gives rise to improved forecasts of holdout aggregate penetration. The extended models do not necessarily fit *in-sample* aggregate penetration better, since the likelihood function for the extended models includes not only the aggregate penetration data and the survey adoption data, but also the social interactions data. Therefore, less emphasis is put in these models on fitting aggregate penetration data.

<table>
<thead>
<tr>
<th>Number of aggregate penetration data points used for calibration</th>
<th>Type of model</th>
<th>Model with highest log marginal density</th>
<th>MSE on holdout penetration data</th>
<th>MAPE on holdout penetration data</th>
<th>MSE on calibration penetration data</th>
<th>MAPE on calibration penetration data</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>original KG</td>
<td>0.211</td>
<td>76.531</td>
<td>0.031</td>
<td>8.604</td>
<td></td>
</tr>
<tr>
<td></td>
<td>extended Extended AIM</td>
<td>0.006</td>
<td>11.096</td>
<td>0.035</td>
<td>11.655</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>original AIM</td>
<td>0.068</td>
<td>42.434</td>
<td>0.055</td>
<td>14.130</td>
<td></td>
</tr>
<tr>
<td></td>
<td>extended Extended AIM</td>
<td>0.006</td>
<td>10.866</td>
<td>0.047</td>
<td>12.193</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>original KG</td>
<td>2.209</td>
<td>254.477</td>
<td>0.021</td>
<td>5.901</td>
<td></td>
</tr>
<tr>
<td></td>
<td>extended Extended KG</td>
<td>0.193</td>
<td>74.021</td>
<td>0.059</td>
<td>12.135</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>original KG</td>
<td>0.242</td>
<td>82.569</td>
<td>0.014</td>
<td>6.168</td>
<td></td>
</tr>
<tr>
<td></td>
<td>extended Extended KG</td>
<td>0.052</td>
<td>37.425</td>
<td>0.132</td>
<td>17.248</td>
<td></td>
</tr>
</tbody>
</table>

### Table 4: Parameter estimates (8 aggregate penetration data points used for calibration).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate in best-fitting extended model (extended KG)</th>
<th>Estimate in best-fitting original model (original KG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_1$</td>
<td>0.007</td>
<td>0.008</td>
</tr>
<tr>
<td>$q_1$</td>
<td>0.092</td>
<td>0.506</td>
</tr>
<tr>
<td>$p_2$</td>
<td>0.407</td>
<td>0.409</td>
</tr>
<tr>
<td>$q_2$</td>
<td>0.213</td>
<td>0.517</td>
</tr>
<tr>
<td>$\theta_p$</td>
<td>0.954</td>
<td>0.656</td>
</tr>
<tr>
<td>$\theta_q$</td>
<td>0.870</td>
<td>0.776</td>
</tr>
<tr>
<td>$a$</td>
<td>0.011</td>
<td>--</td>
</tr>
<tr>
<td>$M$</td>
<td>0.895</td>
<td>0.155</td>
</tr>
</tbody>
</table>

Table 4: Parameter estimates (8 aggregate penetration data points used for calibration).
<table>
<thead>
<tr>
<th>Number of aggregate penetration data points used for calibration</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original model with highest log marginal density</td>
<td>KG, 156.136, 0.211</td>
<td>AIM, 151.274, 0.068</td>
<td>KG, 143.530, 2.209</td>
<td>KG, 136.787, 0.242</td>
</tr>
<tr>
<td>Original model with second highest log marginal density</td>
<td>AIM, 156.889, 0.003</td>
<td>MIM, 151.669, 0.204</td>
<td>AIM, 148.837, 0.006</td>
<td>AIM, 137.902, 0.663</td>
</tr>
<tr>
<td>Original model with third highest log marginal density</td>
<td>MIM, 157.128, 0.289</td>
<td>KG, 152.009, 0.081</td>
<td>MIM, 150.266, 0.167</td>
<td>MIM, 146.358, 0.002</td>
</tr>
<tr>
<td>Extended model with highest log marginal density</td>
<td>Extended AIM, 262.120, 0.006</td>
<td>Extended AIM, 256.217, 0.006</td>
<td>Extended KG, 260.232, 0.193</td>
<td>Extended KG, 257.542, 0.052</td>
</tr>
<tr>
<td>Extended model with second highest log marginal density</td>
<td>Extended KG, 271.381, 0.077</td>
<td>Extended KG, 264.877, 0.066</td>
<td>Extended AIM, 266.660, 0.286</td>
<td>Extended AIM, 263.323, 0.844</td>
</tr>
<tr>
<td>Extended model with third highest log marginal density</td>
<td>Extended MIM, 295.059, 1.123</td>
<td>Extended MIM, 291.669, 0.597</td>
<td>Extended MIM, 284.938, 0.971</td>
<td>Extended MIM, 280.404, 0.649</td>
</tr>
</tbody>
</table>

