No. 044/2020/MKT

# Diffusion of Innovations: Modeling, Estimation, and Normative Developments

Dipak C. Jain*
Department of Marketing
China Europe International Business School (CEIBS)

Trichy V. Krishnan
Department of Marketing
Great Lakes Institute of Management

Sophie Feng
Department of Marketing
Monash Business School

December 2020

* Corresponding author:
1. Dipak C. Jain (dipakcjain@ceibs.edu). Address: 699 Hongfeng Road, Pudong, Shanghai, China, 201206

# Diffusion of Innovations:
# Modeling, Estimation, and Normative Developments


**Dipak C. Jain***

**Trichy Krishnan****

**Sophie Feng*****

***President (European), China Europe International Business School, Shanghai, China**

****Professor of Marketing, Great Lakes Institute of Management, Chennai, India**

*****Faculty, Monash University, Melbourne, Australia**



**January 2021**

**ABSTRACT**

Understanding the adoption process of a new product and modeling the resulting diffusion pattern over time has been of great interest to both academicians and practitioners for the past 50 years. The keen interest in modeling diffusion of innovations would continue to grow in the new digital economy also, perhaps more vigorously given the increased speed with which information spreads on social media.

We are presenting a research manuscript that has three purposes. The first is to give a treatise on the interesting model developments that have happened in modeling diffusion of innovations. Noting that diffusion modeling is one of the strong theory-based quantitative investigations that has evolved in the marketing literature, we explain how theoretical modeling by various researchers has advanced the knowledge on new product sales growth.

Second, given that no theoretical model can claim to be valid unless it is supported strongly by empirical data, we discuss in detail the development of estimation procedures as they pertain to offering validation support to the diffusion of innovation models. From the application perspective of diffusion models, we also look at the forecasting implications and normative developments, e.g., optimal marketing mix policies.

Third, acknowledging that there have been comprehensive review papers in the diffusion of innovations area, including: Mahajan and Muller (1979), Mahajan and Wind (1986), Mahajan, Muller and Bass (1990), Mahajan, Muller and Bass (1993), Mahajan, Muller and Wind (2000), and the latest one being Meade and Islam (2006). Our objective is not to <u>just</u> do another review article on diffusion models but rather concentrate on emerging areas such as: multigeneration diffusion models, social media related product diffusion, and repeat purchase behavior, to name a few. Our focus is on building strong analytical models with theoretical insights and in-depth empirical analysis.

Our work is likely to be more useful for the PhD students and researchers interested in new product adoption and diffusion. We focus our discussion on the following three areas: ***model formulations***, ***estimation methods*** and ***normative implications***.

## INTRODUCTION

Rogers (1983) conducted several field experiments and found that a new product was not getting adopted immediately by all but rather gradually, in a systematic way. He also found that the adoption process was largely influenced by word of mouth (WOM) spreading from adopters to potential adopters. Rogers theorized that a new product diffuses through five groups of users namely: *innovators*, *early adopters*, *early majority*, *late majority* and *laggards*, in that order. Of the 5 groups, the innovator group, Rogers argued, adopted the product first by looking at how the product was useful to them and its role in their lifestyle. Then they make a decision of whether to adopt. Once they decide to adopt the product, the other four groups are influenced by WOM and chose to gradually adopt over time.

Frank M Bass (1969) gave a mathematical structure to the Roger's findings after collapsing the last four groups into one group and calling it "imitator group". Thus, Bass took the Roger's finding to mean that a small fraction of a market adopted a new product without getting influenced by WOM and that the rest of the market adopt it after getting enough information through WOM. He called the two groups *innovators* and *imitators* respectively. He used this theory to develop a mathematical expression for the adoption of a new product by an individual and derived a diffusion model that described the periodic sales growth of the product at the market level. He further offered strong support to his model through fitting it on eleven new consumer durable goods' sales growth data.

The rather strong theoretical foundation and empirical support helped the Bass (1969) model to set off an unprecedented research movement in Marketing and elsewhere. This resulted in the Bass model of the 1969 article, which was published in Management Science, getting the maximum ever citation count in Management Science. Till today, it holds that honor.

## 1: Diffusion Models

Bass (1969) asked how a given individual at time t would go about deciding to adopt a new product introduced at time 0. Bass took the Rogers' main finding on "innovators and imitators" dichotomy in the marketplace and transformed that to mean that the individual at time t would be influenced partly by innovative thinking and partly by WOM i.e. imitating tendency. He captured the individual's new product adoption decision in a hazard framework as follows.

$$\frac{f(t)}{1-F(t)} = p + qF(t) \tag{1}$$

where f(t) is the probability that the individual would adopt the new product at time t, and [1-F(t)] is the probability he had not adopted the product until time t. Thus, if we take t is the random variable of "time to adoption", f(t) is the pdf (Probability Density Function) and F(t) is the cdf (Cumulative Distribution Function). The left-hand side function is thus the *hazard* that an individual would adopt a new product at time t given that she has not adopted that product until time t. The right-hand side has two components: a) "q F(t)" captures the WOM influence of previous adopters on the individual, where q is the parameter that captures the strength of the WOM influence generated by F(t), b) the parameter p captures all those non-WOM forces influencing the individual's decision. These include advertising, salesforce efforts (especially in a B2B context), brochures and information on the website. Bass referred p and q as the coefficients of innovation and imitation respectively.

