Integrating the Social Network to Diffusion Model and Evaluation of the Value of Hubs in the Adoption Process

Jacob Goldenberg,
The Jerusalem School of Business Administration, Hebrew University, Jerusalem, Israel 91905. msgloden@huji.ac.il]

Oded Lowengart, Daniel Shapira
The Guilford Glazer School of Management, Ben-Gurion University, Beer Sheva, Israel, 84105. [OdedL@bgu.bgu.ac.il,shapirad@bgu.ac.il]

Acknowledgments

We would like to thank to thank Eyal Biyalogorsky, Yaniv Dover, Barak Libai, Andrew Stephen and Christophe Van den Bulte for their helpful comments and suggestions. This research was supported by the Israel Science Foundation, The authors are grateful to the Yeshaya Horowitz Association through the Center for Complexity Science and the Kmart International Center of Marketing and Retailing; the Davidson Center, Hebrew University of Jerusalem for supporting this research.

Integrating the Social Network to Diffusion Model and Evaluation of the Value of Hubs in the Adoption Process

Abstract

In this paper we analytically study the effect of social hubs on the penetration of new products. Aggregating individual-level social network considerations to the market level, we identify boundary conditions for hubs' effect on diffusion. Our results demonstrate that seeding hubs has a differential accelerating effect on diffusion measured by the additional net present value (NPV) of potential future sales. On the basis of closed-form solutions, we find that where consumers' decisions to purchase a new product are almost entirely induced by word-of-mouth communications, seeding a small number of hubs whose social-connectedness is about 10 times greater than that of ordinary individuals, may help initiate a valuable diffusion process in which the NPV is increased by several tens of percentage points. On the other hand, seeding such highly connected hubs adds less than 1% to the NPV. Tapping into a category of social influence that is characterized by the number and intensity of social ties, we find that a hub's "area of influence" has greater impact on NPV than its tie intensity. Focusing on the evolution of adoption in a segment of hubs, we show that the product life cycle in this segment is about two to three times shorter than the life cycle in the entire market. We find that the ratio of hub-to-non-hub degree has the most significant impact on reducing life cycle length, and its effect exceeds other effects (i.e., the average proportion of hubs among individuals' neighbors, the intensity of external influence, or word-of-mouth communications). We examine the proposed analytical framework using empirical data from an online social network.

Key words: Social networks, social hubs, diffusion, influentials.

Introduction

Influential people, or *influentials*, are individuals with disproportional influence on other individuals. They are thought to have three important traits: 1) they are blessed with an exceptional persuasive power (they are perhaps even charismatic), 2) they know a lot (i.e., are experts), and 3) they have an exceptionally high degree (a large number of social ties). Influentials are not necessarily superior to others in all these traits and, consequently, may vary in the way they influence others. Recent interest in the third aspect of influentials (their degree) has increased, perhaps due to the growing interest in social networks. While influentials' persuasiveness and expertise are somewhat subjectively measured (Iyengar et al. 2008), degree is more objectively measurable because the number of ties can, in principle, be counted. People with an exceptionally large degree, independent of other influential traits, are also known as *social hubs* (Goldenberg et al. 2009b).

The differential importance of influentials is not entirely clear from the literature. For example, Watts and Dodds (2007) argue that influentials' role in the diffusion process is marginal. Katona, Zubcsek and Sarvary (2009) report findings of degree effect (individuals with high degree adopt sooner) and consistently with Watts and Dodds (2007) they found that individuals' probability of adoption tend to decrease with the average number of neighbors (although these neighbors were not hubs). In contrast, empirical evidence shows that social hubs adopt relatively early and significantly accelerate the diffusion process (Goldenberg et al. 2009b). Generalization based on these two examples, therefore, is rather limited.

In this paper we develop a comprehensive framework that is sufficiently general to allow the integration of social networks, and social hubs in particular, into the dynamics of diffusion, and to identify boundary conditions for influentials' effect on product adoption. While previous works investigate the role of social hubs using computer simulations (e.g, Watts and Dodds, 2007) or with empirical validations through regression models (e.g., Katona, Zubcsek and Sarvary, 2009; Goldenberg, et al 2009), the proposed study incorporates analytically social networks in the traditional diffusion dynamics framework and enables the evaluation of the order of magnitude of the monetary value of social hubs, and the conditions under which this value might change. Such evaluation is accomplished analytically and is based on closed-form solutions.

The proposed framework can be used as managerial decision support tool for conducting efficient viral marketing strategies. It may help managers making decisions whether and how much to invest in persuading influentials to purchase new products or tracking their consumption behavior depending on both hubs and target markets and products attributes. Our main results indicate that hubs accelerate the diffusion of innovations. Specifically, we find that significant economic value accrues when consumers' new product purchase decisions are almost entirely based on word-of-mouth communications (i.e., external influence is very low, compared to an "average product" with P=0.03 and Q=0.38; see Sultan, Farley, and Lehmann (1990)). In such cases, the Net Present Value (NPV) of the diffusion process is increased by several tens of percentage points for a single hub (assuming a 10% interest rate). Furthermore, we show that the adoption rates inside the hub segment are about two to three times greater than those in the remainder of the market. That is, the new product growth rate in the hub segment can be used, at an early stage of the process, as an indication of the subsequent course of diffusion in the entire market. Under average word-of-mouth and marketing efforts a single hub adds less than 1% to the NPV (which is still a relatively strong effect considering that this is a contribution of one individual).

The structure of the paper is as follows: In the next section we present the background to our framework; In the following we evaluate the order of magnitude of social hubs' impact on the diffusion dynamics in the entire market where we control social hubs' time of adoption; Then, we relax this condition, and numerically and empirically study the acceleration effect in the dynamics of diffusion among social hubs. Toward this end we develop an analytical framework that integrates social networks into the dynamics of diffusion as described in details in Appendix A. In the concluding section we discuss the implications of our findings.

Background

There is growing agreement in the literature about the fundamental role played by social networks in the dissemination of information to consumers, channel members, and suppliers (Achrol and Kotler 1999; Iacobucci 1996; Rosen 2002; Valente 1995; Shaikh et. al 2006; Van den Bulte and Wuyts 2007). Research has also linked social network properties to the success of marketing actions such as pricing and promotion strategies (Mayzlin 2002; Shi 2003; Stephen and Toubia 2009a). Although much of the empirical research in this general area has focused on relatively small networks (see Houston et al. 2004 for a review), strength of ties (Brown and Reingen 1987; Rindfleisch and Moorman 2001), and social capital (Ronchetto et al. 1989), it is a plausible generalization that word-of-mouth dynamics is strongly influenced by network structure and features such as degree distribution, clustering coefficient, and density distribution. Thus, network structure can be used to characterize a specific consumer group – the influentials- and their role in diffusion dynamics.

Research suggesting that a relatively small number of individuals have substantial influence on the opinions and decisions of the majority of a population can be traced back at least 50 years (Katz and Lazarsfeld 1955). Current literature on opinion leaders is relatively broad and has extended to a variety of areas, including marketing, public opinion, health care, communication, education, agriculture, and epidemiology. There is broad agreement that opinion leaders can have a major impact on opinion formation and change, and that a small group of influential opinion leaders may accelerate or block the adoption of a product. Although confusion remains in the literature, several related yet distinct sub-categories of influence have been identified: “opinion leaders" are thought to have expertise in a specific area, while "hubs" (Barabasi 2003; Goldenberg et al. 2009b; Valente 1995) are individuals with a large number of social ties leading to a heavy tailed degree distribution of the network (Barabasi 2003; Stephen and Toubia 2009b).

Being connected to many interconnected people confers an information advantage through access to information earlier than the average network member. Burt (1997) suggested that the number of linked individuals to which the individual is connected determines an individual's social capital. Weimann (1994) suggested that centrally positioned scholars (i.e., scientific opinion leaders) determine the direction of scientific progress since innovations adopted by central figures are more widely accepted by other members of the profession. Opinion leaders in a field tend to be inter-connected, thus creating a powerful "invisible college" that dominates the adoption or the rejection of new scientific models, ideas, and methods. Keller and Berry (2003) discuss people who influence others and their relatively large numbers of social links. In Barabasi and Albert's (1999) scale-free model, the distribution of individual node degrees often follows a power law, and a small number of nodes dominate network connectivity due to their extremely great number of ties (these individuals with the highest degree are considered hubs). They show that individuals in a network remain connected thanks to the role of hubs, even when a large number of links are broken (disconnected). Hubs were also found to play a central role in a computer virus infection process (Goldenberg et al. 2005). Thus, scholars concur that individuals with high degrees (hubs) are central in any network.

Nonetheless, mixed results on the contribution of hubs have appeared in the literature on new innovation diffusion. In a simulation-based study, based on Watts (2002), Watts and Dodds (2007) suggest that under most conditions, large cascades of influence can be driven by a critical mass of easily influenced individuals rather than by hubs. They generated computer simulations in which cascades are created by an initial activation of an individual. The chaintype-reaction eventually results in a cascade that is also based on certain properties of the neighboring individuals. In essence, this is an endogenous-driven process that is not directly time-dependent, in contrast to classical diffusion process assumptions. They also note that their results do not exclude the possibility that hubs can be important and that examination of the role of hubs requires more careful specification and testing than it has received so far. Recently, Trusov, Bodapati, and Bucklin (2008) examined an online social network and found that the average network member is influenced by a small number of other members, and in turn influences only a small number of others. In addition, strong heterogeneity was observed: A small proportion of users participated in a substantial share of the influential dyads identified in the network (the impact of a small subset of network members was greater by a factor of eight than that of most other network members). They did not find that having many links (high degree) makes users influential per se. However, their research focused on network activity rather than adoption processes.

A connection between influence and high-connectivity was presented by Goldenberg, Han, Lehmann, and Hong (2009b), who showed that adoption by hubs accelerates the growth evolution of a network, or a diffusion process. A distinction between innovator hubs (primarily affect the speed of adoption) and follower hubs (affect market size) is made. Furthermore, hub adoption might serve as a useful predictor of eventual product success where a small sample of hubs can provide reasonable predictions in very early stages of the PLC. This study, however, did not consider marketing efforts (external force). Similarily, Katona, Zubcsek and Sarvary (2009) found that potential adopters who are connected to many adopted neighbors have a higher probability to adopt in an online social network. They also found that the average influential power of individuals decreases with the total number of their contacts, Their data, according to their report, doesn't include a sufficient number of hubs, so the question about such influence is still relevant.

In sum, the literature on network communications, social connectivity through opinion leaders, and network propagation indicates that hubs may affect dissemination processes on different levels and in different aspects, although the exact nature and limits of these effects are not entirely clear. Nonetheless, it has been shown that the pace of information transfer or product adoption in a network is affected by the number of ties between influentials and non-influentials, the degree of interconnectivity of non-influentials, and the intensity of ties (i.e., intensity of influence). Therefore, social hub adoption patterns may provide firms with early estimations of growth rates.

Social hubs' impact on diffusion dynamics in the entire market

In this section we assess the importance of social hubs and measure the acceleration of the diffusion process by estimating the added NPV of sales. By "seeding" a social hub at time τ, we control for hub activation time and examine how hubs affect new product penetration into the entire market. We take into consideration parameters such as a hub's *area of influence* (i.e. percentage of all potential adopters who are directly connected to a hub), *intensity of ties* (i.e. the intensity of a hub's influence on her neighbors), and the *timing of hub activation*, τ, applying classical diffusion dynamics assumptions to the entire market.

Consider a market comprising M potential adopters, where k is the typical degree of the network - the order of magnitude of most network members' degrees. We assume that external influence and intensity of word-of-mouth communications among individuals do not vary over time and are identical for all network members, and only adopters of the innovation convey word-of-mouth communications. The probability per unit of time that a potential adopter is persuaded by external influence to adopt is p, and q is the probability per unit of time that a potential adopter is persuaded to adopt by word-of-mouth communications by her neighbor (an adopter of the product). Let us further consider a social hub that is seeded to convey word-of-mouth communications at time τ, subsequent to the new product launch. The intensity of a hub's influence on her peers is denoted by v, the probability per unit of time that a potential adopter who is directly connected to a hub will be persuaded by a hub to adopt the innovation. The number of individuals who are directly connected to a hub (the hub's neighbors) is denoted by h, where $h >> k$. The model defines two market segments: S, a segment comprising the hub's immediate neighbors, and $\bar{S}$, a complementary segment of S comprising all individuals in the network who are not directly connected to a hub ("non-neighbors")[1]. The S segment consists of a fraction χ of the market potential where $\chi = \frac{h}{M}$ and is also termed a hub's "area of influence." We estimate the average percentage of actual adopters (excluding hubs) among a potential adopter's neighbors using the percentage of actual adopters in the entire market. Application of the analytical framework is presented in Appendix A and results in the following system of two coupled Ordinary Differential Equations that describe the adoption dynamics in each segment as shown in Appendix B:

$$(1) \quad \frac{d\mu_S(t)}{dt} = \begin{cases} (1-\mu_S(t))(P+Q\mu(t)) & if \quad t < \tau \\ (1-\mu_S(t))(P+v+Q\mu(t)) & if \quad t \geq \tau \end{cases} \text{ and } \frac{d\mu_{\bar{S}}(t)}{dt} = (1-\mu_{\bar{S}}(t))(P+Q\mu(t)).$$

Here, $\mu_S(t)$ and $\mu_{\bar{S}}(t)$ denote the proportion of actual adopters in segments S and $\bar{S}$, respectively, as a function of time t, with the initial condition $\mu_S(t=0) = \mu_{\bar{S}}(t=0) = 0$, where $t = 0$ is the new product launch time. The aggregate-level coefficients of diffusion, P and Q, are associated with the external and internal individual-level influence via the following relations: $P = p$ and $Q = kq$, respectively. Equation system 1 is coupled through $\mu(t)$, the proportion of actual adopters in the entire market at time t which is obtained by the weighted average $\mu(t) = \chi\mu_S(t) + (1-\chi)\mu_{\bar{S}}(t)$.

