Journal of Business Research 165 (2023) 114054

Contents lists available at ScienceDirect

Journal of Business Research

journal homepage: www.elsevier.com/locate/jbusres

Peak sales time prediction in new product sales: Can a product manager rely on it?

Trichy V. Krishnan [a], Shanfei Feng [b,*], Dipak C. Jain [c]

[a] Rajagiri Business School, Cochin, Chennai, Rajagiri Valley P. O, Kakkanad, Cochin 682 039, Kerala, India
[b] Department of Marketing, Monash Business School, Monash University, PO Box 197, Caulfield East, VIC 3145, Australia
[c] European President and Professor of Marketing, CEIBS, 699 Hongfeng Road, Pudong, Shanghai 201206, PR China

ARTICLE INFO

Keywords:
New product diffusion
Peak sales time
Prediction accuracy

ABSTRACT

Managers dealing with new products need to forecast sales growth, especially the time at which the sales would reach the peak, known as the peak sales time (T*). In most cases, they only have a few initial years' data to predict T*. Although product managers manage to predict T*, there is no method to date that can predict T* *accurately*. In this paper, we develop a new metric based on the diffusion modeling framework that can help in assessing the prediction accuracy of T*. This metric is built on the premise that observed sales growth is affected both by the force that systematically varies with time and by the non-systematic random forces. We show that the two forces must be carefully combined to assess if a predicted T* is accurate enough. In addition, we empirically prove the efficacy of the proposed metric.

1. Introduction

Until the 1990s, much of academic research in marketing revolved around proposing a theoretical model to explain an observed marketing phenomenon and proving the efficacy of the model with the empirical data pertaining to the phenomenon. Theory was the principal focus, an in-sample empirical fit being secondary. A classic example is the Bass model (Bass, 1969) where a diffusion theory was proposed to explain the new product sales-growth pattern, which was empirically tested with multiple data sets. In such theory-based research, the predictive ability of the model was not the major focus. However, in recent decades, with the advent of real-time big data, there has been a shift in managers' thinking: the predictive ability of a model has become more important. In the light of the shift, modelers have been employing analytical tools such as ANN (artificial neural networks), AI (artificial intelligence), and ML (machine learning), and developing model-based algorithms primarily focused on their predictive ability. In other words, the objective of developing a model has shifted from pure theoretical prediction or explanatory aspects to feasible predictive performance.

Applying this shift in the perspective of a manager to the case of new product marketing, managers would like to predict, for example, the success of a new product before launch or the time when its sales would reach the peak. In our research, we focus on peak sales time, which is defined as the point in time when sales grow to a peak and then level off at some magnitude lower than the peak (Bass, 1969, p. 215). It is typically denoted by T*.

Interestingly, marketing researchers have not paid much attention to T* prediction. Fortunately, in the diffusion literature a strong theory for the existence and location of T* has been established. Starting from Mansfield (1961) and then Bass (1969), it has been shown that almost every new durable product's diffusion in a marketplace experiences two factors as their sales grow in the market, one of them being word-of-mouth (WOM) and the other being market saturation. We expect the sales to grow initially (when WOM is stronger than market saturation) and then decline (when market saturation starts dominating), and the transition point is T*. At T* sales reach the peak, and then start to stabilize or decline.

Marketing researchers have developed various models, for example Bass (1969) and Bass, Krishnan and Jain (1994), over the past six decades to capture these two factors in the diffusion process, and a few of them yield a closed-form solution, i.e., a direct function for T*. If we have such a diffusion model that gives T* as a direct function of time, we can theoretically "link" the initial periods' sales data to T* in two steps. In the first step, the model is regressed on the initial years' data (i.e., pre-peak sales data) to estimate the parameters of the model, and in the second step, the parameter estimates are used to estimate T* using the

* Corresponding author.
E-mail addresses: krishnan61@gmail.com (T.V. Krishnan), shanfei.feng@monash.edu (S. Feng), dipakcjain@ceibs.edu (D.C. Jain).

https://doi.org/10.1016/j.jbusres.2023.114054
Received 16 June 2021; Received in revised form 17 May 2023; Accepted 19 May 2023
Available online 28 May 2023

functional form of T*. Thus, a closed-form solution is good to have for prediction purposes because it avoids unnecessary complications like cascading errors encountered in stepwise iterative prediction.

However, a challenge faced here is that the number of initial years' data is usually small—anywhere between 5 and 10—and hence any noise associated with those few data points would seriously affect the accuracy of T* prediction. Note that noise in a data point refers to those disturbances that are not captured by the diffusion model. Given this challenge, a key question that arises is: *Would that T* estimate be accurate enough for a manager to plan for the new product launch and act on it?*

The question of accuracy becomes challenging because there are two forces influencing the growth dynamics. One is the *diffusion force*, which is captured by an analytical model, and the other is the *market noise*. Unless the manager understands the relative influence of each of these two factors, it is difficult to judge the accuracy of the predicted T*. We propose a theoretical metric that captures the relative impact of these two forces.