Recognizing that f(t) is actually dF(t)/dt, the differential equation (1) can be solved to yield:

$$F(t) = \frac{1-\exp[-(p+q)t]}{1+\frac{q}{p}\exp[-(p+q)t]} \tag{2}$$

This is the probability an individual might have adopted by time t. The pdf is obtained by differentiating F(t) with respect to t, resulting in:

$$\boldsymbol{f(t)} = \frac{[\frac{(\boldsymbol{p}+\boldsymbol{q})^2}{\boldsymbol{p}}]\boldsymbol{e}^{-(\boldsymbol{p}+\boldsymbol{q})\boldsymbol{t}}}{[\mathbf{1}+\frac{\boldsymbol{q}}{\boldsymbol{p}}\boldsymbol{e}^{-(\boldsymbol{p}+\boldsymbol{q})\boldsymbol{t}}]^2} \qquad (3)$$

Multiplying equation [2] by market potential, m, we get:

$$CS(t) = m\,\frac{1-\exp[-(p+q)t]}{1+\frac{q}{p}\exp[-(p+q)t]} \qquad (4)$$

In equation [4], m is the market potential i.e. the total number of adopters in the target market, and CS(t) is the cumulative number of adopters until time t.[1] Interestingly, Mansfield (1961) proposed a model that was same as equation [1] but in deriving a closed form solution, Mansfield made an assumption that yielded a different model (i.e. a model without p) which became a simple logistic regression equation.

The sales growth pattern of a new product is shown in Figure 1:

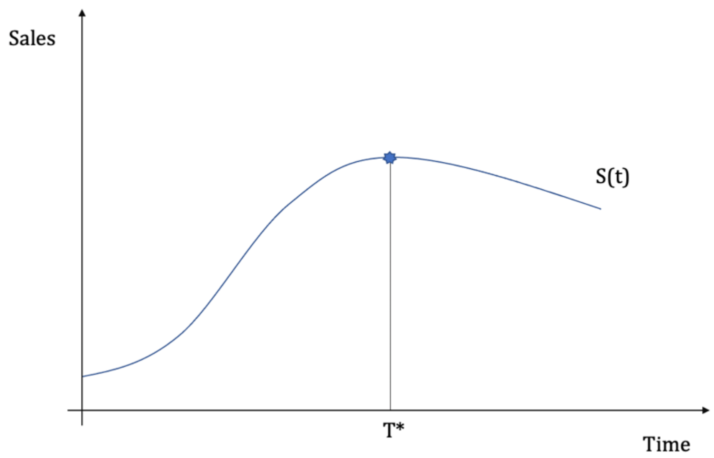



Figure 1: Sales Growth Curve

In Figure 1, T* is the time to peak sales. To find T*, differentiating f(t) where f(t) is given in equation (3), and setting the result equal to zero will yield:

[1] For a durable good like a smart phone or a laptop, the number of adopters would equal the sales of the product because an adopter would buy *only* one unit in general. This is true of adopting a service product like signing up for a club, cell phone services, Facebook or WhatsApp. Hence, for durable goods and services, we can say that CS(t) represents Cumulative Sales up to t.

$$T^* = \frac{1}{p+q}\ln(\frac{q}{p}) \tag{5}$$

$$S(T^*) = \frac{m(p+q)^2}{4q}$$

Substituting T* in equation (2), we get:

$$F(T^*) = \frac{(q-p)}{2q}$$

And cumulative sales will be:

$$CS(T^*) = mF(T^*) = \frac{m(q-p)}{2q} = \frac{m}{2}(1-\frac{p}{q})$$

Given that p is normally smaller than q, cumulative sales can be approximated as,

$$CS(T^*) = \frac{m}{2}$$

Therefore, cumulative peak sales of a new product will be half the total market potential.

Noting that S(t) = m f(t) and CS(t) = m F(t), where m is the market potential, equation [1] becomes:

$$S(t) = mp + (q-p)CS(t) - (q/m)CS^2(t) \tag{6}$$

Equation [5] was used by Bass (1969) to estimate the model with sales data in eleven consumer durables, where he found a high R-square indicating an excellent fit of the model) to the annual sales data[2]. RCA, the leading Color TV manufacturer of those times, and WSJ mentioned about the usefulness of the Bass model for forecasting purposes.

### 2.1 Model Extensions: Role of Marketing Mix variables

After the rather very successful empirical demonstration of the diffusion model by Bass in 1969 and acknowledgement of the same by the industry, there was surprisingly no follow up work for almost 6 years. Robinson and Lakhani (1975) of Kodak used the Bass model to derive optimal pricing policies for their new cameras. An interesting thing is that the original Bass model (1969) doesn't have price in their function, and so Robinson and Lakhani (1975)

[2] Equation [4], which is a better equation to use in estimation, was introduced 15 years later by Srinivasan and Mason (1986).

inserted the price in a rather *ad-hoc* way in the Bass model equation [1] and proceeded to derive the optimal dynamic pricing policy for a new camera. We use the phrase *ad-hoc* for the reason that the authors didn't test the model with empirical data. Later, other researchers carried out empirical analysis but couldn't find support for this model[3].

In fact, in the two decades following the publication of the Bass model, only a few research articles offered empirically supported models while investigating the role of marketing mix variables in the Bass model. Almost all of the models followed the Robinson and Lakhani (1975) approach in that their main focus was on deriving optimal policies without empirically justifying them with empirical data[4].

Bass (1980) proposed that the marginal cost decreasing over time with sales volume would encourage firms to decrease the price accordingly, and used that principle to infer the role of price. Horsky and Simon (1983) incorporate advertising in the Bass model by making the coefficient of innovation (p) a function of advertising. They tested their model empirically as well as derived the optimal advertising policy.

Horsky (1990) proposed that income-wage distribution would affect the market potential while Kalish (1985) used the reservation price framework to explain the role of price on adoption. Although the theoretical frameworks used in these two models are plausible, the empirical exercise carried out are difficult to replicate.

Jain and Rao (1990) focused on empirically testing out if price should be added to the Bass model to affect the market potential or the rate of adoption.

[3] Although one could propose a model that logically explains a marketing process, providing good empirical support for the model is important to justify the mechanism underlying the model because there could be many such competing models one could come up with. In such a case, strong empirical support for a model would provide the researchers and the managers enough confidence in the model and in the mechanism proposed in the model.

[4] Later we will show how optimal policies depend critically on the model one uses to derive those policies although all models try to explain the same marketing process.

Two research articles had a more comprehensive approach to incorporate marketing mix variables in the Bass diffusion model: Bass, Krishnan and Jain (1994) and Bass, Jain and Krishnan (2000).

