1. Introduction

At the end of 2008, 4 billion people around the world were using mobile phones (ITU, 2008; The Economist, 2009). Launched in 1981 in Scandinavia, mobile phone service has become a part of everyday life for more than half of the world’s population residing in 211 countries. Moreover, in several developed nations, the mobile phone has reached a penetration level that now exceeds 100%, with consumers adopting more than one handset, more than one phone number, and possibly more than one provider. The massive penetration of mobile telephony is not exceptional — many commonly used products and services, such as DVDs, personal computers, digital cameras, online banking, and the Internet, were unknown to consumers three decades ago. As firms invest continually in innovation, this influx of new products and services is expected to continue into the future.

The spread of an innovation in a market is termed “diffusion”. Diffusion research seeks to understand the spread of innovations by modeling their entire life cycle from the perspective of communications and consumer interactions. Traditionally, the main thread of diffusion models has been based on the framework developed by Bass (1969). The Bass model considers the aggregate first-purchase growth of a category of a durable good introduced into a market with potential m. The social network into which it diffuses is assumed to be fully connected and homogenous. At each point in time, new adopters join the market as a result of two types of influences: external influences (p), such as advertising and other communications initiated by the firm, and internal market influences (q) that result from interactions among adopters and potential adopters in the social system. The Bass model states that the probability that an individual will adopt the innovation — given that the individual has not yet adopted it — is linear with respect to the number of previous adopters. The model parameters p, q, and m can be estimated from the actual

© 2010 Elsevier B.V. All rights reserved.
The proliferation of newly introduced information, entertainment, and communication products and services and the development of market trends such as globalization and increased competition have resulted in diffusion processes that go beyond the classical scenario of a single market monopoly of durable goods in a homogenous, fully connected social system. The diffusion modeling literature since 1990 has attempted to extend the Bass framework to reflect the increasing complexity of new product growth. Table 1 provides an overview of the main changes in research focus over the past two decades.

One of the fascinating shifts of focus described in Table 1 is an in-depth discussion of the various types of internal influences involved in the diffusion process. In the original article by Bass, as well as in many of the diffusion studies that followed it, the internal parameter $q$ was interpreted as representing the influence of word of mouth between individuals. Recent contributions to the diffusion modeling literature have reexamined this interpretation to identify and discuss other types of social interactions. On the basis of these recent developments, we believe that the definition of diffusion theory should be revised. The traditional perception of diffusion as a theory of interpersonal communication (Mahajan, Muller & Bass, 1990; Mahajan, Muller & Wind, 2000) should be extended to encompass social interdependence of all kinds (Goldenberg et al., 2010; Van den Bulte & Lilien, 2001). We therefore define diffusion of innovation as follows:

**Innovation diffusion is the process of the market penetration of new products and services, which is driven by social influences. Such influences include all of the interdependencies among consumers that affect various market players with or without their explicit knowledge.**

We discuss two types of additional social influences (besides word-of-mouth communications) that have garnered recent interest: network externalities and social signals.

**Network externalities** exist when the utility of a product to a consumer increases as more consumers adopt the new product (Rohls, 2001). Network externalities are considered to be direct if utility is directly affected by the number of other users of the same product, as in the case of telecommunication products and services such as fax, phone, and e-mail. Network externalities can also be indirect if the utility increases with the number of users of another, complementary product. Thus, for example, the utility to a consumer of adopting a DVD player increases with the increased penetration of DVD titles (Stremersch & Binken, 2009; Stremersch, Tellis, Franses & Binken, 2007). Interpersonal communication is not necessarily needed for network externalities to work. Potential adopters can find out about the penetration level of a new product from the media or simply by observing retail offerings. For example, during the transition from videotape to DVD, a consumer had merely to walk into a Blockbuster movie rental store and observe the amount of aisle space devoted to VHS vs. DVD to understand that DVDs were about to become the new standard. We elaborate on network externalities in Section 2.2.

**Social signals** relate to the social information that individuals infer from adoption of an innovation by others. Through their purchases, individuals may signal either social differences or group identity (Bourdieu, 1984). These signals are transmitted to other individuals, who follow the consumption behavior of people in their aspiration groups (Simmel, 1957; Van den Bulte & Joshi, 2007; Van den Bulte & Wuyts, 2007). Social signals operate vertically and horizontally. A vertical social signal indicates the status of the adopter. Recent research indicates that the competition for status is an important growth driver, sometimes more important than interpersonal ties, and that the speed of diffusion increases in societies that are more sensitive to status differences (Van den Bulte & Stremersch, 2004). Social signals are also transmitted horizontally to indicate group identity. Adoption of an innovation by people in a given group signals to members of that group to adopt and to members of other groups who want to differentiate to avoid adoption (Berger & Heath, 2007, 2008). While social signals can be transmitted via word of mouth and/or advertising, neither is a necessity. These signals are observed by potential adopters who infer from them the social consequences of adoption.

We note that a distinction should be made between social signals and other types of signals, such as functional signals. Functional signals contain information regarding the market perception of the functional attributes of a product, such as its quality or the amount of risk involved in adopting it, whereas social signals contain information regarding the social consequences of adopting the product, including the social risk of adopting the innovation. An important question is whether inclusion of social influence and network externalities as internal influences contradicts the Bass framework. Traditional applications of the Bass framework have interpreted internal influence in terms of word-of-mouth and personal communications (Mahajan et al., 1990). However, this interpretation is not dictated by the model itself, which does not specify the drivers of social contagion. Thus, the consumer interactions of network externalities and social influence certainly fit the framework, as do other possible growth drivers, as long as they imply that the probability of purchase increases with the number of previous adopters.

In spite of growing evidence of the importance of personal communication in product adoption, an alternative research branch has emerged. This branch argues that the major driver of growth of new products is consumer heterogeneity rather than consumer interaction. The heterogeneity approach claims that the social system is heterogeneous in innovativeness, price sensitivity and needs, leading to heterogeneity in propensity to adopt. Thus, innovators are the least patient in adopting, whereas laggards are the most patient. In such models, patience is often inversely related to product affordability, consumer willingness to pay, or reservation price (Bemmaor, 1994; Golder & Tellis, 1998; Russell, 1980; Song & Chintagunta, 2003). The dynamics of market volume are determined by the shape of the distribution of “patience” in the face of falling prices. If incomes are log-normally distributed in the population, then growth is S-shaped (Golder & Tellis, 1997). This line of research implies that the current approach of diffusion-based research has overemphasized the influence of word-of-mouth communication (Van den Bulte & Lilien, 2001, and Van den Bulte & Stremersch, 2004). Fig. 1 illustrates the range of possible drivers of new product diffusion, arranged according to the level of direct interpersonal communication they involve.

Our objective in this paper is to review the interaction-based diffusion literature published in the past decade and analyze how it has broadened its scope to describe the richness of consumers’ internal influences so as to bring these influences in a unified way into the diffusion framework. We do not aim in this paper to cover the entire diffusion literature; for that, we refer the reader to recently published diffusion overviews (Mahajan et al. 2004; Venkatesan, Krishnan and Kumar 2004; Lilien et al., 2000; Sultan, Farley and Lehmann 1990; and Van den Bulte and Lilien 1997).

---

**Table 1**

<table>
<thead>
<tr>
<th>Previous focus</th>
<th>Complemented with current focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word of mouth as driver</td>
<td>Consumer interdependencies as drivers</td>
</tr>
<tr>
<td>Monotonically increasing penetration curve</td>
<td>Turning points and irregularities in the penetration curve</td>
</tr>
<tr>
<td>Temporal industry-level analysis</td>
<td>Spatial</td>
</tr>
<tr>
<td>Aggregate or segment-based models</td>
<td>Brand-level analysis</td>
</tr>
<tr>
<td>Fully connected networks</td>
<td>Individual-level models</td>
</tr>
<tr>
<td>Products forecasting</td>
<td>Partially connected and small-world networks</td>
</tr>
<tr>
<td>Services</td>
<td>Managerial diagnostics</td>
</tr>
</tbody>
</table>
Rather, we feel that there is a need for a review paper that integrates the modeling efforts of various types of interpersonal influences into a single framework and reviews studies that have explored the manifestations of these influences within and across markets and brands.

We start by discussing consumer influences within a single market in Section 2. We discuss issues such as modeling the social network, network externalities, takeoffs and saddles, and technology generations. Consumer influences across markets and brands are discussed in Section 3, where we relate to cross-country influences, differences in growth across countries, and effects of competition on growth. In Section 4, we suggest topics for further research. Fig. 2 illustrates the section flow. Table 2 conveys a summary of the focus of the main research efforts in each literature stream, as well as the corresponding directions for further research.

2. Diffusion within markets and technologies

In this section, we discuss four of the seven most influential diffusion-related areas studied in the past decade. These areas — social networks, network externalities, takeoffs and saddles, and technology generations — concern effects within a single market or technology.