Table 5: Detailed results from field study. For each model, the first number reported is - log marginal density, the second is the MSE on holdout penetration data. The benefit of using social interactions data comes in great part from an improved ability to select, based on in-sample fit, the model that will produce the best holdout penetration forecasts. The extended models do not dominate the original models in a Pareto sense. The three extended models make different assumptions on the diffusion process and on the generation of social interactions. One should not necessarily expect the same data to improve the forecasts made by different models that make different assumptions. Rather, one should expect the additional data to improve the forecasts made by a model only to the extent that this model captures well the underlying data generating process.
Figure 1: Actual versus fitted penetration curves with 5, 6, 7, and 8 aggregate penetration data points used for calibration. The time period is on the x-axis and the proportion of households in the market who purchased the product for the first time in each period is on the y-axis. The solid line is the actual penetration, the lighter dotted line is the penetration obtained from the extended model with the highest in-sample log marginal density, and the darker dotted line is the penetration obtained from the original model with the highest in-sample log marginal density. The vertical line represents the number of aggregate penetration data points used for calibration. Only the points to the left of this line are part of the calibration data.
Appendix 1: General Model

We present here a more general model that nests the models presented in the main body of the paper and that relaxes some of the assumptions that were made in order to nest extant models. We consider two segments in the population ("innovators" and "imitators"). The conditional hazard rates in the innovators and imitators segments are modeled as follows for a consumer indexed by $i$:

$$h^1(p^1_i, q^1_i, \{r^1_{t-\tau}\}_{\tau=0,1,2,\ldots}) = 1 - (1 - p^1_i)(1 - q^1_i)\sum_{\tau=0}^1 a^1_{\tau, t-\tau}$$

$$h^2(p^2_i, q^2_i, \{r^2_{t-\tau}\}_{\tau=0,1,2,\ldots}) = 1 - (1 - p^2_i)(1 - q^2_i)\sum_{\tau=0}^1 a^2_{\tau, t-\tau}$$

The parameter $a^\tau_{\tau}$ captures the effect of a recommendation made $\tau$ periods ago by a consumer in segment $\tau$ to a consumer in segment $\tau$. The number of recommendations of different types received in period $t$ are given by:

$$r^1_{t-\tau} \sim \text{Binomial}(\text{ties}^1_{t-\tau}, \sum_{\tau=0}^1 a^1_{\tau, t-\tau} f^1_{t-\tau})$$

$$r^2_{t-\tau} \sim \text{Binomial}(\text{ties}^2_{t-\tau}, \sum_{\tau=0}^1 a^2_{\tau, t-\tau} f^2_{t-\tau})$$

Closed-form expressions for the aggregate diffusion process in the innovator segment are as follows:

$$f^1_t = (1 - F^1_t) \sum_{p^1, q^1, \text{ties}^1, f^1, f^2} p^1 \cdot q^1 \cdot \text{ties}^1 \cdot f^1 \cdot f^2 \cdot \frac{h^1(p^1, q^1, \{r^1_{t-\tau}\})}{[h^1(p^1, q^1, \{r^1_{t-\tau}\})]}$$

$$P(\{r^1_{t-\tau}\}_{\tau=0,1,2,\ldots} | \text{ties}^1, f^1, f^2) = \prod_{\tau=0}^1 \left( \frac{\text{ties}^1}{\text{ties}^1_{t-\tau}} \left( \sum_{\tau=0}^1 a^1_{\tau, t-\tau} f^1_{t-\tau} \right) \right)^{f^1_{t-\tau}} (1 - \sum_{\tau=0}^1 a^1_{\tau, t-\tau} f^1_{t-\tau})^{1 - f^1_{t-\tau}}$$