In this research, we contribute to the diffusion literature in the following ways. First, we focus on prediction of peak sales time T*. Although T* is strategically important for a firm to make decisions concerning production, inventory level, and marketing, it has not received much attention from researchers in the field of marketing. As presented in Table 1, diffusion researchers have focused on developing theoretical models that could explain the observed sales growth data patterns of new durable goods. When the in-sample fit analysis provided sufficient evidence to support their proposed model, researchers used one-step-ahead forecasting for additional support (Lattin and Roberts, 1989; Young, 1993; Meade and Islam, 1995)[1].

Two other interesting findings have emerged from Table 1: (i) Simple models are generally better than complex models for sales forecasting. (ii) A forecast that is a weighted combination of forecasts from a few carefully chosen models outperforms the forecast from any individual model in the group (Meade and Islam, 1995). However, there is no clear procedure that can be developed from the findings to extrapolate the sales forecasts to T* prediction.[2]

The second contribution *is a metric that we propose to enable managers to assess the accuracy of the predicted T**. We develop a theoretical formulation for the metric using the diffusion framework. Interestingly, the proposed metric is not limited to any specific diffusion model. However, we apply it to the well-known Bass model for demonstration purposes, allowing managers to use this framework on any diffusion model for a given application. We further use multiple data sets to offer empirical support for the proposed metric's ability to assess the accuracy of a predicted T*. Investors and product managers can use this metric and make more informed decisions on their new product marketing strategies.

The rest of the paper is organized as follows. In Section 2, we present the importance of knowing T* in new product marketing and show how the proposed framework contributes to the literature. In Section 3, we develop the metric to assess the accuracy of a predicted T* and derive a functional form for it. In Section 4, we provide empirical support and in Section 5, we provide guidelines useful to managers derived from our empirical data. In Section 6, we conclude the paper with key insights and propose directions for future research in this area.

2. Importance of knowing T*

In the 2004 Management Science article, Frank Bass, writing a commentary on his hugely successful model, mentioned the following (Bass, 2004, p.1834):

> "… *[In 1966] I decided to try my luck at forecasting the sales of color television…Sales data were available for three years: 1964, 1965 and 1966… The result was a forecast that color television sales would peak in 1968 at 6.7 million units…Industry people were more optimistic… As it turned out, color television did peak in 1968 but at a level slightly lower than my forecast… The industry had built capacity for 14 million color picture tubes and there was substantial economic dislocation following the sharp downturn in sales following the 1968 peak…*"

The above incident, as explained by Bass, clearly shows the importance of accurately predicting the peak sales time (T*) for a firm, especially for its production and operations planning department. If we notice carefully, the incident talks about over-capacity for picture tubes, an important vendor-component of the television those days. This implies that an inaccurate forecasting of T* can affect not only the company marketing the new product, but also the whole set of vendors supplying various components to the company.

The Bass (1969) article set the stage for a large stream of research to start flowing, which is continuing until now unabated. Particularly, the T* prediction that Bass boldly made for the color television helped in this cause immensely. For an excellent summary of the various modeling developments and ideas in new product sales research, see Meade and Islam (2006).

Ho, Savin, and Terwiesch (2002) analyzed in detail the operations and production issues concerning a new product and derived optimal pre-production level, production capacity, and launch time for the new product. They assumed that T* was known and thereby derived the optimal policies. Negahban and Smith (2018) extended the optimal production capacity issue to the case of multiple generations, which further supported the importance of knowing T* for each generation.

One of the important frameworks that academicians and practitioners use is the product life cycle. It has been demonstrated that a firm should change its marketing strategies when a new product enters the maturity stage, i.e., close to T*. (Kotler and Keller, 2016) This is important also from a competition perspective because at the maturity stage, the competitors start indulging in price-based competition, leading to lower profits for all the players in the marketplace. This further suggests that the firms should predict T* carefully.

