MARKETING SCIENCE
Vol. 31, No. 4, July–August 2012, pp. 689–712
ISSN 0732-2399 (print) | ISSN 1526-548X (online)

http://dx.doi.org/10.1287/mksc.1120.0711

Network Traces on Penetration: Uncovering Degree Distribution from Adoption Data

Yaniv Dover
Yale School of Management, Yale University, New Haven, Connecticut 06520,
yaniv.dover@yale.edu

Jacob Goldenberg
Jerusalem School of Business Administration, Hebrew University, Jerusalem, Israel 91905; and
Graduate School of Business, Columbia University, New York, New York 10027, msgolden@huji.ac.il

Daniel Shapira
Guilford Glazer Faculty of Business and Management, Ben-Gurion University, Beer Sheva, Israel 84105,
shapirad@bgu.ac.il

We show how networks modify the diffusion curve by affecting its symmetry. We demonstrate that a network's degree distribution has a significant impact on the contagion properties of the subsequent adoption process, and we propose a method for uncovering the degree distribution of the adopter network underlying the dissemination process, based exclusively on limited early-stage penetration data. In this paper we propose and empirically validate a unified network-based growth model that links network structure and penetration patterns. Specifically, using external sources of information, we confirm that each network degree distribution identified by the model matches the actual social network that is underlying the dissemination process. We also show empirically that the same method can be used to forecast adoption using an estimation of the degree distribution and the diffusion parameters at an early stage (15%) of the penetration process. We confirm that these forecasts are significantly superior to those of three benchmark models of diffusion.

Our empirical analysis indicates that under heavily right-skewed degree distribution conditions (such as scale-free networks), the majority of adopters (in some cases, up to 75%) join the process after the sales peak. This strong asymmetry is a result of the unique interaction between the dissemination process and the degree distribution of its underlying network.

Key words: social networks; diffusion of innovation; diffusion models; forecasting
History: Received: December 20, 2009; accepted: January 17, 2012; Eric Bradlow and then Preyas Desai served as the editor-in-chief and Christophe Van den Bulte served as associate editor for this article. Published online in *Articles in Advance* April 10, 2012.

1. Introduction

The structure of a network, i.e., the pattern of connections among its nodes, significantly affects how information and new products diffuse (Goldenberg et al. 2009b, Hill et al. 2006, Katona and Sarvary 2009, Katona et al. 2011, Mayzlin 2002, Newman et al. 2006, Shaikh et al. 2006, Van den Bulte and Wuyts 2007). Because network structure is typically invisible, the few diffusion models that incorporate network structure do so by resorting to rough assumptions (e.g., division into two markets) that oversimplify the diffusion process and its dynamics. One of these oversimplifying assumptions in the literature is that networks generally exhibit scale-free degree distributions,[1] despite the fact that other types of distributions have been documented in studies of social networks, including Gaussian/Poissonian distributions in offline social groups (Amaral et al. 2000), online forums, and active email networks (Liben-Nowell and Kleinberg 2008, Newman et al. 2002, Yeung 2005), as well as lognormal distributions in a Web linkage structure and in online social networks (Gomez et al. 2009, Pennock et al. 2002, Stutzbach and Rejaie 2005). Here, we model the network structure—specifically, its degree distribution—and we investigate its effect on the adoption curve.

Because the magnitude and speed of the contagion process, as well as the shape of the adoption curve, depend strongly on the degree distribution of the underlying network (as we show in §3 and validate in §6), knowledge of the degree distribution of a given network is important for marketers to effectively isolate network effects from other factors affecting contagion.

[1] A degree distribution is the probability distribution of the number of links per node over the entire network.

Using analytical modeling, numerical simulations, and empirical studies, we show that the network degree distribution significantly affects the temporal form of the penetration curve. The adoption growth rate depends on the average number of neighbors and the variance of the degree distribution. In cases where the network degree is significantly skewed (e.g., the common scale-free distribution), the penetration curve is asymmetrical. Initial growth of sales is accelerated, and the sales peak occurs early in the process, whereas sales decline occurs at a slower rate and extends over a longer period. Such asymmetry has implications for estimating market potential and for generating reliable forecasts early in the penetration process.

It has previously been shown that adoption curve asymmetry may result from asymmetric adopter influence (Van den Bulte and Joshi 2007), heterogeneous first-purchase timing (Bemmaor 1994, Bemmaor and Lee 2002), or nonuniform properties of word-of-mouth processes (Easingwood et al. 1983). In §3, we show how adoption curve asymmetry may be explained by the underlying network's degree distribution, which leaves traces on each penetration process. Analysis of the heterogeneity of connectivity improves the ability to identify the degree distribution and allows us to correct contagion estimations and enhance forecasting accuracy.

Previous works have used models to show that a given pattern of adoption can be explained by network effects (although alternative explanations can, in principle, be provided). In this paper, we propose and empirically validate a unified network-based growth model that links network structure and penetration patterns. Using external sources of information, we confirm that each network degree distribution identified by the model matches the actual social network that is underlying to the dissemination process (see §6). We show that the course of diffusion is an indirect measure of the degree distribution of the underlying network, owing to the manner in which the network affects the adoption pattern. More specifically, the paper offers the following contributions:

1. We show analytically that in contagion-based dissemination over an undirected[2] random network[3] (with up to an average level of clustering),[4] an increase in degree heterogeneity increases contagion, even if the average number of ties remains constant.

2. We show analytically and by simulation (§§3 and 4) that highly skewed degree distributions lead to more highly skewed adoption curves. This is validated empirically in §§6.1.2 and 6.2.

3. We propose (in §5) a method for estimating the type of degree distribution of the network (i.e., scale-free, normal, lognormal, or uniform), as well as its two first moments, based solely on early-phase penetration data. We validate this empirically (in §6.1.2) and using simulated data (in §6.1.1), implying that a network leaves its traces on the resulting penetration pattern.

4. On the basis of the previous three contributions, we show that such uncovered network information can be incorporated into a unified growth model to generate more accurate forecasts of adoption at a very early stage in the penetration process, when such forecasts have significant managerial implications (see §6.3).

The proposed model operates well within a set of boundary conditions defined in §7. The main simplifying assumptions of the model are symmetrical ties (undirected network with identical tie strength across nodes), no clustering, constant marketing efforts over time and space, and lack of contagion decay. We analyze the model's sensitivity to these assumptions in §7. Model accuracy is reduced when the network is fragmented (either disconnected or with an extremely high degree of clustering) or when the dissemination process exhibits a flat curve (typically associated with extremely low contagion or with failures).

The remainder of this paper is as follows. Section 2 is a review of past research and the background for our work. In §3, we develop the analytical baseline for our analysis of the influence of degree heterogeneity on contagion. In §4, we show how degree heterogeneity can lead to high skewness in adoption curves. We propose a model for estimating network degree distribution using aggregate penetration data in §5. This model is tested in §6.1 using simulated data (in §6.1.1). In a more stringent test, we use real microlevel network data to validate the ability of the method to correctly estimate the underlying network (in §6.1.2). In §6.2, we further test the model using a variety of empirical cases. The value of network information is explored in §6.3 by comparing forecasts generated by this model to the performance of three alternative models—the gamma/shifted Gompertz model (Bemmaor 1994, Bemmaor and Lee 2002), the nonuniform influence model (Easingwood et al. 1983), and the Bass model (Bass 1969)—using only data from the early stages of the diffusion process. Finally, in §7, we discuss the boundary conditions of the model, and we summarize our conclusions in §8.

[2] An undirected network is a network in which the links are undirected; i.e., a link between any nodes A and B is also a link between B and A.

[3] A random network is a network in which the links between nodes are determined randomly.

[4] The clustering coefficient of a given node is defined as the ratio of the number of actual links to the number of potential links in a node's neighborhood. The average clustering coefficient is the average of the node-level clustering coefficients over the entire network.

2. Background

Recent work highlights the importance of knowledge of network structure to marketers. Network structure has been shown to have significant effects on the flow of information and influence (Katona and Sarvary 2009, Stephen and Toubia 2010). For example, a category of consumers known as "influentials" has been identified as one source of network effects (for an early observation, see Katz and Lazarsfeld 1955; for a generalized and quantified model, see Van den Bulte and Joshi 2007; and for a related field experiment study, see Hinz et al. 2011). In addition to traits such as expertise, influentials' primary source of influence is their extraordinarily great number of ties to other individuals (high degree). Whereas the actual influence of opinion leaders' degree distribution has been questioned by Watts and Dodds (2007), recent empirical evidence confirms that influentials accelerate the diffusion process (e.g., Goldenberg et al. 2009b) and that a high indegree is correlated with earlier adoption (Iyengar et al. 2011, Katona et al. 2011). It has further been shown that network effects have concrete economic value (Gupta et al. 2006); i.e., network links add monetary value to a market. Stephen and Toubia (2010) show that specific network structures, such as those including individuals who have a single link each, might even diminish economic market value. In another application, Hill et al. (2006) show that knowledge of the consumers' network structure can significantly improve the firm's ability to predict consumers' likelihood to purchase. Finally, Shaikh et al. (2006) use a generalized diffusion model to evaluate the structure of small-world networks and find that structure has an important impact on the temporal aspects of new product diffusion. They demonstrate that disregard of network effects may lead to incorrect interpretation of penetration data, and they conclude that knowledge of network structure plays a substantial role in the estimation of contagion parameters.

Recently, models of diffusion have been revised to capture the effect of network structure on new product diffusion. Typically, these models encounter one common obstacle: networks of social influence are usually invisible or extremely hard to map (Ebbes et al. 2010). In view of this inherent challenge, marketers and researchers generally assume that dissemination processes propagate over the entire visible (overt) network. Despite its widespread support, this is an oversimplified assumption, as the dissemination, in fact, occurs over a subset of this overt network. We define this subset as the *active network*, and we argue that it does not necessarily share the same properties or structure as the overt network (similar to Stumpf et al. 2005). For example, in a network characterized by a scale-free distribution, a diffusion process might spread only among a subset of nodes (the actual adopters, i.e., the active network) characterized by a Gaussian network degree distribution. Indeed, Internet chain letters have been found to propagate in a narrow but very deep tree-like pattern, continuing for several hundred steps, rather than fanning out widely and reaching many people in very few steps, as would be expected according to small-world principles (Liben-Nowell and Kleinberg 2008).

There are related works on the estimation of unknown network properties from other known properties (e.g., number of links, link probability) using maximum-likelihood methods (Garlaschelli and Loffredo 2008, Ramasco and Mungan 2008). Recently, Trusov and Rand (2011) used a combination of Bayesian model selection techniques and a large number of aggregate data sets over a single population to infer the "universal" social network associated with that population. On the micro level, Braun and Bonfrer (2010) developed a method to uncover the hidden dyad-level interdependence between consumers. Our approach is different, mainly because we attempt to use a single set of aggregate, early-phase penetration data to calculate the degree distribution of the active social network. We then use that information to predict the market potential and adoption curve in those early phases.

3. The Impact of Degree Heterogeneity on Contagion

Consider the simplest case of social influence, in which individuals affect one another equally, the market is homogeneous for both external and internal influences, and both effects are constant in time. These assumptions lead to the Bass differential equation (Bass 1969):

$$\frac{dN(t)}{dt} = \left(P + Q \cdot \frac{N(t)}{M}\right) \cdot (M - N(t)) = MP + (Q - P)N(t) - \frac{Q}{M}N(t)^2. \quad (1)$$

Here, $N(t)$, M, P, and Q are the cumulative number of adopters at time t, population size (market potential), external force coefficient, and the internal force coefficient, respectively. The notation for the variables used throughout are listed in Table A.1 in the appendix.

For the early stages of the process (i.e., for low values of t), the cumulative number of adopters is relatively small in comparison to the market potential, M, so that $N(t) \ll M$, yielding the following first-order linear approximation:

$$\frac{dN(t)}{dt} = MP + (Q - P)N(t) - Q\left(\frac{N(t)}{M}\right)N(t) \approx MP + (Q - P)N(t). \quad (2)$$

In its early stages, the adoption process is characterized by exponential growth of the following form:

$$\frac{dN(t)}{dt} \propto e^{(Q-P)t}. \tag{3}$$

As the process progresses (i.e., for high values of t), $N(t) \sim M$, and hence $\delta(t) = M - N(t) \ll M$, yielding the following linear approximation:

$$\begin{aligned}\frac{dN(t)}{dt} &= \left(P + Q\left(1 - \frac{\delta(t)}{M}\right)\right)\delta(t) \approx (P+Q)\delta(t) \\ &= (P+Q)(M - N(t)).\end{aligned} \tag{4}$$

Thus, toward the conclusion of the process, the adoption rate declines exponentially:

$$\frac{dN(t)}{dt} \propto e^{-(Q+P)t}. \tag{5}$$

Both slopes are exponential and therefore become linear on a log scale.

In a scenario where the internal force is greater than the external force ($P \ll Q$), the temporal dynamics for a fully connected market (no network is assumed) are symmetrical around the peak, and the absolute value of the exponent in the growth phase is approximately equal to that in the decline phase. As will be shown, the existence of a network degree distribution disrupts this symmetry. Furthermore, the deviations from the assumed symmetry provide indications of the network degree distribution.

Consider a random network, i.e., a network in which the nodes are connected randomly, with an equal probability for any pair of nodes to be connected, up to a limit of k neighbors for each node (Erdös and Rényi 1959). The parameter k, i.e., the number of neighbors or network degree of a node, is retrieved from the network degree distribution probability mass function P_k. In each time step Δt, adopters have a probability $q\Delta t$ of influencing their neighbors (persuading them to adopt the product). Potential adopters also have a probability $p\Delta t$ of being influenced by external forces (marketing forces) to adopt the product. We also assume that both internal and external influence rates are completely homogeneous in time and space and that the network is undirected. Assuming that internal and external forces are independent, the probability that a potential adopter with x neighbors who have already adopted the product will adopt the product in time step Δt is linear and takes the form $(p + xq)\Delta t$. In the linear model, the dynamics of the expected rate of adoption over a network in the continuous limit (Δt approaches zero) are given by

$$\frac{dN(t)}{dt} = \sum_x H_x(t) \cdot (p + xq). \tag{6}$$

Here, $H_x(t)$ denotes the number of potential adopters of order x, i.e., consumers who have not yet adopted but have exactly x neighbors who have already adopted the product at time t. This is effectively an individual-level diffusion model that incorporates the network effect through the effect of degree heterogeneity. The hazard rate is greater for individuals with greater values of x, implying that adoption rates may vary under different degree distributions. One of the immediate properties of (6) is that the adoption rate increases as individual exposure to adopters increases (potential adopters with more adopter friends adopt faster, *on average*).