In Appendix A we show that the proportion of actual adopters in the entire market evolves according to a special case of the *Ricatti Equation* (Polyanin and Zaitsev 2003), to give:

(2) $$\frac{d\mu(t)}{dt} = (1-\mu(t))(P_{ef}(t) + Q\mu(t)) \text{ where } P_{ef}(t) = \begin{cases} P & if \quad t < \tau \\ P + \dfrac{v\chi}{(1-\chi)e^{v(t-\tau)} + \chi} & if \quad t \geq \tau \end{cases}$$

with the initial condition: $\mu(t = 0) = 0$, leading to a closed-form solution:

(3) $$\mu(t) = \begin{cases} \dfrac{1 - e^{-(P+Q)t}}{1 + \dfrac{Q}{P} e^{-(P+Q)t}} & if \quad t < \tau \\ \mu_\tau + (1-\mu_\tau) \dfrac{\chi A(1 - e^{-(P+v+Q)(t-\tau)}) + (1-\chi)(1 - e^{-(P+Q)(t-\tau)})}{\chi(A + Be^{-(P+v+Q)(t-\tau)}) + (1-\chi)(1 + Ce^{-(P+Q)(t-\tau)})} & if \quad t \geq \tau \end{cases}$$

where $\mu_\tau \equiv \mu(t = \tau) = \dfrac{1 - e^{-(P+Q)\tau}}{1 + \dfrac{Q}{P} e^{-(P+Q)\tau}}$, $A = \dfrac{(P+Q)(P+v+Q\mu_\tau)}{(P+Q\mu_\tau)(P+v+Q)}$, $B = \dfrac{Q(1-\mu_\tau)(P+Q)}{(P+Q\mu_\tau)(P+v+Q)}$

and $C = \dfrac{Q(1-\mu_\tau)}{P+Q\mu_\tau}$.

As can be seen, prior to hub activation (i.e., when $t < \tau$), the penetration process follows Bass model dynamics (Equations 2-3). At time τ, a hub is "seeded" and impacts her social environment. Seeding a hub whose neighbors comprise χ percent of the market with tie intensity coefficient v, is analogous to applying additional external force with power v on χ percent of the entire market (Equation 1), and the entire market reaction is indicated by Equation 2. Immediately after hub seeding (at time $t = \tau$), the effective incremental external force is $v\chi$. As a result individuals who are directly affected by the seeded hub (neighbors who have not yet adopted) tend to adopt earlier than other potential adopters. Thus, over time the proportion hub's neighbors of in the population of potential adopters declines so the effective increment to the external force that is applied on the entire market decays as well.

Next, consider a set of hubs that are simultaneously seeded, with either no overlap or complete overlap in their area of influence. In the case of non-overlapping areas of influence, a hub's area of influence χ is applicable to all seeded hubs, and therefore represents the fraction of the market that directly interacts with a seeded hub. Conversely, in the case of completely overlapping areas of influence, hubs' tie intensity v denotes the total area of influence of the entire set of seeded hubs. In general, when areas of influence partially overlap, the parameter χ can be associated with the share of the potential market that is directly exposed to the influence of at least one seeded hub, and the hubs' tie intensity coefficient v represents the average impact of hubs on individuals. As a result, Equations 2-3 can also be used to evaluate the order of magnitude of the hub effect on the dynamics of new product growth in the case of a set of simultaneously seeded hubs (see Appendix B for a detailed explanation).

When market potential is constant over time, hubs affect consumers' time of adoption, such that they accelerate product penetration. To assess the order of magnitude of the economic value of the acceleration effect, we compare the NPV of a penetration process with a hub (or a group of hubs) with tie intensity coefficient v and area of influence χ, seeded at time τ, to the NPV of a process without the contribution of the seeded hubs (similarly to Goldenberg et al. 2007; Hogan et al. 2005)):

(4)
$$\frac{\Delta NPV}{NPV_0} = \frac{\int_{\tau} \frac{\dot{\mu}(t) - \dot{\mu}_0(t)}{(1+R)^t} dt}{\int_{0} \frac{\dot{\mu}_0(t)}{(1+R)^t} dt}$$

Here $\dot{\mu}(t) \equiv \frac{d\mu(t)}{dt}$, where $\mu(t)$ is the fraction of adopters in the entire market at time t, assuming that a hub, or a number of hubs, with tie intensity coefficient v and area of influence χ are seeded at time τ with diffusion coefficients P and Q, as presented in Equation 3. On the other hand, $\dot{\mu}_0(t) \equiv \frac{d\mu_0(t)}{dt}$, where $\mu_0(t)$ is the percentage of adopters at time t when a classical diffusion with coefficients P and Q takes place, resulting in $\mu_0(t) = \frac{1 - e^{-(P+Q)t}}{1 + \frac{Q}{P} e^{-(P+Q)t}}$.

We numerically calculate the relative increment of the diffusion NPVs due to hubs' influence given by Equation 4, for various combinations of parameters (see below). We estimated the integrals utilizing the rectangle rule, while using their Riemann sums at a daily resolution to create an appropriate partition of time (Davis and Rabinowitz 1984).

Hubs may differ in their impact on their neighbors. We distinguish between "conventional" hubs and "intensive" hubs, based on the intensity of their ties to their respective neighbors. Thus, the tie intensity coefficient of conventional hubs may satisfy the approximation $v \sim q$, whereas the tie intensity coefficient of intensive hubs meets the condition $v >> q$. Here, $q = \frac{Q}{k}$ is the individual-level internal force coefficient, and k is the typical network degree. The number of neighbors of a non-hub can range from several dozens to more than one hundred, so typical values of k would lie between $k \sim 10$ and $k \sim 100$ (Trusov et al. 2008; Trusov and Rand 2009; Dover et al. 2009).

The contribution of a single hub to the aggregate hub area of influence χ is of an order of 1% or less[2]. Figure 1 presents the relative increments in the NPV of a diffusion process resulting from hub seeding for different products representing different combinations of diffusion coefficients P and Q, where P ranges from 0.001 to 0.05 and Q ranges from 0.01 to 1.5 in annual units. We consider the two levels of hub tie intensity: a conventional hub with tie intensity coefficient $v = \frac{Q}{10}$ (shown by blue curves) and an intensive hub with tie intensity coefficient $v = Q$ (shown by the red curves), assuming hub impact is of an order of magnitude equal to that of the total market word-of-mouth impact on a single individual. In this analysis, hub area of influence is $\chi = 1\%$, to emulate the effect of a single seeded hub; interest rate is 10%; and seeding takes place at new product launch time ($\tau = 0$). As can be seen from Figure 1, the incremental NPV due to seeding conventional and intensive hubs reaches maximum values of around 8% and 40%, respectively. For a typical product with $P = 0.03$ and $Q = 0.38$ in annual units (see Sultan et al. 1990), added NPV of conventional and intensive hubs is 0.2% and 1.4%, respectively. Alternatively, seeding a single ordinary individual yields a maximum NPV increment of about 0.8% (over all product types), yet less than 0.01% for a typical product[3].

Not surprisingly, in all cases, increasing external market force yields a decrease in the relative NPV increment of the diffusion process because it diminishes the impact of hubs (see Figures 1A). This can be attributed to the fact that hub influence is analogous to incremental external influence. When the original level of external influence P is increased, the relative contribution of a hub to the effective external force is reduced (see Equation 2). In contrast, when the value of P approaches zero, the only way to trigger diffusion is by seeding hubs; as a result, the relative NPV increment due to hub seeding increases to infinity. Hubs influence, therefore, is stronger when external force is lower (e.g., less effective marketing efforts).

Less intuitive is the decrease in the relative NPV increment when Q takes large values (above approximately 0.3 in annual units), as shown in Figure 1B. Even though hub intensity is created by word-of-mouth, the acceleration effect caused by hub activities diminishes after a certain level of Q. This non-intuitive result is attributed to the fact that when the word-of-mouth effect is originally strong, the probability of a potential adopter to adopt is high even without taking into account the effect of the seeded hubs, so their relative influence becomes less important in persuading other people to purchase a new product.

Insert Figure 1 about there

Figure 2 illustrates the effect of hub attributes (intensity of ties and area of influence), and timing of hub seeding on the relative NPV increment. We consider two penetration processes: (a) a typical product with $P = 0.03$ and $Q = 0.38$, and (b) a process with low external influence, represented by $P = 0.001$ and $Q = 0.38$. Since the source of hub tie intensity is word-of-mouth communications, we evaluate the hub tie intensity coefficient v as a fraction of Q , the coefficient of imitation of the entire market. We investigate the relative NPV increment as a function of the hub tie intensity v (Figure 2A) and the hub area of influence χ (Figure 2B) while holding the other variables constants. We refer to a case in which time of seeding coincides with the new product launch ($\tau = 0$), and interest rate is 10%. Naturally the relative NPV increment is an increasing function of both v and χ [4]. We find that relatively weak external market force effects result in an increase in NPV several tens of percentage points for low-P products compared to an increase of just a few percents at the most for typical products.

It is interesting to find that the impact of the area of influence, χ , of a seeded hub is stronger than its tie intensity, v . This feature is illustrated in Table 1, where we present results of a sensitivity analysis of the NPV for these effects when seeding is performed at new product launch, i.e., $\tau = 0$. The rows in this table represent product types, and the columns represent different combinations of hub tie intensities and areas of influence, which are variations of $v\chi$. Naturally, a greater value of $v\chi$ (two right columns in Table 1) results in a rise in relative NPV increments in both product types. This increase, however, is greater when area of influence χ increases (while increasing the product $v\chi$) compared to a similar relative change in hub tie intensity (to reach the same value of $v\chi$ - second vs. third columns). This result is attributed to the fact that hubs establish their broad influence through their neighbors. Thus, even an extremely intensive hub would affect just a few neighbors in case where the number of friends is relatively low. Namely, in order to generate a macro-type effect on the dynamics of diffusion, the most important attribute of the seeded hubs is their area of influence.

Insert Table 1 about here

Another dimension that potentially affects the hubs' role is seed timing. We, therefore, analyzed the change in NPV due to late seeding, several years from launch time, subject to an area of influence $\chi = 1\%$, seeded hub tie intensity $v = \frac{Q}{10}$, and interest rate $R = 10\%$, as presented in Figure 2C. As expected, postponement of hub seeding reduces the NPV increment. In addition to late proceeds stemming from late purchases that affect this result, the impact of hub seeding is diminished. In this case, since the diffusion process is underway and a greater number of consumers have already adopted the product, the impact of seeding hubs (which, as mentioned, is analogous to applying supplementary external market force on the hub area of influence) is reduced.

Insert Figure 2 about here

. In the analysis presented above, we considered hubs that are seeded to broadcast word-of-mouth communications to a market comprising ordinary consumers. Cases in which the market itself consists of hubs and ordinary consumers can be represented by large values of "natural" word-of-mouth coefficient Q. As a result, the NPV increment generated by seeding a given number of hubs in the market is reduced, as explained above and illustrated in Figure 1B, which also illustrates the order of magnitude of the relative NPV increments.

We now expand our analysis to include effects of consumer characteristics by examining a heterogeneous market comprising hubs and ordinary individuals (non-hubs). In the next section, hub activity is determined endogenously through the dynamics of diffusion (rather than exogenously by seeding) where our objective is to compare the dynamics of adoption in the segment of hubs with the dynamics in the segment of ordinary consumers.

The dynamics of diffusion dynamics among hubs

In this section we focus on hubs, and how hubs, as a group, are affected by the penetration of a new product, compared to the adoption process among ordinary individuals. Unlike the previous analysis where timing of hub activation was an independent variable, here we study the dynamics of diffusion over hubs; Consequently, hub activation becomes now a dependent variable. First, we explore the dynamics of new product penetration into a heterogeneous market featuring a defined degree distribution. Then, to evaluate the order of magnitude of the acceleration effect among the hubs, we reduce our model to a two-population model: hubs and non-hubs. Again, we first look at the problem at the individual level, consider social network factors, and then aggregate the effects to the market level. We compare Product Life Cycles (PLC's) in the two populations using an acceleration factor that is defined by the ratio of the PLCs in the non-hub and hub segments.