2.1. Diffusion in social networks

The social network, or social system, is the substrate onto which an innovation propagates. The gradual decrease in the efficiency of offline advertising (e.g., Evans, 2009) and the development of online social networks such as Facebook have led firms to approach directly the social networks of customers in their target markets and to invest marketing efforts in enhancing the internal influences in those networks. In order to successfully enhance such influences, a better understanding is required as to how the structure and dynamics of the social network influence the diffusion process.

A fundamental question regarding diffusion within social networks is how the social network structure influences product growth. This question has not yet been answered theoretically but has been explored in some empirical studies (see Van den Bulte & Wuyts, 2007, for an overview). Much of the research attention has been on the roles of central individuals (influentials, social hubs) on the overall growth process (e.g., Goldenberg, Han, Lehmann & Hong, 2009; Iyengar et al., in press). Research on the role of network structure in diffusion is still in its infancy, mainly because of the lack of data. New methods that enable large-scale sampling and analyses of online networks will probably boost further research (see, for example, Dorogovtsev & Mendes, 2003; Jackson, 2008). From the modeling perspective, the basic question is how to incorporate the social network into the diffusion model. The implicit assumption of the Bass model and of most of its extensions is that the social system is homogenous and fully connected. Therefore, the adoption process can be appropriately represented at the aggregate level. Aggregate diffusion models have advantages as well as downsides (Parker, 1994): On one hand, they are parsimonious and require few data for parameter estimation and forecasting; on the other hand, they provide little intuition as to how individual market interactions are linked to global market behavior. Since the extensive recent research on social networks has revealed that they are neither homogenous nor fully connected (Kossinets & Watts, 2006), diffusion research is gradually extending its focus from the aggregate level to an individual-level perspective.

One well-known approach for describing individual adoption decisions and tying them to aggregate outcomes is agent-based modeling, which describes the market as a collection of individual elements (units, agents, or nodes) interacting with each other through connections (links). The behavior (in our case, adoption) of...
each individual element is determined by a decision rule. Neural networks, cellular automata, and small-world models are examples of agent-based modeling techniques. A typical agent-based model is the cellular automata of Goldenberg, Libai and Muller (2001a; 2001b). Each unit represents an individual consumer and has a value of “0” if it has not yet adopted the product and “1” if it has. Potential adopters (“units”) adopt due to the combination of external influences (parameter $p$) and internal influences (parameter $q$) in a manner similar to the Bass framework.

Such a model overcomes some of the limitations of aggregate-level diffusion models. First, by establishing a connection between individual-level influences and aggregate effects, it enables the researcher to better relate individual-level marketing activities to firm performance, which is measured at the aggregate level. The response to marketing activities can be applied in the transition rule of the individual agents, and the performance is the aggregation of the decisions over the entire network. Second, this modeling approach enables one to distinguish between interdependencies: for example, Goldenberg et al. (2010) used it to explore network外部ities by adding a threshold to the decision rule. Third, this model allows for heterogeneity by enabling individual susceptibility to influences ($p_i$ and $q_i$) to differ among units or by setting different link structures for each unit. The model can be adapted to include almost any aspect of heterogeneity, including individual responsiveness to price and advertising (Libai, Muller & Peres, 2005), presence of negative word of mouth (Goldenberg, Libai, Muller & Moldovan, 2007), intrinsic consumer innovativeness (Goldenberg, Libai & Muller, 2002), presence of heavy users and connectors (Kumar, Petersen & Leone, 2007), and individual roles in the social network — that is, hubs, connectors, and experts (Goldenberg et al., 2009).

Another advantage of agent-based models is their ability to take into account the spatial aspect of diffusion. Using a variation of agent-based models called small world, Goldenberg et al. (2001a) studied spatial issues in the market through the relative influences of strong and weak ties. They showed that, consistent with Wuyts, Stremersch, Van den Bulte and Franses (2004) and Rindfleisch and Moorman (2001), the cumulative influence of weak ties has a strong effect on the growth process. Garber, Goldenberg, Libai and Muller (2004) suggested a measure of the spatial density of adoption to predict a product’s success or failure.

Conceptually, aggregate diffusion models represent the overall results of individual-level processes. Therefore, in order to create a coherent market representation, the individual-level and aggregate market formulations should be equivalent. However, the equivalence of the two formulations is not straightforward. Previous studies have proposed methods of aggregating individual-level behavior based on assumptions regarding customer heterogeneity and calculation of the time to adoption (Chatterjee & Eliashberg, 1990; Van den Bulte & Stremersch, 2004). In the specific case of cellular automata, Goldenberg et al. (2001a) demonstrated this equivalence by relating the parameters $p$ and $q$ to the adoption hazard function and presenting simulations that showed that the individual-level probability of adoption generates diffusion curves with $p$ and $q$. The relationship between the Bass model and agent-based models was also investigated by Rahmandad and Strem (2008) and by Fibich et al. (2009). Shaikh et al. (2006) showed how product adoption by units in a small-world network can be aggregated to create the Bass model with some relatively simple assumptions. However, the interface between the individual level and the aggregate level still lacks a closed formulation and needs further exploration.

### 2.2. Diffusion and network externalities

The dynamics of network externalities have received considerable attention in the past two decades (see Stremersch et al., 2007). There is as yet no consensus about the effect of network externalities on growth rate. Conventional wisdom suggests that network effects drive faster market growth due to the increasing returns associated with
such processes (Nair, Chintagunta & Dubé, 2004; Tellis, Yin & Niraj, 2009). However, networks can also create the opposite effect, slowing growth with what is labeled “excess inertia” (Srinivasan, Lilien & Rangaswamy, 2004). Early in the product life cycle, most consumers see little utility in the product as there are few adopters; they may take a “wait-and-see” approach until there are more adopters. Hence, diffusion early on may be very slow and occur primarily among the few consumers who see utility in the product despite a lack of adoption by others. Overall, the process can be characterized by a combination of excess inertia and excess momentum, i.e., slow growth followed by a surge (Van den Bulte & Stremersch, 2006).

This growth pattern can occur via various types of network externalities. In the case of direct network effects, which apply to fax, e-mail, and other communication products, the number of adopters drives utility directly because the utility of the product increases with the number of adopters. (A fax machine is not useful if almost no one else has one.) Regarding indirect network effects, as in the case of hardware and software products, an increase in utility may occur through market mediation (e.g., amount of compatible software applications), which in turn is a function of the number of adopters. Consumers generally wait to adopt hardware until there is enough software to make it worthwhile. In addition, in the case of competing standards, early adopters take the risk of choosing a standard that will eventually “lose” and become obsolete, so many consumers wait until it is clear which is the winning standard and, more importantly, which standard or platform will no longer be supported.

The impact of network externalities on growth rate can be determined by the source of the externalities under examination, namely, global or local. Under global externalities, a consumer takes into account an entire social system when evaluating the utility of a product in terms of the number of adopters, whereas under local externalities, a consumer considers adoption in relation to his close social network. Research is gradually moving from considering only global externalities towards exploring local externalities (Binken & Stremersch, 2009). The marketing decisions of the firm can influence the types of externalities that are relevant: growth of competing standards will probably invoke a global effect since the “verdict” on what eventually becomes the winning standard depends on the total number of users. In contrast, a family program of a mobile phone service provider might evoke a local effect since it involves only the local social system.

The challenge in integrating network externalities into diffusion models stems from the multiple effects of previous adopters on the rate of growth. Previous users are expected to accelerate growth due to interpersonal effects, including word of mouth and imitation, which typically reduce both risk and search costs. Yet, the mere adoption of previous adopters increases network externalities, consequently enhancing growth. The literature on the modeling of diffusion of innovations generally does not separate the two factors, and a single parameter for internal influence is used to capture the effects of both interpersonal communications and network externalities (Van den Bulte & Stremersch, 2004). Goldenberg et al. (2010) separated the two effects by drawing on the literature on collective action to relate network effects to threshold levels of adoption by individual consumers. They found that network externalities have a “chilling” effect on initial growth, which is then followed by a surge of enhanced diffusion. These findings are further discussed in Gatignon (2010), Rust (2010) and Tellis (2010).

2.3. Takeoffs and saddles

In the past decade, a stream of literature has emerged that examines turning points in the product life cycle that are not included in the classic smooth-adoption curve (see illustration in Fig. 3). We focus here on research dealing with two turning points in the product life cycle: takeoff, which occurs at the beginning, and saddle, which occurs during early growth. The classic Bass model starts with spontaneous adoption by an initial group of adopters but does not provide explanations for the mechanisms that lead to this initial adoption, or takeoff. Studies on takeoff focus on this initial stage and explore the market’s behavior and the interface between adoption and the start of communication interactions.