The *total* number of potential adopters at a given time is

$$\sum_x H_x(t) = M - N(t). \tag{7}$$

Formulated in this manner, the Bass equation can be viewed as a special case of (6). Assuming that all individuals are connected to all other individuals, the order of all potential adopters at time t is $x = N(t)$, where $H_x(t) = M - N(t)$. Thus, $dN(t)/dt = (M - N(t))(P + (Q/M)N(t))$, where $P = p$ and $Q = Mq$.

In Appendix §A.1, we describe this model of dissemination formally on a generalized random network, which can be used as an approximation for the adopter network underlying the diffusion process (see also the sensitivity analysis of the model's assumptions in §7). We assume that the maximal degree in this network is significantly smaller than the size of the network (which is a realistic assumption for social networks). We calculate the functions $H_x(t)$ together with a set of conditional probabilities of the form $f_{k|x}(t)$ (where $f_{k|x}(t)$ is the conditional probability that a potential adopter of order x at time t has a network degree of k; see Appendix §A.1). The conditional probabilities $f_{k|x}(t)$ and the functions $H_x(t)$ are then introduced as the solutions to a closed system of coupled ordinary differential equations. The initial conditions for these equations are determined exclusively by network size (i.e., market potential M) and network degree distribution. To extract information for the estimation procedure, we analyze the adoption process under two phases: (1) the growth regime in early stages of penetration and (2) the decline regime in later stages of the process.

Under the following analysis, penetration grows exponentially in the initial stages of the process, and its slope is determined not only by the magnitude of contagion but also by the network degree distribution. Because the structure of the network is random, the probability that two neighbors of a specific node that adopted the product are neighbors themselves is negligible in the early stages of the process (for nonfragmented networks that are sufficiently large

relative to the maximal network degree). We therefore find that the adoption rate in the early stages of the process takes the following form (see Appendix §A.1 for details):

$$\frac{dN(t)}{dt} = (M - N(t))p + qH_1(t) \approx Mp\left(1 - \frac{k_{\text{avg}}q}{\tilde{Q} - 2p}\right)e^{-pt} + Mp\frac{k_{\text{avg}}q}{\tilde{Q} - 2p}e^{(\tilde{Q}-3p)t}. \quad (8)$$

Here, $\tilde{Q}$ is the effective internal force coefficient. All high-order terms that include $H_x(t)$ for $x \geq 2$ are dropped.

Based on (8), the effective internal force coefficient $\tilde{Q}$ is determined by the ratio of the second moment of the network degree distribution to the first moment, giving

$$\tilde{Q} = q\left\{\frac{\sum_{k=k_{\min}}^{k_{\max}} k^2 P_k}{\sum_{k=k_{\min}}^{k_{\max}} kP_k} - 2\right\} = q\left(k_{\text{avg}} + \frac{\sigma^2}{k_{\text{avg}}} - 2\right), \quad (9)$$

where

$$k_{\text{avg}} = \sum_{k=k_{\min}}^{k_{\max}} kP_k \quad \text{and}$$

$$\sigma^2 = \sum_{k=k_{\min}}^{k_{\max}} (k - k_{\text{avg}})^2 P_k$$

are, respectively, the average and variance of the network degree distribution. For the derivation of (8), the assumption of a strong contagion process yields $\tilde{Q} \gg p$. Thus, when $\tilde{Q} > 3p$ (e.g., Farley et al. 1995), the rate of adoption exhibits exponential growth:

$$\frac{dN(t)}{dt} \propto e^{(\tilde{Q}-3p)t}. \quad (10)$$

That is, for a given average, greater *network degree variance leads to greater exponential growth*. This occurs because the presence of highly connected individuals accelerates growth.

For an intuitive perspective on the rationale behind this growth expression, consider the simpler case of $p = 0$. In this case, in the initial noninteractive stage, each exposed consumer has, on average, only one adopter neighbor who potentially induces her adoption. Therefore, if an individual with network degree k adopts the innovation in the initial stages of the process, the number of exposed consumers increases by $k - 2$, as she has $k - 1$ neighbors who have not yet adopted the new product, but also, by becoming an adopter of the innovation, she herself is removed from the group of exposed consumers. Taking into account that nodes with a greater number of links have a greater probability, on average, of being exposed to a spreading adoption process, each new adoption adds $\sum_{k=k_{\min}}^{k_{\max}} (k-2)\tilde{P}_k = \sum_{k=k_{\min}}^{k_{\max}} k\tilde{P}_k - 2$ individuals to the number of exposed consumers, where $\tilde{P}_k$ is the probability mass function for the degree distribution among exposed consumers, given by the probability that the new adopter is a neighbor of another node with degree k. The probability mass function of the distribution of the degrees of neighbors in random networks is known to be (Albert and Barabási 2002)[5]

$$\tilde{P}_k = \frac{k \cdot P_k}{k_{\text{avg}}}, \quad (11)$$

where P_k is the probability mass function for the degree distribution and k_{avg} is the average degree for the purpose of normalization. (Note that this probability is not the original probability P_k that a certain node in the network has degree k. Consequently, the probability that this node is included in another node's sample is skewed at higher degrees.) The average increase in the number of exposed consumers per current adopter in a single time step is $\sum_{k=k_{\min}}^{k_{\max}} (k-2)\tilde{P}_k = \sum_{k=k_{\min}}^{k_{\max}} k\tilde{P}_k - 2$, and hence, the total increase in the number of exposed consumers in a single time step is

$$\frac{dH_1(t)}{dt} = \left(\sum_{k=k_{\min}}^{k_{\max}} k\tilde{P}_k - 2\right) \cdot qH_1(t) = \left(k_{\text{avg}} + \frac{\sigma^2}{k_{\text{avg}}} - 2\right)qH_1(t) \equiv \tilde{Q}H_1(t). \quad (12)$$

In the case of $p = 0$, the adoption rate becomes $dN(t)/dt = qH_1(t) \propto e^{\tilde{Q}t}$. Adding external influence ($p > 0$) moderates the slope of the exponential growth, as indicated by (8).

Equation (9) is the analytical measure of the dependence of effective internal force on network degree heterogeneity. The coefficient depends on the average number of neighbors and the standard deviation of the degree distribution. As a result, diffusion is accelerated when the network is highly skewed.[6] A highly skewed distribution can change the average and standard deviation by several orders of magnitude and consequently affect the internal force coefficient. Thus, the same product diffusing on different networks might exhibit significantly different $\tilde{Q}$s (but not necessarily different qs). To assess the

[5] Owing to the assumption of network randomness, the actual neighbors of adopters' degree distribution are not different from the neighbors of randomly chosen nodes.

[6] Note that the effect of heterogeneity in (9) is normalized by connectivity. Therefore, $\tilde{Q}$ decreases with connectivity for $k_{\text{avg}} < \sigma$ and increases with connectivity for $k_{\text{avg}} > \sigma$. Nonetheless, in the very early stages of the process, overall growth increases with the connectivity in both cases. Namely, $(\partial/\partial k_{\text{avg}})(dN/dt) > 0$ when t is small, as indicated by Equation (8).

error generated by ignoring network degree distribution, consider the simplest estimate of contagion magnitude using the Bass model. In this case, the effect of the internal force (Q) is underestimated because the model multiplies the dyad-level, q coefficient by the entire market potential, M, rather than by the network, q, structure multiplier, $k_{\text{avg}} + \sigma^2/k_{\text{avg}} - 2$.

4. Degree Heterogeneity Leads to Highly Skewed Adoption Curves

By the final stage of the adoption process (large t), almost all individuals have adopted, and most remaining nonadopters are linked exclusively to adopters. As a result, in almost all cases, the order (number of neighbors who have already adopted) of a nonadopter is equal to his or her network degree.[7] Therefore, at this stage, the sole impact of product adoption by an individual with network degree k is a reduction in $H_k(t)$, the number of potential adopters of order k. Recalling that the probability per time step of a potential adopter of order k to adopt is $kq + p$, we obtain

$$\frac{dH_k(t)}{dt} = -(kq + p) \cdot H_k(t), \tag{13}$$

resulting in $H_k(t) \propto e^{-(kq+p)t}$. The rate of adoption in the final stages of the process can be obtained from (5) as follows:

$$\frac{dN(t)}{dt} = \sum_k H_k(t) \cdot (p + kq) \approx \sum_k C_k e^{-(kq+p)t}, \tag{14}$$

where C_k are time-independent coefficients. Namely, the rate of adoption is given by the sum of time-decaying exponentials (representing different populations with different degrees), which is dominated by the slowest-decaying exponent. The adoption rate declines according to the following rule:

$$\frac{dN(t)}{dt} \propto e^{-(k_{\min} q + p)t}, \tag{15}$$

where $k_{\min}$ denotes the smallest degree in the network. (The dynamics throughout the decline stage are derived directly from the general equations of diffusion on random networks in Appendix §A.1.) The growth and the decline slopes are shown to be, respectively,

$$x_1 = q\left(k_{\text{avg}} + \frac{\sigma^2}{k_{\text{avg}}} - 2\right) - 3p, \tag{16}$$

$$x_2 = -(k_{\min} \cdot q + p). \tag{17}$$

[7] At this stage, most nonadopters are linked exclusively to adopters, so for a given nonadopter, the potential number of adopter-neighbors approaches the nonadopter's network degree.

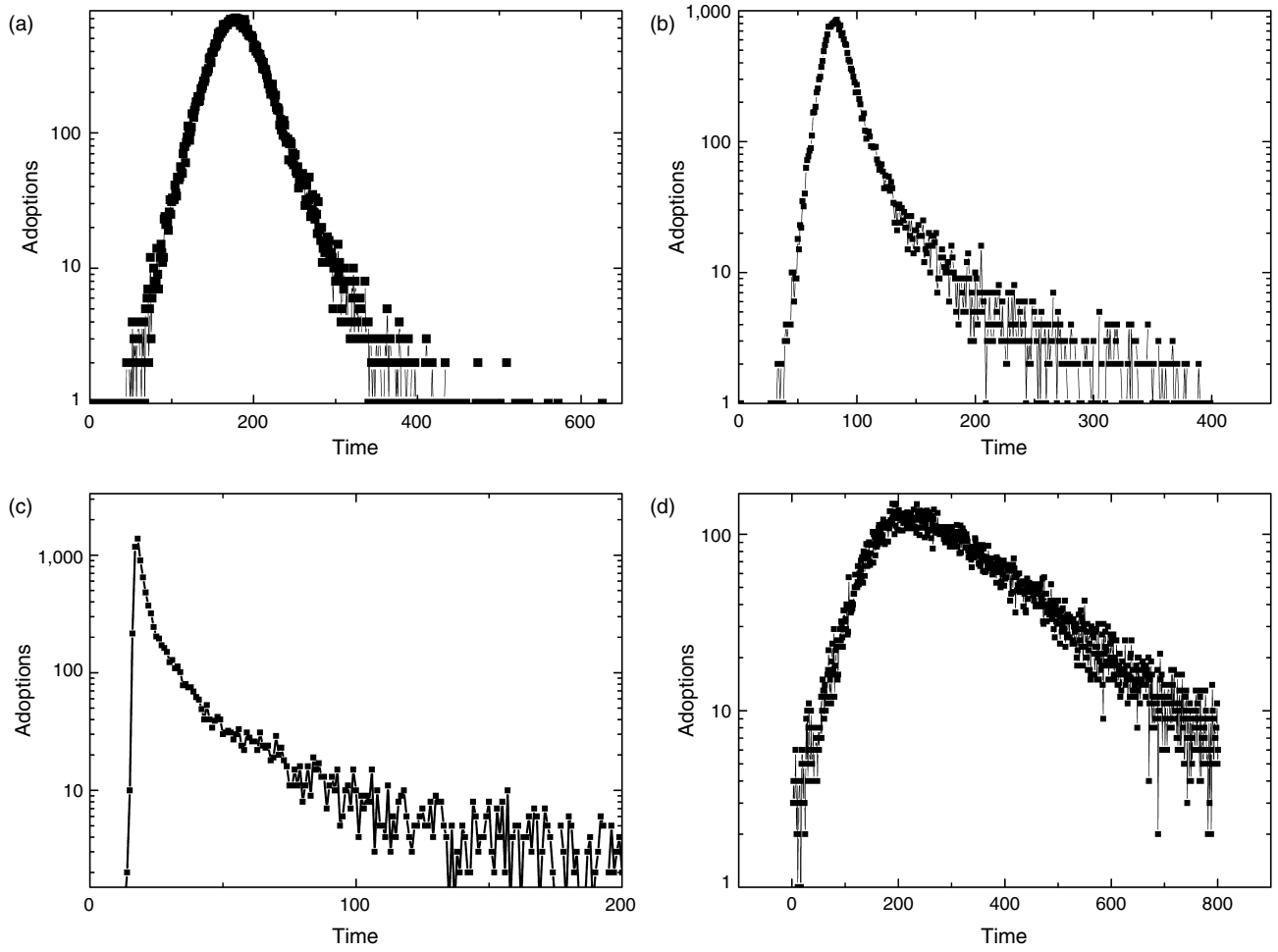
Here, k_{avg} and σ^2 are the mean and the variance of network degrees, respectively, and $k_{\min}$ is the smallest degree in the network. In the case of degree distributions with moderate tails (e.g., Gaussian, Poissonian), the smallest and greatest network degrees are of similar orders of magnitude, resulting in $k_{\text{avg}} + \sigma^2/k_{\text{avg}} = O(k_{\text{avg}}) = O(k_{\min})$ (here, $O(\cdot)$ denotes an order of magnitude).

Considering the case in which external influence p is small, the absolute values of the growth and decline slopes tend to be very similar (e.g., for the Gaussian/Poissonian case) and generate a relatively *symmetric* curve of adoption. In contrast, degree distributions that are highly skewed to the right (e.g., scale-free distributions) and span several orders of magnitude (where $k_{\text{avg}} + \sigma^2/k_{\text{avg}} \gg k_{\min}$) create *asymmetric* adoption curves, in which the slope of growth is much steeper while the slope of the decline exhibits a long-lasting temporal tail. To illustrate the asymmetrical functional form of the penetration curve, we simulated adoption on a 100,000-node network for different classes of networks. Adoption processes under each of four common classes of network degree distributions—Gaussian, uniform, lognormal, and scale-free—are given in Figures 1(a)–1(d), respectively, (for empirical examples of these networks distributions, see Amaral et al. 2000, Barabási and Albert 1999, Liben-Nowell and Kleinberg 2008, Limpert et al. 2001, Newman 2005, Newman et al. 2002, Yeung 2005).