Consider a heterogeneous market comprising several segments differentiated by individual network degree (i.e., the k^{th} sub-market comprises all the individuals with a network degree of *k*). We assume that external influence *p* and intensity of the word-of-mouth communications among individuals *q,* are time-independent, identical for all potential consumers, and operate independently. That is, individuals differ only in their network degree, implying that all the hubs in the market are conventional hubs[5]. We further assume that all individuals have the same probability of being part of the market potential, ρ and we also assume that, on average, the proportion of adopters among an individual's neighbors does <u>not</u> depend on her network degree. We use the analytical framework presented in Appendix A to derive a coupled system of first order ODEs (see Appendix C for the derivation):

(5) $$\frac{d\mu_k(t)}{dt} = (\rho - \mu_k(t))(p + kq\sum_{l=l_{\min}}^{l_{\max}} \tilde{f}_l \mu_l(t)) \quad ; \qquad k = 1,2,\ldots,k_{\max}$$

$\mu_k(t)$ is the percentage of adopters among individuals with network degree k. Here $\tilde{f}_l$ denotes the mean percentage of an individual's neighbors with degree l. That is, equation system 5 is coupled through the average proportion of adopters among individuals' neighbors, given by the weighted sum $g(t)=\sum_{l=l_{\min}}^{l_{\max}}\tilde{f}_l\mu_l(t)$. In general, $\tilde{f}_l$ is <u>not</u> given by f_l, the proportion of individuals with degree l in the entire population, since the distribution of neighbors' degrees is skewed towards high degrees (the probability of being linked to a hub is greater than the percentage of hubs in the market). For example, it follows from Albert and Barabasi (2002) that in a random network, $\tilde{f}_l=\frac{lf_l}{(l)}$, where (l) is the averaged network degree.

Let $t=0$ be the new product launch time, and the initial conditions of the above system are $\mu_k(t=0)=0$ for each sub-market k. The following solution for each sub-market is obtained (see Appendix C for a detailed elaboration):

(6) $$\mu_k(t)=\rho\left(1-e^{-(pt+kqG(t))}\right)$$

where $G(t)=\int_0^t g(t')dt'$ is the *generating function* of all the solutions $\mu_k(t)$ and satisfies the following first-order differential equation:

(7) $$g(t)=\frac{dG(t)}{dt}=\rho\left(1-\sum_{l=l_{\min}}^{l_{\max}}\tilde{f}_l e^{-(pt+lqG(t))}\right)$$

with the initial condition $G(t=0)=0$. The percentage of total adopters in the entire market is a weighted sum of the fractions of adopters in all the sub-markets. Consequently,

(8) $$\mu(t)=\sum_{l=l_{\min}}^{l_{\max}}f_l\mu_l(t)$$

where f_l is the ratio of the number of individuals with network degree l to the total market size[6].

Equations 6-8 describe the dynamics of new product growth in a heterogeneous market and imply that the existence of hubs accelerates the penetration of an innovation into the market, as hubs tend to adopt earlier. Since $G(t=0)=0$ at launch time, it is clear from Equation 7 that for any $t>0$, $\frac{dG(t)}{dt}>0$, and hence the generating function $G(t)$ increases monotonically. Thus $qG(t) > 0$ implies that the greater the network degree k, the greater the percentage of adopters μ_k in the corresponding sub-market at any given time t (as indicated by Equation 6). Furthermore, the dynamic equations imply that the diffusion process in the entire market accelerates as the number of hubs in the market increases, since the probability of an individual to be linked to a hub increases. Therefore, the distribution of the degrees of neighbors $\tilde{f}_l$ is skewed towards high network degrees so the exponential terms in the sum $\sum_{l=l_{\min}}^{l_{\max}} \tilde{f}_l e^{-(pt+lqG(t))}$ with high decaying rates are more heavily weighted and boost the growth rate $\frac{dG(t)}{dt}$ of the generating function (see Equation 7). As a result, $G(t)$ and hence all the solutions $\mu_k(t)$ assume greater values (see Equation 8). In addition, increasing the number of hubs in the market also skews the degree distribution in the network f_l. Thus, the total percentage of adopters in the market $\mu(t)$ increases with the number of hubs. In Appendix C we show that similar patterns of adoption can also be found in cases where the underlying network is not symmetric.

The empirical verification of our modeling approach is based on data of a group of 37,704 individuals affiliated with a specific web-supported interest group that is part of a large online social network. The number of ties in the large network and time of adoption are known for each group member at a daily resolution. In Figure 3A we present the retrospective curve of adoption for our group. This curve represents the product life cycle pattern (the drops between days 300 and 400 are due to website crashes). The distribution of network degrees is displayed in Figure 3B. The degrees are distributed according to a heavy-tailed distribution spanning over four orders of magnitude. Most individuals have between 101 to 1,000 links in the larger online network (75.1%). Some individuals have less than 100 links (19.7%), while 1% of the individuals in our data set have less than 10 links. The remainder (5.2%) have more than

1,000 links. The mass points around the degree 3,000 are due to an exogenous limitation defined by the online network organizers. The mean and median network degrees are 339 and 229, respectively. The standard deviation is 353, where 4.6% and 1.8% of the group have a network degree that exceeds 2 and 3 standard deviations from the mean, respectively.

For each network degree, the percentage of individuals in the data set who have not yet adopted group membership can be deduced at any given time t. The fraction of non-adopters in the sample with a certain degree is, in fact, a good measure for the unexhausted market potential of the sub-market defined by that degree. That is, the data can provide empirical curves of the unexhausted potential $\eta(k,t)$ as a function of network degree k and time t. From Equation 6, we expect:

(9) $$\eta(k,t) \equiv \frac{\rho - \mu_k(t)}{\rho} = e^{-(a(t)k + b(t))}$$

where $a(t) = qG(t)$, encapsulating the local effects in the process (i.e. the word-of-mouth communications disseminated by an individual's neighbors through the network) and $b(t) = pt$, representing the global effects caused by external forces[7]. For any given time t, the values of $a(t)$ and $b(t)$ can be measured by calculating logarithmic regression equations for corresponding empirical curves, where network degree k is the independent variable and the percentage of the unexhausted potential market $\eta(k,t)$ is the dependent variable. The slope and intercept of each logarithmic regression equation are $-a(t)$ and $-b(t)$, respectively. Since the generating function $G(t)$ is positive and monotonically increasing, we anticipate that the more advanced time t, the steeper and more negative the regression equation slope. Furthermore, as $\frac{d^2a(t)}{dt^2} = q\frac{d^2G(t)}{dt^2} = q\frac{dg(t)}{dt} = q\sum_{l=l_{\min}}^{l_{\max}} \tilde{f}_l \frac{d\mu_l(t)}{dt} > 0$, drawing the coefficient of the local effects $a(t)$, which is in fact the absolute value of the slope as a function of time, should yield a convex curve. At the same time, the intercept $-b(t)$ becomes more negative because the coefficient of the global effects $b(t)$ is expected to grow linearly over time.

Five empirical curves of the unexhausted market potential percentage as a function of network degree at five points in time are shown in Figure 3C (the figure is drawn on a semi-logarithmic scale so that exponential relations between variables, through logarithmic regression, are displayed as straight lines). To avoid misrepresentation due to the discrete nature of the problem, we only consider network degrees that define sub-markets of more than 30 individuals. As predicted, as time advances, the slope becomes steeper and the value of the intercept provided by the logarithmic regression becomes more negative.

In Figure 3D, we present R^2, goodness of fit as a function of time. Close to product launch time, slopes are close to zero and hence the R^2 measures are close to zero as well. Gradually, over time, the slopes become steeper and the values of the corresponding R^2 increase to 0.5. The subsequent decline in R^2 is due to the extraction of the sub-market potentials from the regression equation: Since the number of individuals in each sub-market is discrete, for any given sub-market there is a certain point in time after which the unexhausted market potential is zeroed out and cannot take part in the logarithmic regression. In practice, the empirically measured logarithmic values of unexhausted market potential percentage in each sub-market are bounded and hence the observed curve is somewhat distorted.

Insert Figure 3 about here

In Figure 3E we draw the evolution of the coefficient of local influence, $a(t)$, given by the absolute values of the logarithmic regression slopes. In accordance with our calculation, $a(t)$ is an increasing convex function. At some point, (around day 370), $a(t)$ reaches a plateau and even tends to decline thereafter. This, again, is due to effects related to the discrete nature of the market. Sub-markets that exhaust their potential cannot take part in the logarithmic regression as the logarithm of the unexhausted sub-market potential becomes undefined. This results in an effective upper boundary on the slopes in the logarithmic regressions.

The dynamic evolution of the coefficient of global influence $b(t)$ is represented in Figure 3F. As expected, growth is evident, but the rate of growth significantly exceeds the linear increase predicted by our model. Specifically, the empirical data does not conform to our assumption of time-constant external influence $p(t) = p$. Apparently, all individuals who consider joining the group are exposed to the total number of group members. It is therefore reasonable to assume that a group's attraction increases with group size. Thus, we modified external influence to account for this externality effect, incorporating an additional term that is proportional to the total cumulative number of adopters. We adjusted our model such that $p(t) = p_0 + p_1 N(t)$, where p_0 and p_1 are constants and $N(t)$ is the total cumulative number of adopters at time t. Replication of the analytical calculations using the model adjusted for external influence does not change the form of Equation 9 or the conclusions derived earlier (see Appendix C), although the coefficient of global influence now assumes the following form: $b(t) = \int_0^t p(t')dt' = p_0 t + p_1 \int_0^t N(t')dt'$.

Since the cumulative number of adopters $N(t)$ can be retrieved from the empirical data at a daily resolution for any given time t, we also have an empirical measure for its integral which can be approximated by the appropriate Riemann sum: $\int_0^t N(t')dt' \approx \sum_{t'=0}^{t} N(t')\Delta t$ where $N(t')$ is the cumulative number of actual adopters at time t' where $0 \le t' \le t$), and $\Delta t = 1_{day}$. Therefore, we perform a two-dimensional linear regression to verify the adjusted model for the dynamics of the coefficient of global effects $b(t)$, where the independent variables are the time indices t and the integrals of the cumulative number of actual adopters (see Figure 3F). As can be seen from this figure, global effects coefficient values predicted by the adjusted model (red curve) nicely fit the empirical measurements (black curve). The R^2 of this regression is 0.987.

In the remainder of this section we estimate the order of magnitude of the acceleration of diffusion effect among the hubs in the market. For better exposition, we reduce the model to a simpler case with two segments, hubs, and non-hubs, differentiated by their number of ties. (similarly to Lehmann and Esteban-Bravo 2006; Muller and Yogev 2006; Steffens and Murthy 1992; Van den Bulte and Joshi 2007). Note, however, that we do not make any assumption about individual innovativeness but rather analyze the two groups solely based on their network degree.

Consider a market that comprises a small segment of hubs with network degree h and a large segment of non-hubs with network degree k, where $h >> k$. In an analogy to the individual-level interpretation of Bass dynamics, let $P = p$ be the coefficient of innovation, and $Q = kq$ the macro-level coefficient of imitation. The dynamics of diffusion in each segment can be reduced from Equation 6 to give:

(10) $$\mu_k(t) = \rho\left(1 - e^{-(Pt+QG(t))}\right) \text{ and } \mu_h(t) = \rho\left(1 - e^{-(Pt+\Gamma QG(t))}\right)$$

where $\mu_k(t)$ and $\mu_h(t)$ are the fractions of adopters among non-hubs and hubs, respectively. Hub degree factor $\Gamma = \frac{h}{k}$ is an estimate of the relative social connectedness of the hubs in comparison to non-hubs. As discussed above, the dynamics of diffusion in both segments is determined by the generating function, $G(t)$, that is given by the solution of a reduced form of the ODE in Equation 7, which corresponds to a two-population case with the initial condition $G(t = 0) = 0$. Namely,

(11) $$\frac{dG(t)}{dt} = \rho\left(1 - e^{-(Pt+QG(t))} + \tilde{f}e^{-(Pt+QG(t))}(1 - e^{-(\Gamma-1)QG(t))})\right)$$

where $\tilde{f}$ denotes the mean percentage of hubs among any individual's neighbors. If we consider the usual case where most individual's neighbors are non-hubs, $\tilde{f}$ becomes a small parameter of the problem, such that $\tilde{f} << 1$. In Appendix D we show that linearization of Equation 11 yields a solution of the form $G(t) = G_0(t) + \tilde{f}G_1(t) + O(\tilde{f}^2)$ where the degree factor Γ does not appear in the leading term $G_0(t)$ and its effect on the next term $\tilde{f}G_1(t)$ is very small since $\tilde{f}$ is a small parameter of the problem and G_1 (unlike G_0) is a bounded function. Thus we can conclude that the generating function $G(t)$ is imperceptibly dependent on the hub degree factor Γ.