$$P(p^1, q^1, \text{ties}^1 | \text{not\_adopted}^1, f^1, f^2) = \frac{P(\text{not\_adopted}^1 | p^1, q^1, \text{ties}^1, f^1, f^2) g(p^1, q^1, \text{ties}^1)}{\sum_{p^1, q^1} P(\text{not\_adopted}^1 | p^1, q^1, \text{ties}^1, f^1, f^2) g(p^1, q^1, \text{ties}^1)}$$

$$P(\text{not\_adopted}^1 | p^1, q^1, \text{ties}^1, f^1, f^2) = \sum_{\text{ties}^1} (1 - p^1_{t-\tau}) (1 - q^1_{t-\tau})^{f^1_{t-\tau}} P(\{r^1_{t-\tau}\}_{\tau=0,1,2,\ldots} | \text{ties}^1, f^1, f^2)$$

Similarly, we have the following in the imitators segment:

$$f^2_t = (1 - F^2_t) \sum_{p^2, q^2, \text{ties}^2, f^1, f^2} p^2 \cdot q^2 \cdot \text{ties}^2 \cdot f^1 \cdot f^2 \cdot \frac{h^2(p^2, q^2, \{r^2_{t-\tau}\})}{[h^2(p^2, q^2, \{r^2_{t-\tau}\})]}$$

$$P(\{r^2_{t-\tau}\}_{\tau=0,1,2,\ldots} | \text{ties}^2, f^1, f^2) = \prod_{\tau=0}^1 \left( \frac{\text{ties}^2}{\text{ties}^2_{t-\tau}} \left( \sum_{\tau=0}^1 a^2_{\tau, t-\tau} f^2_{t-\tau} \right) \right)^{f^2_{t-\tau}} (1 - \sum_{\tau=0}^1 a^2_{\tau, t-\tau} f^2_{t-\tau})^{1 - f^2_{t-\tau}}$$

$$P(p^2, q^2, \text{ties}^2 | \text{not\_adopted}^2, f^1, f^2) = \frac{P(\text{not\_adopted}^2 | p^2, q^2, \text{ties}^2, f^1, f^2) g(p^2, q^2, \text{ties}^2)}{\sum_{p^2, q^2} P(\text{not\_adopted}^2 | p^2, q^2, \text{ties}^2, f^1, f^2) g(p^2, q^2, \text{ties}^2)}$$

$$P(\text{not\_adopted}^2 | p^2, q^2, \text{ties}^2, f^1, f^2) = \sum_{\text{ties}^2} (1 - p^2_{t-\tau}) (1 - q^2_{t-\tau})^{f^1_{t-\tau}} P(\{r^2_{t-\tau}\}_{\tau=0,1,2,\ldots} | \text{ties}^2, f^1, f^2)$$
The above system of equations provides a closed-form expression of the marginal penetration in period $t$ in each segment, given the marginal penetrations in the previous periods. Marginal penetration in any period is obtained recursively, without using any simulation or numerical approximation. The extended AIM in the paper is obtained as a special case of this general model in which:

- $\alpha_{r,t} = \alpha_{r,t-1} = \alpha_{r-1,0} = 0$ for $r > 0$ (only recommendations from period $t$ influence adoption at $t$)
- $a_{r,t} = a_{r,t-1} = a_{r-1,j}$ for all $(r,t,k,j)$ (uniform influence over time)
- The probability mass function $g$ is concentrated at one point (homogeneous parameters in each segment)
- Ties are symmetric

The extended mixed influence is obtained as a special case if we assume further that one of the segments is empty (e.g., $\theta = 0$), and the extended KG model is obtained if we assume the existence of 4 independent sub-markets.
The AIM. In the case of the AIM, the expected hazard rates are as follows:

\[ E[h(r_j)] = \sum_{ties} (1-(1-p)(1-q)) P(r_j|\text{ties})P(\text{ties}) = 1-(1-p)\sum_{ties} \sum_{ties} (\cdots)(r_j) P(r_j|\text{ties})P(\text{ties}) \]

\[ = 1-(1-p)\sum_{ties} \sum_{ties} P(r_j|\text{ties})P(\text{ties})+(1-p)\sum_{ties} \sum_{ties} (\cdots)(r_j) P(r_j|\text{ties})P(\text{ties})-(1-p)\sum_{ties} \sum_{ties} (\cdots)(r_j) P(r_j|\text{ties})P(\text{ties}) \]

\[ = p+q(1-p)E(\text{ties}) \cdot a \cdot F_{t+1} - (1-p)\sum_{ties} \sum_{ties} (\cdots)(r_j) P(r_j|\text{ties})P(\text{ties}) \]

The above expression shows that the first-order linear approximation in q of the expected value of the hazard rate is equal to the hazard rate of the discretized MIM, with p^{Bass} = p and q^{Bass} = q(1-p)E(\text{ties})a. (The last term in the above expression only contains terms of order 2 and above in q).