Many new products and services launched by firms today are generally improved versions of earlier forms and today's products will be replaced by newer offerings soon. Familiar examples include, Apple's iPhone and iPad, Samsung's Galaxy, Kellogg's breakfast cereals, and Pepsi's healthy snacks. Of the many issues surrounding multi-generational products, market entry timing decision for a new generation has continued to attract the attention of researchers and practitioners over the recent years (Wilson and Norton, 1989; Mahajan and Muller, 1996; Prasad, Bronnenberg and Mahajan, 2004; Kalish and Lilien, 1986; Moorthy and Png, 1992; Krankel, Duenyas and Kapuscinski, 2006; Qin and Nembhard, 2012; Negahban and Smith, 2018; Jiang, Qu and Jain, 2019; Schwarz and Tan, 2021; Bersch, Akkerman and Kolisch, 2021). Timing of introduction of a new generation is of strategic importance because a new generation has the potential to cannibalize the sales of the current product, and, given the large investment in developing a new product, there might be financial risks associated with premature or delayed introduction (Mahajan and Muller, 1996).

Researchers have found that the optimal time for introducing a new generation depends on when the current generation is going to peak. Hence, an accurate prediction of T* would lead to better financial

[1] For example, Bass, Krishnan, and Jain (1994) developed the Generalized Bass Model to explore the role of marketing mix variables on diffusion, tested their model with three data sets, and then showed how the generalized model outperformed the Bass model both in fit and one-step-ahead forecasts. It is important to note that forecasting was not their focus, but a method to show superiority of their model over the Bass model.

[2] Can sales forecast over multiple periods in the future lead to T* prediction automatically? Not really, because one-step-ahead forecasting is not enough to infer the maturity phase. Even if one were to iteratively extend the one-step-ahead forecast to future periods, i.e., moving from one period to the next till we reach the peak sales and beyond, the cascading forecasting errors would make the exercise less reliable.

Table 1
Review of Sales Forecasting in the Diffusion Literature.

Research Articles	Focus	Data Length	Type of Product	Key Results
Young, 1993	One-step-ahead sales forecasts and 3-step-ahead sales forecasts	Data length can be inferred to be around 15.	Not mentioned	Bass model performs better if data have only pre-peak sales data.
Bass, Krishnan, and Jain, 1994	One-step-ahead sales forecasts	10 to 13 periods	Household consumer durables	Generalized Bass Model performs better than Bass model in forecasting, but it has additional explanatory variables.
Meade and Islam, 1995	Sales forecast over 10–11 years	30 years for national data or 72 quarters for regional data	Adoption of telephones (PBX and PBAX systems)	Simple models (models with 2 or 3 parameters) are better than complicated models (models with 4 parameters).
Hardie, Fader, and Wisniewski, 1998	Forecasting of weekly cumulative trials	52 weeks (13 – 26 weeks calibration period)	Consumer Packaged goods	Simple models are better for forecasting, while complex models are better for calibration data fit.
Meade and Islam, 1998	Comparison of different models in terms of model fit and one-step-ahead sales forecasts	Varied (14–46 time periods)	Multiple product categories	Forecast obtained by combining forecasts from different models weighted by fit and stability of the constituent models performs better than forecasts of individual models.
Bass, Jain, and Krishnan, 2000	One-step-ahead sales forecasts	10–13 periods	Household consumer durables	Proportional Hazard model outperforms Generalized Bass model in forecasting.
Bewley and Griffiths, 2003	Forecasting penetration	13 years	CD penetration in 12 countries	Box-cox transformation variant of the FLOG (flexible logistic) model is better for forecasting.
Trusov, Rand and Joshi, 2013.	Pre-launch sales forecasts	Daily data for 4–8 months	Simulated data and Facebook apps	If market conditions are stable, network structure extracted from historical data can help improve predictions.
Toubia, Goldenberg and Garcia, 2014.	Forecast penetration	Uses individual-level data, tracking 398 consumers for weeks on social interaction	Consumer durables	Proposed method does better in forecasting than the traditional models but social interactions data need to be collected, which may not be readily available.
Xiao and Han, 2016	One-step ahead and k-step ahead forecasts using network analysis, also called agent-based model (ABM), that are built using social media and big data	Multiple periods	Consumer durables	ABM network predicts better than traditional models (Bass, Weibull, etc.) in one-step ahead forecasts but they are not manager-friendly yet.
Ramírez-Hassan and Montoya-Blandón, 2020	Pre-launch sales forecast (one-step-ahead and total market size).	0 to 11 time periods for performance comparison	Simulated data sets; consumer durables and natural gas market	Information on similar products or homogeneous markets can be used to forecast sales (and even pre-launch forecast) using Bayesian approach.

Table 2
Products Used in the Empirical Analysis.