To compare the dynamics of new product growth in the hub segment to the adoption process among non-hubs, we investigate the relation between product life cycle times in both segments. Specifically, given a certain threshold θ, we define the product life cycle as the time that elapses until the product is adopted by θ percent of the total population ($\theta < \rho$, where ρ denotes the ratio of the market potential to the total population size.) That is, product life cycle durations T_h and T_k in the hub and non-hub segments, respectively, are defined as the times at which $\mu_k(t = T_k) = \mu_h(t = T_h) = \theta$. Thus, using Equation system 10, the following equality holds:

(12) $$PT_k + QG(T_k) = PT_h + \Gamma QG(T_h) = -\ln\left(1 - \frac{\theta}{\rho}\right)$$

We compare the diffusion processes in the two segments, utilizing an acceleration factor given by the product life cycle duration ratio, defined as $A = \frac{T_k}{T_h}$. If, for example, the duration of a product life cycle in the hub segment is one-half of the duration in the non-hub segment, we may infer that the diffusion process in the hub segment is accelerated by a factor of 2 (i.e., $A = 2$). In Figure 4, this effect is illustrated by numerical measures of the acceleration factor in various calculations. We retrieved these numerical measures by applying numerical solutions of Equation 12, where the values of the generating function were calculated numerically on the basis of the exact dynamic Equation 11, using the fourth-order Runge-Kutta method (Bogacki and Shampine 1989).

Figures 4A and 4B illustrate the impact of hub characteristics. Naturally, both hub degree factor Γ and mean percentage of hubs among individuals' neighbors $\tilde{f}$ increase as hub network degree increases. For clarity of presentation, we examine the acceleration effect as a function of Γ (Figure 4A) and $\tilde{f}$ (Figure 4B) separately while holding the other variables constant. That is, each curve in Figure 4A corresponds to the acceleration factor A as a function of Γ given a constant value of $\tilde{f}$ and vice versa, while each curve in Figure 4B represents the dependence of the acceleration factor A on $\tilde{f}$ for a constant value of Γ. We used a typical product with $P = 0.03$ and $Q = 0.38$ in annual units (Sultan et al. 1990). The threshold for product life cycle completion was defined as $\theta = 0.85\rho$ (i.e. 85% of the market potential), a stage at which only laggards have not yet adopted the product (Rogers 1963). We further assume that the market potential consists of the entire population; consequently, $\rho = 1$.

Naturally, increasing the number of hubs or hub ties in the population increases $\tilde{f}$ and accelerates the diffusion process in both segments. We do find, however, that the diffusion acceleration among hubs is greater than the acceleration in the non-hub segment, since the acceleration factor increases monotonically with both hub degree factor Γ and the mean percentage of hubs among individuals' neighbors, $\tilde{f}$ (see Figures 4A and 4B). Increasing the number of hub ties merely has an indirect effect on non-hubs in the sense that it enhances the adoption process due to word-of-mouth communications generated by hubs who have already adopted the new product. Furthermore, such reinforcement also has a direct effect on the hubs themselves. Increasing the absolute number of ties in the hub segment also increases this segment's exposure to new product adopters, and hence the product life cycle is reduced more significantly in this segment.

In Figures 4C and 4D we evaluate the acceleration effect for different types of products using social hubs whose degree is 10 times greater than the network degree of non-hubs ($\Gamma = 10$), where the mean percentage of hubs among neighbors is $\tilde{f} = 5\%$. The threshold for the product life cycle is $\theta = 0.85\rho$, where $\rho = 1$. Figure 4C presents the acceleration factor A as a function of the coefficient of innovation P for a given coefficient of imitation Q (each curve corresponds to different value of Q). As long as P is small relative to Q, the acceleration factor A increases with P. At a certain point, the acceleration factor reaches maximum and decreases with P thereafter. Obviously, the more intensive the external market force that determine P, the more rapidly new product penetration occurs in both segments. In the region of small P's, any increase in P has a greater relative effect on the hub segment since the initiation of the adoption process is advanced in time. As the rate of growth is much greater in the hub segment (hubs have a greater chance than non-hubs of being connected to adopters), the ratio of the time to product life cycle completion in the hub segment to the time to product life cycle completion in the non-hub segment increases. Namely, the acceleration factor $A = \frac{T_k}{T_h}$ increases with P when P is small. This trend is reversed when the values of the coefficient of innovation, P, are large relative to the coefficient of imitation Q. In this case, the adoption process is dominated by external influence. As a result, word-of-mouth effects become less important and the difference between the product life cycle durations in the two segments is reduced.

Similar behavior is observed when the acceleration factor A is plotted as a function of the coefficient Q , for a given P (see Figure 4D). Analogous to the previous case, when values of Q are small, the acceleration effect increases until it reaches maximum and thereafter decreases with Q when Q is large. When Q is small, there is a considerable difference between word-of-mouth effects in the two segments. Since the Q coefficient in the hub segment is multiplied by the hub degree factor Γ , the probability of early adoption is significantly greater in the hub segment compared to the non-hub segment, as long as Q is small. However, when Q is sufficiently large, the word-of-mouth effect become significant among non-hubs, and the probabilities of adoption in both segments become more similar and approach 1. Hence, the ratio of the product life cycle in the non-hub segment to the product life cycle in the hub segment decreases.

Insert Figure 4 about here

In summary, we find that even under the most conservative assumptions concerning hub characteristics (i.e. distinct number of ties, no difference in innovativeness), hubs tend to be early adopters compared to other individuals who are not as socially connected (without anything assumed about their innovativeness). We also find that the number of hubs' social ties makes the greatest contribution to the acceleration of the dynamics of diffusion in the hub segment vs. the entire market, almost independent of product type. As a result, the hub segment can serve as an indicator of new product penetration success. If, for example, the hub network degree is about 10 times the degree of non-hubs, product life cycle duration in the hub segment is 1.7 to 2.7 times shorter than in the non-hub segment, which represents most of the market. Therefore, marketers can predict the market’s future response to a new product by tracking hub segment adoption in the present.

If the percentage of adopters among hubs reaches a value of θ, at least that fraction of the entire population will eventually adopt the new product. Here we assume that the fraction of potential adopters among hubs is about the same as among non-hubs. The time required for the proportion of adopters in the total population to reach the threshold θ is given by the time required to reach the same θ in the hub segment multiplied by the acceleration factor. It should be noted that, in general, the percentage of potential adopters in the entire population ρ, which affects the acceleration factor, is unknown. To demonstrate this effect, in Figure 4E we present the acceleration factor A as a function of ρ and we define the threshold parameter $\theta = 50\%$. We use a typical product (i.e. $P = 0.03$ and $Q = 0.38$ in annual units), and assume that the mean percentage of hubs among neighbors is $\tilde{f} = 5\%$. The different curves correspond to different values of hub degree factor Γ. It is apparent that the acceleration effect decreases with potential market size.

We can better understand this feature by drawing an analogy between the augmentation of the market potential (while holding threshold θ constant) and the reduction of the threshold θ. Both T_h and T_k, in which the percentage of adopters reaches the threshold θ in the hub and non-hub segment, respectively, are reduced. However, the slope of new product growth in the hub segment is steeper than the corresponding slope in the non-hub segment market (i.e. the adoption curve of the hubs is more rigid compared to the adoption curve of non-hubs). Therefore, as T_h and T_k are defined as the intersection points of the product life cycle threshold θ with the hub and non-hub cumulative adoption curves, respectively, lowering the threshold will result in a more sizeable advance in time of the intersection with the non-hub segment, and the acceleration factor $A = \frac{T_k}{T_h}$ will exhibit a decreasing function. It should be noted that our numerical test shows that the decline is quite moderate. Effectively, from certain point onward, the acceleration factor does not depend on the size of the market potential.

We conclude our study by applying an empirical estimation of the order of magnitude of the acceleration effect in hub segment penetration. Obviously, our data represents an adoption process over a multiple-degree market rather than the two-population model used in our evaluation of the acceleration effect. Nevertheless, we can separately examine the adoption process in each segment wherein all individuals have the same network degree, and compare it to the diffusion process in the entire market, using an appropriate two-population model in which the segment of ordinary individuals comprises the entire population. Thus, for any given segment with network degree h we measured the acceleration factor, A, the ratio of the product life cycle in the entire market T to that in the h^{th} segment, T_h, $A = \frac{T}{T_h}$, and compared to the value retrieved numerically from the two-population model by means of Equations 10-12. The h^{th} segment degree factor is $\Gamma = \frac{h}{k}$ where k is the average network degree in the entire market, and the mean percentage of individuals with network degree h among any individual's neighbors is given by the random network approximation $\tilde{f} \approx \Gamma f$, where f is the proportion of the h^{th} segment in the entire market (see Albert and Barabasi 2002). To characterize our product, we used $P = 0.33$ and $Q = 3.85$ in annual units, which are the results of the OLS estimation using the Bass model in the entire population. Completion of the product life cycle is defined here as the time at which the percentage of adopters in a given segment reaches 85%.

In Figure 5A we plot the acceleration factor versus the segment degree factor. The yellow noisy curve denotes the empirical measures of the acceleration factor; the black curve is the resulted smoothing of the empirical curve, and the blue curve represents the estimation of the acceleration factor provided by the two-population model. Our empirical data set confirms the acceleration effect in the hub segment. When the segment degree factor Γ is greater than 1, the segment network degree is greater than the market average, and the acceleration factor A tends to be greater than 1 as well. Furthermore, despite the presence of noise (the larger Γ, the noisier the data, since segments with a large network degree comprise a small number of individuals), it can be seen that both the empirical and the expected measures of the acceleration factor are of the same order of magnitude.

Despite the similarity in orders of magnitude, it is apparent that our theoretical approximations overestimate the acceleration effect for high values of Γ. Figure 5A implies that the naïve assumption that the greater the number of ties, the more strongly an individual is affected by word-of-mouth communications, which result in an increased tendency to early adoption, does not hold for high degrees. On the contrary, we find that the acceleration effect even decreases from a certain point onward. This phenomenon is more explicitly illustrated in Figure 5B where the segment product life cycle is plotted against its network degree. The yellow noisy curve designates the product life cycle measures, and the black and the blue curves denote the smoothed curve of the empirical data and the model predicted curve, respectively. Around a network degree of 1,500, the diminishing tendency of the product life cycle is replaced by a moderately increasing movement. One explanation for the diminishing acceleration effect among individuals with a large network degree could be capacity constraint. If there are an extremely large number of social ties, it is reasonable to assume that the quality of a hub’s ties with her neighbors is relatively low, and as a result these neighbors are less affected by hubs' word-of-mouth communication. In other words, it seems that people who amass a great number of social ties may neglect some proportion of their ties. This may lead to a more realistic model that assumes that the micro-level coefficient of interaction between two individuals, q, depends to some extent on network degree; Specifically, q can be assumed to be a decreasing function of the degrees of each pair of connected individuals.

Insert Figure 5 about here

Another digression of the empirical findings from the model is reflected in segments that are defined by a low network degree. In Figure 5B we find that the model overestimates product life cycles for low network degree segments, in which we anticipated a slower adoption processes compared with the adoption process of the entire market. Empirical findings, however, indicate that product life cycles of low-degree segments are similar to the entire market product life cycle. We attribute this feature to the externality effect in our data, as discussed earlier. That is, as all individuals are exposed to the total number of the adopters in the market, the dynamics of new product growth becomes increasingly similar over the product life cycle. When the network degree in a certain segment becomes sufficiently large, the local effect becomes enormous and reduces the product life cycle in this particular segment relative to the entire market.

Discussion

We analyze the dynamics of innovation diffusion taking into account the effect of the underlying social network. In particular we develop an analytic framework to better understand the effect of hubs on a diffusion process and identify the boundary conditions under which hubs affect diffusion. This framework is based on a two-stage model where first we analyze individual-level adoptions with social network ties and an external effect, and we aggregate the results to the market level. This analysis allows us to address the effect of hubs through NPV calculations, by considering variations in hub (tie intensity and number of ties), product (innovation and imitation parameters), and consumer (degree distribution) characteristics. Taking these variations into account we quantify the role of hubs in the diffusion process. Using our approach, we utilize the hub segment to develop an early prediction of adoption, and show that the product life cycle in the hub segment is about 2 or 3 times shorter than the product life cycle in the non-hub segment: Thus we are able to explain empirical findings concerning early prediction (Goldenberg et al. 2009b, Katona et al. 2009).