The same theoretical result applies to the KG model (diffusion in each segment follows the MIM) and to the AIM. In the case of the AIM, the expected hazard rates are as follows:

\[ E[h^2(r_j)] = 1-(1-p)\sum_{ties} \sum_{ties} (\cdots)(r_j) P(r_j|\text{ties})P(\text{ties}) \]

\[ = p_1 + q_1(1-p_1)a \cdot E(\text{ties})F_{t+1} - (1-p_1)\sum_{ties} \sum_{ties} (\cdots)(r_j) P(r_j|\text{ties})P(\text{ties}) \]

\[ = p + q(1-p)E(\text{ties}) \cdot a \cdot F_{t+1} - (1-p)\sum_{ties} \sum_{ties} (\cdots)(r_j) P(r_j|\text{ties})P(\text{ties}) \]

Therefore, the first-order linear approximation in q1 and q2 of the expected value of the hazard rate of the extended AIM is equal to the hazard rate of the discretized AIM, with p^{AIM} = p, q^{AIM} = q, d^{AIM} = d, E(\text{ties}) = E(\text{ties}), p^{AIM} = p_2, q^{AIM} = q_2(1-p_1)a \cdot E(\text{ties}) + a^2 \cdot E(\text{ties}), w^{AIM} = a \cdot E(\text{ties}) + a^2 \cdot E(\text{ties}).
Appendix 3: Calibration of the Extended MIM

We describe here the likelihood contribution of each piece of survey data.

Number of recommendations received \((P(r_{it} | ties_{i}))\): Let \(r_{it}\) refer to the (latent) number of recommendations received by consumer \(i\) in period \(t\).

\[
P^{\text{EXT-MIM}}(r_{it} | ties_{i}) = \sum_{\{r_{il}: l=1...5\}} P^{\text{EXT-MIM}}(\{r_{il} \}_{l=1...5} | ties_{i}) = \sum_{\{r_{il}: l=1...5\}} \prod_{t=1}^{5} \left( f_{r_{il}}(ties_{i}) \right)^{r_{il}} (1 - f_{r_{il}}(ties_{i}))^{ties_{i} - r_{il}}
\]

where the cumulative penetration \(F_{t-1}\) is given by using the closed-form expression in Equation (4).

Adoption conditional on recommendations received \((P(y_{i} | r_{i}, ties_{i}))\): Based on the conditional hazard rate from Equation (1), the cumulative hazard rate conditional on \(r_{i}\) is directly obtained as: 10

\[
P^{\text{EXT-MIM}}(y_{i} = 1 | r_{i}) = m(1 - (1 - p)^{r_{i}}(1 - q)^{ties_{i}})
\]

Number of recommendations given conditional on adoption and recommendations received \((P(g_{i} | y_{i}=1,r_{i}, ties_{i}))\): Let \(g_{it}\) be the (latent) number of recommendations given by respondent \(i\) in period \(t\). Equation (2) implies that for all periods following adoption we have: \(g_{i} \sim \text{Binomial}(ties_{i}, a)\). Suppose that consumer \(i\) adopted the innovation in period \(t_{i}\). The total number of recommendations given by this consumer over periods \(t_{i}+1\) to 5 is therefore given by the following binomial distribution:

\[
g_{i} \sim \text{Binomial}(ties_{i} - (5 - t_{i}), a).\]