As can be seen in Figure 1, each network type imposes a typical imprint on the adoption curve, which is easily identifiable on a semilog scale (and less easily identifiable on the more commonly used linear scale). In cases where the network degree distribution has negligible skewness, such as in Gaussian or Poissonian distributions, only a minor temporal tail of the adoption rate time series is created by the "smallest-degree" nodes, which are the last to participate in the process. The result is a largely symmetric pattern with an exponentially decaying tail in the final stages of the dissemination process (see Figure 1(a)). In the case of uniform distribution, the tail is more accentuated and longer lasting in relation to the curve itself (see Figure 1(b)), owing to the relatively larger population of smallest-degree nodes.

The lognormal degree distribution has a longer tail, emanating from the peak itself, with several decaying exponentials (because of the large population of smallest-degree nodes) and an accelerated growth rate because of the degree distribution's heavy skewness to the right, representing highly connected nodes (see Figure 1(c)). Finally, the scale-free case is dominated by the smallest-degree nodes. Its tail is the longest and demonstrates a single exponential decay following sharp rapid growth (see Figure 1(d)).

Figure 1 Adoption Patterns for Different Network Types: (a) Gaussian, (b) Uniform, (c) Lognormal, and (d) Scale-Free



In the following section, we use the effect of degree heterogeneity on the adoption curve to demonstrate how the model can be used to extract the network's degree distribution from a given adoption curve.

5. A Procedure for Estimating Network Degree Distribution Using Penetration Data

The procedure we use to estimate a network's degree distribution is based on the simplified assumption that diffusion occurs on an approximately random network. To model the adoption process, we use a parsimonious agent-based model (e.g., Garber et al. 2004). In §7, we analyze the robustness of our model by relaxing our assumptions of an unclustered network and of linear effects of adopter-neighbors on potential adopters' probability of adoption.[8] We consider four classes of networks: Gaussian/Poissonian, uniform, lognormal, and scale-free.[9]

Each network degree distribution comprises two parameters. Thus, the estimation involves a total of four parameters: two network parameters and the external and internal forces (p and q). Because the analytical solution is not in closed form (see Appendix §A.1), we use simulations with constraints for the purpose of fitting. We iterate over simulation

[8] We find that the proposed method works for low and moderate levels of network clustering but not for high-level clustering, which is the less frequent case. We also find that the linear assumption of the hazard model does not affect the efficacy of the method as long as the average network degree is lower than the concavity factor of the process. See §7 for a detailed explanation.

[9] It should be noted that some networks cannot be perfectly associated with any "pure" category. In hybrid networks, degree distributions can span more than one category (for example, a Gaussian distribution for small degrees and a heavy scale-free tail for large degrees). We use here only four types of networks, but broadly speaking, one can use any set of networks, including any number of types in that set for specific purposes. The method can identify the network of the set that is closest to the degree distribution of the active network underlying the adoption process. In our data, we do not find any evidence of the existence of a distinct hybrid distribution.

runs fitted to empirical data, using maximum goodness-of-fit measures as criteria for reaching a solution. Therefore, the estimation process entails two stages: (1) estimation of parameter constraints and (2) fitting simulations to empirical adoption curves (using the constraints).

5.1. Stage 1: Estimating the Constraint

In this stage, we estimate three indicators from the adoption curve: the initial pre-takeoff adoption rate (see below), the growth slope (16), and the decline slope (17).[10] We use these expressions to reduce the four-dimensional (four parameters) solution space to a two-dimensional space to facilitate our search.

5.1.1. Growth and Decline Slopes. The adoption process modeled here is subject to the two constraints, expressed by slopes x_1 and x_2, as indicated by (16) and (17). We then separate the external influence from the network degree distribution by retrieving empirical measures of the adjusted growth and decline slopes, and we revise (16) and (17) to define the following two-equation system and solve for the three unknown parameters π_1, π_2, and q:

$$z_1 = q\left(k_{\text{avg}} + \frac{\sigma^2}{k_{\text{avg}}} - 2\right) \tag{18}$$

and

$$z_2 = qk_{\min}. \tag{19}$$

Here, $z_1 = x_1 + 3p$ and $z_2 = -x_2 + p$ are the adjusted slopes of growth and decline, respectively (satisfying $z_1 > z_2 > 0$), and $k_{\text{avg}} = k_{\text{avg}}(\pi_1, \pi_2)$, $\sigma^2 = \sigma^2(\pi_1, \pi_2)$, and $k_{\min} = k_{\min}(\pi_1, \pi_2)$ are functions of the two network degree distribution parameters. In Table B.1 in the online appendix (at http://dx.doi.org/10.1287/mksc.1120.0711), we provide the explicit form of these functions for each network class. Equations (18) and (19) enable us to express two unknown parameters as a function of the third unknown parameter for all network classes (with the exception of the scale-free network class, where only one unknown parameter can be analytically solved using the remaining two unknown parameters).

5.1.2. Pre-Takeoff Period. We can estimate the external influence p. In the pre-takeoff stage,[11] a very early stage of the process, the rate of adoption is approximately $dN(t)/dt \sim (dN/dt)|_{t=0} = Mp$. The external force process is dominant in this stage, and therefore p can be extracted if M is known.[12] If there is no pre-takeoff stage, i.e., early sales are already high and the peak is reached quickly, one can measure $(dN/dt)|_{t=0}$, as it is also approximately equal to $M \cdot p$ in this initial stage; it can be assumed that the first single unit of time contains a negligible number of social interactions.

5.2. Stage 2: Fitting Simulations to Data

After estimating the constraints ((18) and (19)), we use a simulation to fit for the final parameter. We use an agent-based model to simulate the process of a new product's penetration over a given network. We then use the adoption pattern generated by the simulation as a fitting function for the data. Using equations that describe sections of the pattern (see (16) and (17)), we narrow the fit to a fit of one or two parameters and thus significantly improve convergence accuracy. This estimation procedure is conducted for each network class. The estimated active network is defined as the parameter set that produces the highest goodness-of-fit score (in terms of R^2 measures) of the simulated network's noncumulative adoption curve to the empirical adoption data.

In the following sections, we present the results of four studies that evaluate the accuracy and validity of the proposed approach and its contribution to forecasting accuracy.

6. Model Evaluation

6.1. Recovering Network Parameters

6.1.1. Simulations Study. In this study, we test the proposed model on a large range of network degree distributions. Sensitivity of methods is commonly tested on simulated data. Simulations also provide the ability to carry out controlled tests on wider ranges of parameters than are typically available in field data. For this purpose, we conducted computer simulations to generate diffusion scenarios on various network degree distributions. Each scenario produced an adoption curve that was subsequently used as input for testing the estimation method. We evaluated the accuracy of the method by comparing the estimated network degree distribution with the distribution generated in the simulation.

In total, we tested our model on 40 growth processes (i.e., 10 parameter sets for each of the four degree-distribution classes). To represent diverse dynamic scenarios, we tested sets over a wide range of values of q (between 0.0005 and 0.5) and p (between 0 and q). For each generated growth process, 10 fitting procedures were conducted. The fitted adoption curve parameters with the best resulting R^2 values were considered to be the estimated solution parameter set. (We also tried a greater number of fits per generated network and found that 10 fits are sufficient to

[10] See Appendix §A.3 for the practical operation of identifying these stages of the diffusion process.

[11] For definitions of takeoff, see Appendix §A.3.

[12] In the case of forecasting, M is not known, and we have five parameters. This case is discussed in §6.3.

Table 1 Goodness-of-Fit Indicators for the Reconstruction Method Over Simulated Data: Average and Standard Deviation of R^2 and RMSS

Network class	$(R^2)_{avg}$	$(R^2)_{sd}$	$RMSS_{avg}$	$RMSS_{sd}$
Gaussian	0.95	0.03	0.94	0.02
Uniform	0.93	0.02	0.96	0.01
Lognormal	0.92	0.04	0.89	0.03
Scale-free	0.97	0.02	0.98	0.02

Note. sd, standard deviation.

obtain the greatest degree of accuracy.) For each distribution type, the goodness of fit between the estimated and the real distribution, measured according to the R^2 indicator and the normalized root mean square statistic (RMSS; following Jedidi et al. 1997), is given in Table 1.

We find that the average fit of the actual distribution of the simulation-generated surrogate set to the distribution estimated using the proposed method is greater than 90% in all cases and with small deviations (see Table 1). In this case, we also find that the estimation accuracy of the diffusion and network parameters in the proposed model ranges from 6% to 14%[13] (with a standard deviation ranging from 6% to 17%). This finding suggests that it is possible to extract the network type and estimate its parameters for a large range of variables, given a single adoption pattern. In all tested cases, we demonstrate that the proposed model provides estimations that are very close to the actual degree distribution.

In such a simulation-based approach, the model under investigation is tested using other models. Real-life phenomena include a richer set of mechanisms and noise. To address this, we present the results of the studies in the following sections.

6.1.2. Empirical Study. In this study we aimed to achieve two important objectives. First, we applied a more stringent test to the proposed model to demonstrate that the estimated degree distributions are close to actual degree distributions that are known from an external source. Second, we attempted to elucidate whether this procedure leads to a unique solution or whether different degree distribution parameters or even different network categories can generate identical adoption curves.

We used data from Friendster (http://www.friendster.com), an online social network with more than 115 million users (as of 2008). The advantage of this online social network as a data source is that it provides documentation of the time at which each user joins a group, and network data are largely in the public domain. Using only the social group membership cumulative adoption data, we estimated the network degree distribution for each social group. In Table 2, we list the results of the identification and estimation of the active networks for eight diffusion data sets in which the active network is known to us. For each known active network (corresponding to users' group memberships), we calculated fit to each of the four representative classes of degree distribution (Gaussian/Poissonian, uniform, lognormal, and scale-free). We calculated cross entropy (Kullback 1997, Garber et al. 2004), the more intuitive R^2 measure *between distributions*, and RMSS (e.g., Jedidi et al. 1997).

Table 2 shows, for each network, the lowest cross entropy, the highest R^2 for the fit between distributions, and the highest RMSS (i.e., the closest distributions based on these fit statistics). For the first network (denoted A), the best fit was obtained for a

Table 2 Measures of Goodness of Fit Between Estimated and Directly Mapped Network Degree Distributions: R^2, Cross Entropy, and RMSS

	Scale-free	Gaussian/ Poissonian	Lognormal	Uniform
Network A (scale-free)				
Cross entropy	0.13	1.20	0.92	1.13
Distributions R^2	0.98	0.30	0.69	0.75
RMSS	0.99	0.34	0.70	0.74
Network B (Poissonian)				
Cross entropy	1.44	1.40	1.46	1.51
Distributions R^2	0.42	0.95	0.36	0.18
RMSS	0.45	0.96	0.40	0.20
Network C (scale-free)				
Cross entropy	0.22	1.42	1.25	1.36
Distributions R^2	0.78	0.18	0.51	0.45
RMSS	0.79	0.19	0.55	0.46
Network D (Poissonian)				
Cross entropy	1.36	1.02	1.47	1.40
Distributions R^2	0.54	0.81	0.50	0.74
RMSS	0.56	0.85	0.50	0.76
Network E (scale-free)				
Cross entropy	0.40	1.67	1.57	1.11
Distributions R^2	0.75	0.29	0.67	0
RMSS	0.76	0.31	0.72	0
Network F (scale-free)				
Cross entropy	0.69	1.99	1.70	2.39
Distributions R^2	0.75	0.34	0.38	0.25
RMSS	0.77	0.37	0.32	0.28
Network G (Poissonian)				
Cross entropy	1.45	1.11	1.67	1.78
Distributions R^2	0.58	0.95	0.24	0.33
RMSS	0.62	0.96	0.32	0.35
Network H (scale-free)				
Cross entropy	0.13	1.04	0.99	1.23
Distributions R^2	0.96	0.43	0.56	0.70
RMSS	0.97	0.45	0.58	0.72

[13] Estimation accuracy is defined as the percentage deviation of the estimated parameter value from the known parameter value.

scale-free degree distribution. The first two moments of the reconstructed network were[14] $k_{\min} = 13.00 \pm 2.00$ and $\alpha = 2.22 \pm 0.15$, compared with the actual values of the active network,[15] $k_{\text{trans}} = 12.21 \pm 2.00$ and $\alpha = 2.59 \pm 0.23$. The second network (B) fit best to a Poissonian degree distribution. The parameter of the reconstructed network was $k_{\text{avg}} = 2.00$ versus $k_{\text{avg}} = 2.00$ in the real network, i.e., an accurate match. The third network (C) parameters are $k_{\min} = 4.00 \pm 1.00$ and $\alpha = 2.25 \pm 0.18$, compared with the real active network values of $k_{\text{trans}} = 5.10 \pm 2.00$ and $\alpha = 2.66 \pm 0.40$. The fourth network (D), identified as a network with a Poisson degree distribution, was shown by the proposed model to have $k_{\text{avg}} = 5.00$ versus the empirical $k_{\text{avg}} = 7.00$. The fitted and empirical parameters for the remaining networks and more details are given in Table A.2 of the appendix. Finally, the R^2 and RMSS values for the distributions show good fits for networks A, B, G, and H (R^2 values of 0.98, 0.95, 0.95, and 0.96, and RMSS values of 0.99, 0.96, 0.96, and 0.97, respectively).

For networks C–F, high R^2 values (0.75–0.81) and high RMSS values (0.76–0.85) were obtained, with cross-entropy measures superior to those of the alternative tested networks.

Overall, the model correctly identified the underlying network degree distributions in all cases,[16] and the error in the estimation of the actual parameter values ranged from 5% to 15% in all cases but one (in which the error was 28%).

6.2. Fitting Adoption Curves

To address situations that are more realistic than those described in §6.1.1 and that involve various types of noise and other error factors, we tested real-life adoption patterns. We sought to include only data sets that met the following criteria:

1. The growth process has an identifiable (dominant) peak, with relatively smaller fluctuations.

2. The time series has sufficient resolution to allow differentiation between patterns: the pattern comprises at least 50 points of data.