#	Product (j)	First-use Decade	Observed T* (Year …)	No. of Data Sets (T^{*}–5 + 1)	No. of Usable Data Sets
1	Answering machine	1980	9	5	4
2	VCR	1980	20	16	13
3	Desktop PC	1990	12	8	7
4	Fax machine	1990	11	7	7
5	Cordless phone	1980	14	10	10
6	DVD player	2000	9	5	5
7	Digital camera	2000	10	6	5
8	CD player	1990	16	12	11
9	Cell phone	1990	15	11	8
10	Digital TV	2000	13	9	9
11	Cell-Finland	1980	12	8	8
12	Internet-China	2000	12	8	8
13	Cellphone-China	1990	14	10	10
14	Printer	1990	13	9	4
15	Cellphone-India	2000	10	6	6
16	iPhone	2000	9	5	5
17	Fitbit	2010	7	3	3
18	Smartphones	2000	12	8	7

Sum = 146 Sum = 130.

decisions and enhance the firm's profitability. An example that illustrates the role created by T* is presented by Jain and Rao (2020) regarding the launch timing of revised textbooks. Jain and Rao (2020) provides a simple framework for determining the optimal timing for a new textbook revision, while the current version would still be selling in the secondary markets as used-text books. The publisher would like to have the current version in the market long enough to recover as much profit as possible realizing the negative effects of the used book sales. Jain and Rao (2020) suggests that before launching the textbook, publishers should estimate the size of the potential market for the new version. Since by the peak sales time, approximately 50 % of the market would have purchased the book, Jain shows that the new and revised edition be launched at peak sales time or anytime soon after that. This example clearly highlights the importance of knowing, i.e., predicting T* in determining the optimal market entry timing for successive generations of new products.

It is rather intriguing that despite the importance of knowing or predicting T* in the areas of Operations and Strategic Marketing, research articles simply assume that T* is known and go about deriving optimal strategies. If firms use a T* that is not accurate, they are likely to encounter such major issues as faced by the color television industry in the last century. In this research, our objective is to propose a metric that will help managers assess if a predicted T* is accurate and hence reliable.

3. Analytical model formulation: Assessing the accuracy of predicted T*

Our model development rests on the following premise: *if early years' sales data carry a strong "signal" and, moreover, are affected least by "random noise", one would be able to make a more accurate prediction of T* using the evolution process of those sales data points*. In a given data set, we can observe *how far the signal overwhelms the noise*, prompting us to claim that a high Signal-over-Noise ratio might indicate a high predictive ability.

In the context of the sales growth of new products, the signal is the diffusion force, which is powered by the *voice* of the adopters through WOM. Hence, we paraphrase the Signal-over-Noise as Voice-over-Noise (VON) for the purposes of our research.

A key question is: how to measure this VON ratio?

We need a metric that can give us the VON ratio in a given data set *as* *it pertains to the predictive ability*. To devise the new metric, we judiciously use the diffusion theory that underlies the new product sales growth pattern and the noise level in the data set.

3.1 Lemma: With a product experiencing a non-linear, systematic sales growth that has a natural peak, the time to peak and the average growth rate until that peak are related to each other uniquely.

Proof: Let r denote the average periodic growth rate in the interval [0, T*]. Let S(0) be the sales at the launch time, i.e., t = 0, and S(T*) be the sales at the peak, T*. Then, the sales growth from S(0) to S(T*) can be expressed as follows.

$$S(T^*) = S(0) \bullet (1+r)^{T^*} \tag{1}$$

Taking the natural log of both sides of Equation (1) and rearranging the terms, we get:

$$Ln(1+r) = \frac{1}{T^*} Ln\left(\frac{S(T^*)}{S(0)}\right) = \frac{k}{T^*} \tag{2}$$

where $k = \mathrm{Ln}[\mathrm{S}(T^*)/\mathrm{S}(0)]$. For a given value of k, the rate of growth r and T* are directly related to each other, although inversely.

Some important points are to be noted from this lemma.

1. We are not making any inference on what S(0) or S(T*) is. No mathematical formulation or link between the two is assumed at this point.
2. We assume the existence of a peak due to the S-shaped curve commonly observed in the sales of new products (Chandrasekaran and Tellis, 2018).
3. The growth rate, r, assumes near-monotonicity of sales growth.

3.1. New product sales growth rate, r

Let us look at the forces that influence the observed sales growth rate, *r*. As explained earlier at the beginning of this section, we focus on two forces, namely, Voice of the adopters and the Noise in the data set. Let us call them simply the Voice(V)-Force and Noise(N)-Force We explain each force now.

V-Force: It has been well documented, starting from as early as the 1962 findings of Everett Rogers(Rogers,1962), that WOM plays a critical role in the new product adoption rate. As more and more people adopt the product, WOM increases leading to more adoptions, and thus WOM monotonically grows with time. However, given that there are only a finite number of first-time adopters for a given product, with every adoption the yet-to-adopt market shrinks, and this shrinking happens monotonically with time. Thus, at a given point in time in the [0, T*] interval, the growth rate is influenced by two forces, namely, WOM and yet-to-adopt market size, both of which are monotonic functions of that time but in opposite directions. We call these forces that vary systematically over time V-Forces.