Integrating out over the possible values of \(t_{i}\) gives:

\[
P^{\text{EXT-MIM}}(g_{i} | y_{i} = 1, r_{i}, ties_{i}) = \sum_{t_{i}=1}^{5} \left(ties_{i} - (5-t_{i})\right)^{a} P^{\text{EXT-MIM}}(\text{adopted in } t_{i} | y_{i} = 1, r_{i}, ties_{i})
\]

where:

\[
P^{\text{EXT-MIM}}(\text{adopted in } t_{i} | y_{i} = 1, r_{i}, ties_{i}) = \frac{P^{\text{EXT-MIM}}(\text{adopted in } t_{i} | r_{i}, ties_{i})}{\sum_{t_{i}=1}^{5} P^{\text{EXT-MIM}}(\text{adopted in } t_{i} | r_{i}, ties_{i})},
\]

and

\[
P^{\text{EXT-MIM}}(\text{adopted in } t_{i} | r_{i}, ties_{i}) = \sum_{\{r_{il} : \Sigma_{l} r_{il} = ties_{i}\}} P^{\text{EXT-MIM}}(\{r_{il} \}_{l=1...5} | r_{i}, ties_{i})
\]

We estimated this model using Bayesian MCMC (Rossi and Allenby, 2003), with the following uninformative priors: \(\sigma^2 \sim IG(r_{0}, s_{0})\) with \(r_{0} = s_{0} = 1\), \(p\), \(q\) and \(a\) uniform on \([0,1]\), and \(m\) uniform on \([0,1]\). The Metropolis-Hastings algorithm was used for all the parameters, except for \(\sigma\) which was drawn directly from its (inverse-gamma distributed) conditional posterior distribution. We used 300,000 MCMC iterations, using the first 250,000 as burn-in and saving 1 in every 10 draws. Convergence was assessed informally through time-series plots of the parameters. The original MIM was estimated using a similar procedure, with the same uninformative priors on \(\sigma\) and \(m\) as above and diffuse proper priors (on \(R^{+}\)) for \(p\) and \(q\). Similar priors and numbers of draws were used for the other models as well (AIM and KG).

---

10 The hazard rate from Equation (1) applies to consumers who are potential adopters. This hazard rate must be multiplied by \(m\), the proportion of potential adopters in the market, to model the hazard rate of a consumer from a representative consumer panel (which is not limited to potential adopters).
Appendix 4: Calibration of the extended AIM

Following Van den Bulte and Joshi (2007), we limit ourselves to two special cases of the AIM: a) one in which the innovators and imitators segments are independent (w=0 in the original AIM; \( t_i = 1 \rightarrow 2 = a^{1 \rightarrow 2} = 0 \) in the extended AIM); and b) one in which innovators are not subject to social influence (\( q_g = 0 \) in the original AIM; \( q_g = a^{1 \rightarrow 1} = t_i = 0 \) in the extended AIM). This allows avoiding the problematic computation of Gaussian hypergeometric functions in the original AIM (Van Den Bulte and Joshi 2007, page 412). Following Van Den Bulte and Joshi (2007), we retain the special case that provides the best fit with the calibration data (measured in our case using the log marginal density). We describe here the likelihood contribution of each piece of the survey data.

Number of recommendations received \((P(r_i \mid \text{ties}_i))\):

\[
P_{\text{EXT-AIM}}(r_i \mid \text{ties}_i) = P_{\text{EXT-AIM}}(r_i \mid \text{innovator, ties}_i)\theta + P_{\text{EXT-AIM}}(r_i \mid \text{imitator, ties}_i)(1 - \theta)
\]

Special case a): \( P_{\text{EXT-AIM}}(r_i \mid \text{innovator, ties}_i) = \sum_{\{t_i, h_i, j_i, s_i, a_i, g\}} \prod_{r_i=1}^{5} \left( \text{ties}_i^{r_i=1} \right) (a^{1 \rightarrow 1} F_{i=1}^{1})^{\text{ties}_i^{r_i=1}} (1 - a^{1 \rightarrow 1} F_{i=1}^{1})^{\text{ties}_i^{r_i=1}-w_i}
\]

\[
P_{\text{EXT-AIM}}(r_i \mid \text{imitator, ties}_i) = \sum_{\{t_i, h_i, j_i, s_i, a_i, g\}} \prod_{r_i=1}^{5} \left( \text{ties}_i^{r_i=2} \right) (a^{2 \rightarrow 2} F_{i=2}^{2})^{\text{ties}_i^{r_i=2}} (1 - a^{2 \rightarrow 2} F_{i=2}^{2})^{\text{ties}_i^{r_i=2}-w_i}
\]

where \( \text{ties}_i \) = \( \text{ties}_i^{r_i=1} \) for innovators and \( \text{ties}_i \) = \( \text{ties}_i^{r_i=2} \) for imitators.