Bass, Krishnan and Jan (1994) turned the puzzle upside down and asked: Why does the Bass model fit without decision variables? Analyzing the unique structure of the Bass model, the researchers found that the only explanatory variable, i.e. t in the Bass model (see equation [3]) was somehow capturing also the impact of the decision variables such as price and advertising over time. This is a classic case of the widely known statistical principle that if omitted variables are correlated with the included variables the model would provide a good fit to the data even in the absence of those omitted variables. This is of course good news for forecasting purposes because the forecaster now doesn't have to worry about the other variables but guess the values for the three parameters Bass model, namely, {p, q, m} and use equation [4] and forecast the sales growth into future. However, for controlling the sales growth process, managers need a modified Bass model that is theoretically sound and empirically proven. It was provided by Bass, Krishnan and Jain (1994). The model is given below.

$$\frac{f(t)}{1-F(t)} = [p + qF(t)]\, x(t) \tag{7}$$

where x(t) was called marketing efforts at t (or current marketing efforts), which was formulated as:

$$x(t) = 1 + \beta_1 \frac{1}{\mathrm{Pr}(t)} \frac{dPr(t)}{dt} + \beta_2 \frac{1}{\mathrm{adv}(t)} \frac{d\, adv(t)}{dt} \tag{8}$$

where $\frac{1}{\mathrm{Pr}(t)} \frac{dPr(t)}{dt}$ is the percentage change in price at time t, and $\frac{1}{\mathrm{adv}(t)} \frac{d\, adv(t)}{dt}$ percentage change in advertising at time t, and the new parameters $\{\beta_1, \beta_2\}$ are respectively the impact of those price and advertising variables. This was called "Generalized Bass Model" because when price and adv change at more or less a constant rate over time, these two factors would be a constant, yielding x(t) to be some constant. In such a case, equation [6] would become observationally equal to the Bass model [1] [5].

[5] In order to accommodate the fact that advertising is likely to have a carry-over effect implying that an occasional drop in advertising might not mean negative influence on sales the variable $\frac{d\, adv(t)}{dt}$ in expression [7]

Solving the differential equation [7] with [8], we get:

$$F(t) = \frac{1 - \exp[-(p+q)X(t)]}{1 + \frac{q}{p}\exp[-(p+q)X(t)]} \tag{9}$$

where,

$$X(t) = t + \beta_1 Ln\frac{Pr(t)}{Pr(0)} + \beta_2 \frac{adv(t)}{adv(0)} \tag{10}$$

GBM has found wide support in the academia thanks to its theoretical appeal on explaining why the Bass model is so successful without including any decision variables.

Bass, Jain and Krishnan (2000) proposed another interesting framework that stems from the hazard model literature. They claim that the Bass (1969) model represents the basic hazard process while the marketing variables act on that basic process like in a proportional hazard model. This is as given below.

$$h(t\,|x(t)) = h(t)\,\exp[\beta_1 \mathrm{Pr(t)} + \beta_2 \mathrm{adv(t)}] \tag{11}$$

where h(t) on the righthand side is the base level hazard function as given by Bass model (equation [1]) and h(t|x(t)) on the left hand side is the hazard that includes the impact of price and adv. This model has proved to be more versatile than GBM, especially in estimating the role of decision variables on the diffusion.

---

was devised to have the lower bound of zero. This of course can be modified to suit different conditions of advertising expenditure policies.

A list of models proposed in the literature with empirical support and normative implications is presented in Table 1.

**Table 1**
**Diffusion of Innovations: Model Extensions**

| Research | Modeling Approach | Specific Formulation |
|---|---|---|
| Robinson and Lakhani (1975) | Experience Curve and Price-Normative Model | Price affects the rate of adoption multiplicatively. |
| Bass (1980) | Experience Curve and Reservation Price –Empirical Analysis | Price acts multiplicatively on the rate of adoption. |
| Dolan and Jeuland (1981) | Price – Normative Model | Price has a multiplicative effect on the adoption rate. |
| Bass and Bultez (1982) | Price – Normative Model and Empirical Analysis | Price and rate of adoption have a multiplicative relationship. |
| Horsky and Simon (1983) | Advertising – Normative Model and Empirical Analysis | Advertising effects the coefficient of innovation |
| Tang and Thomson (1983) | Advertising – Normative Model | Advertising affects the diffusion coefficients |
| Kalish (1983) | Price – Normative Model | Price affects the market potential and the rate of adoption multiplicatively |
| Kalish (1985) | Price and Advertising – Normative Model | Price affects the market potential, the rate of adoption is controlled by advertising. |
| Simon and Sebastian (1987) | Advertising – Empirical Analysis | Advertising effects the coefficient of imitation |
| Horsky (1990) | Income-Wage Distribution – Normative Model and Empirical Analysis | Existence of income-price effect on market potential is hypothetical and tested. |
| Jain and Rao (1990) | Price – Empirical Analysis | Price effects market potential and rate of adoption |
| Jones and Ritz (1991) | Distribution – Empirical Analysis | Market potential is a function of distribution growth. |
| Krishnan, Bass and Jain (1994) | Price and Advertising – Empirical Analysis | Accelerated Hazard Model Formulation |
| Krishnan, Jain and Bass (1999) | Price and Advertising – Normative Model | Advertising and price affect rate of adoption |

| | | |
|---|---|---|
| Bass, Jain and Krishnan (2000) | Price and Advertising – Empirical Analysis | Proportional Hazard Model Formulation |
| Danaher, Hardie and Putsis (2001) | Price – Empirical Analysis | Proportional Hazard Framework |
| Krishnan and Jain (2006) | Advertising – Normative Model | Advertising affects rate of adoption |

**2.2 Modeling other aspects**

There are many other aspects that are of interest to academicians and practitioners as well concerning a new product. We list a few of them here first and take up for discussion one by one to explain how researchers handle them.