When Bass (1969) formulated his famous diffusion model to accommodate these two forces in an elegant mathematical framework, he obtained new product sales as a function of just one independent variable: time *t*. As an empirical proof of how the V-Force lends such power to the time variable, Bass (1969) and, later, scores of other researchers fitted the Bass model to sales data of hundreds of different new products and almost always found the fit to be very good, yielding significant estimates for the Bass model parameters and proving that the natural time explains the underlying pattern of the sales growth of a new product well.

Bass, Krishnan, and Jain (1994) wondered how the Bass model was able to successfully explain the sales growth of a new product without including marketing mix variables. They came up with the explanation that in many cases, the effect of marketing mix variables got subsumed in the time variable. This can happen, for example, when:

a. Firms decrease price at a constant periodic rate owing to technological improvements and/or economies of scale.
b. Firms decide on their advertising budget as a constant percentage of the sales revenue of the previous year.

Thus, the V-Force also includes all those effects that vary systematically with time and hence could be empirically captured by the Bass Diffusion Model.

N-Force: The second force that influences sales growth occurs nonsystematically with time. This includes the effect of marketing mix variables that are sporadic and other random factors. The nonsystematic marketing mix variables include promotions offered for a certain period to clear the inventory and seasonal events like Black Friday during Thanksgiving. Stores offer special promotions and also try to get the attention of consumers through displaying the new product at the beginning of an aisle or in a focal place within the retail store. We call this the N-force to indicate its *nonsystematic* nature[3].

To summarize, the V-Force has three components, namely, WOM, yet-to-adopt market (i.e., untapped market potential), and time-varying marketing mix variables. The V-Force gives the trend to the sales growth rate. Since the N-Force has no systematic pattern, it can be treated as a discrete noise term at a given point in time. Given that both the forces influence the sales growth rate, we next express them in a mathematical form.

3.2. Expressing the components of r in mathematical terms

Let us look at the two sales data points of interest, S(0) to S(T^*), which determine the growth rate. First, consider S(0), which has $S_V(0)$ and $S_N(0)$ as its components, where, $S_V(0)$ is the effect of the V-Force and $S_N(0)$ is the effect of the N-Force on sales at time 0. Since the N-Force is a nonsystematic discrete noise term, we assume it to be a random variable. Noting that the annual sales of a new consumer durable grow from hundreds in the launch period to millions at T*, it is prudent to include the nonsystematic component multiplicatively, i.e., proportionally to the systematic component (Van den Bulte and Lilien, 1997; Venkatesan, Krishnan, and Kumar, 2004; Niu, 2006). Tian et al. (2013) have demonstrated analytically that a multiplicative error model is better than an additive formulation, especially when large-range data sets are involved, which is true of typical diffusion data sets. Consequently, we propose.

$$S(0) = S_V(0) \bullet S_N(0) \tag{3}$$

Taking the log on both sides of both the expressions in Equation (3), we get:

$$Ln[S(0)] = Ln[S_V(0)] + Ln[S_N(0)] \tag{4}$$

In Equation (4), $Ln[S_N(0)]$ follows the normal distribution when we assume $S_N(0)$ to follow a log-normal distribution with mean 1 and variance $\exp(\sigma^2) - 1$, in line with the multiplicative inclusion of noise. The corresponding normal distribution will have mean $(-\sigma^2/2)$ and variance σ^2.

Similarly, we can split S(T^*) as follows.

$$Ln[S(T^*)] = Ln[S_V(T^*)] + Ln[S_N(T^*)] \tag{5}$$

where $Ln[S_N(T^*)]$ is normal error with mean $(-\sigma^2/2)$ and variance σ^2.

From Equations (4) and (5), we get:

[3] It is perhaps important to note that researchers (Niu, 2006; Qin and Nembhard, 2012) use specific stochastic processes to study the inherent uncertainty in diffusion rate, which reinforces the importance of considering more carefully the noise in the sales data rather than treating them just from the perspective of regression. Further, these studies have not extended their results to help managers evaluate the reliability of T* prediction although Qin and Nembhard (2012) attribute the uncertainty of predicted T* to model selection.

$$Ln\left(\frac{S(T^*)}{S(0)}\right) = Ln\left(\frac{S_V(T^*)}{S_V(0)}\right) + Normal(0, 2\sigma^2) \tag{6}$$

because the mean of the differences between the two normal error terms that are IID is zero, and the variance of that difference is the sum of the two variances[4].