Golder and Tellis (1997) defined takeoff time as the time at which a dramatic increase in sales occurs that distinguishes the cutoff point between the introduction and growth stage of the product lifecycle. The importance of takeoff time to the firm is quite clear: a rapid increase in sales requires substantial investments in production, distribution, and marketing, which most often involve considerable lead time to put into place successfully. Golder and Tellis (1997) applied a proportional hazard model to data that included 31 successful innovative product categories in the U.S. between 1898 (automobiles) and 1990 (direct broadcast satellite media). They found that the average time to takeoff for categories introduced after World War II was six years and that average penetration at takeoff was 1.7% of market potential. Price at takeoff was found to be 63% of the original price. Other studies have investigated factors that influence time to takeoff. Accelerating factors are price reduction, product category (brown goods such as CDs and television sets take off faster than white goods), and cultural factors such as masculinity and a low level of uncertainty avoidance (Foster, Golder & Tellis, 2004; Tellis, Stremersch & Yin, 2003).

Takeoff as such does not require any consumer interaction. Instead, it results from heterogeneity in price sensitivity and risk avoidance—as the innovation’s price is reduced, adoption becomes associated with less risk, and the product takes off. Therefore, if one believes that both heterogeneity and communication play a role in new product adoption, takeoff is an excellent example of an interface point: heterogeneity is dominant prior to takeoff, but consumer interactions become the driving force immediately afterwards. As already stated by Mahajan et al., 1990 review, there is a need for a comprehensive theory that delves deeper into early market growth prior to takeoff.

Following takeoff, diffusion models predict a monotonic increase in sales up to the peak of growth. However, in some markets a sudden decrease in sales may follow an initial rise. This decrease in sales was observed by Moore (1991), who denoted it as the chasm between the early and main markets, and this concept was later formalized and explored by Mahajan and Muller (1998); Goldenberg et al. (2002); Golder and Tellis (2004); Muller and Yogev (2006); Van den Bulte and Joshi (2007); Vakratsas and Kolarski (2008); and Libai, Mahajan and Muller (2008). Goldenberg et al. (2002) referred to the phenomenon as a “saddle” and defined it as a pattern in which an initial peak predates a trough of a substantial depth and duration that is followed by increased sales that eventually exceed the initial peak.

While a saddle can be attributed to causes such as changes in technology and macroeconomic events, it can also be explained by consumer interactions. Golder and Tellis (2004), as well as Chandrasekaran and Tellis (2006), have claimed that the saddle phenomenon can be explained using the informational cascade theory. Small shocks to the economic system such as a minor recession can
temporarily decrease the adoption rate, and the decrease is magnified through the informational cascade. Another explanation is based on heterogeneity in the adopting population and its division into two distinct groups. If these two groups adopt the innovation at widely differing rates and have weak communication between them, sales may show an interim trough (Goldenberg et al., 2002; Muller & Yoge, 2006; Van den Bulte & Joshi, 2007). These findings demonstrate how combining heterogeneity and consumer interactions explains a phenomenon that does not fit the typical bell-shaped sales curve.

2.4. Technology generations

In theory, the basic diffusion process is terminated by a decay of the number of new adopters and saturation of the market potential. In practice, however, products are often substituted with newer generations of products with more advanced attributes. New product growth across technology generations has garnered growing interest among marketing scholars such as Bass and Bass (2001, 2004), Mahajan and Muller (1996), and Norton and Bass (1987, 1992). A major issue examined by these researchers is whether diffusion accelerates between technology generations. This question has theoretical importance for forecasting as projections regarding the growth of advanced generations of a product must often be made during the early stages of product penetration or before launch and are thus based on using diffusion parameters from previous generations. Theoretically, this question is important, because it deals with dependency within a sequence of diffusion processes and, more broadly, with rigidity of the social system across generations. Does the social system learn to improve its adoption skills across generations or does it start each diffusion process from scratch? If it has learning capabilities, how strong and how category-specific are they?

As Stremersch et al. (2010) pointed out, the literature offers contradicting answers to the question of whether diffusion accelerates across technology generations. The key finding (or assumption) of several studies across multiple product categories is that growth parameters are constant across technology generations (Bass & Bass, 2004; Kim, Chang & Shocker, 2000; Mahajan & Muller, 1996; Norton & Bass, 1992, 1987). Bayus (1994), for example, used a proportional hazard model to analyze the diffusion of four generations of personal computers and concluded that the average product lifespan did not decline over time. This was true even when moderating variables (such as year of entry and technology used) were included. Exceptions to this premise were provided by Islam and Meade (1997) and Pae and Lehmann (2003), who demonstrated that the results can be an artifact of the difference in the length of time covered by the data series used in the analysis; their findings were subject to criticism by Van den Bulte (2004).

In a contradiction to the stability of growth parameters across generations, there is a great body of evidence suggesting that the overall temporal pattern of diffusion of innovation accelerates over time (Van den Bulte & Stremersch, 2004; Kohli, Lehmann & Pae, 1999). A recent analysis by Van den Bulte (2000; 2002) found conclusive evidence that such acceleration does indeed occur. Van den Bulte investigated the issue of acceleration by adopting the Bass model with the internal influence parameter ($p$) set to zero and running the model on 31 product categories in consumer electronics and household products. The average annual acceleration between 1946 and 1980 was found to be around 2%. Exceptions to this generalized finding are rare (Bayus, 1994) and contested on the grounds of estimation bias and invalid inference (Van den Bulte, 2000).

These two research streams form an intriguing paradox: It seems that, in the same economy, an acceleration of the diffusion of innovations over time should be reflected in an acceleration of diffusion of technology generations that succeed one another; however, the diffusion rates of sequential technology generations remain constant. A resolution to the paradox was suggested recently by Stremersch et al. (2010), who noted constant growth parameters across generations but a shorter time to takeoff for each successive generation. They investigated whether the faster takeoff of successive generations is due to the passage of time or to the generational effect. They defined technology generation as a set of products similar in customer-perceived characteristics and technology vintage as the year in which the first model of a specific technology generation was launched commercially. Using a discrete proportional hazard model in 12 product categories, Stremersch, Muller, and Peres found that acceleration in time to takeoff is due to the passage of time and not to generational shifts. Thus, time indeed accelerates early growth, whereas generational shifts do not.

The issue of technological substitution has raised questions related to heterogeneity in the adopting population. Goldenberg and Oreg (2007) proposed a redefinition of the “laggards” concept; they suggested that laggards from previous product generations may often become innovators of the latest generation because of leapfrogging. Thus, for example, in the early days of the MP3 revolution, an early adopter of MP3 could be a user of a Walkman cassette player who did not adopt CD technology and decided to upgrade by leapfrogging to an MP3 player. Hence, early adopters of MP3 players were not necessarily innovative; some may have been leapfrogging laggards from previous generations. Thus, firms should approach them with the appropriate marketing mix tools and not treat them as innovators.

The entry of a new technology generation complicates the growth dynamics and generates consumer-related processes that are not observed in single-generation diffusion. First, the entry of a new generation is usually considered to increase the market potential. In addition, customers can upgrade and replace an old technology with a new one. On the other hand, individuals who belong to the increased market potential might decide eventually to adopt the older generation of the product and, hence, cannibalize the new technology's potential. If there are more than two generations, adopters can skip a generation and leapfrog to advanced versions. This means that the entry of a new technology reveals heterogeneity in the adopting population that was not realized in a single generation scenario. Surprisingly, none of the diffusion models have yet offered a comprehensive treatment of these dynamics. Studies so far have focused on one or two of these aspects, such as upgrading (Bass & Bass, 2001, 2004; Norton & Bass, 1992) and cannibalization (Mahajan & Muller, 1996), but a unified theoretical treatment of the subject is still required.

Normative decisions are also influenced by intergenerational dynamics. Several papers have investigated optimal pricing decisions under technological substitution (Padmanabhan & Bass, 1993; Danaher, Hardie & Putsis, 2001; Lehmann & Esteban-Bravo, 2006). However, with the exception of Lehmann and Esteban-Bravo (2006), these studies did not address the dynamics of specific groups of adopters.

3. Diffusion across markets and brands

In this section, we discuss the remaining three of the seven research areas that, in our view, are the most significant in terms of innovative diffusion research in the past decade. These three areas, which concern various cross-market and cross-brand effects, are cross-country influences, differences in growth across countries, and effects of competition on growth.

3.1. Cross-country influences

A number of papers since 1990 have followed the global trend of focusing on multinational product acceptance and have extended the traditional single-market scope to explore problems and issues
related to international diffusion (see Dekimpe, Parker and Sarvary (2000a) for a review).