**Micro-modelling Adoption**

Although Bass model and its extensions were found to be successful and useful in describing the sales growth process of a new durable, it is interesting that they don't describe the adoption of a new product at an individual level. Bass model is actually founded on the individual level adoption process but by assuming that all individuals are the same, the individual level differences get washed out by the aggregate model. In other words, the aggregate model ignores the rich information contained in the differences among individuals with respect to how they view a given new product. This is a little troubling because we don't set to gain insights that could answer questions such as "are there individuals who are more risk averse than others so a manager could approach them with a different strategy?"

In order to address this critical issue Chatterjee and Eliashberg (1990) advanced an interesting model that goes to the heart of the individual's decision regarding adopting a new product. They start with a utility model for an individual that builds on how far the information gathered by the individual affects his perception of the product's expected performance, and augment that utility with a price function. The individual is modeled to update his utility in a Bayesian manner as more and more information bits come to him, and upon reaching a threshold the individual is said to adopt the product. There is unobserved stochasticity in the information process, represented by Weiner process, and this creates the heterogeneity among the individuals. This leads to a nicely structured individual-level adoption model, which when aggregated gives Bass model under certain conditions. The authors did a pilot study with Wharton MBA students to demonstrate how their model could be used and estimated, leading to rich insights that could not be obtained from using aggregate models.

**Diffusion Models with Repeat Purchase**

Even though Bass acknowledges that his model is concerned with initial purchases only, the Bass Model is frequently applied to fit sales data that counts both initial purchases and repeat purchases. Such applications are reasonable when repeat purchases are infrequent or negligible, which might be the case for durable products during the early stages of product sales.

For products that have frequent upgrades or multiple versions, repeat purchases can constitute a significant proportion of sales even in the early stages of product diffusion, thus limiting the application of the Bass Model. When significant repeat purchases exist, the fitting and forecasting performance of the Bass Model may degrade, particularly when the sales trend begins to assume an asymmetrical shape after the peak point of diffusion. Even if the model fits and predicts well in some situations despite the existence of considerable repeat purchases, the parameter estimates will be different from those obtained using only initial purchases data.

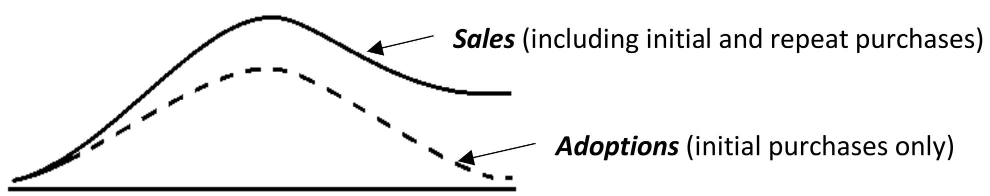
In the prior literature, researchers have proposed model extensions to account for repeat product purchases made to replace existing product units or for adopting multiple product units. Olson and Choi (1985) propose a model assuming that sales comprise only adoptions and replacements, and that the replacement hazard function follows the Rayleigh distribution (Papoulis and Pillai 2002). The Olson-Choi Model is developed for cases in which the number of products in use is available.

Kamakura and Balasubramanian (1987) propose a similar model that generates long-term forecasts by incorporating the adoption and replacement components of sales. Benefiting from a more flexible hazard function to capture replacement purchases, their model is applicable with or without data for replacement sales. Their model uses information from similar products when data for replacement purchases is not available. Steffens and Balasubramanian (1998) further advance the modeling of replacement sales by allowing the distribution of the service life of replaced products to change over time.

Focusing on the PC processor industry, Gordon (2009) introduces a dynamic structural model that explicitly considers product replacement decisions under uncertain future product quality and price. Gordon’s model concentrates specifically on product replacements due to obsolescence as a result of product upgrade releases. The model requires highly specific data to function as it uses a composite dataset including sales, ownership, price, and product quality. Due to data limitations, Gordon’s model does not consider multi-unit ownership purchases.

The aforementioned models have their focus on adoptions and replacement purchases. There are other models from the prior literature that have considered other components of sales, in particular multi-unit ownership purchases, in addition to adoptions and replacements. Dodson and Muller (1978) propose a model that captures the type of asymmetric sales trend illustrated by the solid curve in Figure 2 without decomposing sales into multiple components. The market they consider is composed of three groups, i.e., those who are not aware of the product, those who are aware of the product but have not made a purchase, and those who have purchased the product. Although the portrayed interaction between adopters and non-adopters is insightful, their model cannot be operationalized if the data for the different market divisions is unattainable.

**Figure 2. Sales vs. Adoptions**



Based on econometric and simulation models, Bayus, Hong, and Labe (1989) incorporate first time sales, replacement sales, additional-unit sales, and institutional sales into their analysis of color television set sales. As stated by Steffens (2003), Bayus et al.'s model is developed to perform well over short terms.

Steffens (2003) presents a model for sales resulting from multi-unit ownership purchases based on the Bass Model. His model differs from Bayus et al.'s in that it imposes a saturation level on multi-unit ownerships, which makes the model applicable for longer time frames.

Researchers have also developed models for repeat purchases of products for a specific industry. For instance, Lilien, Rao, and Kalish (1981) propose a highly specialized model to project sales of prescription drugs as a function of a focal pharmaceutical company's own detailing effect, competitors' detailing effect, and word of mouth. In a subsequent study, Rao and Yamada (1988) provide further empirical support for the model by Lilien, Rao, and Kalish, then propose an alternative method for developing priors for the new drug's parameters, and  demonstrate how the parameters can be updated after sales data becomes available. Because Lilien, Rao, and Kalish's and Rao and Yamada's models are specifically designed for prescription drugs, they cannot be used to predict sales of general product categories.

A detailed review of the literature suggests that there is a need for a comprehensive model that can be fit to aggregate sales data that records both initial purchases and repeat purchases including both replacements and multi-unit ownerships.