3. The process involves more than several thousands of adopters, to allow for a distinct dissemination pattern.

We collected a total of 17 data sets representing adoption processes in diverse fields, including CD sales, online movie penetration (based on search query volume), petition signing rates, and enrollment in online thematic groups. Despite the inherent noise and potentially high degree of interference caused by external events in these real-life cases, the dissemination curves reconstructed using the proposed model fit the real-life data with a relatively high degree of goodness of fit[17] (R^2 in the range of 90%–98%).

Figure 2 presents data on cases grouped according to the underlying network degree distribution. In Figure 2(a), we present sales of an audio CD (Moe and Fader 2001) (Gaussian pattern; $R^2 = 0.97$). Figure 2(b) presents the daily number of queries for the term "Cloverfield," referring to a movie that was exceptionally credited for its efficient online viral marketing campaign (a heavily skewed tail distribution, lognormal; $R^2 = 0.98$). Because the campaign was conducted mainly on Internet websites, including YouTube, the lognormal pattern we observed is not surprising (Limpert et al. 2001). Figure 2(c) represents the search volume for the movie *The Kite Runner* in a similar manner. In contrast to the previous case, this penetration process exhibits a low skew distribution (Gaussian pattern, $R^2 = 0.94$). Finally, membership in a thematic social group on Friendster, comprising 74,500 users, is depicted in Figure 2(d) ($R^2 = 0.91$), and the distribution is identified as scale-free. In the latter case, the adoption process represents the timing and number of users that opted to join a specific thematic group (specifically, a group of fans of a certain category of TV shows) among the groups offered by Friendster. In addition, we performed the proposed estimation procedure on each of the four degree distributions (scale-free, Gaussian, lognormal, and uniform) as if it were the actual network. The accuracy obtained using this identification method is presented in Table 3 for fits of all 17 cases (see Table B.2 in the online appendix for details of each case).[18]

The columns in Table 3 present R^2 fit results for different network degree distribution types. In all cases,

[14] The estimations for the parameters of the scale-free case were calculated using numerical fits to the data, using a method described by Newman (2005). The error ranges were taken to be one standard deviation from the fitted value.

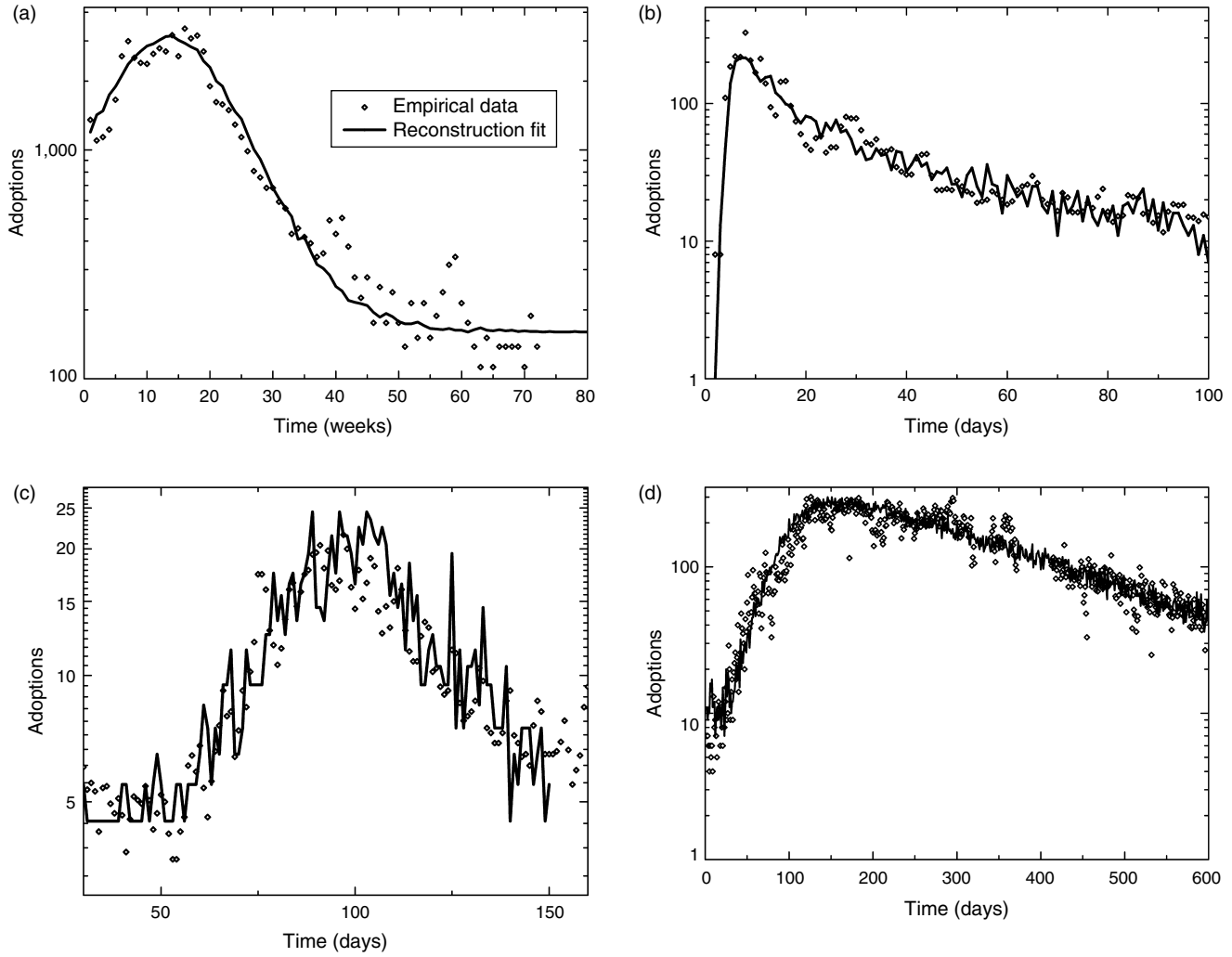
[15] Because empirical scale-free distributions do not have a clear equivalent to a cutoff minimum degree, we selected the maximum point on the left-hand side of the distribution (denoted as k_{trans}, the degree in which the transition from power law to zero takes place) as a comparison point for the model estimation of $k_{\min}$.

[16] The eight diffusion sets were randomly chosen, and they apparently contain no uniform or lognormal distributions. This may be a property of the specific community and activity selected. Because these two network types are reported to exist (and have been found in the studies presented in §6.2), we do not rule out the possibility that they may exist.

[17] The fits throughout this subsection were calculated between the simulated and the empirical noncumulative adoption curves.

[18] To compare the network model to existing diffusion models, we also fitted the same set of empirical adoption curves to three other existing diffusion models (in fact, these are the same models we use in §6.3 in the forecasting study): (1) the Bass model (Bass 1969), (2) the gamma/shifted Gompertz model (Bemmaor 1994, Bemmaor and Lee 2002), and (3) the nonuniform influence model (Easingwood et al. 1983). We conclude from the results of the fittings in Table B.5 in the online appendix that the network model outperforms the other three diffusion models for this data set.

Figure 2 Fit to Empirical Data: (a) CD Sales, (b) the Movie *Cloverfield*, (c) the Movie *The Kite Runner*, and (d) an Online Thematic Social Group



the R^2 fit with the identified degree distribution is 90% or greater and is distinctly different from the data's fit to any alternative degree distribution. Perhaps surprisingly, almost one-half (45%) of all cases represent non-scale-free active networks, suggesting that the common assumption, according to which diffusion occurs over a scale-free network, may be overused. The existence of a nonnegligible population of consumers with above-average degrees, as is the case in heavily skewed tail distribution, accelerates adoption by generating a rather rapid exposure of large portions of the market to the product. The right-hand column in Table 3 presents the asymmetry ratio, defined as the ratio of the post-peak percentage of adopters to the pre-peak percentage of adopters (e.g., for the symmetric case, this ratio is 1).

It can be seen that in all cases, the asymmetry ratio is greater than 1, and among scale-free networks, the ratio is greater than 3 in the majority of cases. This implies that in scale-free networks, only 25% of the total market typically adopts before the sales peak is achieved. Therefore, firms that concentrate marketing efforts on generating pre-peak sales are in effect ignoring 75% of their potential market. Furthermore, estimation of the market potential can be biased if the network degree distribution is not taken into account (for an estimation of this bias, see Appendix §A.2). We show in §6.3 that this has significant implications for penetration forecasting accuracy.

6.3. Forecasting

The purpose of this study is to assess the value of the information concerning the underlying network. Although there might be different applications for this knowledge (such as improved marketing decisions, strategizing in buzz programs, etc.), we demonstrate managerial value by testing forecasting accuracy. The problem of forecasting penetration early in the process is well known (Chandrasekaran and Tellis 2007, Van den Bulte and Lilien 1997). From a

Table 3 Empirical Goodness-of-Fit Results: R^2 Values and Asymmetry Ratios for All Data Sets and Networks

	Scale-free	Lognormal	Uniform	Gaussian	Asymmetry ratio
Identified scale-free networks					
Friendster online group ("Korean Drama Fanz")	0.91	0.68	0.70	0.45	3.93
Friendster online group ("Emo Is Love")	0.94	0.77	0.81	0.53	2.36
Friendster online group ("ABSCBN")	0.91	0.70	0.45	0.59	3.37
Friendster online group ("Muse")	0.95	0.72	0.88	0.68	5.25
Friendster online group ("LinkinPark")	0.95	0.83	0.72	0.81	1.93
Friendster online group ("Handsome and Pretty")	0.95	0.83	0.54	0.31	1.78
Friendster online group ("Registered Nurse")	0.90	0.25	0.40	0.81	3.42
Friendster online group ("Islam")	0.93	0.75	0.64	0.22	3.41
Friendster online group ("Dota")	0.94	0.89	0.78	0.64	2.33
Identified uniform networks					
Petition signers ("Free our friends in Iraq...")	0.90	0.86	0.95	0.52	2.36
Petition signers ("12 Year Old Jobs")	0.81	0.86	0.95	0.57	7.51
Petition signers ("Save Wonderfalls")	0.81	0.63	0.91	0.78	2.77
Identified Gaussian networks					
Search volume for "Kite Runner" (after launch)	0.68	0.61	0.79	0.94	1.39
Friendster online group ("American Idol")	0.66	0.78	0.86	0.90	1.02
CD sales ("DINK")	0.81	0.87	0.63	0.97	1.33
Friendster online group ("Starbuckerz")	0.65	0.63	0.73	0.90	1.07
Identified lognormal networks					
Search volume for "Cloverfield" (after launch)	0.82	0.98	0.71	0.4	4.01

managerial perspective, it is difficult to obtain sufficient market information at an early enough stage in the diffusion process to produce useful forecasts. Furthermore, if the market potential M is not provided by an external source, forecasting may be plagued with large-scale errors. Recently, it was shown that even a simple truncation of diffusion data can lead to overestimation of parameters (Van den Bulte and Iyengar 2011).

To test the value of a unified approach to diffusion, we use the proposed model to predict diffusion parameters at different stages of the diffusion process and compared predictions to results of three well-known models: (1) the Bass model (Bass 1969), (2) the gamma/shifted Gompertz model (Bemmaor 1994, Bemmaor and Lee 2002), and (3) the nonuniform influence model (Easingwood et al. 1983). The forecasting method is similar to the reconstruction procedure we used previously (§§6.1 and 6.2; see also the online appendix), but in this case, the network degree distribution recovered at an early stage in the process is used to forecast the entire adoption curve and to generate an estimation of total market potential.

For this study, we used all 20 adoption cases (all data sets used in prior studies). For each case, we estimated the diffusion parameters as well as the first two moments of the network distribution, based solely on early penetration data (see below). We then generated an adoption forecast by estimating the remainder of the penetration curve.[19] Similarly, we used the same early data with three alternative models to predict penetration.

The prediction results of all four models are presented for four cases (first four cases in Table 3) in Figures 3 and 4 as an illustration (all 20 cases are reported in Tables B.3 and B.4 of the online appendix, and the mean performance is reported in Table 4). The market potential forecasting percentage error[20] is presented as a function of forecast timing (presented as penetration percentage) in Figure 3.

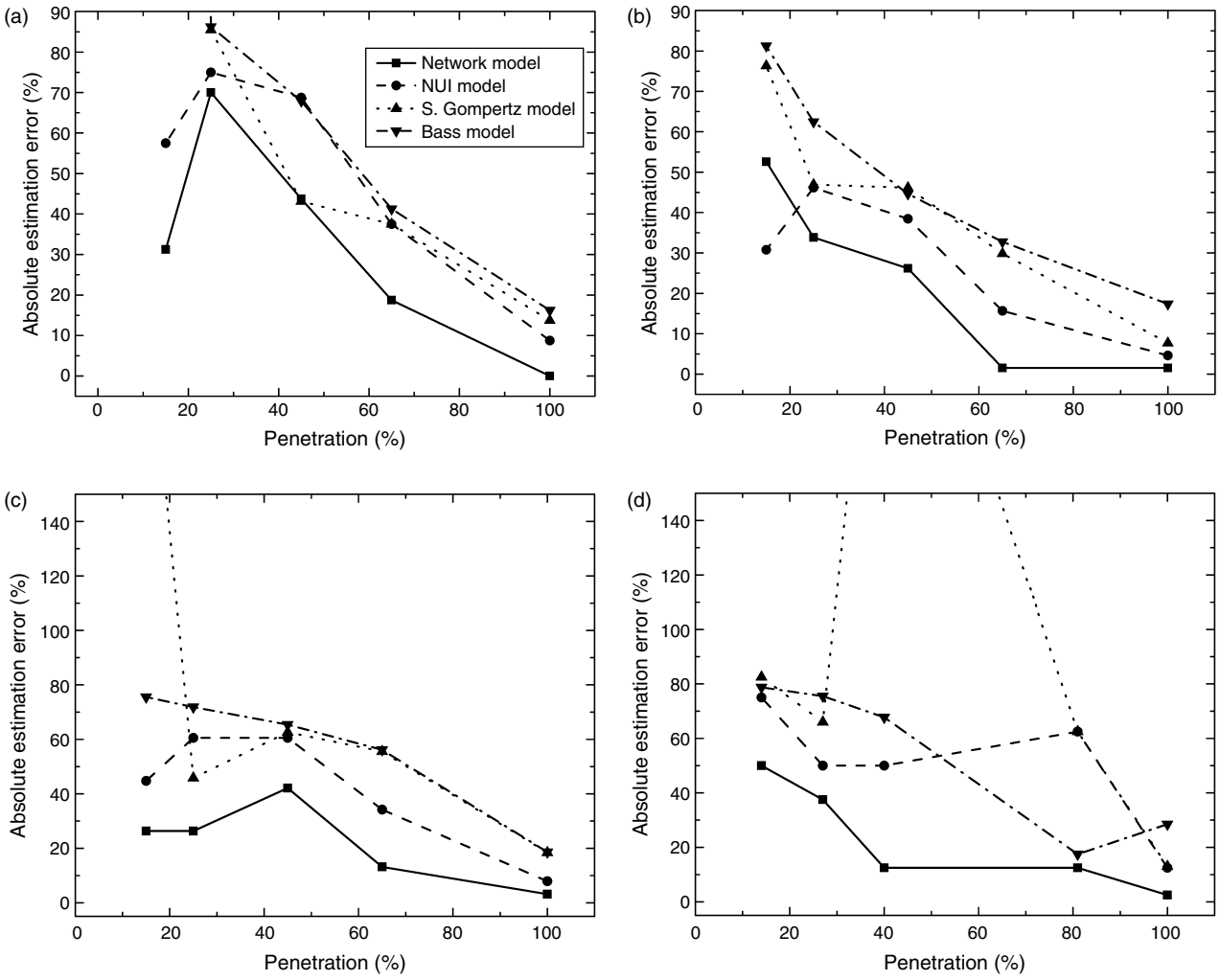
Inclusion of the network degree distribution in the forecasting model significantly improves market potential forecasting accuracy and enables predictions to be made earlier in the penetration process, compared with the three remaining models, for most cases and most time points.