A key issue in multinational diffusion, especially with respect to order of entry, is the mutual influence of diffusion processes in various countries. One of the major findings of the studies to date on cross-country influences (with a few exceptions, such as Desiraju, Nair and Chintagunta (2004) and Elberse and Eliaishberg (2003)) is that the entry time lag has a positive influence on the diffusion process; that is, countries that introduce a given innovation later show a faster diffusion process (Tellis et al., 2003; Dekimpe, Parker & Sarvary, 2000b;2000c; Ganesh, Kumar & Subramaniam, 1997; Takada & Jain, 1991) and a shorter time to takeoff (Van Everdingen, Fok & Stremersch, 2009). This cross-country influence has been called the lead-lag effect; however, influence can be multidirectional.

Several papers have modeled multi-market diffusion with cross-country influences (Van Everdingen, Aghina & Fok, 2005; Kumar & Krishnan, 2002; Ganesh et al., 1997; Putsis, Balasubramanian, Kaplan & Sen, 1997; Eliaishberg & Helsen, 1996; Ganesh & Kumar, 1996). Generalizing these models, a generic cross-country influence model may take the following form (\(x_i\) is the proportion of adopters in country \(i\)):

\[
\frac{dx_i(t)}{dt} = (p_1 + q_x(t) + \sum_{j \neq i} \delta_{ij}x_j(t) - (1 - x_i(t)).
\] (1)

The parameter \(\delta_{ij}\) represents cross-country effects between country \(i\) and another country \(j\). Cross-country effects can be a result of two types of influence mechanisms: weak ties and signals. Weak ties come from adopters in one country who communicate with nonadopters from other countries (Wuyts et al., 2004; Rindfleisch & Moorman, 2001). However, even without communicating with or imitating other individuals, nonadopters are influenced by diffusion in other countries. In other words, the level of acceptance of the innovation in one country acts as a signal for customers in other countries, reducing their perceptions of risk and increasing the legitimacy of using the new product. While several studies stated explicitly that the dominant effect was due to communication (Putsis et al., 1997; Eliaishberg & Helsen, 1996; Ganesh & Kumar, 1996), others explored the effect without relating it to a specific mechanism (e.g., Dekimpe et al., 2000b;2000c; Takada & Jain, 1991).

Note that although the distinction between weak ties and signals has evident managerial implications, the commonly used aggregate models of the type presented in Eq. (1) do not distinguish between the two effects; both are represented through the parameters \(\delta_{ij}\). Further research is required to estimate the relative roles of word of mouth and non-communication signals in cross-country spillover and to study their relative effects on the overall diffusion process. Individual-level models might be able to make this distinction.

Understanding cross-country influences is also valuable in the context of normative managerial decisions in multinational markets. Some studies have explored entry strategies — i.e., the question of whether a firm should enter all of its markets simultaneously (a “sprinkler” strategy) or sequentially (a “waterfall” strategy). Kalish, Mahajan and Muller (1995) built a game-theoretic model for two brands and suggested that the waterfall strategy is preferable when conditions in foreign markets are unfavorable (slow growth or low innovativeness), competitive pressure is low, the lead-lag effect is high, and fixed entry costs are high. Libai et al. (2005) extended this question to explore responsive budgeting strategies in which firms dynamically allocate their marketing efforts according to developments in the market. Many other questions are still waiting to be answered, and issues such as regulation (addressed by Stremersch & Lemmens, 2009), international competition, and the optimal marketing mix of growing international markets can be further explored.

### 3.2. Growth differences across countries

Research on the evolution of multi-markets reveals a noteworthy aspect of diffusion that deals with heterogeneity among different social systems in which the same product is adopted. If various social systems adopt the same product in different ways, understanding these differences will clarify the diffusion process within each country. A large number of studies published during the last two decades have focused on describing and explaining such inter-country differences. Key findings are summarized in Table 3. These studies have generally focused on differences in the diffusion parameters \(p\) and/or \(q\) (e.g., Van den Bulte, 2002; Dekimpe, Parker & Sarvary, 1998; Helsen, Jedidi & DeSarbo, 1993; Takada & Jain, 1991), the ratio \(q/p\) (Van den Bulte & Stremersch, 2004), time to takeoff (Tellis et al., 2003) and duration of the growth stage (Stremersch & Tellis, 2004).

The salient result emerging from all of these papers is that diffusion processes vary greatly among countries, even for the same products or within the same continent (Ganesh, 1998; Mahajan & Muller, 1994; Helsen et al., 1993). In addition to measuring the differences among growth processes, these studies have also investigated the country-specific sources of these differences. Besides product entry time, discussed above, the underlying factors can be divided into cultural sources and economic sources.

**Cultural sources** — These relate to the country’s cultural characteristics and values. Takada and Jain (1991) found that the diffusion parameter \(q\) is higher in countries that are high-context and homophilous (such as Asian Pacific countries) relative to countries such as the U.S. that are low-context and heterophilous. High-context refers to a culture in which much of the information conveyed through a communication resides in the context of the communication rather than in its explicit message, and homophilous implies that communication takes place among individuals who share certain characteristics. Similarly, Dekimpe et al. (2000b;2000c) that population heterogeneity has a negative effect on both time to adoption and the probability of transition from nonadoption to partial or full adoption in a country. Similar findings are described in Talukdar, Sudhir and Ainslie (2002).

**Dwyer, Mesak and Hsu (2005) used Hofstede (2001) dimensions of national culture and found positive relationships between \(q\) and collectivism (vs. individualism), masculinity (assertiveness and competitiveness as desired traits), and high power distance (the extent to which the less powerful members of institutions and organizations in a country expect and accept that power is distributed unequally). Their findings were supported by Van den Bulte and...**

### Table 3

<table>
<thead>
<tr>
<th>Source of difference</th>
<th>Influence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entry time mix</td>
<td>Mostly positive</td>
</tr>
<tr>
<td>Marketing mix</td>
<td>Positive (brown goods: CD players, TVs, camcorders, etc. penetrate faster)</td>
</tr>
<tr>
<td>Entry time lag</td>
<td>Positive</td>
</tr>
<tr>
<td>Product type</td>
<td>Positive</td>
</tr>
<tr>
<td>(brown goods vs. white goods)</td>
<td></td>
</tr>
<tr>
<td>Existence of competition</td>
<td>Positive</td>
</tr>
<tr>
<td>Market-related</td>
<td>Positive</td>
</tr>
<tr>
<td>Regulation</td>
<td>Negative</td>
</tr>
<tr>
<td>Demographic</td>
<td>Mixed</td>
</tr>
<tr>
<td>and cultural</td>
<td>Positive</td>
</tr>
<tr>
<td>Population heterogeneity</td>
<td></td>
</tr>
<tr>
<td>Population growth rate</td>
<td>Direction: less uncertainly</td>
</tr>
<tr>
<td>No. of population centers</td>
<td></td>
</tr>
<tr>
<td>Hofstede’s dimensions</td>
<td>Positive</td>
</tr>
<tr>
<td>Economic</td>
<td>Positive</td>
</tr>
<tr>
<td>Wealth (GDP, income per capita)</td>
<td></td>
</tr>
<tr>
<td>Media availability</td>
<td>Mixed</td>
</tr>
<tr>
<td>Income inequality</td>
<td>Positive</td>
</tr>
<tr>
<td>Regulation</td>
<td>Negative</td>
</tr>
</tbody>
</table>
Stremersch (2004). In addition, Dwyer et al. (2005) found that short- 
term orientation was positively associated with q, but they did not 
observe a significant negative relationship between q and uncertainty 
avoidance.

Economic sources – The influences of many macroeconomic 
variables have been studied, yielding two main empirical general-
izations: First, the wealth of the country (usually measured by gross 
domestic product (GDP) per capita, but also by lifestyle, health status, 
and urbanization) has a positive influence on diffusion (Desiraju et al., 
2004; Talukdar et al., 2002; Dekimpe et al., 2000b;2000c; Putsis et al., 
1997; Helsen et al., 1993). Note that wealth is not necessarily 
equivalent to general welfare. For example, Van den Bulte and 
Stremersch (2004) found a positive influence between the Gini index 
for inequality and the ratio q/p. This finding is consistent with cultural 
findings concerning positive associations with power distance and 
masculinity. A second generalization is that access to mass media 
(usually operationalized by the penetration of TV sets) has a positive 
influence on the diffusion parameter p (Stremersch & Tellis, 2004; 
Talukdar et al., 2002; Tellefsen & Takada, 1999; Putsis et al., 1997).

In pharmaceutical markets, regulation was found to influence 
and Stremersch and Tellis (2004) distinguished between the 
influences of cultural and economic factors on penetration stages. 
They found, for example, that cultural factors influence time to 
takeoff, whereas economic factors influence growth.