Jiang, Aslan and Jain (2019) proposed a model termed as the Generalized Diffusion Model with Repeat Purchases (GDMR). The development of this model is based on a branch of mathematics named fractional calculus. Specifically, the model generalizes the fundamental differential equation governing the Bass Model, and employs a non-integer integral operator with flexible order, thereby rendering the Bass Model a special case of the extended model. The GDMR adopts an approach that is different from those in the prior literature in that it captures the sales growth rate using a non-integer order integral equation, rather than an integer-order differential equation as used in the prior literature.

**Multi-Generation Diffusion of Innovations Models**

Technological advances fuel the development of new products and services. Examples are abundant. Decades ago, black-and-white TV was replaced by color TV, which is now ceding market share to HDTV. Even within the HDTV category, newer models are continuing to emerge, with the most recent variety 3D-capable, due to even more technologies. In the cellular phone market, the earliest generation was only equipped with basic calling features, the following generation enhanced to include cameras, media players, etc., while the newest generation, called smart phones, allow users to surf the Web, and run more sophisticated applications. The same phenomenon also exists in the software market, where vendors keep releasing new versions to meet users' ever-increasing appetite for functionalities and take advantage of improvements in hardware technologies. The Microsoft Windows and Office lines of products are two well-known examples, with new version typically introduced every few years.

The diffusion of successive product generations has been well studied in the prior literature. Most of the existing multigeneration diffusion models are inspired by the seminal Bass model (Bass 1969). Among them, the model proposed by Norton and Bass (1987) (NB model for short) is often credited as the pioneering work in describing multigenerational diffusion. The NB model assumes that each generation has its own market potential and market penetration process, and adopters of earlier generations can shift to newer generations. After Norton and Bass, several other notable multigeneration diffusion models have been proposed. Speece and MacLachlan (1995) extend the NB model to incorporate the influence of pricing and test it with multigenerational data for fluid milk packaging technologies. Mahajan and Miller (1996) develop a model that captures the number of systems-in-use for each generation and use it to study the optimal market entry timing for successive generations. Jun and Park (1999) combine the diffusion effects and choice effects and propose two integrated models: the Type 1 model distinguishes first-purchased demand and upgrade demand while the Type 2 model does not. Kim et al. (2000) propose a dynamic market growth model that captures not only the diffusion of multiple generations within the same product category, but also the complementarity and competition present by related product categories. Danaher et al. (2001) develop a two-generation model that includes both first-time sales and periodic renewals. By selecting appropriate adoption time distributions, their model can also incorporate the impact of market mix variables. More recently, Jiang (2010) proposes a simple two-generation model to analyze the optimal free offer policy for successive software versions.

Despite the progress made in the last two decades, a review of the literature reveals that the NB model remains the most tested and extended multigeneration diffusion model to date. We believe that the desirable mathematical properties (e.g., offering closed-form expressions, parsimonious, and continuous-time based) of the model plays a key role behind its popularity. However, the NB model is not applicable to all business scenarios. This is primarily because when counting the number of adopters who substitute an old generation with a new generation, the NB model does not differentiate those who have already adopted an earlier generation and those who are first-time adopters of any generation. In their study, Norton and Bass do acknowledge the existence of two different types of substitutions but admit that their model does not differentiate them (Norton and Bass 1987, p. 1074). Without such differentiation, the NB model cannot be used to estimate the number of cross-generation repeat purchases, nor can it help forecast future demand or project revenue for certain business scenarios (e.g., when revenue is generated through both product sale and after-sale service). Jiang and Jain (2010) propose a Generalized Norton-Bass (GNB) model that overcomes this limitation while retaining the desirable mathematical properties of the NB model. As we will demonstrate later, the proposed model offers greater flexibility in parameter estimation, forecasting, and revenue projection for a wider range of scenarios.
A couple of other areas that need more attention are a) *multinational diffusion models* b) *stochastic diffusion models* and c) *successive new product generations models*. We summarize the research work in these areas in Tables (2), (3) and (4), respectively. Researchers interested in these areas can build on the work highlighted in the respective tables.

**Table 2: Multinational Diffusion Models**

| Research Article | Focus |
|---|---|
| Gatignon, Eliashberg and Robertson (1989) | Modeling Multinational Diffusion Patterns |
| Takada and Jain (1991) | Cross-National Analysis of Diffusion of Consumer Variables in Pacific Rim Countries |
| DeKimpe, Parker and Sarvary (1998) | Estimation of International Diffusion Models |
| Eliashberg and Helson (1996) | Modeling Lead-lag Effects in Global Markets |
| Jain and Maesincee (1997) | Lead-lag Time Effects on Global Product Diffusion |
| DeKimpe, Parker and Sarvary (2000) | Multimarket and Global Diffusion |
| Talukdar, Sudhir and Ainslie (2002) | Investigating New Product Diffusion Across Products and Countries |
| Kuman and Krishnan (2002) | Multinational Diffusion Models |

**Table 3: Stochastic Diffusion Models**

| Research Article | Focus |
|---|---|
| Eliashberg and Chatterjee (1986) | Stochastic Issues in Innovation Diffusion Models |
| Boker (1987) | A Stochastic First Purchase Diffusion Model |
| Niu (2002) | A Stochastic Formulation of the Bass Model of New-Product Diffusion |
| Niu (2006) | A Piecewise-Diffusion of New-Product Demands |
| Raman and Jain (2010) | A Stochastic New Product Growth Model |