Adoption conditional on recommendations received \((P(y \mid r, \text{ties}_i))\):

\[
P_{\text{EXT-AIM}}(y = 1 \mid r, \text{ties}_i) = m[1 - (1 - p_1)^{\text{ties}_i^{r_i=1}} (1 - q_1)^{\text{ties}_i^{r_i=2}}] \cdot P_{\text{EXT-AIM}}(\text{innovator} \mid r, \text{ties}_i) + [1 - (1 - p_2)^{\text{ties}_i^{r_i=1}} (1 - q_2)^{\text{ties}_i^{r_i=2}}] \cdot P_{\text{EXT-AIM}}(\text{imitator} \mid r, \text{ties}_i)
\]

Number of recommendations given conditional on adoption and recommendations received \((P(g \mid y=1, r, \text{ties}_i))\):

\[
P_{\text{EXT-AIM}}(g \mid y = 1, r, \text{ties}_i) = P_{\text{EXT-AIM}}(g \mid y = 1, r, \text{ties}_i, \text{innovator}) \cdot P_{\text{EXT-AIM}}(\text{innovator} \mid y = 1, r, \text{ties}_i)
\]

\[
+ P_{\text{EXT-AIM}}(g \mid y = 1, r, \text{ties}_i, \text{imitator}) \cdot P_{\text{EXT-AIM}}(\text{imitator} \mid y = 1, r, \text{ties}_i)
\]

\[
P_{\text{EXT-AIM}}(g \mid y = 1, r, \text{ties}_i) = \frac{P_{\text{EXT-AIM}}(r, y = 1 \mid \text{innovator, ties}_i) \cdot \theta}{P_{\text{EXT-AIM}}(r \mid \text{innovator, ties}_i) \cdot \theta + P_{\text{EXT-AIM}}(r \mid \text{imitator, ties}_i) \cdot (1 - \theta)}
\]

\[
P_{\text{EXT-AIM}}(g \mid y = 1 \mid \text{innovator}) = P_{\text{EXT-AIM}}(r \mid \text{innovator, ties}_i) (1 - (1 - p_1)^{\text{ties}_i^{r_i=1}} (1 - q_1)^{\text{ties}_i^{r_i=2}})
\]

\[
P_{\text{EXT-AIM}}(g \mid y = 1 \mid \text{imitator}) = P_{\text{EXT-AIM}}(r \mid \text{imitator, ties}_i) (1 - (1 - p_2)^{\text{ties}_i^{r_i=1}} (1 - q_2)^{\text{ties}_i^{r_i=2}})
\]

42
Special case a): $P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}} = \sum_{l=1}^{s} \left( \text{ties}_{l}^{y_i=1} (5-t_i) \right) \left( 1 - a^{t_i} \text{ties}_{l}^{y_i=1} (5-t_i) \right) a^{t_i} P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}}$ \\

$P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}} = \sum_{l=1}^{s} P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}}$ \\

$P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}} = \frac{\sum (1 - p_i) (1 - q_i) \text{ties}_{l}^{y_i=1} (5-t_i) a^{t_i} P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}}}{P_{\text{EXT-AIM}}^{y_i=1,ties_{i innovate}}}$ \\

$P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}} = \frac{\sum (1 - p_i) (1 - q_i) \text{ties}_{l}^{y_i=1} (5-t_i) a^{t_i} P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}}}{P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}}}$ \\

Special case b): $P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}} = \sum_{l=1}^{s} \left( \text{ties}_{l}^{y_i=1} (5-t_i) \right) \left( 1 - a^{t_i} \text{ties}_{l}^{y_i=1} (5-t_i) \right) a^{t_i} P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}}$ \\

$P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}} = \sum_{l=1}^{s} P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}}$ \\

$P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}} = \frac{\sum (1 - p_i) (1 - q_i) \text{ties}_{l}^{y_i=1} (5-t_i) a^{t_i} P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}}}{P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}}}$ \\

$P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}} = \frac{\sum (1 - p_i) (1 - q_i) \text{ties}_{l}^{y_i=1} (5-t_i) a^{t_i} P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}}}{P_{\text{EXT-AIM}}^{y_i=1,r_i,ties_{i innovate}}}$
Appendix 5: Calibration of the extended KG Model

We describe here the likelihood contribution of each piece of survey data.