Several managerial implications can be drawn from this study. First, marketers can evaluate the order of magnitude of social hubs economic value as part of viral marketing strategies design. Second, managers can identify cases where their firms may benefit from seeding the market with hubs. Third, optimizing parameters such as the number of hubs and timing of their seeding (assuming there is a cost in identifying and enrolling such highly-connected individuals) can be done. For example, if a firm plans to launch a product with typical diffusion parameters, seeding a small number of hubs (less than 10) that are only 10 times more socially connected than ordinary consumers, adds around 1% to the NPV of the diffusion process. Under the same hub characteristics, but in a case of a product with low external force, a similar seeding generates a significant increase in potential NPV that reaches several tens of percentage points. This incremental NPV estimation method offers alternative resource allocation strategies to managers. Fourth, beyond comparing hub seeding and other marketing efforts (e.g., advertising), managers can estimate the boundaries for their investment in seeding hubs to accelerate the penetration of a new product.

Furthermore, through calculation of the diffusion acceleration factor within social hubs' segments, quantifying the hubs' tendency to become early adopters of new products is possible. Firms can utilize the hubs acceleration factor approximation for early estimation of the entire market potential. Specifically, tracking the consumption behavior of social hubs, marketers can estimate in advance low boundary values of the entire market potential and use the hubs acceleration factor to evaluate the time it will take for the entire market to reach this boundary. These estimations can be used as a baseline for early predictions of new products success or failure.

In this context it should be noted that this paper provides an evaluation for *the order of magnitude* of the differential importance of social hubs in the diffusion of innovation rather than exact estimations for all kinds of network topologies and consumers types. In order to make our study more thorough we considered simple (yet representative) cases. Our analyses are mostly based on classical diffusion dynamics assumptions. For example, we assume market forces that do not vary over time and identical for all individuals with the same network degree. We also approximate the impact of consumers' environment on her decision to adopt by the average interaction generated by the representative sample of actual adopters among individuals' neighbors at any given time. That is, we do not consider local effects as well as stochastic effects in the adoption process itself that create the variance in those interactions. Therefore, further research can explore the case of networks with high clustering coefficient and consumer type heterogeneity while taking into account the stochastic nature of the course of adoption. Addressing the general case of repeat-purchase products (rather than new product penetration alone) would also provide further insight about the role of hubs in a diffusion process.

References

Achrol, Ravi S. and Philip Kotler (1999), "Marketing in the Network Economy," Journal of Marketing, 63, 146-163
.

Albert, Reka and Alber Laszio Barabasi (2002), "Statistical mechanics of complex networks," Reviews of Modern Physics, 74 (1), 47-97.

Barabasi, Alber Laszio and Reka Albert (1999), "Emergence of Scaling in Random Networks," Science, 286 (5439), 509-512
.

Barabasi, Alber Lazsio (2003), Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life: Plume Books.

Bogacki, Przemyslaw and Lawrence F. Shampine (1989), "A 3 (2) pair of Runge-Kutta formulas," Appl. Math. Letters, 2, 1-9.

Brown, Jacqueline Johanson and Peter H. Reingen (1987), "Social Ties and Word-of-Mouth Referral Behavior," The Journal of Consumer Research, 14 (3), 350-62.

Burt, Ronald S. (1997), "The Contingent Value of Social Capital," Administrative science quarterly, 42 (2), 339-65.

Davis, Philip J. and Philip Rabinowitz (1984), Methods of numerical integration. Orlando: Academic Press.

Dover, Yaniv, Jacob Goldenberg and Daniel Shapira (2009), "Uncovering social network structures through penetration data", presented in the INFORMS Marketing Science Conference, Ann Arbor, Michigan.

Goldenberg, Jacob, Barak Libai, Sarit Moldovan, and Eitan Muller (2007), "The NPV of bad news," International Journal of Research in Marketing, 24 (3), 186-200

Goldenberg, Jacob, Oded Lowengart, and Daniel Shapira (2009a), "Zooming In: Self-Emergence of Movements in New Product Growth," Marketing Science, 28(2), 274-292.

Goldenberg, Jacob, Han Sangman, Donald R. Lehmann, and Jae W. Hong (2009b), "The role of hubs in the adoption process," Journal of Marketing, 73 (2), 1-13.

Goldenberg, Jacob, Yaniv Shavitt, Eran Shir, and Sorin Solomon (2005), "Distributive immunization of networks against viruses using the "honey-pot" architecture," Nature Physics 1(3), 184-88.

Hogan, John E., Katherine N. Lemon, and Barak Libai (2005), "Quantifying the Ripple: Word-of-Mouth and Advertising Effectiveness," Journal of Advertising Research, 44(3), 271-280.

Houston, Mark, Mike Hutt, Christine Moorman, Peter H. Reingen, Aric Rindfleisch, Vanitha Swaminathan, and Beth Walker (2004), "Marketing Networks and Firm Performance," Assessing Marketing Strategy Performance, 227-268
.

Iacobucci, Dawn (1996), Networks in Marketing: Sage Publications.

Iyengar, Raghuram, Thomas Valente, and Christophe van Den Bulte (2008), "Opinion Leadership and Social Contagion in New Product Diffusion," working paper, University of Pennsylvaia.

Katona, Zsolt., Peter. Zubcsek, and Miklos. Sarvary (2009), "Network Effects and Personal Influences: Diffusion of an Online Social Network". Working paper.

Katz, Elihu and Paul Felix Lazarsfeld (1955), Personal influence; the part played by people in the flow of mass communications. Glencoe, Ill.: Columbia University. Bureau of Applied Social, Research ,Free Press.

Keller, Edward B. and Jonathan L. Berry (2003), The influentials. New York: Free Press.
Lehmann, Donald and Mercedes Esteban-Bravo (2006), "When giving some away makes sense to jump-start the diffusion process," Marketing Letters, 17 (4), 243-54.

Mayzlin, Dina (2002), "The Influence of Social Networks on the Effectiveness of Promotional Strategies," Yale School of Management Working Paper.

Muller, Eitan and Guy Yogev (2006), "When does the majority become a majority? Empirical analysis of the time at which main market adopters purchase the bulk of our sales,"

Technological Forecasting and Social Change, 73 (9), 1107-1120

Polyanin, Andrei and Valentin F. Zaitsev (2003), Handbook of exact solutions for ordinary differential equations. Boca Raton: Chapman & Hall/CRC.

Rindfleisch, Aric and Christine Moorman (2001), "The Acquisition and Utilization of Information in New Product Alliances: A Strength-of-Ties Perspective," Journal of Marketing, 65(2), 1-18.

Rogers, Everett M. (1963), "The adoption process: Part I [Electronic version]," Journal of Extension, 1 (1), 16-22.

Ronchetto, John R., Michael D. Hutt, and Peter H. Reingen (1989), "Embedded Influence Patterns in Organizational Buying Systems," The Journal of Marketing, 53 (4), 51-62

Rosen, Emmanuel (2002), The Anatomy of Buzz: How to Create Word-Of-Mouth Marketing: Doubleday Business.

Shaikh, Nazrul. I., Arvind Rangaswamy and Anant Balakrishnan (2006), "Modeling the Diffusion of Innovations Through Small-World Networks". Working paper.

Shi, Mengze (2003), "Social Network-Based Discriminatory Pricing Strategy," Marketing Letters, 14 (4), 239-256.

Steffens, Paul R. and Deeksha N. P. Murthy (1992), "A mathematical model for new product diffusion: The influence of innovators and imitators," Mathematical and Computer Modeling, 16 (4), 11-26.

Stephen, Andrew T. and Olivier Toubia (2009a), " Deriving Value from Social Commerce Networks " Journal of marketing research, (forthcoming).

Stephen, Andrew T. and Olivier Toubia (2009b), " Explaining the Power-Law Degree Distribution in a Social Commerce Network" Social Networks, 31(4), 262-270.

Sultan, Fareena, John U. Farley, and Donald R. Lehmann (1990), "A Meta-Analysis of Applications of Diffusion Models," Journal of Marketing Research, 27 (1), 70-77.

Trusov, Michael, Anand V. Bodapati, and Randolph E. Bucklin (2008), "Determining Influential Users in Internet Social Networks," working paper.

Trusov, Michael and William Rand (2009), “Identifying Network Properties from Aggregate Data,” presented in the INFORMS Marketing Science Conference, Ann Arbor, Michigan.

Valente, Thomas W. (1995), Network models of the diffusion of innovations: Hampton Press, Cresskill, NJ.

Van den Bulte, Christophe and Yogeh V. Joshi (2007), "New Product Diffusion with Influentials and Imitators," Marketing Science, 26 (3), 400-21.

Van den Bulte, Christophe and Stefan Wuyts (2007), Social Networks and Marketing: Marketing Science Institute.

Watts, Duncan J. (2002), " A simple model of global cascades on random networks " Proceedings of the National Academy of Sciences (PNAS), 99 (9), 5766-71.

Watts, Duncan J. and Peter S. Dodds (2007), "Influentials, Networks, and Public Opinion Formation," Journal of Consumer Research, 34 (4), 441.

Weimann, Gabriel (1994), The influentials : people who influence people. Albany: State University of New York Press.

Notes

1 Hubs are excluded from this segmentation since they are activated exogenously.

2 Considering a market comprising several hundred thousand consumers, in which the hubs' number of neighbors is about 10 times the ordinary individuals' number of neighbors so h lies between 100 and 1000 the hub area of influence $\chi = \frac{h}{M}$ represents less than 1% of the market.

3 Non-hubs differ from conventional hubs merely in their area of influence, which is about $\frac{h}{k}$ times smaller than hubs' area of influence. . In our calculations we assumed that hubs have ten times the number of ties of non-hubs, namely $\frac{h}{k} \sim 10$.

4 The increase is convex because it has an upper bound determined by the maximal NPV increment accomplished when the entire market adopts at the hub seeding time (at the limit of an extremely large v and $\chi = 100\%$).

5 While there are studies that are based on the assumption that the number of individuals that affects influentials opinion is less than the number of individuals they have impact on (Van den Bulte and Joshi 2007) other studies find empirical evidence for early innovation adoption in positive correlation with the number of individuals' adopting neighbors (Katona et al 2009) and network degree (Goldenberg et al. 2009b). Therefore we consider the conservative assumption of symmetric underlying social network.

6 In Appendix C we show that if all individuals have the same network degree (a homogeneous market), the dynamics of new product growth (described by Equations 6-8) is reduced to a special case of the Bass model.

7 Because our empirical data provides the total number of individuals for each network degree k who will eventually adopt, we do not have to estimate the fraction of the market potential ρ.

8 This assumption may not hold when the maximal degree value is close to the size of the whole population. For example, if there is an individual with network degree $M-1$ where M is the population size., any other individual will be her neighbor with probability 1. Hence in case she adopts the new product her contribution to the proportion of actual adopters among the neighbors of another individual with network degree k is $\frac{1}{k}$. That is, her contribution to each sub-market is different. Thus we also assume that the order of magnitude of the maximal degree value is much less than the size of the population so that such "surface effects" become negligible.

Table 1: Relative Change in NPV under Various Combination of Seeded Hub Tie Intensity and Area of Influence

Product Characteristics	Intensity Number of Ties		
	Low intensity, few ties	High intensity, few ties	Low intensity, many ties
	$v=\frac{Q}{10}$, $\chi=1\%$	$v=Q$, $\chi=1\%$	$v=\frac{Q}{10}$, $\chi=10\%$
Typical product $P=0.03$, $Q=0.38$ (annual units)	0.2%	1.4%	2.3%
Product with low external influence $P=0.001$, $Q=0.38$ (annual units)	7.6%	30.1%	44.0%

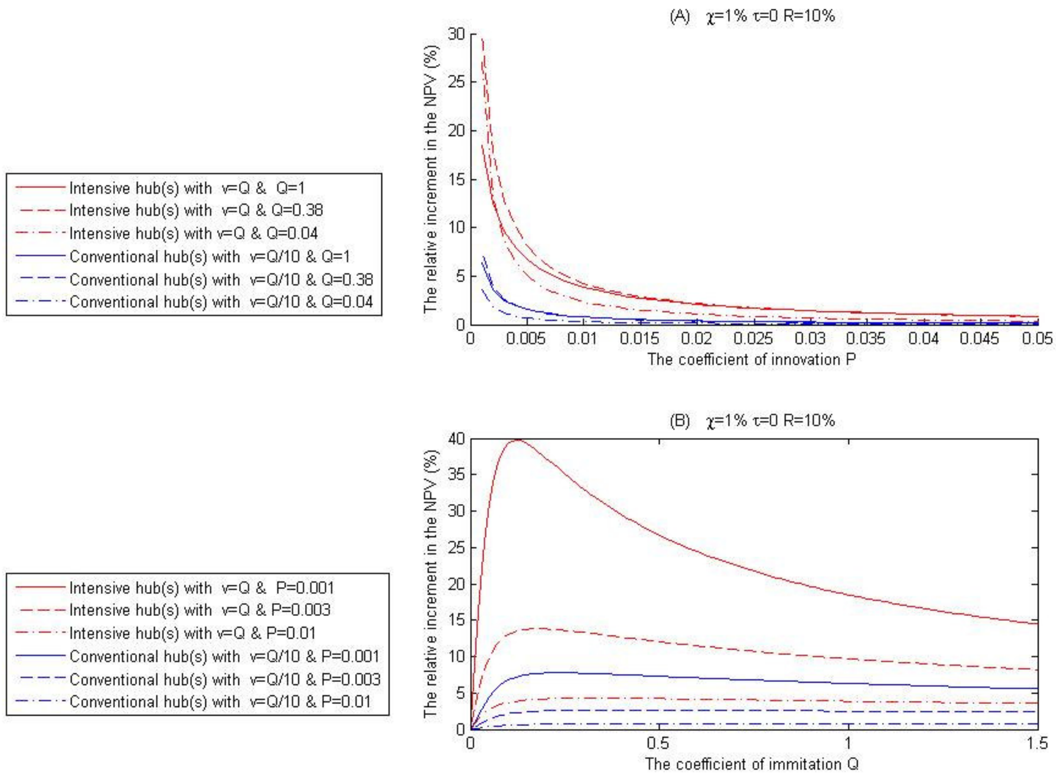

***Figure 1*.** The relative increments in the NPV of the diffusion process resulting from hub seeding for products characterized by different combinations of diffusion coefficients P and Q . The blue curve corresponds to conventional hubs, while the red curves represent intensive hubs.