**Table 4: Results of Successive New Product Generations Research**

| Research Studies | Research Focus | Product/Service Analyzed (Examples) |
|---|---|---|
| Wilson and Norton (1989) | Optimal entry timing for a line extension of a new product | Publishing firm's decision on introducing hard cover and paperback versions |
| Moorthy and Png (1992) | Sequential vs. Simultaneous introduction of new product extensions | a. IBM PS/1, PS/2 Game Computer<br>b. Minolta single lens reflex cameras<br>c. Volvo 6-cyllinder 760 and 4-cyllinder 740 models |
| Mahajan and Muller (1996) | Optimal timing of successive generations of technological innovations | d. Four successive generations of IBM mainframe computers |
| Morgan, Morgan and Moore (2001) | Quality and Time to Market Trade-offs Made for Multiple Product Generation Launches | Auto, medical devices, cellular phones and integrated circuit chips |
| Prasad, Bronnenberg and Mahajan (2004) | Timing strategies for new product versions through multiple channels | Timing of video release after a theatrical release of a movie |
| Krankel, Duenyas and Kapuscinski (2006) | Successive Product Introductions Timings under Demand Diffusion and Technology Improvements | IBM's mainframe computer |
| Mehra, Seidmann and Mojumder (2014) | Optimal upgrade intervals for successive software versions | Product lifecycle management of packaged software |
| Jiang, Qu and Jain (2019) | Optimal market entry timing for successive generations of technological innovations | Technology products (e.g. computers, Microsoft Office 365, software licenses) |

## 3: Estimation Methods

However logically correct a theory is, the world would not accept it wholeheartedly unless it gets empirical support. This is very much true for a marketing model like the Bass model because of its practical application in business. Empirical support means that a proposed theoretical model is able to explain the underlying real phenomenon happening in the marketplace. In order to do that, we first gather data of the phenomenon from the real world, and statistically see if the proposed model “fits” with the data. We do the empirical test with data sets from multiple categories. In this chapter, we will analyze the various statistical procedures proposed in the literature to empirically test the Bass model and a few other extended models.

### 3.1 OLS (Ordinary Least Squares)

Bass (1969) proposed the fundamental diffusion model in a differential and solved the resulting differential equation to get a closed form expression for the Sales in time domain. However, for empirical testing he chose to use equation [1] and used Ordinary Least Squares (OLS) method after converting equation [1] to a linear model as follows.

$$\frac{f(t)}{1-F(t)} = p + qF(t)$$

$$S(t) = m\,f(t) = m\,[1-F(t)][p + qF(t)]$$

$$S(t) = mp + (q-p)CS(t) - (q/m)\,CS^2(t)$$

Note that in equation [9], CS(t) is cumulative sales at t and is m F(t), while sales is S(t) which is m f(t). It is easy to see that expression [9] is a linear equation where the dependent variable is S(t) and the independent variable is CS(t) and that the intercept is “m p” and the two slope coefficients are “q-p” and “-q/m”.

$$\boldsymbol{S(t) = \alpha + \beta CS(t) + \gamma CS^2(t)}$$

*Where* $\boldsymbol{\alpha = mp, \beta = q - p \text{ and } \gamma = -q/m}$

*Solving* $\boldsymbol{\alpha, \beta \text{ and } \gamma \text{ in terms of } p, q \text{ and } m}$

$$\boldsymbol{m = \frac{-\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\gamma}}$$

$$\boldsymbol{q = -\gamma m}$$
$$\boldsymbol{and\ p = \frac{\alpha}{m}}$$

Bass fitted this linear model on 11 new products' sales data and found excellent fit. By using sales data of 11 new products to offer empirical support for his model, Bass (1969) showed the versatility of the proposed model to track sales growth of different new products and the robustness of the model. Note that OLS was not, as researchers later showed, a good estimation technique for the inherently non-linear Bass model and yet the fit was excellent, and the parameter estimates had plausible values with expected signs. This further demonstrated the flexibility and robustness of the Bass model.

**3.2 NLS Estimation**

In their 1986 paper, Srinivasan and Mason (SM) offered an elegant non-linear regression to estimate the Bass model. The non-linear least squares (NLS) method used the closed form solution of the Bass model (equations [2] and [4]) i.e. the sales in time domain as follows. The equations are reproduced here for convenience.

$$S(t-1,t) = CS(t) - CS(t-1)$$

*where,*

$$CS(t) = m\ \frac{1-\exp[-(p+q)t]}{1+\frac{q}{p}\exp[-(p+q)t]}$$

The NLS method showed remarkable improvement in the fit and the parameter estimates (less standard error). Most importantly, it enabled much better forecasting ability with the Bass model, which we will explain now.

Jain and Rao proposed the following non-linear formulation of the Bass model

$$S(t-1,t) = (m - CS(t-1))\frac{[CS(t) - CS(t-1)]}{[1 - CS(t-1)]}$$

Where CS(t) is given in equation (4). NLS method was used to estimate the three model parameters, p, q and m.

A model's empirical performance depends on how cleverly it uses the information contained in the data to estimate its parameters. In that context, Jain and Rao (1990) model outperforms many other models. It also uses NLS but in a much more clever way and hence yields better estimates.

### 3.3 Bias in Estimates

Van den Bulte and Lilien (1997) embarked on a rather unique research. They tested if the Bass model parameters estimated from data suffered from any systematic bias induced by the number of data points used in the data set. They used NLS-SM estimation method and used 11 real data sets and many simulated data sets to show that as the number of data points used for estimation increased the m (market potential parameter) estimate was systematically obtaining a higher value, and the q (WOM parameter) was obtaining a lower value. These findings are naturally important from forecasting perspective. However, Krishnan and Feng (2010) showed that inappropriate pooling of various data in the Van den Bulte and Lilien (1997) study might have led to their conclusion on the systematic bias. More research needs to be done in this area to figure out if the problem is for real, and, if yes, if the cause was estimation procedure or something else.

### 3.4 MLE (Schmittlein and Mahajan, 1982):

Noting that each adoption is an outcome of a binomial choice –adopt or not- at a given time for a given individual, they formulated a likelihood function from the Bass model and used MLE to estimate the model parameters from a given data set. Although many researchers prefer to use NLS over MLE mostly for simplicity, the discussion of which estimation procedure is better is still going on. For example, Hardie, Fader and Wisniewski (1998) found maximum likelihood to be noticeably better than NLS when applied to period-by-period sales. However, over shorter series, they found NLS applied to cumulative sales comparable with MLE.