Having expressed the starting point (t = 0) and end point ($t = T^*$) of the sales growth in terms of V-Force and N-Force, we now substitute Equation (6) in the growth rate Equation (2) to get:

$$Ln(1+r) = \frac{1}{T^*}Ln\left(\frac{S(T^*)}{S(0)}\right) = \frac{1}{T^*}Ln\left(\frac{S_V(T^*)}{S_V(0)}\right) + \frac{1}{T^*}Normal(0, 2\sigma^2) \tag{7}$$

The first term on the right-hand side in Equation (7) captures the trend component in the sales growth and so can be adequately explained by a simple but powerful diffusion model like the Bass model. The second term captures the random, discrete spikes on the sales growth curve. The standard deviation of the noise term is $\frac{\sqrt{2}}{T^*}\sigma$. It is rather easy to see that the variance or unpredictability of r depends on how far the righthand side of Equation (7) is dominated by the second term.

3.3. Proposed metric: Voice-over-Noise (VON) ratio

We first define our metric that could be used to assess the reliability of T* as follows. We call it VON, to indicate the relative influence of "V-Force Over N-Force".

$$VON = \frac{Effectsoffactorsthatvarysystematicallywith Time}{Noise(nonsystematiceffect)} \tag{8}$$

$$VON = \frac{\frac{1}{T^*}Ln\left(\frac{S_V(T^*)}{S_V(0)}\right)}{\frac{\sqrt{2}}{T^*}\sigma} = \frac{1}{\sqrt{2}\sigma}Ln\left(\frac{S_V(T^*)}{S_V(0)}\right) \tag{9}$$

VON is the metric we propose to assess the accuracy of T* prediction. As mentioned in the beginning of this section, we have not used any diffusion model to derive our metric expressed in equation [9]. Now, we will apply the Bass Diffusion Model to evaluate equation [9].

3.4. VON for the Bass model (1969)

One can use any proven diffusion model to evaluate $\frac{S_V(T^*)}{S_V(0)}$ and σ. However, what we need is a parsimonious yet powerful model that is robust and well-proven across hundreds of categories in various countries over different eras. All these requirements are met by one model, i.e., the Bass model (1969). It states that the probability of a potential adopter adopting a new product at time t given that they have not adopted it yet directly depends on the WOM generated from the cumulative number of adopters until time t. This gives the following sales function.

$$S(t) = [M - CS(t)]\left[p + q\frac{CS(t)}{M}\right] \tag{10}$$

where *S*(t) is the sales function, *CS*(t) is the cumulative sales until time t representing the influence of previous adopters at *t*, and the three parameters, namely, *p*, *q*, and *M*, are the coefficient of innovation, coefficient of influence of previous adopters, and market potential, respectively. Noting that $S(t) = \frac{dCS(t)}{dt}$, the Bass (1969) model basically becomes a differential equation:

$$\frac{f(t)}{[1 - F(t)]} = [p + qF(t)] \tag{11}$$

where f(t) is the pdf (probability density function) of adoption time, i.e., f(t) = S(t)/m, and F(t) is its cdf (cumulative density functions), i.e., F(t) = CS(t)/m. Assuming F(0) = 0, equation (11) can be solved to yield:

$$F(t) = \frac{1 - \exp(-(p+q)t)}{1 + (q/p)\exp(-(p+q)t)} \tag{12}$$

It is then differentiated to yield the sales pdf and a theoretical function for T*:

$$f(t) = \frac{(p+q)^2}{p}\frac{\exp(-(p+q)t)}{[1 + (q/p)\exp(-(p+q)t)]^2} \tag{13}$$

$$T^* = t : \left\{\frac{df(t)}{dt} = 0\right\} = \frac{1}{p+q}ln\frac{q}{p} \tag{14}$$

Equations (13) and (14) further lead us to:

$$S_V(T^*) = \frac{m(p+q)^2}{4q} and S_V(0) = pm \Rightarrow \frac{S_V(T^*)}{S_V(0)} = \frac{(1+q/p)^2}{4(q/p)} \tag{15}$$

Substituting expression (15) in equation (9) results in:

$$VON = \frac{1}{\sqrt{2}\sigma}Ln\left(\frac{S_V(T^*)}{S_V(0)}\right) = \frac{1}{\sqrt{2}\sigma}Ln\left(\frac{(1+q/p)^2}{4(q/p)}\right) \tag{16}$$

where p, q are parameters of the Bass model to be estimated from the initial years' data and σ is the noise representing what is left unexplained by the Bass model.