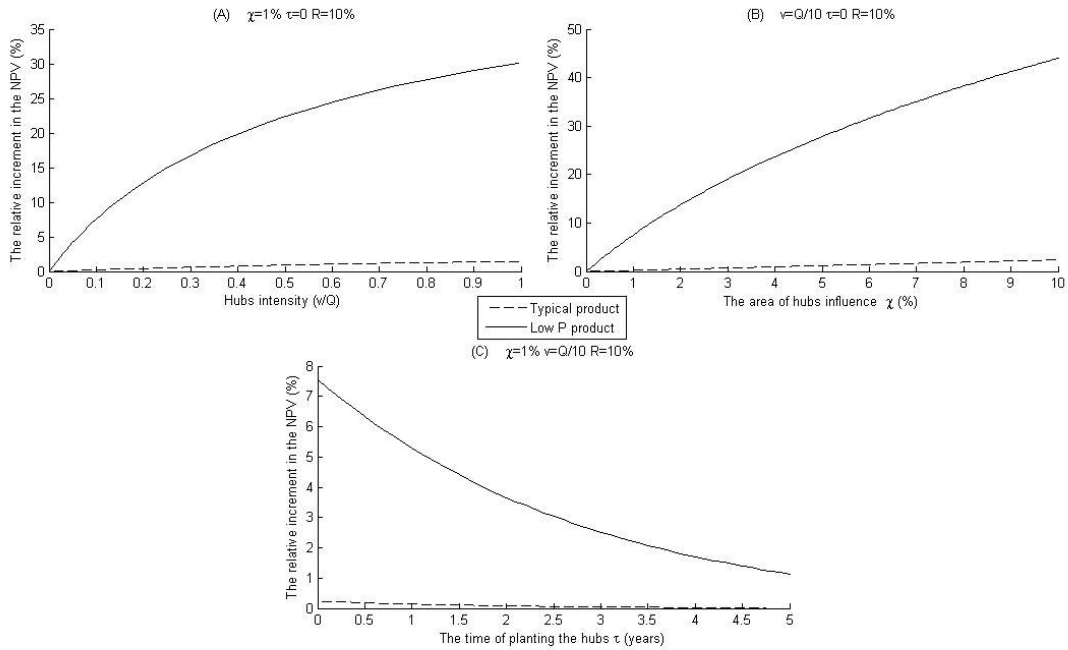

Figure 2. The effect of seeded hub attributes: (A) intensity of interactions (B) area of influence, and (C) timing of the hub seeding on the relative increment of the NPV. In our demonstration we consider two representative penetration processes: A typical penetration process with $P = 0.03$ and, $Q = 0.38$ and a process with low external influence with $P = 0.001$ and $Q = 0.38$.

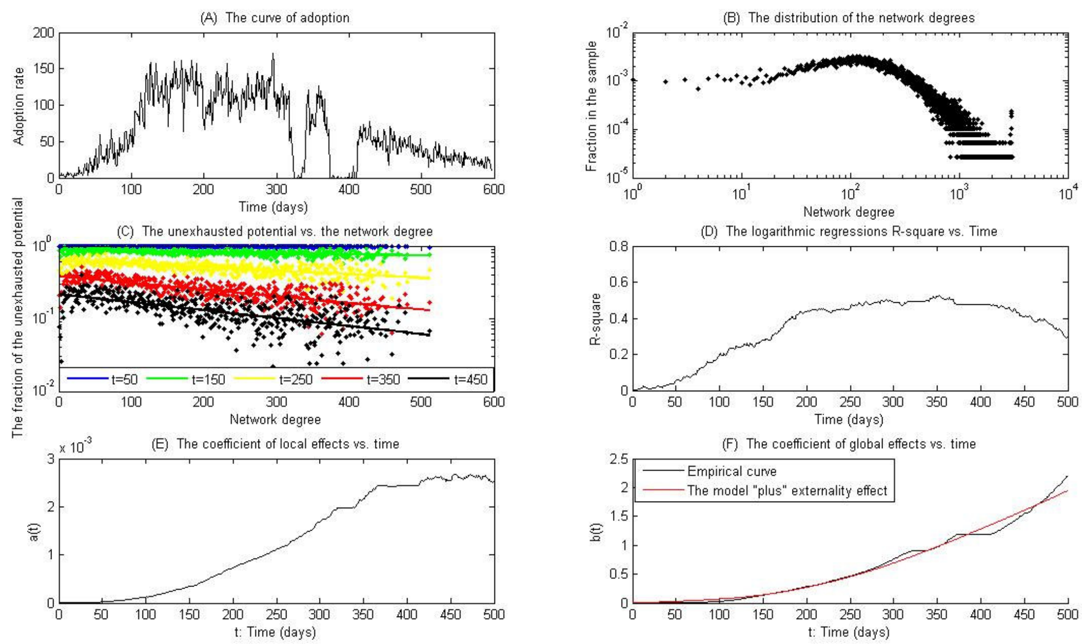

Figure 3. The dynamics of new product growth in a heterogeneous market (empirical results). The data set consists of a group of 37,704 individuals who are members of an online interest group, and can be considered a network. This figure presents: (A) the curve of group membership adoption as a function of time; (B) the degree distribution of the network; (C) the market potential percentage as a function of the network degree at five different points in time (straight lines denote the logarithmic regression fits that are based on the data points); (D) the R^2 measure of the goodness of fit of the empirical data to the theoretical model as a function of time; (E) the coefficient of local influence as a function of time; and (F) the coefficient of global influence as a function of time.

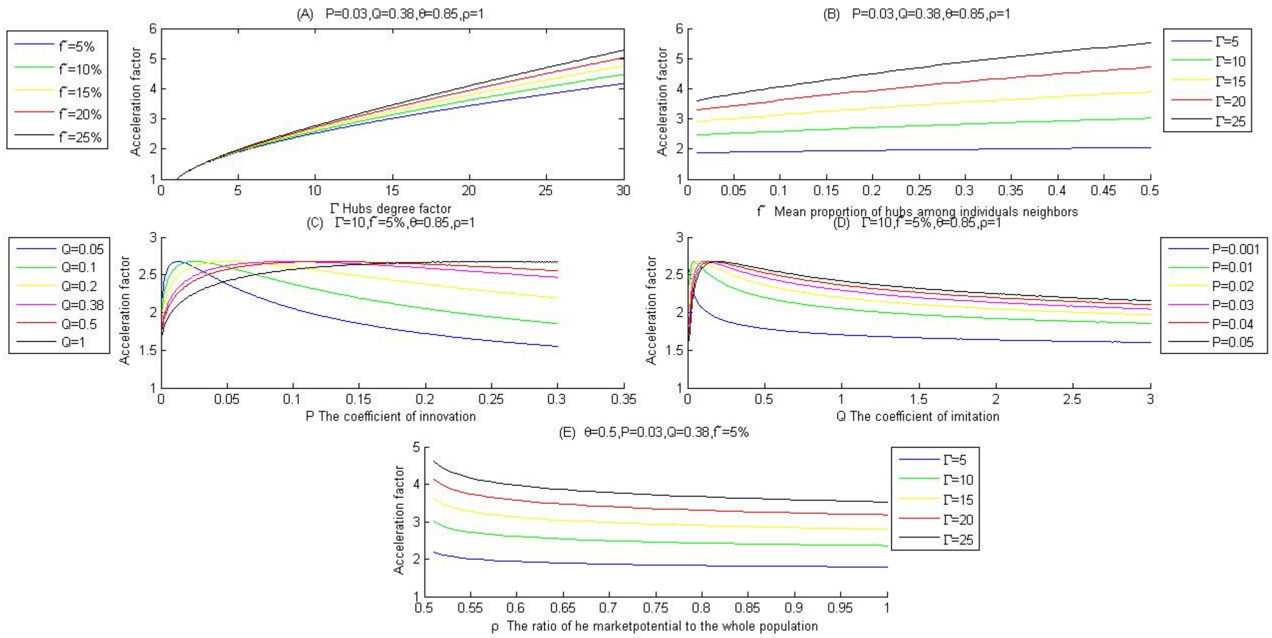

Figure 4.The acceleration effect in a two-population model, measured by the product life cycle acceleration factor in various conditions: (A and B) different combinations of hub degree factor and mean percentage of hubs among individuals' neighbors, considering a typical product with $P = 0.03$ and $Q = 0.38$ in annual units; (C and D) various classes of products distinguished by combinations of coefficients of diffusion P and Q, where hub degree factor is $\Gamma = 10$ and mean percentage of hubs among individuals' neighbors is $\tilde{f} = 5\%$, and (E) various proportions of potential adopters in the entire population for a typical product with $P = 0.03$ and $Q = 0.38$ in annual units, where the threshold for product life cycle completion is $\theta = 50\%$.

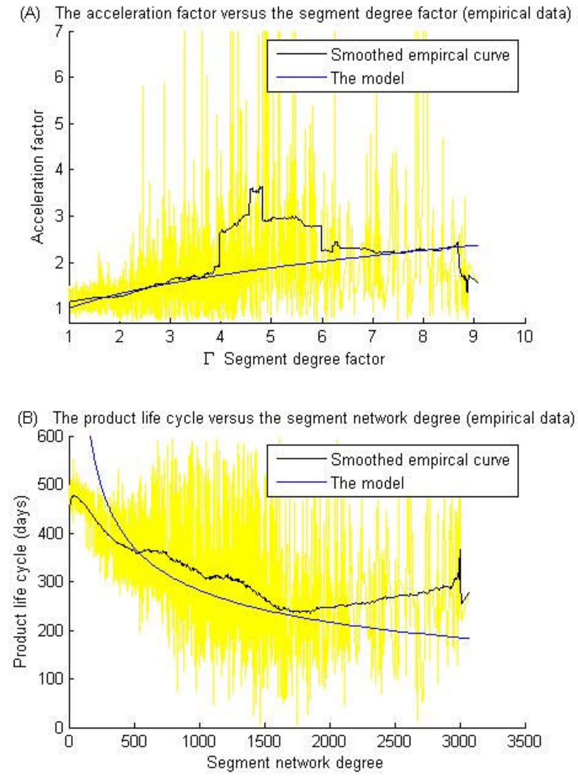

Figure 5. Empirical estimation of the order of magnitude of the acceleration effect in hub segment adoption: (A) the acceleration factor as a function of the hub degree factor Γ and (B) segment product life cycle as a function of segment network degree.

Appendices

A) Integrating the social network to diffusion model – an analytical framework

Our formulation is based on the idea that individual behavior is probabilistic and can be summed to represent the entire market response and dynamics (Goldenberg et al. 2009a). Along these lines and unlike aggregate approaches, our unit of analysis is the individual consumer.