Some efforts have been made to include developing countries in 
inovation diffusion studies, (e.g., Dekimpe et al., 1998;2000b, 
Stremersch & Lemmens, 2009; Van Everdingen et al., 2009); however, 
it remains to be determined whether emerging economies are 
characterized by the patterns and forces that are at work in developed 
economies or whether the theories have to be revised (Mahajan, 
2009, Mahajan & Banga, 2006; Steenkamp & Burgess, 2002).

3.3. The effects of competition on growth

Competitive forces influence the growth of a new product and 
decisions made about it. Although some innovative categories start as 
monopolies, many quickly gain multiple competing brands. Interest-
ingly, the existence of competition influences not only the interac-
tions of customers with the firms, but also the dynamics of the 
interactions of consumers with other consumers. Thus, a competitive 
growing market reveals a range of consumer interdependencies that 
are not present in the case of a monopoly: word of mouth now applies 
to within-brand and cross-brand communications, compatibility 
issues can increase or decrease the effect of network externalities, 
and the entry of a new competitor can act as a signal of the quality of a 
product.

Despite the richness of competitive phenomena in growing 
markets, the early diffusion literature dealt with either growth of 
monopolies or category-level growth. The marketing literature on 
competition has investigated mainly mature markets, with some studies of competitive effects in growing markets (see Chatterjee, 
Eliashberg and Rao (2000) for a review). Only in the last two decades 
have researchers combined these two research streams and incorpo-
rated some of the competitive effects into diffusion modeling.

Fig. 4 illustrates various competitive effects that modify and 
influence the growth process and that do not exist in monopolies. 
These competitive effects relate to customer flow and information 
flow. Regarding customer flow, the firms compete on two fronts. The 
first is acquisition of adopters from the market potential before their 
competitors, and the second is prevention of customer churn and 
acquisition of customers who churn from other brands. Competitors 
might be legal brands or illegal brands; hence, a factor of coping with 
piracy is also at play. The information flow under competition becomes 
more complex since consumers’ communication exists within and across 
brands. Among these effects are the influences of competition on the 
category growth rate and communication transfer within and between 
brands.

At the category level, an intriguing empirical question is whether 
competition enhances or delays category growth. Generally, compe-
tition has been found to have a positive effect on diffusion parameters 
(Kaufmann & Techasanasoontorn, 2005; Van den Bulte & Stre-
mersch, 2004; Kim, Bridges & Srivastava, 1999; Dekimpe et al., 1998). 
An exception was observed by Dekimpe et al. (2000c), who showed 
that an existing installed base of an old technology negatively affected 
studied the impact of late entrants on the diffusion of incumbent 
brands. Using data on diffusion of minivans and cellular phones in 
several U.S. states, they found that the effect varied across markets. In 
some markets, the market potential and the internal communication 
parameter q increased with the entry of an additional brand, whereas 
in other markets only one of these parameters increased. None of 
these studies provides explanations of the mechanisms underlying 
the acceleration. One potential explanation is that acceleration results 
from greater marketing pressure on the target market. Kim et al. 
(1999) have implicitly suggested that the number of competitors 
constitutes a signal of the quality and long-term potential of the 
product, which may result in acceleration. Alternatively, the positive

![Fig. 4. Competitive effects on focal brand A of competition for market potential, piracy, cross-brand communication, and churning customers to and from the competitors.](image-url)
effect may be a result of reduction in network externalities (Van den Bulte & Stremersch, 2004). Clearly, further research is needed here.

Turning to the brand level, constructing a brand-growth model requires discussing several conceptual issues. A basic issue is the extent to which internal influence mechanisms operate at the brand level. Some regard brand adoption as a two-stage process in which consumers first adopt the category and then choose the brand (Givon, Mahajan & Muller, 1995; Hahn, Park, Krishnamurti & Zoltners, 1994) based on factors other than internal communication such as promotion activities, price deals, and special offers. In spite of the intuitive rationale of this approach, attempts to use it have been rare, partially because it requires high-quality individual-level data. The development of service markets and increased use of Customer Relationship Management (CRM) systems by service providers can facilitate data availability and promote the use of these types of models. Landsman and Givon (2010) used banking data to investigate the growth of financial products, and Weerahandi and Dalal (1992) used business-to-business data for fax penetration. Such a two-stage model is an excellent platform for incorporating customer heterogeneity (Lee, Lee & Lee, 2006; Jun & Park, 1999) into communication-based approaches. The two-stage model enables an explicit investigation of the link between personal characteristics, product attributes, individual choice behavior, and aggregate growth. Intensive research is still needed to build firm theoretical and empirical infrastructures for choice-based growth models.

Although the relationships between brand-level and category-level adoption have not yet been clearly identified, the main body of literature has assumed that internal dynamics are important at the brand level, and, therefore, a Bass-type model can be used to model brand choice. Mapping the communications flow in the market, one can say that a potential customer adopts brand i as a result of the combination of two optional communication paths: within-brand communication with adopters of brand i and/or cross-brand communication with adopters of other brands. Cross-brand communication can influence a consumer’s choice of a brand in two ways: the consumer may receive negative information about the competing brands or the consumer may receive information about the category from adopters of other brands and subsequently adopt brand i because its marketing mix is most appealing. Generalizing from the competitive diffusion models published so far, a diffusion equation for multiple brands that explicitly presents both communication paths can take the following form (adopted from Libai, Muller and Peres (2009a) and Savin and Terwiessch (2005)): 

$$\frac{dN_i(t)}{dt} = (p_i + q_i \frac{N_i(t)}{m} + \sum_{j \neq i} \delta_{ij} \frac{N_j(t)}{m})(m-N(t))$$

where N is the total number of adopters (N = N_i + N_j), and $\delta_{ij}$ represents the cross-brand influences. The parameter m is the overall market potential. Many of the existing brand-level diffusion models are special cases or variations of this generic model. Some assume that within-brand communication equals cross-brand communication, namely, $\delta_{ij} = q_i$ (Krishnan et al., 2000; Kim et al., 1999; Kalish et al., 1995). Their underlying assumption is that there is no relevance to the brand ownership of the individual who spreads the information. Other modelers, such as Mahajan, Sharma and Buzzell (1993), assume that all communications are brand-specific ($\delta_{ij} = 0$).

The relationship between the Bass model and the summation of the equations of individual brands in each category should be further studied. For example, when $\delta_{ij} = q_i$, summing the equations for all brands yields the category equation of the Bass model; however, when cross-brand communication is not equal to within-brand communication, the summation results in a model different from the Bass model. Two studies have tried to examine systematically the distinction between within- and cross-brand communication: Parker and Gatignon (1994) and Libai et al. (2009a). For consumer goods (Parker and Gatignon) and cellular services (Libai, Muller, and Peres), the studies concluded that both within-brand and cross-brand influences exist.

Note the formative resemblance between Eqs. (1) and (2): Both describe consumer interdependencies; Eq. (1) represents influences of consumers from other markets, whereas Eq. (2) deals with influences from customers of competing firms. As argued above, $\delta_{ij}$ can represent any type of interdependency, and further research should be done to separate the effects of word of mouth, signals, and network externalities. Empirical and behavioral research is needed to elucidate what determines the ratio of within-brand to cross-brand communications — it might be a combination of the characteristics of the innovation, the social system, and/or the nature of competition.

A conceptual difference between the two equations relates to market potential. The multinational Eq. (1) describes diffusion processes that operate in separate markets in which each firm draws from its own market potential. In the competitive scenario of Eq. (2), the assumption is that both firms compete for the same market potential. Some studies have relaxed the assumption of a joint potential and assumed that brands can develop independently, with each brand having its own market potential (Parker & Gatignon, 1994). In that case, the market potential $m_i$ in Eq. (2) should be modified to $m_i$, i.e., the equation should now be:

$$\frac{dN_i(t)}{dt} = (p_i + q_i \frac{N_i(t)}{m_i} + \delta_{ij} \frac{N_j(t)}{m_j})(m_i-N_i(t)).$$

Note that Eq. (3) requires careful treatment and interpretation. If one assumes that the market potentials of the two brands do not overlap, then the brands do not compete for the attention and wallets of the same potential consumers. On the other hand, if one assumes that the market potentials of the brands do overlap and that the total market potential $m = \sum m_i$ then this overall market potential overestimates the true potential since the intersections should be subtracted from the overall count. Mahajan et al. (1993) investigated the market potential issue in the context of Polaroid’s lawsuit against Kodak, the latter having been accused of patent violation and of attracting Polaroid’s customers to a new brand of digital camera. By breaking the nonadoption pool $m - N$ into sub-pools according to the market potential of each brand, the researchers concluded that Kodak took about 30% of its customers from Polaroid’s potential buyers. However, at the same time, Kodak expanded the market for Polaroid since about two thirds of Polaroid’s sales would not have occurred if Kodak had not entered the market.