For example, using only data from the 15% penetration mark, the network-based model improves

[19] To simplify the process, we exclusively used scale-free degree-distribution networks for the fitting procedure. We find that using only scale-free networks produces good forecasting estimations. A scale-free distribution, for high-power law exponents, can also approximate a Gaussian distribution, which proves to be useful when trying to forecast diffusion curves in general. In a separate study, not reported here for the sake of brevity, we found that using the scale-free distribution with our data sets, in forecasting, achieves better results than when using Gaussian distributions.

[20] Because the forecasting estimation was performed using early-stage data for diffusion curves of processes that have reached a certain market potential, the error of the early-stage estimations was taken to be the absolute percentage of the deviation from that market potential (based on knowledge of the subsequent stages of these processes).

Figure 3 Comparison of Estimated Market Potential Forecasting Errors of the Network-Based Model and Three Benchmark Models for Online Groups: (a) Korean Drama Fanz, (b) Emo Is Love, (c) ABSCBN, and (d) Muse



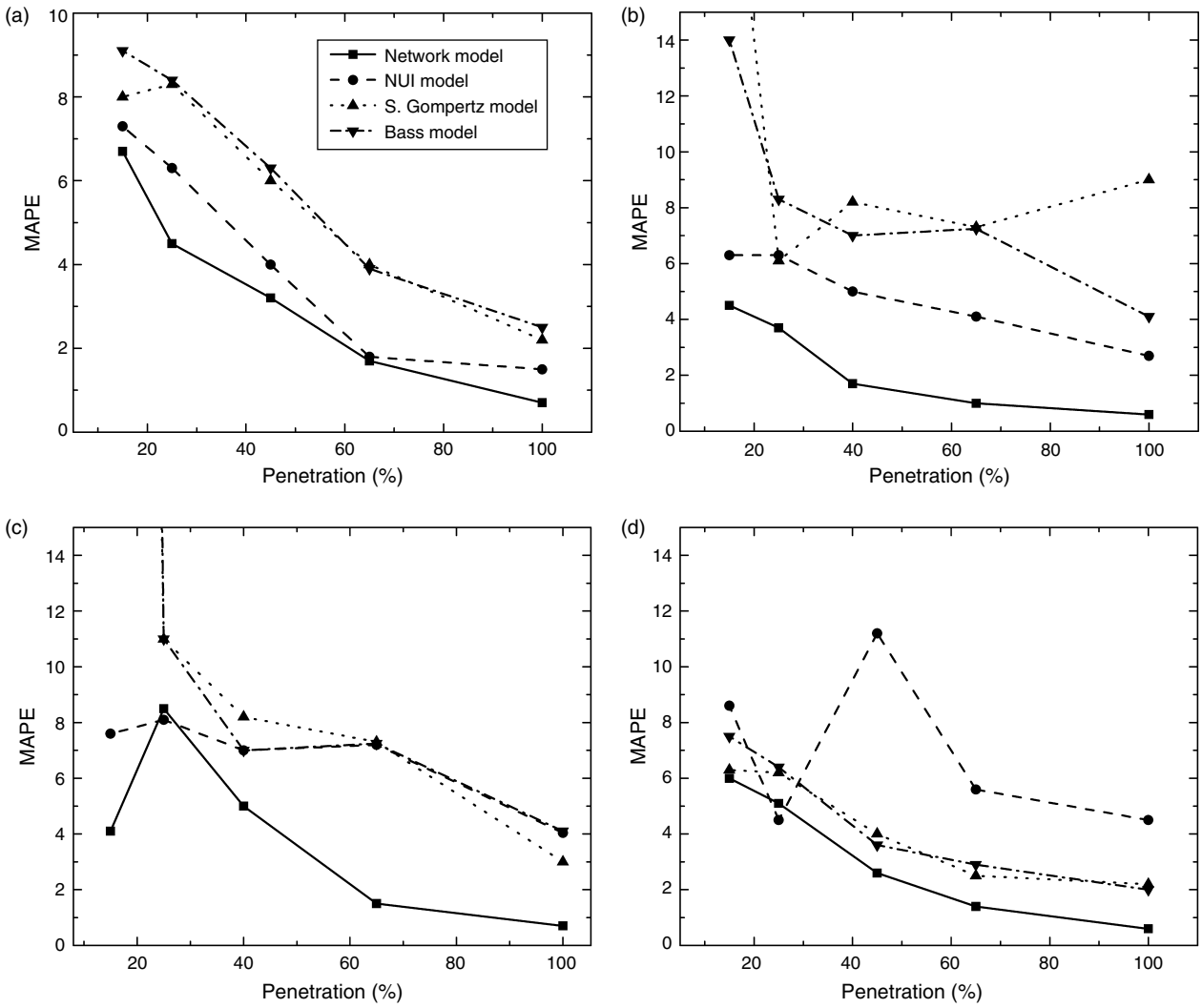
Note. NUI, nonuniform influence; S. Gompertz, gamma/shifted Gompertz.

forecasting accuracy by at least 25% over other models in most cases (in two of the cases, accuracy improves by approximately 80%). The consistently better results of the network model suggest that it carries valuable added information.

At the early 15% penetration mark, the network model provides the lowest mean absolute percentage errors (MAPEs)[21] among all four model forecasts (see Figure 4), significantly improving our ability to forecast the full penetration curve. In three of the four cases, at 15% penetration, the network model generates a MAPE that is substantially lower than that of the next best method (which varies from case to case). Using data from the 25% penetration mark, the network model is almost converged on the forecasted curve with mean squared error levels that are close to the final fit, based on the entire curve, in all four cases. Again, with the exception of isolated points, the network model improves the MAPE for forecasted adoption curves. In Table 4, we report the average percentage of forecasting error and average MAPE for all 20 cases (for detailed results, see Tables B.3 and B.4 in the online appendix). On average, the error percentage of the estimated market potential generated by the network model is 15% smaller in magnitude than that of the second-best model—the nonuniform influence model—and 57%–65% smaller than errors produced by the other models. The average MAPE of the network model is more than 70% lower than that of the nonuniform influence model and more than 85% lower than that of the other models. We conclude that the network model generated more accurate predictions of the market potential and penetration curve shape than did the other models examined in the study.

[21] The MAPE was calculated between the forecasted noncumulative adoption curve and the actual noncumulative adoption curve.

Figure 4 **Comparison of MAPE of Estimated and Real Future Penetration Curves Using the Network-Based Model and Three Benchmark Models for Online Groups: (a) Korean Drama Fanz, (b) Emo Is Love, (c) ABSCBN, and (d) Muse**



Note. NUI, nonuniform influence; S. Gompertz, gamma/shifted Gompertz.

Table 4 **Comparison of Average Performance of Forecasting Models**

Network class	Network model	Bass model	Gamma/shifted Gompert model	Nonuniform influence model
Average % of error	34.4	99.5	91.9	50.0
Average MAPE	3.3	26.2	26.7	11.9

7. The Boundary Conditions of the Model

The proposed model is based on several simplifying assumptions that may limit its validity and applicability. Below, we discuss the model's sensitivity to these assumptions and outline the model's boundary conditions.

7.1. Continuous Process

One of the restrictive conditions of the model is that it is applicable only to continuous processes in which the diffusion parameters and market size are sufficiently large to produce a continuous penetration curve that meets the three criteria defined in §6.2. This model is not applicable to discontinuous and/or flat diffusion curves; contagion models, in general, are not meant to capture such dynamics. Flat penetration curves (without a dominant peak) are thought to represent market failures, in which social contagion plays a marginal role at best. We performed several checks to investigate the boundaries of the model. On simulated data, for small values of q and p, the diffusion curves are discontinuous, and the model performs poorly. For example, for a network in which the average degree is 10, daily coefficients of q that are smaller than $5e-5$ and of p that are around $1e-7$ lead to a flat diffusion curve. Our tests confirm that model performance diminishes in such cases. We observe a similar reduction in performance in the three benchmark diffusion models used. We conclude that the

model does not perform well in cases of market failures, discontinuous data, lack of contagion effects, or highly dominant external force.

7.2. Clustering

Another limitation of the model is its disregard of network clustering. Many real-life social networks exhibit some level of clustering and are therefore not always as random as typically assumed. The results presented in §6.3 suggest that the model's predictive performance is not diminished by a real-life clustering effect, but an evaluation of the drawbacks of using the random network approximation in the case of clustered networks is warranted. The local clustering coefficient of a given node is defined as the ratio of the number of actual links to the number of potential links in a node's neighborhood (Watts and Strogatz 1998). The clustering coefficient of an entire network (the network average clustering coefficient) is the average of all local clustering coefficients.

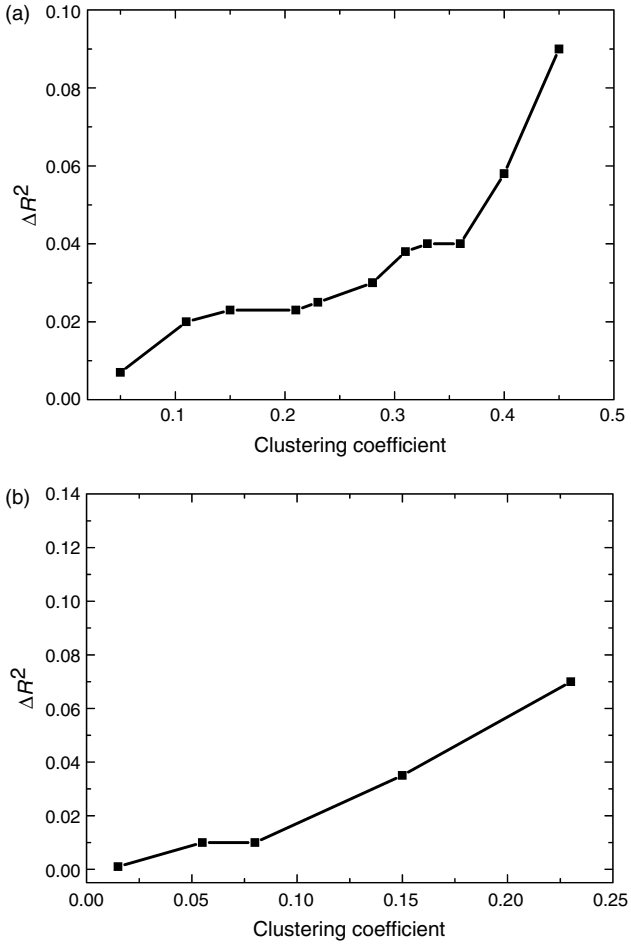
Degree distribution and network clustering are two strongly interrelated network qualities, as confirmed in a study by Volz (2004). For a given degree distribution, the magnitude of global network clustering coefficient has an upper boundary. It was also shown that clustering coefficients greater than approximately 0.3 effectively induce fragmented networks (Volz 2004). Diffusion over a fragmented network produces curves that may have several peaks and a high degree of sales volatility (as we also see in our simulations). As a result, fragmented networks do not meet the criteria specified in §6.2. For a sensitivity analysis of the clustering effect, given a specific degree distribution, we generated clustering in the networks while maintaining each network's specific degree distribution. This was done while simultaneously maintaining the requirement that the network is nonfragmented, meaning that it is not composed of disconnected subnetworks. The greatest clustering coefficients obtained for scale-free and Poissonian networks were 0.45[22] and 0.23, respectively (which is consistent with Mislove et al. 2007). This is also consistent with reported average values of social network clustering coefficients of 0.26, with a standard deviation of 0.17 (Mislove et al. 2007, Newman et al. 2002). Furthermore, for all eight networks used in §6.1.2, clustering coefficients were smaller than 0.17. Therefore, we limited our sensitivity analysis to networks with clustering coefficients no greater than 0.45, a point at which fragmentation is already considerable.

For each clustering coefficient and for each adoption curve predicted by the proposed model, we evaluated the difference between the R^2 obtained with the clustering coefficient and that obtained with no clustering coefficient. The results for the scale-free case are presented in Figure 5(a). The differences range from several percentage points (∼2.5% for the most common clustering coefficient) to 9%.

In the case of a Poissonian degree distribution network, the deviation from the no-clustering case becomes marked at a clustering coefficient of 0.23 (see Figure 5(b)). We confirmed that network fragmentation is also high at these clustering coefficient levels (0.23 and above).