4. Empirical analysis

4.1. Forming meta set for analysis

We use pre-peak sales data of 18 product categories based on data availability, the time they were launched (before and after 2000), and of time savings and/or entertainment value to the consumers. Every product gives multiple data sets as explained below. Consider a new product whose T* is well beyond 5 years. At the end of the 5th year, the manager can use the 5 years' sales data to estimate the parameters of the Bass model (p, q, and m) and use those estimates to predict T*. Then, at the end of the 6th year, the manager can use the 6 years' sales data to obtain another prediction for T*. The manager can keep coming up with a prediction for T* every year until they reach the T*. Assuming that this product hits peak sales in Year 9, we see that the manager can make five T* predictions, one in Year 5, next in Year 6, …, and the last in Year 9. Thus, every product gives multiple data sets. The products in our analysis range from cordless phones of the 1980s to the smartphones of the 2000s.

Each subset is a sample element in our analysis. Thus, we collected 146 subsets (the sum of column 6 in Table 2) in all. We fitted the Bass model on each sample element for this screening, and out of the 146 subsets, 16 had problems with convergence or suffered from the takeoff phenomenon (Golder and Tellis, 1997),[5] and hence were removed, yielding us a net sample size of 130. Column 7 of Table 2 gives the number of usable sample elements. We call this the metaset.

4.2. Estimating T*, noise, and VON on the metaset

Having formed the metaset, we now characterize each of its sample

[4] The IID assumption can be challenged when, for example, consumers form rational expectations about a forthcoming season-sale and postpone their purchases, which would create a dip in pre-season times leading to a bigger spike in the subsequent season-times. However, this is not a major issue if we use annual or semi-annual data, which we use in our empirical analysis.

[5] Three of the 18 products (the answering machine, DVD players, and the digital camera) exhibited a clear takeoff period, which refers to the fact that some new products have sluggish sales trends in the introductory years and a rapid growth after reaching a breakpoint (Golder and Tellis, 1997).

elements on three dimensions: the T* it predicts, the noise it has, and the value of VON it holds. Consider product j whose T^* is T_j^*. Forming data subsets from the 5th year, we will get ($T_j^* - 5 + 1$) subsets. The Bass model was estimated on each subset using the estimation equation (Srinivasan and Mason, 1986) for a given subset i as follows:

$$Sales_j(t) = (m_{ji})(F_{ji}(t) - F_{ji}(t-1))$$

where $Sales_j(t)$ is the *observed* sales of product j at time t, m_{ji} is the market potential parameter specific to subset i of product j, and $F_{ji}(u)$ is the cdf (cumulative density function) as given in Equation (12). We included log-normal error multiplicatively in the above expression. Hence, taking the log on both sides, we get:

$$Ln(Sales_j(t)) = Ln(m_{ji}) + Ln(F_{ji}(t) - F_{ji}(t-1)) + \varepsilon\ N(0, \sigma^2) \tag{17}$$

We used the nonlinear least-squares (NLS) procedure in SAS for the estimation.[6]

For each sample element in the metaset, the estimated Bass model parameters, namely $\{\widehat{p_{ji}}, \widehat{q_{ji}}, \widehat{m_{ji}}\}$, and their respective standard errors were in turn used to get estimates of:

1. $\widehat{\sigma_{ji}}$, the noise estimation, the nonsystematic component in the data points,

2. $\widehat{VON}_{ji}$(Equation (16)), the proposed metric to assess the accuracy of the T^* estimate, and.

$\widehat{T^*_{Ji}}$, the peak sales time as evaluated from the subset's estimates (see Equation (14)).[7]

Note that while the actual T_j^* is the same for all the subsets pertaining to product j, the predicted peak sales time, $\widehat{T^*_{Ji}}$, may come out to be different for different subsets in the product j.

A key question here is "how high should the VON be to claim that the corresponding predicted T* is accurate and reliable?" We will answer this question in Section 5. We will now see how far our claim regarding VON holds across the 18 categories we tested.

4.3. Estimating the impact of VON on the accuracy of predicted T*

Having characterized each element in the metaset in terms of (a) the T* it predicts, i.e., $\widehat{T^*}$, and (b) VON, our objective is to assess whether VON is positively associated with the accuracy of the predicted T*, i.e., the deviation of predicted T* from the actual-T*. Accordingly, we use the following regression on the metaset to estimate this influence of VON over the accuracy of T* prediction.

$$Ln(Y_{ji}) = \beta_0 + \beta_1 Ln(X_{ji}) \tag{18}$$

where $Y_{ji} = \left|\widehat{T^*_{Ji}} - T_j^*\right| * 100/T_j^*$, which measures the deviation in T*-prediction obtained with using subset i of product j, and $X_{ji} = VON_{ji}$, which measures the VON of the corresponding product-subset. We expect β_1 to be negative, indicating that a larger VON would ensure a more accurate T* prediction.

We now consider two additional competing factors to find out if VON could remain important even after considering other possible causes of the deviation in T* prediction.