Within- and cross-brand influences occur even among brands that do not directly compete. Joshi, Reibstein and Zhang (2009) consider a brand extension of a high-status market that develops a new, lower-status version of the product. The existing high-status market has a positive influence on the new market, but the reciprocal social influence of the new, low-status market on the old market is negative. The example given is Porsche’s entry into the SUV market: the target customers of the Porsche SUV were metrosexual males, who were negatively influenced by the profile of the existing adopters of the category, the suburban “soccer moms”. Modeling of this phenomenon was achieved by setting $\delta_{12}$ and $\delta_{21}$ of Eq. (3) to be positive and negative, respectively.

In addition to competing for market potential, firms can compete for one another’s existing customers for multi-purchase products such as consumer goods and services. Brand switching, also termed attrition, defection, or churn, is a major concern in many innovative industries. Industry data imply that in the U.S. mobile phone industry, the average annual attrition rate in 2005 was 26.2% (World Cellular Information Service), while the average attrition rate for U.S. companies in the 1990s was estimated at 20% (80% retention, Reichheld, 1996). Attrition and its consequences have been discussed in the CRM literature on mature markets. However, recent studies
have demonstrated that customer attrition can have a substantial effect on growing markets (Gupta, Lehmann & Stuart, 2004; Hogan, Lemon & Libai, 2003; Thompson & Sinha, 2008). Since most of the studies in the diffusion literature deal with durable goods, researchers have generally modeled the diffusion of services as if they were durable goods and have not examined customer switching (Krishnan et al., 2000; Lilien, Rangaswamy, and Van den Bulte, 2000; Jain, Mahajan & Muller, 1991). Similarly, a few studies have attempted to incorporate churn into the diffusion framework (Libai, Muller & Peres, 2009b; Givon, Mahajan & Muller, 1997; Hahn et al., 1994).

4. Directions for further research

From its inception, diffusion modeling has aimed to offer a comprehensive description of the life cycle of innovative products. In this paper, we have documented how technological developments and changes in the nature of innovations have extended the scope of classical diffusion questions. Future innovations are expected to broaden this scope still further and reveal growth patterns not previously observed. Taking the mobile phone industry example from the introduction, we expect patterns such as long-term coexistence of multiple technology generations, numerous services offered by various (sometimes competing) providers on the same equipment, and an increasing role of global considerations in the adoption process. The enhanced penetration of communication and other technological innovations in emerging economies, with their particular constraints and needs, provides a rich substrate for the development of such patterns. To remain timely and abreast of market trends, research in diffusion modeling will have to expand its horizons. In this section we propose potential directions for that expansion.

4.1. Diffusion, social networks, and network externalities: future directions

The overall economic outcome of diffusion processes is usually measured at the aggregate level. However, firms’ marketing activities often take place at the individual level. Such activities have increasingly been aimed at influencing the internal dynamics of the market (influential programs, buzz campaigns, etc.): maximizing their effectiveness requires a transition from an aggregate-level to an individual-level perspective. We believe that this transition is a promising future area of diffusion research.

Although modeling of individual adoption decisions started in the 1970s, exploration of those decisions in the context of a growing market and through the lens of individual-level diffusion remains in an early stage of development. This is mainly because it is difficult to simultaneously map networks, collect individual-level data, and track diffusion. However, the need for just such a study is increasing. The online medium offers better opportunities for such research by generating new types of individual-level data via blogs, CRM systems, and sites like LinkedIn and Facebook.

To effectively investigate individual adoption decisions, researchers should elaborate on the individual-level models by separating the adoption process into a hierarchy of effects (awareness, consideration, liking, choice, purchase, and repeat purchase), integrating into each stage findings from behavioral studies. The choice stage, for example, has thus far been explored mainly through pre-purchase experiments, such as conjoint experiments, but there are very few works (e.g., Landsman & Givon, 2010) that incorporate the choice stage into a diffusion model. Another important stage in the hierarchy of effects is the repeat purchase, which influences sales rather than adoption; this is a major source of revenue in many service and goods industries. Applying repeat-purchase individual-level mechanisms, as well as developing models for sales rather than for adoption, can assist in understanding the relative roles of repeat purchase vs. initial adoption in the diffusion process and the influences of these factors on growth and long-term profitability. Initial efforts were made by Prins, Verhoef and Frances (2009) in the context of telecommunications services. Individual-level models should also be extended to allow for flexibility in determining the unit of adoption. Traditionally, that unit has been the individual customer; however, adoption decisions are not always made by a single individual. For example, for home computing and communication products (cable TV; Wii vs. Sony PlayStation), the adoption decision is made by several household members. Innovations in business markets have to be adopted by the entire buying center and used by multiple individuals in the organization. In models, such phenomena can be interpreted either as network externalities or as examples of collective decision-making.

The structure of the social network is another factor that should be taken into account when modeling individual-level decisions, because it directly influences the speed and spatial pattern of diffusion and, as a result, the marketing decisions of the firm. If, for example, the social system is composed of isolated “islands” that communicate little with one another, the firm should launch the product separately on each such island in order to create global diffusion, whereas for other network structures, the firm might be better off enhancing internal communications. Thus, researchers should delve deeper into the structure of the specific social system involved and its influence on growth. Another important characteristic of a social network, which can considerably influence the diffusion pattern, is clustering, an element that distinguishes social networks from random graphs. Clustering means that if customer X is directly connected to Y and Y is directly connected to Z, then there is an increased probability that X is also directly connected to Z. There is little research so far that relates the degree of clustering in a given social network to the speed of diffusion of an innovation within that network. For example, Goldenberg et al. (2001a) investigated the effects of weak ties on the speed of diffusion but did not relate weak ties to clustering. Siaihk et al. (2006) studied the role of clustering, but not on the speed of diffusion. Clustering is expected to be associated with several effects: Within a cluster, the speed of diffusion should increase with the level of clustering. Between clusters, clustering should strengthen the role of weak ties since information, once entered into a cluster with a high clustering coefficient, will not likely leave the cluster unless pulled out by a weak tie. Given the availability of agent-based models for theoretical studies and network data for empirical studies, such issues appear to be ripe for new research.

One of the notable marketing phenomena of recent times is firms’ attempts to impact their customers’ word-of-mouth processes. Marketers invest considerable resources in spreading information via word-of-mouth agent campaigns, referral reward programs, programs to identify and impact influencers, online communities, viral marketing campaigns, and a range of other programs (Godes & Mayzlin, 2009; De Bruyn & Lilien, 2008; Ryu & Feick, 2007; Dellarocas, 2003). A common means of word-of-mouth initiative is sampling, which is heavily used for cosmetics, food, and online software distribution. In the case of software, a sample takes the form of a free copy of the software, either with reduced features or for a limited time. The complex nature of word-of-mouth dynamics, including difficulties in following the spread of the effects of word of mouth and a lack of established ways to measure its effects, makes the financial justification of word-of-mouth programs a pressing issue, especially since such initiatives compete for resources with traditional marketing efforts. Using agent-based modeling along with empirical verification via actual social networks, researchers are beginning to investigate approaches to quantifying the effects of word-of-mouth programs (see Watts and Dodds (2007) and Libai et al. (2009c)). One potential means of quantifying the value of a member of such a program is to observe and measure that person’s ripple effect, i.e., the number of others that he or she affects directly as well as second- and third-degree “infections”. Little work has been done so far in terms of empirically measuring the effects of such programs in general and in relation to social networks in particular.
An additional network-related issue that we believe should garner more research attention is network externalities. The empirical literature on network externalities, surprisingly, lacks evidence on individuals’ adoption threshold levels. For adoption to occur in the presence of network externalities, a potential adopter has to overcome two barriers. First, the consumer has to be convinced via the communication process that the product provides good value. Second, the consumer must be assured that the number of other adopters is such that the network product will indeed supply its potential value, i.e., it surpasses the consumer’s individual threshold level. The shape of the distribution of the thresholds within a population is of utmost importance to the speed of diffusion. To see this, imagine a distribution in which every person needs exactly one other individual who have already adopted the product in order to adopt the product as well. Under such a distribution, no one will adopt the product since all consumers would be eternally waiting for someone else to adopt and initiate their adoption process.

Given that social threshold modeling is already well grounded in the sociological literature on collective action, one would imagine that the issue of the distribution of thresholds is by now well established. Unfortunately, this is not the case. We know of only a few (partial) empirical verifications of Granovetter (1978) original claim that the distribution is truncated normal: Ludemann (1999) and Goldenberg et al. (2010). Taking on the network perspective requires disentangling the effects of each of the three types of social influences — word-of-mouth, signals, and network externalities. In aggregate models such as the Bass model, all three influences are embedded in the parameter q. While some researchers have attempted to isolate these influences, more knowledge is needed for managers to exercise some control over them. For example, different types of word-of-mouth (e.g., within- and cross-brand, positive vs. negative) should be investigated; models should separate various signals from word-of-mouth (e.g., weak ties vs. cross-country signals, or cross-brand influences), and the signaling effects of marketing mix variables should be studied. In addition, more empirical studies are needed to find ways to identify, isolate, and measure the three types of influences, their antecedents, and their influences on market growth.