We were unable to generate networks with greater clustering coefficients that conformed to the degree-distribution and nonfragmentation requirements. We conclude that the effect of network clustering on the explanatory power of the model under the criteria in §6.2 is within 1% to 4%. Comparing this with

Figure 5 The Effect of Network Clustering on Goodness of Fit to Adoption Curves: (a) Scale-Free Case and (b) Poissonian Case



[22] We find in our simulations that the limit on the average clustering coefficient regarding fragmentation (0.3; see Volz 2004) changes in some cases, depending on network characteristics. Therefore we used the maximum average clustering coefficients attainable in the simulations, which was as high as 0.45 for the scale-free case and 0.23 for the Poissonian case, to cover all scenarios.

the results in Tables 2 and 3, an error term of this magnitude, which is small in relation to the magnitudes of goodness of fit exhibited in these tables, is not expected to substantially affect the results for nonfragmented networks with low to moderate levels of clustering.[23] We can conclude that moderate levels of clustering, which are most common, do not fragment the network and therefore have little impact on the diffusion process. High clustering coefficients attest to fragmented networks that do not meet the model's first applicability criterion (see §6.2).

7.3. Concavity

The influence of a consumer's neighbors on the consumer's decision to adopt in this model is linearly proportional to the number of neighbors that are already adopters (see §3). A more realistic assumption may be a concave adoption function in which the probability of adoption diminishes with each additional adopter neighbor. To illustrate the effect of concavity on adoption, we use a log-concave model. The probability of any potential adopter to adopt within a short time interval Δt, given that y of his neighbors have adopted, is $1-(1-p\Delta t)(1-q\Delta t)^y \approx (p+yq)\Delta t$. To model concavity, we substitute the linear y term with the logarithm term:

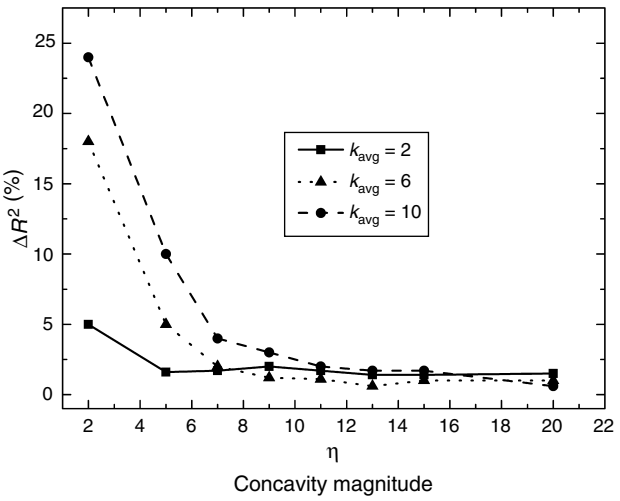
$$P(y) = \left(p + q\log\left(1+\frac{y}{\eta}\right)\right)\Delta t \approx 1-(1-p\Delta t)(1-q\Delta t)^{\log(1+y/\eta)}. \quad (20)$$

Here, η is the concavity parameter. Low values of η imply high-magnitude concavity, and high values of η imply low-magnitude concavity. When the concavity parameter is much greater than the number of an adopter's neighbors (i.e., $\eta \gg y$), the log-concave model converges to the linear model (recall that $\log(1+y/\eta) \approx y/\eta$ when $y/\eta \ll 1$). Therefore, the importance of the overall concavity effect depends on the ratio between η and the average degree (average y). We ran simulations to calculate the effect the concavity assumption has on the results of the model by measuring R^2 differences between the concave and the nonconcave models:

$$P(y) = 1-(1-p\Delta t)(1-q\Delta t)^{\log(1+y/\eta)}. \quad (21)$$

We analyzed the effect of the concavity parameter η versus the average degree of the social network (see Figure 6). For $k_{\text{avg}} \le \eta$, adding the concavity assumption to the model has virtually no effect on the predicted adoption pattern. For $k_{\text{avg}} > \eta$, our original model does not account for the nonnegligible concavity (the adoption pattern also becomes a function of η). We conclude that concavity should be included in the modeling scheme only in cases in which the average degree equals or is greater than concavity (η). For example, $\eta = 20$ implies that the effect of each adopter-neighbor on a potential adopter diminishes when there are more than 20 such adopter-neighbors. In this case, if the average network degree is estimated to be smaller than 20, our model is valid. Although it is beyond the scope of this paper to conduct an empirical investigation of η, theoretically, we assume that $k_{\text{avg}} \le \eta$ in all the studies we performed.

Figure 6 The Effect of Concavity on Goodness of Fit to Adoption Curves



7.4. Declining Propensity to Generate Word of Mouth

Word-of-mouth decay, i.e., a consumer's diminishing propensity to generate word of mouth over time, is a plausible assumption and an emergent subject in research literature (e.g., Berger and Schwartz 2011). We conducted a study to assess the effect of word-of-mouth decay on the model's performance. The diffusion parameters in the proposed model are assumed to be an effective average and therefore constant over time. This assumption does not imply that the diffusion parameters are actually constant in time, but only that they can be represented by an average constant. Still, to check for time-dependency effects, we added time dependency of word-of-mouth generation to the model such that word of mouth decays exponentially over time until a minimal constant level is reached. This minimal level represents effects of the neighbors' long-term exposure to a product even if adopters do not produce word of mouth. We find that results are not affected as long as the minimal constant level of word of mouth is not negligible.

[23] We again note that our claims regarding the role of clustering are limited to cases that meet the criteria specified in §6.2.

7.5. Asymmetric Ties

Another limitation of the model is related to its assumption of symmetric ties (see Van den Bulte and Joshi 2007) and its disregard of asymmetrical ties (which implies that some consumers, regardless of their degree, are more influential than others). In some cases, the effect of asymmetrical ties may be important and may generate multiple peaks, which, in turn, may reduce the model's accuracy. This is an important issue that calls for future research, preferably using individual-level data, including information on tie strength.

7.6. General Identification

Heterogeneity in individual adoption can result from a variety of factors. In this paper, we explored the role of the social neighborhood of the individual—specifically, the individual's degree. Owing to limitations of the data, we were unable to control for other alternative explanations. Although it is possible that alternative models can fit the empirical adoption curves (see §6.2), it is less probable that more than one model will simultaneously fit the adoption curve, provide reasonable estimations of the underlying structure of the network (see §6.1.2), and provide forecasts superior to those of existing benchmark methods (see §6.3). We have also made an effort to include different types of data sets to demonstrate the explanatory power of the model across types of adoption processes.

8. Conclusion

We have shown that in networks whose degree distributions have heavily skewed tails, such as scale-free networks, the pattern and symmetry of the adoption curve tend to differ from those observed in other types of networks. Specifically, we have shown that adoption asymmetry can be substantial (see Table 3), and in most cases, the bulk of the market adopts after the peak. Ignoring such network-based features leads to estimation errors as well as biased forecasting. Correcting for this error is not simple, however, because in most cases, the network degree distribution of the active network effectively involved in the diffusion process is unknown. Detailed individual-level adoption data are available from Web-based network data bases, but the degree distribution of the active network is invisible and difficult to extract.[24] Data on off-line networks are even scarcer. As a result of the inaccessibility of these important network data, marketers have typically resorted to generalized assumptions about the networks underlying the adoption process, resulting in less-than-accurate predictions of market potential and diffusion. One such assumption is the assumption that the active networks always have a scale-free degree distribution. We have shown that in our data set, this assumption is incorrect. Furthermore, almost one-half of the networks we have investigated are, in fact, not scale-free.

We have shown here how it is possible to extract hitherto inaccessible network degree distributions from typically available aggregate-level adoption data, and we have demonstrated how to estimate the parameters of the active network's degree distribution. The proposed approach generates reasonable forecasts using early-stage aggregate-level penetration data of a single product.

Until now, predicting market potential was one of the most difficult challenges of diffusion modeling. One source of the inaccuracy of the predictions generated by current models is penetration curve asymmetry, potentially caused by degree distributions with heavily skewed tails. As shown in §6.3, an integrated model of diffusion and network connectivity that takes this asymmetry into account improves the accuracy of market potential forecasts. Furthermore, because the model links heterogeneity of an individual (consumer)-level trait (degree) to the temporal traits of aggregate consumption (adoption curve), we believe it is a potentially useful approach for future attempts to model aggregate patterns based on heterogeneous micro-behavior.

Finally, in addition to uncovering degree distribution, the model provides further information such as the number of influentials and their degree centralities. A growing body of literature connects consumers' value to the firm to their degree centralities (Goldenberg et al. 2009a, Libai et al. 2010, Stephen and Toubia 2010). Therefore, information on the distribution of consumers' centralities may enhance firms' ability to more accurately assess the potential value of a market, given only a single relevant penetration curve over that market.

Electronic Companion

An electronic companion to this paper is available as part of the online version at http://dx.doi.org/10.1287/mksc.1120.0711.

Acknowledgments

The authors thank Barak Libai, Yoram Louzon, Lev Muchnik, Oded Netzer, Arvind Rangaswamy, Sorin Solomon, and Olivier Toubia for their constructive comments and suggestions. This research was supported by the Israel Science Foundation, The Kmart International Center of Marketing and Retailing; The Davidson Center, Hebrew University of Jerusalem; The Horowitz Association; and the Center for Complexity Science.

Appendix

A.1. The Dynamics of Diffusion on a Generalized Random Network

In this appendix, we formally develop the diffusion dynamics on a generalized random network. The probabilities of a

[24] Extracting the active social influence network from visible individual-level data is not a trivial matter, and success depends on the measured indicators and the quality of the data.

potential adopter to be affected by external influence (e.g., marketing efforts) and word-of-mouth communications generated independently by each one of her actual adopter neighbors at each time step Δt are $p\Delta t$ and $q\Delta t$, respectively. Hence, the expected value of the number of adopters within a short time interval Δt after time t is $\Delta N(t) = \sum_x H_x(t) \cdot (p + xq)\Delta t$. Under the continuous limit,

$$\frac{dN(t)}{dt} = \sum_x H_x(t) \cdot (p + xq). \tag{22}$$

Here, $H_x(t)$ denotes the number of potential adopters of order x, i.e., consumers who have not yet adopted but have exactly x neighbors who *have* already adopted the product by time t. The total number of potential adopters at time t is $\sum_x H_x(t) = M - N(t)$, where M is the market potential. At $t = 0$, the initial conditions are $H_0(t = 0) = M$ and $H_x(t = 0) = 0$ for all $x \neq 0$. If the network is sufficiently large compared to the maximal degree in the network, the network is sparsely connected. Because the network is random and hence does not contain short cycles, we can assume that within a short time interval, Δt, the order of potential adopters cannot increase by more than 1 (i.e., the probability of simultaneous adoption of more than one potential adopter's neighbor in the same short time interval Δt is extremely low). Therefore, the change in the number of potential adopters of order x at time t is given by

$$\Delta H_x(t) = -\Delta N_x(t) - \Delta H_{xx+1}(t) + \Delta H_{x-1x}(t), \tag{23}$$

where $\Delta N_x(t) = H_x(t) \cdot (p + xq)\Delta t$ is the number of potential adopters of order x who become adopters within a short time interval Δt after the time t, and $\Delta H_{xx+1}(t)$ is the number of potential adopters of order x that increase their potentiality order to $x + 1$ as result of their neighbors' decisions to adopt at the time t.

In general, adoption by any potential adopter of order y and degree k increases the potentiality order of her $k - y$ potential adopter neighbors by 1. Thus, on average, the total number of potential adopters that increase their order by 1 within a short time interval Δt after time t is $\sum_y \Delta N_y(t) \sum_k (k - y) f_{k|y}(t)$, where $f_{k|y}(t)$ is the conditional probability that potential adopter of order y at time t has a network degree k. By definition, $f_{k|y}(t) = 0$ for all $y > k$ (the number of the individual's adopter-neighbors is limited by her network degree), and the normalization condition is $\sum_k f_{k|y}(t) = 1$. Because the network is random, ΔH_{xx+1} (the number of potential adopters of order x that increase their order to $x+1$) in Δt is proportional to the number of potential adopters of order x; namely,

$$\Delta H_{xx+1}(t) = \frac{H_x(t)(1 - f_{x|x}(t))}{\sum_z H_z(t)(1 - f_{z|z}(t))} \cdot \sum_y \Delta N_y(t) \sum_k (k - y) f_{k|y}(t).$$

Hence the dynamical evolution of the number of potential adopters of order x described by (23) can be rewritten in the continuous limit as follows:

$$\begin{aligned}\frac{dH_x(t)}{dt} = &-(p + xq + (1 - f_{x|x}(t))w(t))H_x(t) \\ &+ (1 - f_{x-1|x-1}(t))w(t)H_{x-1}(t),\end{aligned} \tag{24}$$

where

$$w(t) = \frac{\sum_y h_y(t)(p + xq) \sum_k (k - y) f_{k|y}(t)}{1 - \sum_z h_z(t) f_{z|z}(t)} \tag{25}$$

is the average number of potential adopters that increase their order of potentiality as a result of a single individual's adoption at the time t, and $h_y(t) = H_y(t)/\sum_z H_z(t)$ is the proportion of potential adopters of order x among the entire population of potential adopters at the time t.