Let s_i be a binary variable that represents the state of adoption of a potential adopter i. That is, $s_i(t)$ takes the value 1 if individual i adopted the innovation before time t and 0 otherwise, so the transition index for individual i who purchases the innovation within a short time interval Δt after time t is given by $\Delta s_i(t) = s_i(t+\Delta t) - s_i(t)$. For each individual i we define $prob(\Delta s_i(t)=1)$, which is the transition probability that individual i will adopt the innovation within a short time interval, Δt, after time t. These probabilities are conditional and depend on the relevant history of the entire market dynamics (i.e., time dependent). If we consider the common assumption of independent market forces, the transition probability for individual i to adopt the innovation in a time interval between time indices t and $t+\Delta t$ is given by:

(A1)

$$prob(\Delta s_i(t)=1) = (1-s_i(t))\left(1-(1-p_i(t)\Delta t)\prod_{j=1}^{M}(1-q_{ij}(t)\Delta t\right) \rightarrow (1-s_i(t))\left(p_i(t)+\sum_{j=1}^{M} q_{ij}(t)\right)\Delta t$$

The impact of the external force (e.g., marketing efforts) on individual i is given by $p_i(t)$, which is the probability per unit of time to be persuaded by external influences to adopt the new product. Internal forces are given by the terms $q_{ij}(t)$, where $q_{ij}(t)$ is the probability per unit of time that individual i is persuaded to purchase the innovation on the basis of word-of-mouth communications originating from j. The term q_{ij} encapsulates the underlying social network architecture and expresses the interactions among its nodes. Market potential is denoted by M. The multiplication term $1-s_i(t)$ guarantees that if i is an actual adopter of the new product, its transition probability of adoption $prob(\Delta s_i(t)=1)$ is nullified (there is no repeat purchase).

We are interested in the penetration dynamics of a new product into consumer segment S in a market. Therefore, the aggregate process is $\Delta N_S(t) \equiv \sum_{i|i\in S} \Delta s_i(t)$, the number of individuals in segment S who adopt the innovation between time t and $t+\Delta t$. The non-cumulative penetration rate $\Delta N_S(t)$ is given by a sum of stochastic variables that can be described by the following equation (see also Goldenberg et al. 2009a):

(A2) $$\Delta N_S(t) = F_S(t)\Delta t + \varepsilon_S(t)$$

where:

(A3) $$F_S(t) = \frac{1}{\Delta t}\sum_{i|i\in S} prob(\Delta s_i(t)=1)$$

is the net market force applied to the entire segment S. Note that the transition probabilities $prob(\Delta s_i(t)=1)$ that appear in Equation A3 are determined by individual-level market forces. The noise in this process is denoted by $\varepsilon_S(t)$.

Next, we evaluate the dynamics of $\mu_S(t)$, the proportion of actual adopters in market segment S at time t. Omitting the noise term in Equation A2 and setting the problem to the continuous limit while dividing both sides by M_S, the segment mass (population size) yields the following (dynamical) equation:

(A4) $$\frac{d\mu_S(t)}{dt} = \frac{F_S(t)}{M_S}.$$

B) The dynamics of new product growth in the presence of hub in the case of controlling the hub activation time.

In this appendix we develop the dynamical model of diffusion in the presence of hub where the time of the hub activation is controlled exogenously. Let us first consider a market that consists just of ordinary consumers. In order to make our study more thorough let us assume that the external influence and the intensity of the word-of-mouth communications among individuals do not vary with time and are identical for all consumers, and also assume that only actual adopters of the innovation apply word-of-mouth communications. In this case Equation A1 (see the previous appendix) takes the form:

(B1) $$prob(\Delta s_i(t)=1)=(1-s_i(t))\left(p+q\sum_{j=1}^{M}k_{ij}s_j(t)\right)\Delta t$$

where p and q are the individual-level coefficients of the external and internal market force respectively. The matrix k_{ij} denotes the social networks' topology where $k_{ij}=1$ if and only if individuals i and j are socially related and $k_{ij}=0$ otherwise. (More explicitly, $p_i(t)=p$ and $q_{ij}(t)=qk_{uj}s_j(t)$ where by definition $k_{ii}=0$.) To facilitate our analysis let us assume that k is a typical number of an ordinary individuals' social ties and also assume that each potential adopter is exposed to the same "representative sample" of adopters. Thus the proportion of actual adopters of the new product among any individual's neighbors can be estimated by the proportion of actual adopters in the entire market such that $\frac{1}{k}\sum_{j=1}^{M}k_{ij}s_j(t)\sim\frac{N(t)}{M}$. Here $N(t)=\sum_{j=1}^{M}s_j(t)$ denotes the total number of actual adopters of the innovation at the time t and M is the market potential. This approximation may be considered as "mean field approximation" in the sense that local effects that make interactions among individuals to be different from one another are neglected, assuming that effectively each individual 'feels" an average word-of –mouth effect as if it is coming from the entire population.

Now, suppose that at a certain time τ a hub is seeded to apply word-of-mouth communication on her environment. The hub is characterized by large number of social ties h ($h >> k$) where the hubs' influence on his peers is quantified by an intensity coefficient v. Namely, following the hub seeding $v\Delta t$ denotes the probability that a potential adopter who is socially related to the hub will be persuaded by the hubs' word-of-mouth to adopt the innovation within a time period Δt after the time t. As a result, after the hub activation the market is divided into two segments: the segment of hubs' neighbors S and the complementary segment $\overline{S}$ that embraces all the individuals who are not directly connected to the hub. Hence Equation B1 can be rewritten as follows:

(B2) $$prob(\Delta s_i(t)=1)=\begin{cases}(1-s_i(t)\left(p+v+kq\dfrac{N(t)}{M}\right)\Delta t & if \quad i\in S \,\&\, t\geq\tau\\ (1-s_i(t))\left(p+kq\dfrac{N(t)}{M}\right)\Delta t & Otherwise.\end{cases}$$

Along the lines drafted in Appendix A, we calculate the net market forces $F_S(t)$ and $F_{\overline{S}}(t)$ applied on the segments S and $\overline{S}$ respectively by substituting the expressions given in Equation B2 in Equation A3 to give:

(B3) $$F_S(t)=\begin{cases}(M_S-N_S(t))\left(p+kq\dfrac{N(t)}{M}\right) & if \quad t<\tau\\ (M_S-N_S(t))\left(p+v+kq\dfrac{N(t)}{M}\right) & if \quad t\geq\tau\end{cases}$$

(B4) $$F_{\overline{S}}(t)=(M_{\overline{S}}-N_{\overline{S}}(t))\left(p+kq\frac{N(t)}{M}\right)$$

where $M_S=h$ and $M_{\overline{S}}=M-h$ denote the size of the population in the segments S and $\overline{S}$ respectively and $N_S(t)$ and $N_{\overline{S}}(t)$ are the cumulative adoptions in these segments (where $N(t)=N_S(t)+N_{\overline{S}}(t)$). Next, the aggregate-level dynamics of diffusion in both segments is derived using Equations B3 and B4 in the framework presented in Equation A4 to obtain:

(B5) $$\frac{d\mu_S(t)}{dt}=\frac{F_S(t)}{M_S}=\begin{cases}(1-\mu_S(t))(P+Q\mu(t)) & if \quad t<\tau\\ (1-\mu_S(t))(P+v+Q\mu(t)) & if \quad t\geq\tau\end{cases}$$

(B6) $$\frac{d\mu_{\bar{S}}(t)}{dt} = \frac{F_{\bar{S}}(t)}{M_{\bar{S}}} = (1-\mu_{\bar{S}}(t))(P+Q\mu(t))$$

where $\mu_S(t) = \frac{N_S(t)}{M_S}$ and $\mu_{\bar{S}}(t) = \frac{N_{\bar{S}}(t)}{M_{\bar{S}}}$ are the proportions of actual adopters in the segments S and $\bar{S}$ respectively and $P = p$ and $Q = kq$ are the aggregate-level coefficients of the diffusion. Equations B5 and B6 are coupled through the proportion of actual adopters in the entire market:

(B7) $$\mu(t) = \frac{N(t)}{M} = \chi\mu_S(t) + (1-\chi)\mu_{\bar{S}}(t)$$

where $\chi = \frac{M_S}{M} = \frac{h}{M}$ is the hub area of influence and the initial conditions are

$\mu_S(t=0) = \mu_{\bar{S}}(t=0) = 0$, where $t = 0$ is the new product launch time.

Substitution of Equations B5 and B6 in the temporal derivative of Equation B7 in the case of $t < \tau$ yields a penetration process that evolves according to the Bass dynamics so that:

(B8) $$\frac{d\mu(t)}{dt} = \chi\frac{d\mu_S(t)}{dt} + (1-\chi)\frac{d\mu_{\bar{S}}(t)}{dt} = (1-\mu(t))(P+Q\mu(t))$$

and thus:

(B9) $$\mu(t) = \mu_S(t) = \mu_{\bar{S}}(t) = \frac{1-e^{-(P+Q)t}}{1+\frac{Q}{P}e^{-(P+Q)t}}$$

On the other hand, in the case of $t \geq \tau$ (i.e. after the time of hub seeding) Equations B5 and B6 take the form:

(B10) $$\frac{d\mu_S(t)}{dt} = (1-\mu_S(t))(P+v+Q\mu(t))$$

(B11) $$\frac{d\mu_{\bar{S}}(t)}{dt} = (1-\mu_{\bar{S}}(t))(P+Q\mu(t))$$

with the initial conditions $\mu(t=\tau) = \mu_S(t=\tau) = \mu_{\bar{S}}(t=\tau) \equiv \mu_\tau$ where $\mu_\tau = \frac{1-e^{-(P+Q)\tau}}{1+\frac{Q}{P}e^{-(P+Q)\tau}}$.

Namely,

(B12) $$\frac{\dot{\mu}_{\bar{S}}(t)}{1-\mu_{\bar{S}}(t)} = \frac{\dot{\mu}_S(t)}{1-\mu_S(t)} - v$$

where $\dot{\mu}_S(t)=\frac{d\mu_S(t)}{dt}$ and $\dot{\mu}_{\bar{S}}(t)=\frac{d\mu_{\bar{S}}(t)}{dt}$. Hence, integrating Equation B12 in time from τ to any given time t we obtain that $1-\mu_{\bar{S}}(t)=(1-\mu_S(t))e^{v(t-\tau)}$. Thus, from Equation B7 it follows that: $1-\mu(t)=\chi(1-\mu_S(t))+(1-\chi)(1-\mu_{\bar{S}}(t)=(\chi+(1-\chi)e^{v(t-\tau)})(1-\mu_S(t))$ so that substitution of Equations B10 and B11 in the temporal derivative of Equation B7 yields:

(B13)
$$\begin{aligned}\frac{d\mu(t)}{dt} &= \chi\frac{d\mu_S(t)}{dt}+(1-\chi)\frac{d\mu_{\bar{S}}(t)}{dt}\\ &= (1-\mu(t))(P+Q\mu(t))+v\chi(1-\mu_S(t))\\ &= (1-\mu(t))(P+\frac{v\chi}{\chi+(1-\chi)e^{v(t-\tau)}}+Q\mu(t))\end{aligned}$$

which is a special case of of the *Ricatti Equation* of the form:

(B14)
$$\frac{dx}{dt}=x^2+f(t)x-Q^2+Qf(t)$$

where $x(t)=-Q\mu(t)$ and $f(t)=Q-P-\frac{v\chi}{\chi+(1-\chi)e^{v(t-\tau)}}$ resulting in a general closed-form solution (of the Equation B14) that can be written as follows (Polyanin and Zaitsev 2003):

(B15)
$$x(t)=-Q+\frac{\Phi(t)}{c-\int\Phi(t)dt}$$

where c is a constant determined by the initial condition and:

(B16)
$$\Phi(t)=\exp(-2Qt+\int f(t)dt).$$

The functions $\int f(t)dt$ and $\int\Phi(t)dt$ are the primitive functions of $f(t)$ and $\Phi(t)$ respectively.

These functions are given by the following indefinite integrals' solutions:

(B17)
$$\begin{aligned}\int f(t)dt &= (Q-P)t-\int\frac{v\chi}{\chi+(1-\chi)e^{v(t-\tau)}}dt\\ &= (Q-P-v)t+\ln\left(1+\frac{1-\chi}{\chi}e^{v(t-\tau)}\right)\end{aligned}$$

and

(B18)
$$\begin{aligned}\int\Phi(t)dt &= \int\exp\left(-(Q-P-v)t+\ln\left(1+\frac{1-\chi}{\chi}e^{v(t-\tau)}\right)\right)dt\\ &= -\left(\frac{1}{Q+P+v}+\frac{(1-\chi)e^{v(t-\tau)}}{\chi(Q+P)}\right)e^{-(Q+P+v)t}\end{aligned}$$

At this point we can substitute Equation B17 in Equation B16 and then substitute Equations B16 and B17 in Equation B15 to derive an explicit solution of $x(t)$ leading to a closed-form solution for the proportion of actual adopters in the entire market $\mu(t)$ at any given time $t \geq \tau$ (recall that $\mu(t) = -\frac{x(t)}{Q}$.) The exact value of the constant c is extracted from the initial condition equation $\mu(t = \tau) = \mu_\tau$. The closed form solution for $\mu(t)$ in the case of $t \geq \tau$ is given in Equation 3.