### 3.5 Augmented Kalman Filter with Continuous State variable and Discrete observations:

AKF(C-D): AKF(C-D) was borrowed from engineering and introduced by Xie, Song, SIrbu and Wang (1997) to offer another estimation procedure for the Bass model. Interestingly, this method uses the differential equation form of the Bass Model and not its closed-form solution, although it is not clear if the superior performance would remain the same if one use AKF(C-D) on the latter form. AKF(C-D) acts on a closed-loop feedback mechanism, where each of the three parameters get updated by the difference between the predicted one-step ahead forecast and the actual sales in the following period. At the start, each parameter is given a prior chosen judiciously inferred from the diffusion of previous analogous products or from meta-analytic studies like Sultan, Farley and Lehmann (1990). These are used on the Bass Model to forecast the first year's sales, and the difference between this forecasted first year sales and the actual first year sales, the parameters get updated through an updating function. And the process gets repeated year by year until all the observations are used up. The authors show the superior performance of the AKF(C-D) by comparing it with OLS, NLS-on-SM and MLE. However, the authors use the accuracy of the one-step ahead forecast as the metric for this comparison, which is a little unfair because other estimation procedures use the "maximizing the overall fit of the model with the data" as their objective while AKF(C-D) uses exactly the one-step ahead forecast accuracy as its objective to update the parameters. Further, the AKF(C-D) procedure assumes observation error and parameter estimates' errors, and it is not clear how far these assumptions affect the performance. Also, if a manager wants to do multi-period ahead forecasting, it might be difficult to employ this procedure because it's basic premise is updating the next period performance. A comprehensive list of estimation procedures is presented in Table 2.

**Table 5**
**Diffusion Models: Parameters Estimation Methods**

| **Estimation Methods** | **Research Articles** |
|---|---|
| Ordinary Least Squares (OLS) | Bass (1969)<br>Bass (1980)<br>Horsky and Simon (1983)<br>Norton and Bass (1987)<br>Simon and Sebastian (1987)<br>Kamakura and Balasubramaniam (1988)<br>Horsky (1990)<br>Jones and Ritz (1991)<br>Bass, Krishnan and Jain (1994) |
| Maximum Likelihood Estimation (MLE) | Schmittlein and Mahajan (1982)<br>Hardie, Fader and Wisniewiski (1998) |

| | |
|---|---|
| Non-linear Least Squares (NLS) | Srinivasan and Mason (1986)<br>Jain and Rao (1990)<br>Van den Bulte and Lilien (1997)<br>Bass, Jain and Krishnan (2000) |
| Hierarchical Bayes (HB) | Lenk and Rao (1990) |
| Augmented Kalman Filter (AKF) | Xie, Song, Sirbu and Wang (1997) |

## 4: Normative Implications of Diffusion Models

Theoretical models like diffusion models will not be appreciated by practitioners unless they could be put to use to address some of the practical managerial issues. Diffusion models can be ideally used to address two such managerially relevant problems: forecasting and optimal marketing mix (pricing and advertising) policy decisions.

### 4.1 Forecasting

Frankly speaking, forecasting is as much an art as it is science. It is especially true with a new product where the consumer himself carries a lot of uncertainty with the product. Further, the sales growth of a new product very much depends on how the product manager employs the strategy in the marketplace, and a forecaster will not have access to those strategies. Having said that, managers would find it immensely useful if we could provide them with a fairly reliable tool to forecast the sales of a new product, stating therein the assumptions made. To that extent, let us now see how we can use the Bass model for the forecasting purposes.

If a new product finds some traction in the marketplace, the product manager might like to know how the sales would grow over time so she could plan the production and distribution activities appropriately. Diffusion models perform their best in this forecasting exercise. Taking a simple approach, if we could somehow get the three main parameters of the Bass model, namely, {p, q, m}, for the given new product, we can easily draw out the annual sales of the product using expressions [4] and [6]. Where do we get these three parameters?

Initially from some analogous products introduced earlier in the same marketplace to the same target segment, and later when some actual sales number of first few periods are collected, using those sales numbers one can getter better estimates for the three parameters. One may wonder why the Bass model is the preferred model for forecasting? The answer to this question is rather straightforward. Recognizing that the Bass model is simple, theory-based, robust and is driven by only one independent variable, time, we can comfortably say that the Bass model is the best model to use for forecasting purposes. We may have other models like GBM that have successfully incorporated the influence of price and advertising, but if we want to use them for forecasting purposes, we have to necessarily predict how those additional variables, price and advertising, also would change over time. On the other hand, Bass model does not require them. Its parameters already reflect their impact implicitly. Secondly, the Bass model is a very robust model. Krishnan and Bass measured this robustness and compared that with robustness of other models.

When the product shows clear signs of success after the first 3 or 4 periods, the manager would be interested to know when the sales growth would hit the peak. i.e. Peak-Sales Time (T*), as presented in equation (5). Knowing the T* for the given product can help a manager to (a) plan the production capacity and schedule, (b) know when to invest in the next new product and (c) decide on the right pricing and promotion strategies to face the competition that is likely to arise near the peak sales time. In other words, forecasting the T* using early years data might be the most important forecasting exercise for a product manager to undertake. Krishnan, Feng and Jain (2020) show how one could go about doing this forecast.

### 4.2 Optimal Pricing

Of all the marketing mix variables, perhaps the max attention is paid to 'price' because it is one element that directly and immediately affects the sales and, further, the impact can be measured. Starting with Robinson and Lakhani (1975), many optimal pricing paths have been advanced in the literature. These include the articles presented in Table 6.