To define a closed system of dynamical equations, we also retrieve the dynamics of the conditional probabilities $f_{k|x}(t)$. $f_{k|x}(t) = H_x^{(k)}(t)/H_x(t)$, where $H_x^{(k)}(t)$ is the number of potential adopters of order x with network degree k. In the case where $H_x(t) \neq 0$,

$$\begin{aligned}f_{k|x}(t + \Delta t) &= \frac{H_x^{(k)}(t) + \Delta H_x^{(k)}(t)}{H_x(t) + \Delta H_x(t)} = \frac{f_{k|x}(t) + (\Delta H_x^{(k)}(t))/H_x(t)}{1 + (\Delta H_x(t))/H_x(t)} \\ &\approx f_{k|x}(t) + \frac{\Delta H_x^{(k)}(t) - f_{k|x}(t)\Delta H_x(t)}{H_x(t)}.\end{aligned} \tag{26}$$

For the same reasons that apply to (23), $\Delta H_x^{(k)}(t) = -\Delta N_x^{(k)}(t) - \Delta H_{xx+1}^{(k)}(t) + \Delta H_{x-1x}^{(k)}(t)$, where $\Delta N_x^{(k)}(t) = \Delta N_x(t) f_{k|x}(t)$ is the number of potential adopters of order x and network degree k that adopt at time t, and $\Delta H_{xx+1}^{(k)}(t) = \Delta H_{xx+1} \tilde{f}_{k|x}(t)$ is the number of potential adopters of order x and network degree k that increase their order to $x + 1$ as result of a neighbor's decision to adopt at the time t. Here, $\tilde{f}_{k|x}(t)$ denotes the conditional probability at time t that the network degree of a potential adopter's neighbor is k, given that the neighbor is a potential adopter of order x. Because a potential adopter of order x and network degree k has $k - x$ connections with other potential adopters, it follows that in the case of a sufficiently large random network (using the distribution of a node's neighbors),

$$\tilde{f}_{k|x}(t) = \frac{(k - x) f_{k|x}(t)}{\sum_{k'} (k' - x) f_{k'|x}(t)}. \tag{27}$$

Substituting the explicit expressions $\Delta H_x^{(k)}(t)$ and $\Delta H_x(t)$ in (26) (recall that $\Delta N_x(t) = H_x(t) \cdot (p + xq)\Delta t$ and $\Delta H_{xx+1}(t) = (1 - f_{x|x}(t))w(t)H_x(t)\Delta t$), while setting to the continuous limit, we find that in the case of $H_x(t) \neq 0$,

$$\begin{aligned}\frac{df_{k|x}(t)}{dt} = &\frac{H_{x-1}(t)}{H_x(t)} w(t)(1 - f_{x-1|x-1}(t))(\tilde{f}_{k|x-1}(t) - f_{k|x}(t)) \\ &- w(t)(1 - f_{x|x}(t))(\tilde{f}_{k|x}(t) - f_{k|x}(t)).\end{aligned} \tag{28}$$

On the other hand, in the case of $H_x(t) = 0$,

$$f_{k|x}(t + \Delta t) = \frac{\Delta H_x^{(k)}(t)}{\Delta H_x(t)} = \frac{\Delta H_{x-1x}^{(k)}(t)}{\Delta H_{x-1x}(t)} = \tilde{f}_{k|x-1}(t). \tag{29}$$

This implies that the population at the new product launch time comprises potential adopters with an order of potentiality $x = 0$ (no one has an adopter neighbor) so that $f_{k|0}(t = 0) = P_k$, where P_k is the probability mass function of the degree distribution of the network, and for each potentiality of order $x > 0$,

$$\begin{aligned}f_{k|x}(t = 0) &= \lim_{\Delta t \to 0^+} f_{k|x}(\Delta t) = \tilde{f}_{k|x-1}(t = 0) \\ &= \frac{(k - x + 1) f_{k|x-1}(t = 0)}{\sum_{k'} (k' - x + 1) f_{k'|x-1}(t = 0)}.\end{aligned}$$

The Growth Stage. In a random network, it is unlikely that any individual's two neighbors are neighbors themselves. Thus, in the relatively early stages of the process, the number of potential adopters that have more than one neighbor who is an adopter is very small compared with the number of potential adopters with either one or no adopter neighbor. Furthermore, as most of the population has a network degree that is greater than 1 and most individuals have either 0 or 1 neighbor-adopters at the initial stage of the adoption process, we can assume that $f_{0|0}(t) \ll 1$ and $f_{1|1}(t) \ll 1$ when t is small. Thus, the penetration dynamics at early stages of the diffusion process (see (22)) takes the following form:

$$\frac{dN(t)}{dt} = (M - N(t))p + q(H_1(t) + 2H_2(t)) + O(H_3), \quad (30)$$

where, according to (24),

$$\frac{dH_0(t)}{dt} = -(p + w(t))H_0(t), \quad (31)$$

$$\frac{dH_1(t)}{dt} = w(t)H_0(t) - (p + q + w(t))H_1(t), \quad \text{and} \quad (32)$$

$$\frac{dH_2(t)}{dt} = w(t)H_1(t) + O(H_2), \quad (33)$$

where

$$w(t) = h_0(t)p\sum_k kf_{k|0}(t) + h_1(t)(p+q)\sum_k (k-1)f_{k|1}(t) + O(h_2)$$

and $\quad h_x(t) = \dfrac{H_x(t)}{M - N(t)}$

(see (25)). Let $X_1(t) = H_1(t) + 2H_2(t) = H_1(t) + O(H_2)$ so that

$$\frac{dN(t)}{dt} = (M - N(t))p + qX_1(t) + O(H_2), \quad (34)$$

where

$$\begin{aligned}\frac{dX_1(t)}{dt} &= \frac{dH_1(t)}{dt} + 2\frac{dH_2(t)}{dt} \\ &= (M - X_1(t) - N(t))p\sum_k kf_{k|0}(t) + X_1(t)(p+q) \\ &\quad \cdot \left(\sum_k kf_{k|1}(t) - 2\right) + O(H_2). \end{aligned} \quad (35)$$

Alternatively,

$$\begin{aligned}\frac{dX_1(t)}{dt} &= (M - N(t))p\sum_k kf_k(t) + X_1(t) \\ &\quad \cdot \left\{q\left(\sum_k kf_{k|1}(t) - 2\right) - 2p\right\} + O(H_2), \end{aligned} \quad (36)$$

where

$$f_k(t) = \sum_x h_x(t)f_{k|x}(t) = \frac{M - X_1(t) - N(t)}{M - N(t)}f_{k|0}(t) + \frac{X_1(t)}{M - N(t)}f_{k|1}(t) + O(h_2)$$

is the network degree distribution among potential adopters at time t. Recall that $f_k(t)$ denotes the ratio of the number of potential adopters with network degree k to the total number of potential adopters at time t, so that $f_k(t) = (MP_k - N^{(k)}(t))/(M - N(t))$, where $N^{(k)}(t)$ is the cumulative number of actual adopters with network degree k, and P_k is the network degree distribution (thus, $M \cdot P_k$ is the total number of individuals in the entire population with network degree k). Because at early stages of the process $N(t)/M \ll 1$, we can apply the following approximation: $f_k(t) = P_k(1 - N(t)/M) - N^{(k)}(t)/M + O((N^{(k)}N)/M^2) + O((P_kN^2)/M^2)$, where $f_k(t=0) = f_{k|0}(t=0) = P_k$, and hence $f_{k|1}(t=0) = \tilde{f}_{k|0}(t=0) = (kP_k)/(\sum_{k'} k'P_{k'}) \equiv \tilde{P}_k$. Therefore, the linearization of (36) produces the following:

$$\begin{aligned}\frac{dX_1(t)}{dt} &= Mpk_{\text{avg}} - pX_2(t) + (\tilde{Q} - 2p)X_1(t) + O\left(\frac{X_2N}{M^2}\right) \\ &\quad + O\left(\frac{k_{\text{avg}}N^2}{M^2}\right) + O(H_2), \end{aligned} \quad (37)$$

where $k_{\text{avg}} = \sum_k kP_k$ and $\tilde{Q} = q(\sum_k k\tilde{P}_k - 2)$. The function $X_2(t) = \sum_k kN^{(k)}(t)$ evolves through the dynamical equation $dX_2(t)/dt = \sum_k k(dN^{(k)}(t)/dt) = \sum_k k\sum_x H_x(t)(p + xq)f_{k|x}(t)$, which can also be linearized to produce

$$\begin{aligned}\frac{dX_2(t)}{dt} &= Mpk_{\text{avg}} - pX_2(t) + (\tilde{Q} + 2q)X_1(t) + O\left(\frac{X_2N}{M^2}\right) \\ &\quad + O\left(\frac{k_{\text{avg}}N^2}{M^2}\right) + O(k_{\text{avg}}H_2). \end{aligned} \quad (38)$$

Thus, the temporal derivatives of (34), (37), and (38) generate the following system of linear and homogeneous ordinary differential equations with constant coefficients:

$$\frac{d^2N(t)}{dt^2} \approx -p\frac{dN(t)}{dt} + q\frac{dX_1(t)}{dt}, \quad (39)$$

$$\frac{d^2X_1(t)}{dt^2} \approx (\tilde{Q} - 2p)\frac{dX_1(t)}{dt} - p\frac{dX_2(t)}{dt}, \quad (40)$$

$$\frac{d^2X_2(t)}{dt^2} \approx (\tilde{Q} + 2q)\frac{dX_1(t)}{dt} - p\frac{dX_2(t)}{dt}, \quad (41)$$

and the initial conditions are $(dN/dt)|_{t=0} = Mp$, $(dX_1/dt)|_{t=0} = Mpk_{\text{avg}}$, and $(dX_2/dt)|_{t=0} = Mpk_{\text{avg}}$. The solution of the subsystem (40) and (41) yields

$$\frac{dX_1(t)}{dt} = A_+e^{\lambda_+ t} + A_-e^{\lambda_- t}, \quad (42)$$

where the λs are the roots of the characteristic polynomial $\lambda^2 - (\tilde{Q} - 3p)\lambda + 2p(q + p)$ such that $\lambda_\pm = \frac{1}{2}(\tilde{Q} - 3p) \cdot [1 \pm \sqrt{1 - 8p(q+p)/(\tilde{Q} - 3p)^2}]$, and $A_\pm = \pm(\lambda_\pm/(\lambda_+ - \lambda_-)) \cdot Mpk_{\text{avg}}$. As a result, consider the case where the mean network degree is much larger than 1 and thus $\tilde{Q} \ll q$, and the aggregate-level word-of-mouth effect is much stronger than the effect of the external influence (e.g., marketing efforts) so that $\tilde{Q} \gg p$. One finds that $\lambda_+ = \tilde{Q} - 3p + O(((p+q)/\tilde{Q})p)$ and $\lambda_- = O(((p+q)/\tilde{Q})p) \ll \lambda_+$, and hence $A_- \ll A_+ \approx Mpk_{\text{avg}}$. Consequently, (42) takes the following form:

$$\frac{dX_1(t)}{dt} \approx Mpk_{\text{avg}}e^{(\tilde{Q} - 3p)t}. \quad (43)$$

The solution of the ordinary differential equation (39) following the substitution of (43) in (39) is given by (see (8)):

$$\frac{dN(t)}{dt} \approx Mp\left(1 - \frac{k_{\text{avg}}q}{\tilde{Q} - 2p}\right)e^{-pt} + Mp\frac{k_{\text{avg}}q}{\tilde{Q} - 2p}e^{(\tilde{Q} - 3p)t}. \quad (44)$$

The Decline Stage. In advanced stages of the diffusion, the majority of the population has already adopted the innovation. Hence, most of the remaining potential adopters are surrounded by adopters and become "perfect holes": for almost all potential adopters, the order of potentiality x and the network degree k are equal. In effect, $f_{k|x}(t) \approx \delta_{kx}$, where δ_{kc} is Kronecker's delta. As a result, the dynamic evolutions of the number of potential adopters with network degree k and the number of potential adopters of potentiality order $x = k$ are identical and are given by reducing (24) as indicated in (13).

A.2. Correction Factor for Asymmetric Diffusion Curves

Let t_0 be the time at which we conduct the forecast of the remaining market potential. We assume that t_0 is measured after the peak. Therefore, the decline rate can be estimated.

If $\dot{N}(t_0)$ is the penetration rate at the time t_0 and $N(t_0)$ is the cumulative penetration until time t_0, then the estimated total market potential is

$$M = N(t_0) + \dot{N}(t_0)\int_{t_0}^{\infty} e^{-\beta(t'-t_0)}dt' = N(t_0) + \frac{\dot{N}(t_0)}{\beta},$$

where β is the decline rate. Therefore, if we take the Bass model as a benchmark with decline rate β_B, the ratio of the estimated total market potential values is

$$\frac{M}{M_B} = \frac{N(t_0) + \dot{N}(t_0)/\beta}{N(t_0) + \dot{N}(t_0)/\beta_B}.$$

If we focus on the tails, we find that $\Delta M/\Delta M_B = \beta_B/\beta$. Thus, taking the Bass model as a benchmark, $\beta_B = Q + p$, where Q is the exponential term in the growth stage (which is also the inner force coefficient Q) while $\beta \approx k_{\min}q + p$, where $k_{\min}$ is the minimal degree in the network. Now recall that the growth stage exponential coefficient is given by $Q = q(k_{\text{avg}} + \sigma^2/k_{\text{avg}} - 2)$. We therefore conclude that

$$\frac{\Delta M}{\Delta M_B} = \frac{\beta_B}{\beta} \approx \frac{q(k_{\text{avg}} + \sigma^2/k_{\text{avg}} - 2) + p}{k_{\min}q + p},$$

while in the case of a scale-free network (and also assuming that $k_{\min}q \ll p$), we can approximate this factor to be dependent exclusively on network properties: $\Delta M/\Delta M_B = (k_{\text{avg}} + \sigma^2/k_{\text{avg}})/k_{\min}$. If we wish to incorporate an exponential discount term ρ, the estimated market potential is then given by

$$NPV = \dot{N}(t_0)\int_{t_0}^{\infty} e^{-(\beta+\rho)(t'-t_0)}\,dt' = \frac{\dot{N}(t_0)}{\beta+\rho},$$

and thus

$$\frac{NPV}{NPV_B} = \frac{\beta_B + \rho}{\beta + \rho}.$$

A.3. Description of the Numerical Method Used to Extract Network and Diffusion Parameters

Using Complete Penetration Curve Data Ex Post Facto: Estimations Procedures. We now describe the method of extracting the numerical constraints from the penetration pattern, which is the basis for the network reconstruction method.

Table A.1 Symbols

$N(t)$	Cumulative number of adopters at time t.
$dN(t)/dt$	Rate of adoption at time t.
Q	Aggregate-level internal force coefficient.
P	Aggregate-level external force coefficient.
q	Individual-level internal force coefficient.
p	Individual-level external force coefficient.
H_x	The number of potential adopters of order x, i.e., those who have exactly x neighbors who are adopters of the product at time t.
$f_{k\|x}(t)$	The conditional probability that potential adopters of order x (i.e., with exactly x neighbors who are adopters) at time t have network degree k.
$\tilde{f}_{k\|x}(t)$	The conditional probability at time t that the network degree of a potential adopter's neighbor is k, given that the neighbor is potential adopter of order x.
$\tilde{Q}$	Effective internal force coefficient. This coefficient contains the effect of the network's degree distribution (Equation (9)).
P_k	The network's degree probability mass function.
$\tilde{P}_k$	The probability mass function of the degree of the neighbors of nodes on the network.
k_{avg}, σ	The average and standard deviation of the network degree distribution, respectively.
$\alpha, k_{\min}$	The scale-free exponent and minimal degree of the network degree distribution.
μ, S	The average and standard deviation of the network degree logarithms, respectively (for the lognormal degree distribution).
a, b	The minimal and maximal degrees, respectively (for a uniform degree distribution).
x_1, x_2	The growth and decline exponential slopes, respectively (Equations (16) and (17)).
z_1, z_2	The adjusted exponential slopes of growth and decline, respectively (Equations (18) and (19)).
π_1, π_2	The two parameters of the network degree distribution. Depending on the distribution, these could be k_{avg} and σ (Gaussian) or α and $k_{\min}$ (scale-free). For detailed information, see Table B.1 in the online appendix.