Number of recommendations received (\(P(r_i | ties_i)\)):

\[ P^{\text{EXT-KG}}_{r_i | ties_i} = \sum_{p \in \{p_1, p_2, \ldots, p_T\}} P^{\text{EXT-KG}}(r_i | p, q, ties_i) g(p, q) \]

\[ P^{\text{EXT-KG}}(r_i | p_1, q_1, ties_i) = \sum_{q \in \{q_1, q_2, \ldots, q_T\}} P^{\text{EXT-KG}}(r_i | p_1, q, ties_i) = \sum_{q \in \{q_1, q_2, \ldots, q_T\}} \prod_{i=1}^{s} \binom{\text{ties}_i}{r_i} a_{r_i} \left(1 - a_{r_i + 1}\right)^{\text{ties}_i - r_i} \]

and similarly for the other segments, where \(g(p, q)\) is the probability of the combination \(\{p, q\}\).

Adoption conditional on recommendations received (\(P(y_i | r_i, ties_i)\)):

\[ P^{\text{EXT-KG}}_{y_i = 1 | r_i, ties_i} = \sum_{p \in \{p_1, p_2, \ldots, p_T\}} m[1 - (1 - p)^5 (1 - q)^5] P^{\text{EXT-KG}}(p, q | r_i, ties_i) \]

\[ P^{\text{EXT-KG}}_{y_i = 1 | r_i, ties_i} = \frac{P^{\text{EXT-KG}}(r_i | p, q, ties_i) g(p, q)}{\sum_{p \in \{p_1, p_2, \ldots, p_T\}} P^{\text{EXT-KG}}(r_i | p, q, ties_i) g(p, q)} \]

Number of recommendations given conditional on adoption and recommendations received (\(P(g_i | y_i = 1, r_i, ties_i)\)):

\[ P^{\text{EXT-KG}}_{g_i = 1 | y_i = 1, r_i, ties_i} = \sum_{p \in \{p_1, p_2, \ldots, p_T\}} P^{\text{EXT-KG}}(g_i | y_i = 1, r_i, p, q, ties_i) P^{\text{EXT-KG}}(p, q | y_i = 1, r_i, ties_i) \]

\[ P^{\text{EXT-KG}}_{g_i = 1 | y_i = 1, r_i, ties_i} = \frac{P^{\text{EXT-KG}}(r_i | y_i = 1, p, q, ties_i) g(p, q)}{\sum_{p \in \{p_1, p_2, \ldots, p_T\}} P^{\text{EXT-KG}}(r_i | y_i = 1, p, q, ties_i) g(p, q)} \]

\[ P^{\text{EXT-KG}}(r_i | y_i = 1 | p, q, ties_i) = P^{\text{EXT-KG}}(r_i | p, q, ties_i) (1 - (1 - p)^5 (1 - q)^5) \]

\[ P^{\text{EXT-KG}}_{g_i = 1 | y_i = 1, r_i, p, q, ties_i} = \sum_{i=1}^{s} \binom{\text{ties}_i - (5 - t_i)}{(5 - t_i) - g_i} (1 - a_{g_i + 1})^{(s - (5 - t_i) - g_i)} P^{\text{EXT-KG}}(\text{adopted in } t_i | y_i = 1, r_i, p, q, ties_i) \]

\[ P^{\text{EXT-KG}}_{\text{adopted in } t_i | y_i = 1, r_i, p, q, ties_i} = \frac{\sum_{i=1}^{s} P^{\text{EXT-KG}}(\text{adopted in } t_i | r_i, p, q, ties_i) \binom{\text{ties}_i - (5 - t_i)}{(5 - t_i) - g_i} (1 - a_{g_i + 1})^{(s - (5 - t_i) - g_i)} P^{\text{EXT-KG}}(\text{adopted in } t_i | r_i, p, q, ties_i)}{\sum_{i=1}^{s} P^{\text{EXT-KG}}(\text{adopted in } t_i | r_i, p, q, ties_i)} \]

\[ P^{\text{EXT-KG}}_{\text{adopted in } t_i | y_i = 1, r_i, p, q, ties_i} = \frac{\sum_{i=1}^{s} (1 - p)^{5 - t_i} (1 - q)^{5 - t_i} - g_i - 1 (1 - (1 - p)(1 - q)^5) P^{\text{EXT-KG}}(\{r_i\} | p, q, ties_i)}{P^{\text{EXT-KG}}(r_i | p, q, ties_i)} \]