**Table 6**
**Diffusion Models: Normative Implications**

| Marketing Mix Variable | Research Articles | Normative Findings |
|---|---|---|
| Price<br>(Reference Price, Experience Curve) | Robinson and Lakhani (1975)<br>Bass (1980)<br>Bass and Bultez (1982)<br>Dolan and Jeuland (1981) | Optimal Pricing Policy |

| | | |
|---|---|---|
| | Jeuland and Dolan (1982)<br>Kalish (1983)<br>Kalish (1985)<br>Horsky (1990) | |
| Advertising<br>(Cumulative Advertising) | Dodson and Muller (1978)<br>Horsky and Simon (1983)<br>Simon and Sebastian (1987)<br>Krishnan and Jain (2006) | Optimal Advertising Policy |
| Price and Advertising | Kalish (1985)<br>Krishnan, Jain and Bass (1999) | Optimal Pricing and Advertising Policies |

As mentioned earlier, many of the earlier models were developed by including price variable into the Bass model in a plausible and logical way but without providing empirical support for the same. For example, the optimal policy derived from these models point to something interesting. They recommend that the price should increase as long as the sales keeps increasing and should start decreasing after the sales goes beyond the peak and decrease. In other words, they derive an optimal policy that mimics the sales pattern. It of course makes sense because the sales anyway keep increasing until the peak and hence it would seem prudent to increase the price that time. However, simple economic theory suggests that price increase or decrease should be a function of how sensitive the demand is to price changes. In terms of diffusion, it would mean that price change should depend on how effective the price is in affecting the rate of increase in the sales. Recognizing this, Krishnan, Bass and Jain (1997) used the GBM (1994) and derived the dynamic optimal pricing for a new product. Their results suggest that optimal price would increase initially (as others had suggested) but would start decreasing after some time, not necessarily after peak as others had suggested. The transition time could be even zero implying that the optimal price would decline monotonically right from the introduction time. When exactly the price would transit from increasing to declining trend is a function of sensitivity of demand to price and the future-discount rate.

### 4.3. Optimal Advertising

Compared to the optimal pricing literature, number of researchers worked on deriving optimal advertising is rather limited. There are only three distinct articles in principle: Teng and Thomson (1983), Dockner and Jorgensen (1988) and Krishnan and Jain (2006)[6]. Of the three, the first two papers used the following extended Bass model for deriving their optimal advertising policies.

$$S(t) = [p1 + q1\,F(t) + \{p2 + q2F(t)\}f(a)][M - CS(t)] \qquad (12)$$

In equation [12], f(a) is a function of advertising which acts on one part of the diffusion force represented by innovation and imitation effect, and the rest of the parameters are basically the Bass model. If f(a) is zero, we get the Bass model. Teng and Thomson (1983) used a linear function for f(a) while Dockner and Jorgensen (1988) let it free and be a concave function. Depending upon the relative values of {p2, q2} with respect to {p1, q1}, the authors claim that optimal advertising policy is determined.

Teng and Thomson (1983) results suggest that if advertising is more impactful on imitators optimal advertising should be minimum in the early and final periods and maximum in the middle periods i.e. Min-Max-Min optimal policy. Under the same condition of advertising more impactful on imitators, Dockner and Jorgensen (1988) results suggest a monotonically increasing advertising policy. If the advertising is more impactful on innovators, Dockner and Jorgensen (1988) suggest a monotonically declining advertising policy. A main shortcoming of these results is that the model [equation 10] used was not empirically tested and hence, as mentioned earlier, managers would find it difficult to follow them.

Krishnan and Jain (2006) used the empirically proved GBM to derive their optimal advertising policy. Their results suggest a variety of paths for optimal advertising that depend on how effective is advertising on the diffusion, future discount rate and the ratio of advertising to sales ratio (A/S ratio). They show how different combinations of these three forces result in either increase-decrease or decrease-increase policies, where monotonically increasing is a special case of the increase-decrease policy and the monotonically declining policy is a special case of the decrease-increase policy.

[6] Horsky and Simon (1983) is a special case of the Dockner and Jorgensen (1988) and hence is not considered separately. Similarly, Thomson and Teng (1984) has similar model and results as those of Teng and Thomson (1983) and hence is not considered separately. We will discuss these two papers later, however.

### 4.4. Joint Optimal Pricing and Advertising

It is very challenging to derive jointly optimal price and optimal advertising policies. In Section 4.2 where we discussed optimal pricing policies, all of the results assume that advertising does not interfere with the price-to-sales impact directly, and probably advertising influence the price sensitivity. In Section 4.3 where we discussed optimal advertising policies, the researchers had assumed constant price or price declining at a constant rate. Haodong and Krishnan (2012) derived joint optimal pricing and advertising policies.

## 5. Summary and Directions for Future Research

The purpose of this article is to highlight the developments in the area of diffusion of innovations and its contributions to the new product marketing literature. Specifically, we focus on:

- model formulations that capture the global diffusion dynamics
- robust estimation procedures for empirical analysis
- normative models that yield optimal marketing mix policies

Moving forward, we would like to develop models that take into consideration

a) **The role social media plays in the marketplace**: new product diffusion models need to incorporate ***word of mouse*** in addition to ***word of mouth*** that researchers have dealt with extensively in previous work.
b) **The market entry timing decisions for introducing successive generations**: an area of strategic importance for firms because among other reasons, the introduction of a new generation has the potential to cannibalize the sales of current offerings, thereby affecting the firm's total revenue model. (e.g., Apple's iPhone, Samsung's Galaxy)
c) **Market entry timing strategy for new products starting with quality variation**: new version of better quality than previous version (e.g., high-definition TV vs. standard TV) or new version of poorer quality than previous one (e.g., hard cover textbooks first and paperback edition later)

d) **Product lifecycles are getting shorter**: Given the large investments in developing a new product, there might be financial risks associated with its premature or delayed introduction.
e) **New technological developments**: Impact of artificial intelligence (AI), machine learning (ML), and internet of things (IoT) on new product diffusions.

Successful launch and management of successive generations of technological products is important for firms to sustain its revenue stream and profitability as well as maintain a loyal customer base. Although, product lifecycles may be getting shorter due to rapid technological advancements, but customer lifecycles are becoming longer due to enhanced longevity resulting from health science research and advanced medical innovations. Enriching total customer experience requires firms to focus on R&D for new products and have a proper plan for entry timing for successive product generations. Successful launch of successive product generations also effects product branding as well as corporate branding in achieving a favorable, innovative image in the marketplace.