These results can also be applied to the case where group of hubs (instead of a single hub) is seeded simultaneously at a certain time τ to apply word-of-mouth communication on their environment. Since in general the hubs areas of influence partially overlap the market is divided to more than two segments following the hubs activation. That is, we should define segment for each overlapping configuration. Let s denotes the segment of individuals who are neighbors of exactly s seeded hubs where χ_s is the proportion of that segment in the entire market. Thus assuming that the intensity of influence of s hubs is v_s , the dynamics of the proportion of actual adopters in the s th segment is obtained using similar considerations to those that yield Equation B10 to give:

(B19) $$\frac{d\mu_s(t)}{dt} = (1 - \mu_s(t))(P + v_s + Q\mu(t))$$

where the proportion of actual adopters in the entire market is given by $\mu(t) = \sum_s \chi_s \mu_s(t)$.

Namely, $\mu_s(t) = 1 - (1 - \mu_\tau) e^{-[(P+v_s)(t-\tau) + Q\int_\tau^t \mu(t')dt']}$ and hence

(B20) $$\frac{d\mu(t)}{dt} = (1 - \mu(t)) \left(P + \frac{\sum_s \chi_s v_s e^{-v_s(t-\tau)}}{\sum_s \chi_s e^{-v_s(t-\tau)}} + Q\mu(t) \right).$$

Now by taking the average impact value as an approximation of the seeded hubs impact v_s on each segment s we turn Equation B20 to the form of Equation B13 where

$\chi = \sum_{s|s \neq 0} \chi_s$ is the total seeded hubs' area of influence (which is the share of the potential market that is directly exposed to the influence of at least one seeded hub) and

$v = \frac{1}{\chi} \sum_{s|s \neq 0} \chi_s v_s$ is the average impact that seeded hubs have on potential adopter in their area of influence.

C) The dynamics of new product growth in a multiple degree market

In this appendix we develop the dynamical model of diffusion in a heterogeneous market that consists of a certain network degree distribution. We define the state of adoption of an individual i at the time t by $s_i(t)$ and indicate the topology of the social network by the matrix k_{ij} (where by definition $k_{ii} = 0$ for any individual i) . We assume that the external influence p and the intensity of word-of-mouth communications among individuals q are time independent and are identical for all the potential consumers. Using a binary index r_i to denote whether an individual i is included in the market potential, Equation A1 of the transition probability for individual i to adopt the innovation in a time interval between time indices t and $t + \Delta t$ (see Appendix A) takes the form:

(C1) $$prob(\Delta s_i(t) = 1) = r_i (1 - s_i(t)) \left(p + q \sum_{j=1}^{M} k_{ij} s_j(t) \right) \Delta t$$

where M is the population size. In our analysis we consider a market that comprises of distribution of sub-markets in accordance with the consumers' social network degrees. Namely, all the consumers within a certain sub-market hold the same number of social ties. In our next step we develop the dynamical equation of diffusion for each sub-market.

Let $N_k(t) = \sum_{i|d_i = k} s_i(t)$ denote the total number of actual adopters at the time t within sub-market k where $d_i = \sum_{j=1}^{M} k_{ij}$ is the network degree of the individual i (note that by definition also $N_k(t) = \sum_{i|d_i=k} r_i s_i(t)$); M_k be the k^{th} sub-market size and also suppose that the chance of being in the market potential is equal for all individuals so that $\Pr ob(r_i = 1) = \rho$ for any individual i and hence $\sum_{i|d_i=k} r_i \approx \rho M_k$. Now we can estimate the applied net market force on the k^{th} sub-market at any given time t according to Equation A3 as follows:

(C2)
$$\begin{aligned} F_k(t) &= p\sum_{i|d_i=k} r_i(1-s_i(t)) + q\sum_{i|d_i=k} r_i(1-s_i(t))\sum_{j=1}^{M} k_{ij}s_j(t) \\ &\approx p(\rho M_k - N_k(t)) + q(\rho M_k - N_k(t))kg_k(t) \end{aligned}$$

where

(C3)
$$g_k(t) = \frac{1}{\rho M_k - N_k(t)} \sum_{i|d_i=k} \frac{1}{k} \sum_{j=1}^{M} k_{ij}s_j(t)$$

is the average percentage of actual adopters among potential adopters' neighbors in the k^{th} sub-market. Similarly to the analysis in the previous appendix we assume that on average, people are exposed to a "representative sample" of actual adopters in the sense that the average proportion of actual adopters among individuals' neighbors does not depend on the individuals' network degree and is hence the same in all the sub-markets[8]. Namely for all the sub-markets $g_k(t) \equiv g(t)$. (The sizes of the samples are of course not the same as they are determined by the sub-markets' degrees.) However, unlike the homogeneous case (where all individuals have about the same network degree) here we can not identify the fraction of adopters among an individual's neighbors with the percentage of the new product adopters in the entire population. The proportion of consumers with large number of social ties is larger than their relative weight in the entire population. That is, the chance of being socially related to hub is much larger than the percentage of hubs in the market. Therefore, we approximate the average proportion of actual adopters among potential adopters' neighbors as a weighted average of the form:

(C4)
$$g(t) = \sum_{l=l_{\min}}^{l_{\max}} \tilde{f}_l \mu_l(t)$$

where $\mu_l(t) = \frac{N_l(t)}{M_l}$ denotes the proportion of actual adopters with network degree l and $\tilde{f}_l$ is the mean percentage of an individual's neighbors with degree l. (As mentioned above $\tilde{f}_l$ is in general <u>not</u> given by f_l the proportion of individuals with network degree l in the entire population.) Now we identify $g(t)$ with $g_k(t)$ while substituting Equation C4 in Equation C2 for each sub-market k to obtain

(C5) $$F_k(t) = (\rho M_k - N_k(t))(p + kqg(t))$$

and thus we can write the following coupled system of first order Ordinary Differential Equations:

(C6) $$\frac{d\mu_k(t)}{dt} = \frac{F_k(t)}{M_k} = (\rho - \mu_k(t))(p + kqg(t))$$

compliant with Equation A4. Define $t = 0$ to be the new product launch time, the initial conditions of the above system are $\mu_k(t = 0) = 0$ for each sub-market k. As can be seen, the equation system C6 is coupled through the function $g(t)$ (defined by Equation C4). This coupling can be removed if one divides both sides of Equation C6 for each sub-market k by $\rho - \mu_k(t)$, integrate in time from 0 to t and then take an exponent to obtain:

(C7) $$\mu_k(t) = \rho\left(1 - e^{-(pt + kqG(t))}\right)$$

where $G(t) = \int_0^t g(t')dt'$ is a generating function of all the solutions $\mu_k(t)$. Substituting the solutions from Equation C7 back in Equation C4 we find that the generating function satisfies the first order differential equation that appears in Equation 7. The total proportion of actual adopters in the entire market is given by a weighted sum of the fractions of actual adopters in all the sub-markets as indicated by Equation 8.

Similar patterns of adoption can also be found in cases where the underlying network is not symmetric. Let $q_{kl} = q(k) \cdot w(l)$ be the impact on potential adopter with network degree k

generated by her neighbor who has network degree l and already adopted the innovation Here $q(k)$ denotes the average impact per neighbor on potential adopter with network degree k and $w(l)$ is the relative impact generated by potential adopters' neighbor with network degree l where the normalization condition $\sum_{l=l_{\min}}^{l_{\max}} \tilde{f}_l w(l) = 1$ is satisfied. (The reduction to the symmetric network is the special case where $q(k) \equiv q$ and $w(l) \equiv 1$ for each pair of network degrees k and l.) In the asymmetric network case Equation C7 which expresses the percentage of adopters among individuals with network degree k at any given time t turns to the form $\mu_k(t) = \rho(1 - e^{-(pt + kq(k)G(t))})$ where the temporal derivative of the generating function is given by $g(t) = \frac{dG(t)}{dt} = \sum_{l=l_{\min}}^{l_{\max}} \tilde{f}_l w(l) \mu_l(t)$. For that reason at any given time t providing $kq(k)$ is an increasing function of the network degree k, the proportion of adopters among individuals with network degree k increases with k. Namely, social hubs will tend to be early adopters of the innovation even if their neighbors influence decreases with the hubs' network degree k as long as the apparent rate of the decay is less than $\frac{1}{k}$.

In this context it should be noted that the Bass model is the special case of homogeneous market wherein all individuals have the same network degree k. In this case there is just a single sub-market k which is in fact the entire market. Therefore $\mu(t) = \mu_k(t)$ and $\tilde{f}_k = 1$ (while $\tilde{f}_l = 0$ for all $k \neq l$). Hence the equation system C6 is reduced to the fundamental Bass equation: $\frac{d\mu(t)}{dt} = (\rho - \mu(t))(P + Q\mu(t))$ where $P = p$ and $Q = kq$.

D) The dynamics of new product growth in a two populations market

In this appendix we develop the dynamics of diffusion in the case of a two populations market that consists of two segments: hubs and non-hubs that are distinguished via their network degree. The dynamical evolution of the proportions of actual adopters in both segments are given in Equation system 10. Both solutions are generated by the same generating function $G(t)$ that satisfies the Ordinary differential Equation 11. We assume that the proportion of hubs among individual neighbors is a small parameter of the problem such that $\tilde{f} << 1$ so we can expand the generating function $G(t)$ on the leading terms of $\tilde{f}$ such that $G(t) = G_0(t) + \tilde{f}G_1(t) + O(\tilde{f}^2)$ and hence:

(D1) $$\frac{dG(t)}{dt} = \frac{dG_0(t)}{dt} + \tilde{f}\frac{dG_1(t)}{dt} + O(\tilde{f}^2)$$

On the other hand, substitution of the $G(t)$'s expansion in the right hand of Equation 11 gives:

(D2) $$\frac{dG(t)}{dt} = \rho\left(1 - e^{-(Pt+QG_0(t))}\right) + \tilde{f}\rho e^{-(Pt+QG_0(t)}\left(QG_1(t) + 1 - e^{-(\Gamma-1)QG_0(t)}\right) + O(\tilde{f}^2)$$

Thus, equating the coefficients of equal powers of $\tilde{f}$ yields the following couple of solvable Ordinary Differential Equations:

(D3) $$\frac{dG_0(t)}{dt} = \rho\left(1 - e^{-(Pt+QG_0(t))}\right)$$

with the initial condition $G_0(t=0) = 0$ and

(D4) $$\frac{dG_1(t)}{dt} = \rho Q e^{-(Pt+QG_0(t))}G_1(t) + \rho e^{-(Pt+QG_0(t))}\left(1 - e^{-(\Gamma-1)QG_0(t))}\right)$$

with the initial condition $G_1(t=0) = 0$.

The leading term $G_0(t)$ (see Equation D3) is given by the integration of the Bass model, namely $\frac{d^2G_0(t)}{dt^2} = \rho e^{-(Pt+QG_0(t))}\left(P + Q\frac{dG_0(t)}{dt}\right) = \left(\rho - \frac{dG_0(t)}{dt}\right)\left(P + Q\frac{dG_0(t)}{dt}\right)$ and hence:

(D5) $$G_0(t) = \int_0^t \rho \frac{1 - e^{-(P+Q\rho)t'}}{1 + \frac{Q\rho}{P}e^{-(P+Q\rho)t'}}dt' = \rho t - \frac{1}{Q}\ln\left(\frac{P + Q\rho}{P + Q\rho e^{-(P+Q\rho)t}}\right)$$

The first order term $G_1(t)$ in the power series expansion of the generating function $G(t)$ is the solution of a first order linear Ordinary Differential Equation (see Equation D4) of the form: $\frac{dG_1(t)}{dt} = \varphi_1(t)G_1(t) + \varphi_0(t)$ where $\varphi_1(t) = \rho Q e^{-(Pt+QG_0(t)} = Q\left(\rho - \frac{dG_0(t)}{dt}\right)$,

$\varphi_0(t) = \rho e^{-(Pt+QG_0(t))}(1 - e^{-(\Gamma-1)G_0(t))})$ and $G_0(t)$ is given by Equation D5. This solution takes the form $G_1(t) = ce^{\int \varphi_1(t)dt} + e^{\int \varphi_1(t)dt} \int e^{-\int \varphi_1(t)dt} \varphi_0(t)dt$ where $\int \varphi_1(t)dt = Q(\rho t - G_0(t))$ is the primitive function of $\varphi_1(t)$ and c is a constant determined by the initial condition $G_1(t=0) = 0$ (Polyanin and Zaitsev 2003). Namely,

(D6) $$G_1(t) = e^{\int \varphi_1(t)dt} \int_0^t e^{-\int \varphi_1(t')dt'} \varphi_0(t')dt' = \rho \frac{P + Q\rho}{P + Q\rho e^{-(P+Q\rho)t}} \int_0^t e^{-(P+Q\rho)t'}(1 - e^{-(\Gamma-1)QG_0(t')})dt'$$

where $G_1(t) \xrightarrow[\Gamma\to\infty]{} \frac{\rho}{P}\left(\frac{1 - e^{-(P+Q\rho)t}}{1 + \frac{Q\rho}{P}e^{-(P+Q\rho)t}}\right) < \frac{\rho}{P}$.