The Growth Stage. For the growth stage, our aim is to extract the exponential slope of the adoption rate. We applied the logistic curve rule by fitting part of a logistic curve. We also used maximum sales growth; i.e., we identified the maximum point of sales growth (second derivative of the cumulative adoption) and regressed for the exponential slope in a log-linear space. Assuming an exponential function, we used the "returns" function $dN(t)/N(t)$, which is effectively the slope of the exponential function, or in this case, the growth rate. We expect this function to be constant in the range of constant exponential growth. We then estimated the value of the function (the exponential slope) using the ordinary least squares (OLS) method with different groups of data points of the returns function. The growth stage exponential slope was taken to be the average value across groups.

The Decline Stage. To estimate the decline stage, we also used the OLS method with different groups of data points taken from the post-peak section using a linear fit in a log-linear scale. We found that the results improve dramatically when we also used, in this case, the returns function $dN(t)/N(t)$, which also exhibits an exponential decline toward the end of the diffusion process, coinciding

Table A.2 Data Set Description for §6.1.2

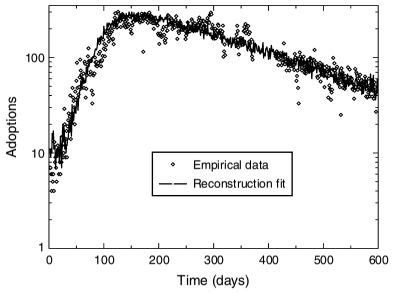
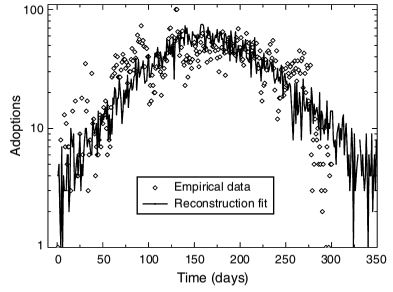
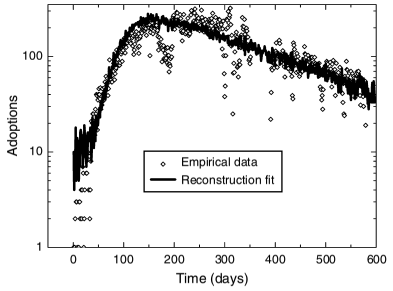
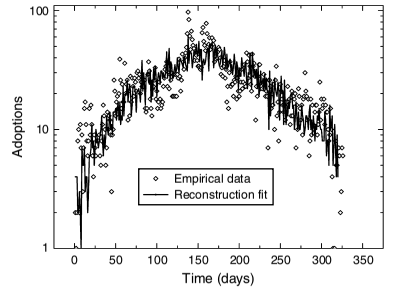
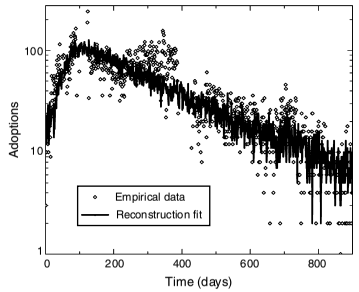
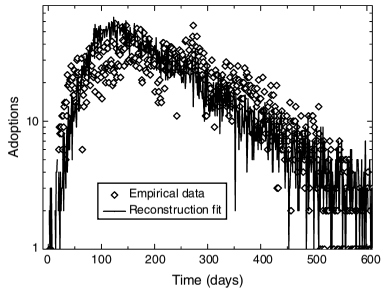
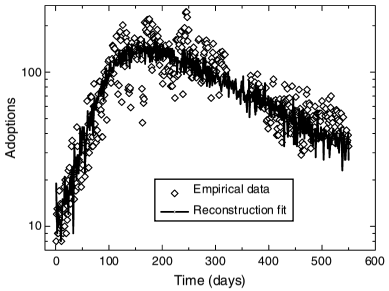
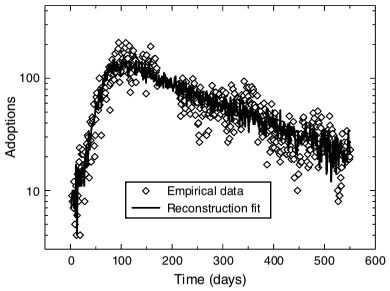
	Data set name and type	Group size	Penetration life cycle (days)	Clustering coefficient	Empirical values of parameters	Values of fitted parameters	Graphs of diffusion curve and fit
Network A (scale-free)	Friendster online group ("Korean Drama Fanz")	74,500	596	0.01	$k_{min} = 13 \pm 2.00$ $\alpha = 2.22 \pm 0.15$	$k_{trans} = 12.21 \pm 2.00$ $\alpha = 2.59 \pm 0.23$	
Network B (Poissonian)	Friendster online group ("Fashionistas")	9,635	301	0.06	$k_{avg} = 2.00$	$k_{avg} = 2.00$	
Network C (scale-free)	Friendster online group ("Emo Is Love")	63,029	620	0.12	$k_{min} = 4.00 \pm 1.00$ $\alpha = 2.25 \pm 0.18$	$k_{trans} = 5.10 \pm 2.00$ $\alpha = 2.66 \pm 0.40$	
Network D (Poissonian)	Friendster online group ("American Idol")	10,622	759	0.14	$k_{avg} = 5.00$	$k_{avg} = 7.00$	

Table A.2 (Cont'd.)

	Data set name and type	Group size	Penetration life cycle (days)	Clustering coefficient	Empirical values of parameters	Values of fitted parameters	Graphs of diffusion curve and fit
Network E (scale-free)	Friendster online group ("Ears Online")	38,323	902	0.09	$k_{\min} = 6.00 \pm 2.00$ $\alpha = 2.10 \pm 0.23$	$k_{\text{trans}} = 9.10 \pm 3.00$ $\alpha = 2.5 \pm 0.10$	
Network F (scale-free)	Friendster online group ("Anime Layouts")	11,123	614	0.13	$k_{\min} = 7.00 \pm 1.00$ $\alpha = 2.37 \pm 0.30$	$k_{\text{trans}} = 4.00 \pm 2.34$ $\alpha = 2.1 \pm 0.3$	
Network G (Poissonian)	Friendster online group ("Starbuckerz")	43,457	1,157	0.05	$k_{\text{avg}} = 3.00$	$k_{\text{avg}} = 4.00$	
Network H (scale-free)	Friendster online group ("ABSCBN")	33,559	1,446	0.17	$k_{\min} = 16.00 \pm 4.00$ $\alpha = 2.02 \pm 0.42$	$k_{\text{trans}} = 8.00 \pm 3.00$ $\alpha = 2.3 \pm 0.10$	

with the exponential decline of the adoption rate. The reason for that was the reduced noise and longer duration of the returns function's tail.

References

Albert R, Barabási AL (2002) Statistical mechanics of complex networks. *Rev. Modern Phys.* 74(January):47–97.

Amaral LAN, Scala A, Barthelemy M, Stanley HE (2000) Classes of small-world networks. *Proc. Natl. Acad. Sci. USA* 97(21):11149–11152.

Barabási AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509–512.

Bass FM (1969) A new product growth for model consumer durables. *Management Sci.* 15(5):215–227.

Bemmaor AC (1994) Modeling the diffusion of new durable goods: Word-of-mouth effect versus consumer heterogeneity. Laurent G, Lilien GL, eds. *Research Traditions in Marketing* (Kluwer Academic Publishers, Boston), 221–229.

Bemmaor AC, Lee J (2002) The impact of heterogeneity and ill-conditioning on diffusion model parameter estimates. *Marketing Sci.* 21(2):209–220.

Berger J, Schwartz EM (2011) What drives immediate and ongoing word of mouth? *J. Marketing Res.* 48(5):869–880.

Braun M, Bonfrer A (2010) Scalable inference of customer similarities from interactions data using Dirichlet processes. *Marketing Sci.* 30(3):513–531.

Chandrasekaran D, Tellis GJ (2007) A critical review of marketing research on diffusion of new products. Malhotra NK, ed. *Review of Marketing* (M.E. Sharpe, New York), 39–80.

Easingwood C, Mahajan V, Muller E (1983) A non-uniform influence innovation diffusion model of new product acceptance. *Marketing Sci.* 2(3):273–295.

Ebbes P, Huang Z, Rangaswamy A (2010) Subgraph sampling methods for social networks: The good, the bad, and the ugly. Working paper, Pennsylvania State University, University Park.

Erdös P, Rényi A (1959) On random graphs I. *Publ. Math. Debrecen* 6:290–297.

Farley JU, Lehmann DR, Sawyer A (1995) Empirical marketing generalization using meta-analysis. *Marketing Sci.* 14(3, Supplement):G36–G46.

Garber T, Goldenberg J, Libai B, Muller E (2004) From density to destiny: Using spatial dimension of sales data for early prediction of new product success. *Marketing Sci.* 23(3):419–428.

Garlaschelli D, Loffredo MI (2008) Maximum likelihood: Extracting unbiased information from complex networks. *Phys. Rev. E* 78(1):015101.

Goldenberg J, Lowengart O, Shapira D (2009a) Integrating the social network to diffusion model and evaluation of the value of hubs in the adoption process. Working paper, Ben-Gurion University, Beer Sheva, Israel.

Goldenberg J, Han S, Lehmann DR, Hong JW (2009b) The role of hubs in the adoption process. *J. Marketing* 73(2):1–13.

Gomez S, Jensen P, Arenas A (2009) Analysis of community structure in networks of correlated data. *Phys. Rev. E* 80:016114.

Gupta S, Mela CF, Vidal-Sanz JM (2006) The value of a "free" customer. Working paper, Harvard University, Cambridge, MA.

Hill S, Provost F, Volinsky C (2006) Network-based marketing: Identifying likely adopters via consumer networks. *Statist. Sci.* 22(2):256–276.

Hinz O, Skiera B, Barrot C, Becker JU (2011) Seeding strategies for viral marketing: An empirical comparison. *J. Marketing* 75(6): 55–71.

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

Jedidi K, Jagpal HS, Desarbo WS (1997) Finite mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Sci.* 16(1):39–59.

Katona Z, Sarvary M (2009) Network formation and the structure of the commercial worldwide Web. *Marketing Sci.* 27(5):764–778.

Katona Z, Zubcsek P, Sarvary M (2011) Network effects and personal influences: Diffusion of an online social network. *J. Marketing Res.* 48(3):425–443.

Katz E, Lazarsfeld PF (1955) *Personal Influence: The Part Played by People in the Flow of Mass Communications* (Transaction Publishers, Glencoe, IL).

Kullback S (1997) *Information Theory and Statistics. Books on Mathematics* (Dover, New York).

Libai B, Muller E, Peres R (2010) Sources of social value in word-of-mouth programs. MSI Reports 10-103, Marketing Science Institute, Cambridge, MA.

Liben-Nowell D, Kleinberg J (2008) Tracing information flow on a global scale using Internet chain-letter data. *Proc. Natl. Acad. Sci. USA* 105(12):4633–4638.

Limpert E, Stahel WA, Abbt M (2001) Log-normal distributions across the sciences: Keys and clues. *Bioscience* 51(5):341–352.

Mayzlin D (2002) The influence of social networks on the effectiveness of promotional strategies. Working paper, Yale University, New Haven, CT.

Mislove A, Marcon M, Gummadi KP, Druschel P, Bhattacharjee B (2007) Measurement and analysis of online social networks. *Proc. 5th ACM/USENIX Internet Measurement Conf.* (ACM, New York), 29–42.

Moe WW, Fader PS (2001) Modeling hedonic portfolio products: A joint segmentation analysis of music compact disc sales. *J. Marketing Res.* 38(3):376–385.

Newman MEJ (2005) Power laws, Pareto distributions and Zipf's law. *Contemp. Phys.* 46(5):323–351.

Newman MEJ, Barabási AL, Watts DJ (2006) *The Structure and Dynamics of Networks* (Princeton University Press, Princeton, NJ).

Newman MEJ, Forrest S, Balthrop J (2002) Email networks and the spread of computer viruses. *Phys. Rev. E* 66:035101.

Pennock DM, Flake GW, Lawrence S, Glover EJ, Giles CL (2002) Winners don't take all: Characterizing the competition for links on the Web. *Proc. Natl. Acad. Sci. USA* 99(8):5207–5211.

Ramasco JJ, Mungan M (2008) Inversion method for content-based networks. *Phys. Rev. E* 77(3):12.

Shaikh NI, Rangaswamy A, Balakrishnan A (2006) Modeling the diffusion of innovations through small-world networks. Working paper, Pennsylvania State University, University Park.

Stephen AT, Toubia O (2010) Deriving value from social commerce networks. *J. Marketing Res.* 47(2):215–228.

Stumpf MPH, Wiuf C, May RM (2005) Subnets of scale-free networks are not scale-free: Sampling properties of networks. *Proc. Natl. Acad. Sci. USA* 102(12):4221–4224.

Stutzbach D, Rejaie R (2005) Capturing accurate snapshots of the Gnutella network. *Proc. IEEE Global Internet Symp.* (*Miami, FL*), 2825–2830.

Trusov M, Rand W (2011) Identifying network properties from aggregate data. Working paper, University of Maryland, Baltimore.

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

Van den Bulte C, Joshi YV (2007) New product diffusion with influentials and imitators. *Marketing Sci.* 26(3):400–421.

Van den Bulte C, Lilien GL (1997) Bias and systematic change in the parameter estimates of macro-level diffusion models. *Marketing Sci.* 16(4):338–353.

Van den Bulte C, Wuyts S (2007) *Social Networks and Marketing* (Marketing Science Institute, Cambridge, MA).

Volz E (2004) Random networks with tunable degree distribution and clustering. *Phys. Rev. E* 70:056115.

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

Watts DJ, Strogatz SH (1998) Collective dynamics of "small-world" networks. *Nature* 393(6684):440–442.

Yeung Y-Y (2005) Macroscopic study of the social networks formed in Web-based discussion forums. *Proc. 2005 Conf. Comput. Support Collaborative Learning—Learning 2005*: *The Next 10 Years*! (International Society of the Learning Sciences, Taipei, Taiwan), 727–731.