Network influences and network externalities have been extensively researched in economics. One of the questions asked there concerns the existence of a tipping point or critical mass necessary to jump-start the growth process. However, the emphasis has been mainly on the state of equilibrium rather than the dynamic path toward that steady state (see, for example, Jackson (2008)). Most marketers will react with discomfort at this notion as a characteristic of real markets. This is not to say that the equilibrium concept is unimportant, but rather that it is incomplete. First, the path to equilibrium is just as important, especially when the path might take an inordinate amount of time. Second, because of the Schumpeterian “creative destruction” process, in which an innovation destructively changes economic and business environments, equilibrium is never reached in most scenarios. On the path to equilibrium, an innovation will disrupt the market, forming a new path to a new equilibrium that also will never be reached.

The switch from aggregate-level modeling to analysis of diffusion from a network perspective requires a revolution in data sources and research tools. In addition to data on individual adoptions, information on social networks should be collected. For online networks, automatic collection methods applied in computer science could be used (Oestreicher-Singer & Sundararajan, 2008, 2009). For offline networks, the task is more complicated. Some firms already insert connectivity information into their CRM databases and survey their user groups and loyalty program members regarding their social connections and other people with whom they plan to share information about new products. In other cases, researchers can collaborate with firms to collect such data. The main tools of investigation will include network simulations such as agent-based models, as well as methods that combine choice and diffusion, such as Markov chains.

4.2. Life cycle issues: future directions

Two critical stages in the life cycle of a new product are its beginning — until the takeoff point — and its substitution by a new technology generation. The buildup to takeoff does not require any consumer interaction; rather, takeoff is a result of heterogeneity in price sensitivity and risk avoidance. Specifically, as the innovation price decreases and becomes associated with less risk, the product takes off. Therefore, if one believes that both heterogeneity and communication play roles in new product adoption, then takeoff is an excellent example of an interface point between these two factors: Heterogeneity is dominant prior to takeoff, whereas consumer interactions become dominant immediately afterwards. Understanding the communication-heterogeneity interplay has evident managerial importance. If, indeed, the time to takeoff is controlled by risk and price perception, firms would do better to invest in these aspects rather than to boost market internal communication. However, takeoff studies so far have been mostly descriptive and have not addressed the underlying mechanisms. There is a need for a comprehensive theory and empirical analysis that investigates this issue directly.

As for technological substitution, while models for the diffusion of technology generations have existed for a while, major questions remain unanswered. The first question relates to the substitution process. According to traditional approaches, the new generation eventually replaces the older generation; however, this is no longer the case. For many products, old and new generations coexist for a long time. In the mobile phone industry, the number of subscribers to analog phones continued to increase long after digital technologies became available. Use of older handset types in emerging economies challenges manufacturers to cope simultaneously with multiple technology generations. The two most frequently cited models of technological substitution, developed by Norton and Bass (1987) and Mahajan and Muller (1996), are restrictive in their treatment of the coexistence of multiple generations. They also provide little insight into other substitution issues, such as leapfrog behavior, and the differences between adopter groups (e.g., new customers joining the category vs. upgraders). Moreover, generational shift at the brand level has not yet been tackled. We call for joint behavioral and modeling research efforts to understand consumer behavior under technological substitution. The findings of such studies should be combined into a unified, comprehensive model.

A second question in the context of technological substitution relates to the timing of the release of a new generation. The common wisdom among practitioners is that the firm should introduce the product as soon as it is available. This rule of thumb is supported by two main studies in the field that indicate that the firm should introduce the new generations either as soon as they are available or never (Wilson and Norton, 1989) or at maturity of the old generation (Mahajan & Muller, 1996). Inclusion of factors such as cannibalization of market potentials and competition among brands might alter these results. Studies in the entertainment industry have investigated optimal timing of the release of a new movie (Lehmann & Weinberg, 2000), but those models are distinct from technological substitution of successive generations. Normative studies regarding the optimal timing of generational entry are called for.

The third question relates to forecasting the adoption of future generations. A multi-generational category enables researchers to use data from previous generations to forecast the diffusion of future generations. In doing so, one might have to resort to semi-parametric and non-parametric models (Stremersch & Lemmens, 2009; Sood, James & Tellis, 2009).

4.3. Cross-country interactions and comparisons: future directions

Current demographic changes are affecting cross-country influences and raising new challenges for global marketers. Diffusion of
innovations in emerging economies is an increasingly important managerial interest, especially in industries such as telecommunications, where market potential in the developed world is about to approach its limit. The World Bank’s report for 2008 indicates that, while annual rates of growth in GDP in developed countries are expected to remain around 2% in coming years, emerging economies will continue to see stable and consistent growth of more than 7% per year (The World Bank, 2008). Thus, emerging economies present rapidly growing potential markets for innovations.

Diffusion of innovations in emerging economies involves unique patterns that are rarely found in developed countries. For example, while standard consumption in the mobile phone industry in the developed world is one user per handset, in many emerging economies the same mobile handset serves several family members, each with a personal SIM card. Sometimes handsets, such as those given to family members or household staff, are restricted so that they can only accept incoming calls or call certain numbers (Chircu & Mahajan, 2009). Consumers in the developed world frequently upgrade to new handsets, whereas in developing regions, there is an active market for secondhand handsets that are used with prepaid phone cards. The patterns of consumption that are ubiquitous in the developing world can act as springboards to larger questions, such as how prepaid services vs. postpaid or package deals affect consumer choices and penetration or how mobile phones can be used as financial intermediaries (Iyengar et al., 2009; Mahajan, 2009; The Economist, 2009).

Although emerging economies usually fall behind developed countries in the propensity to adopt innovations (The World Bank, 2008, p. 5), they are highly responsive to specific innovations that meet their particular needs (Chircu & Mahajan, 2009). As an example, executives of Comverse Technology recalled that early adopters of their innovative product in the mid-1990s—executives of Comverse Technology recalled that early adopters of their innovative product in the mid-1990s—voice mail—were telephone service providers from Africa and India whose subscribers used public phones and needed private voice mail boxes. Use of a single handset by multiple individuals boosted the penetration of prepaid phone cards in emerging economies and led service providers to come up with creative mechanisms for separating billing for calls from a single handset (Mahajan, 2009). To save air time, customers in emerging economies make extensive use of ringtones to transfer messages without answering the call; these customers use many more ringtones and ringtone control functions than the average user in a developed country. Despite the richness of the phenomena and the growing availability of data, the main body of diffusion research rarely relates to diffusion in emerging economies. The unique diffusion patterns in those parts of the world have been cited as anecdotal evidence but have not been incorporated into diffusion models. With the exception of a few papers (e.g., Desiraju et al., 2004), research has not addressed whether such patterns are country-specific or abundant across emerging economies. We do not know whether they are generated by growth drivers that are similar to those already identified for developed countries or what actions firms should take to maximize their profits under such growth patterns.

4.4. Competition and growth: future directions

Managerial decision-making usually applies to the firm’s own brands. As competitive structures become more complex, brand-level decision-making becomes important in optimizing managerial decision-making. We have identified several research questions related to competitive effects that would benefit from applying diffusion modeling.

The first question relates to the scope of competition: Is there a single market potential from which all brands draw, or is it a reasonable working hypothesis that each brand has its own market potential? Since having a distinct market from which to draw customers implies that competitive pressures are relatively low, it seems that when competition is intense, the common market potential hypothesis is more reasonable. However, models of both scenarios should be compared empirically by applying them to a large set of data to resolve this question. The answer to this question is important for determining the true level of competitive intensity in diffusion markets and, consequently, for optimizing firms’ decisions about marketing mix, branding and positioning. For example, if market potentials do not overlap, firms might direct fewer efforts to positioning against competition, be less aggressive in their pricing, and be more willing to cooperate with their competitors in distribution channels.

Second is the question of the influence of competition along the distribution chain. In the mobile phone industry, for example, while competing service providers distribute the same handset model, third parties offer customers real time auto-selection of the network with the best rate, so customers use the services of multiple service providers. Besides the complex cross-firm dynamics involved, such scenarios have important implications for customers’ brand perceptions in such an environment and, in turn, for the brand strategy that will optimize growth. Extending the basic diffusion model to include both multiple layers and competition would improve descriptive and normative investigations of this question.

Third is the still-open question regarding whether brand choice is a one- or two-stage process. If brand choice is a two-stage process in which consumer interactions are dominant in category choice and special offers and advertising are dominant in choosing the brand, then straightforward application of a standard diffusion model to brand-level data is problematic. Although some insights into the brand choice process derive from behavioral studies, diffusion modeling can combine brand choice and individual-level decisions and estimate their relative importance at each stage. Insights from such combined models might be striking in terms of marketing mix decisions: they may aid in establishing direct connections between product attributes, promotion campaigns and brand elements and overall diffusion.

The fourth issue deals with the nature of consumer interactions under competition. Take as an example the launch of the iPhone by Apple, one of the most significant product introductions in the U.S. in 2007. This launch relied heavily on word-of-mouth communication and buzz as opposed to paid advertising. While Apple was working under the assumption that interpersonal communication would help push its product, allowing a relatively limited investment in advertising, it appears that others in the industry reaped the benefits of this cross-brand effect. To quote Verizon spokesperson Michael Murphy, “I would have to think that a rising tide lifts all ships” (Reuters, 2007). The distinction between consumer interactions at the category level and at the brand level has received scant attention so far (Libai et al., 2009a) but is crucial for managing the growth process. A better understanding of such consumer interactions can assist in answering the following questions: To what extent do operations of one brand influence the diffusion of another? What are the implications for the amount and types of advertising used? Which firms, in terms of size, market share, and market potential, will benefit from the category- and brand-level diffusion?

With the preceding example in mind, consider a social network in which a customer spreads positive word of mouth about the iPhone. One can compute the social value of this customer as the indirect benefit that the firm receives not from his direct purchase, but rather from his word-of-mouth activities. The social value stems from the interplay of two sources: acquired customers and acceleration. Acquired customers are customers whom the focal customer helps Apple to acquire, customers whom Apple would not have acquired otherwise because they would have been acquired by competitors. In addition, this customer may cause a different set of customers to accelerate their adoption of the iPhone: those who would have eventually chosen the iPhone anyway. Untangling the effects of acquired customers and acceleration on the growth of new products is necessary for quantifying and measuring the social value of customers and the impact of word-of-mouth initiatives (see Libai et al. (2009c)).
A comprehensive brand-level perspective requires extension of the types of product categories under research. The overwhelming majority of the diffusion papers we have reviewed dealt with diffusion of durable goods and entertainment products. In practice, many products introduced to the market during the past few years have been either services, such as digital cable TV or instant messaging, or combined goods and services, such as mobile phones. Service-related behaviors, such as attrition, multiple purchases, and ongoing word of mouth, should be incorporated into diffusion modeling. First attempts have been made by Gupta et al. (2004) and Libai et al. (2009b). However, modeling should be directed more toward tying diffusion concepts into the CRM literature to describe the influence of relationship-related measures on the growth and valuation of customers and firms.

Additional product-related modeling questions address interdependencies between products. For instance, with Amazon.com offering more than 250,000 music albums as well as various choices of book formats, portfolio decisions become crucial to a new product’s success. Decisions about the allocation of marketing efforts, such as whether to focus on a limited number of blockbusters or to spread efforts over many long-tailed products, are of great managerial interest (Anderson, 2008). Extending the basic diffusion model to include multi-product interactions can contribute to this discussion.

Another characteristic of growth of successful brands today, which we expect to be enhanced in the future, is the existence of parallel “shadow” diffusion processes. Givon et al. (1995) termed shadow diffusion as a diffusion process that accompanies major diffusion growth and influences it, yet is not captured in the adoption or sales data. Although they used the term to describe piracy, we propose extending its scope to describe a wider spectrum of latent parallel diffusion processes that are part of neither sales nor adoption records. Shadow diffusion is salient in the entertainment industry, where films, books, and music CDs are advertised before they are launched so that adoption decisions are made before the product is available (Hui, Eliashberg & George, 2008). Although some appearances of shadow diffusion have been discussed in the current literature, the subject lacks thorough treatment. Future modeling research should describe the variety of shadow processes and measure their influences on the parameters and speed of the main diffusion process.

4.5. Practical applications in specific industries: future directions

A recurring call in previous reviews (Mahajan et al., 1990; Mahajan et al., 2000), which we continue to emphasize here, has been for more research and reports on the actual use of diffusion models in marketing. Diffusion models are employed in two basic ways: to develop a better general understanding of diffusion phenomena (descriptive) and to predict diffusion paths for new technologies (predictive) before there is a significant amount of data available. Though such applications are common in practice — see, for example, Ofek (2005) or Mauboussin (2004) — there is a need to clarify the actual process of using the Bass framework and its extensions to predict sales when data are not available. This applies particularly to industries in which specialized diffusion models might be needed to capture the idiosyncrasies of the industry. Two studies that have demonstrated the issues and challenges of using the Bass framework in real applications to forecast sales and adoption are Lilien, Rangaswamy and Van den Bulte (2000) and Bass, Gordon, Ferguson and Githens (2001). Three industries that seem especially of interest to new diffusion modeling efforts are telecommunications, services, and pharmaceuticals.

Telecom markets form a rich substrate of research opportunities. In most cases these markets are well documented, and thus data are relatively easy to obtain. Many market processes in telecom are regulated; in certain scenarios, this regulation keeps market variables under control, exposing hidden market mechanisms. Note that the distribution structure of the telecom industry is complex: A single 3G telephony end-user application depends on hardware and software manufacturers, service providers, compatibility issues, and global infrastructures, so specialized diffusion models might be called for.

Services are of special interest because of their wide availability and the fact that they differ from durable goods in several important aspects: First, many services involve multiple purchases. This might not have a direct effect on the number of adopters, but it certainly has an influence on the sales function. For many services, recurrent purchasing leads to the development of long-term relationships between customers and service providers; concepts and processes of relationship marketing are therefore of relevance in these cases. Though there have been few studies that model services, there is a need to incorporate CRM into the diffusion framework for models of innovation diffusion in the service sector.

In the pharmaceutical industry, a new wave of research recognizes the unique nature of health care management and especially the idiosyncrasies of its supply chain. Thus, new approaches and models have been formulated for modeling diffusion of pharmaceutical products. See, for example, the special issue of this journal on Marketing and Health (Stremersch, 2008; Vakratsas & Kolsarici, 2008; Grewal, Chakravarty, Ding & Liechty, 2008). Solid pre-launch forecasting methods are still needed, especially due to the pharmaceutical industry’s high investment in R&D in very early stages of the drug development process.

4.6. Afterthought

We have been writing review papers on research in innovation diffusion with other researchers for the past thirty years (Mahajan & Muller, 1979; Mahajan et al., 1990; Mahajan et al., 2000; Muller, Peres and Mahajan, 2009). One could ask, what will the scope and content of a review paper written in 2020 be? The answer to such a question depends on the following:

• The development of new, complex types of product categories. Each such new category will illuminate a different aspect of the diffusion process and raise new research questions. Current examples are the iPhone, which offers over 70,000 apps; the Kindle, with over 360,000 books; and cheap mobile phones with multiple SIM cards for multiple users.

• The emergence of new modes of interdependencies between consumers that connect their utility, consumption and communication patterns. Current examples are social network communities such as Facebook, Amazon book recommendations and the quintessential micro-blogging service Twitter.

• The continuing growth of data from social networking sites, mobile phone service providers, e-mail communications, online communities and search engines such as Google Trends. Such data will open numerous new possibilities for exploring individual adoption decisions and linking them to overall diffusion patterns.

• The integration of cutting-edge modeling tools from other research domains. Current examples are agent-based modeling and network analysis as new tools for modeling growth of new products in networked markets, and nonparametric regressions as forecasting tools.

• The level of firms’ intervention in the internal dynamics of the market. The emergence of “amplified word of mouth,” in which firms use tools such as seeding programs to affect diffusion, is currently a highly popular market trend. Should this trend continue, diffusion studies will pay more attention to optimization questions on social networks. However, should this intervention prove unprofitable or impractical or face legal and regulatory hurdles, researchers and practitioners will pursue other influence mechanisms.
The driving force behind diffusion research, as with any other research field, is an enthusiastic community of inquisitive researchers with a command of cutting-edge research tools, who inspire each other in extending knowledge boundaries and pursuing intriguing questions. We hope that this review has provided an overview of the collective effort of this community in the last decade, as well as a glimpse of what is yet to come.

Acknowledgements

The authors would like to thank the then editor, Stefan Stremersch, the area editor and two anonymous reviewers for a number of thoughtful suggestions and comments. The authors also greatly benefited from the comments and suggestions on previous drafts by Bart Bronnenberg, Deepa Chandrasekaran, Jacob Goldenberg, Towhidul Islam, Trichy Krishnan, Barak Libai, Philip Parker, Ashutosh Prasad, Arvind Rangaswami, John Roberts, Gerard Tellis, Christophe Van den Bulte, and Charles Weinberg.

References

Ch/></div>