4.3.1. Variable 1: Data set length

The first factor we consider is the number of data points in each subset used in the estimation. Noting that we are dealing with small data sets, addition of even one data point is likely to make a significant impact on the estimation and subsequent prediction. In other words, a subset with six data points would be expected to perform better than one with five data points, everything else remaining the same across the two subsets. According to the findings in the extant literature, data set length (i.e., the number of data points used in the estimation) affects the estimates (e.g., Lenk and Rao, 1990; Mahajan and Shama, 1986; Parker, 1994; Srinivasan and Mason, 1986; Van den Bulte and Lilien, 1997). Hence, we conjecture that a longer subset (i.e., one with a higher number of data points) will result in better prediction of T^*. We expect the impact of data set length on the deviation of T* to be negative. We next consider two socioeconomic factors derived from Rogers (1995, p.206).

4.3.2. Variable 2: Social communication channel (Word-of-Mouse)

The communication channels have gotten better over the decades, especially after the Internet was introduced in the late 1980s and the World Wide Web in the mid-1990s. But a breakthrough in WOM communication came in when the social media started going mainstream in the 2000s (Yusuf and Busalim, 2018). The products introduced in the 2000s were likely to have experienced a stronger WOM leading to a larger V-Force, which might have improved the accuracy of T* prediction. Accordingly, we will include in Equation (18) a 'digital WOM' effect (or called the *Word-of-Mouse* effect[8]), a dummy for products from 2000 to signal the extra influence they got from social media, blogs, and other new-age communication channels. We expect the 2000s' products to have a negative effect on the T* deviation or inaccuracy.

Equation (18) is now expanded to include the two competing explanatory variables:

$$Ln(Y_{ji}) = \beta_0 + \beta_1 Ln(X_{ji}) + \beta_2 Ln(L_{ji}) + \beta_3 D_j \tag{19}$$

where, L_{ji} is the number of the data points of product j in the subset i, and D_j assumes value 1 if the product j was launched after the year 2000 and 0 otherwise.

4.3.3. Regression results

The results of regressing Equation (19) on the data set of 130 observations are provided in Table 3.

Table 3
Influence of VON on Predicted T*'s Deviation or Lack of Accuracy.

Parameter	Variable concerned	Estimate	Std. Error	*p*-value
β_0	Intercept	6.5017	0.7188	<0.0001
β_1	VON	−0.5832	0.1389	<0.0001
β_2	Data set length (# of data points)	−1.1177	0.2569	<0.0001
β_3	Dummy (Word-of-mouse)	−0.5759	0.1706	0.0010

[6] We use the log–log version of the estimation equation of Srinivasan and Mason (1986). Researchers have suggested other estimation equations and procedures as well (e.g., Jain & Rao, 1990). We use Srinivasan and Mason's procedure because of its wide acceptance.

[7] Direct evaluation of predicted-T*, Noise, and VON from the $\{\widehat{p_{ji}}, \widehat{q_{ji}}, \widehat{m_{ji}}\}$ will not give us the distributional properties of these estimates. We use the Monte Carlo procedure for this purpose. Let T* be the observed peak sales time for a given data set. We then use the data [0, T*] to estimate the Bass model parameters {p, q, m}. The estimates and their respective standard errors are then used to generate 1000 data sets. With each simulated dataset, we evaluate the error between the sales at each time with the real sales number, leading to an estimate of σ and thereby the VON for that dataset. Repeating this exercise over the 1000 data sets would provide us with the distributional properties of Predicted-T*, Noise, and VON for the chosen real data set. We repeat the whole exercise for subsets of data up to T*-1, T*-2, …, till 5 (as we require at least 5 data points to fit the Bass Model). This is the simulation exercise for one product category. The same simulation was done for all the 18 products we had, giving us a total of 130 subsets for the final regression.

[8] This term is proposed by Dr. Dipak C. Jain to highlight the rule of social media and big data.

VON has a negative and significant impact on the T* deviation even in the presence of two additional variables. This implies that a higher VON would significantly indicate a higher accuracy of the predicted T*. Put differently, the *relative* strength of the V-Force (force varying systematically with time as captured by the Bass model) over the N-Force (Noise) is what gives the proposed metric the power to assess the accuracy of prediction of T*.

What makes this finding even more interesting is that products of the 2000s appear to enjoy the benefits of digital WOM, which not only makes their sales grow at a quicker pace but also enables managers get a more accurate T* prediction (i.e., coefficient β_3 in Table 3 is negative and significant). The data set length (i.e., the number of data points in a subset) has, as argued earlier, a negative and significant impact, implying that the T* prediction accuracy improves when we use a subset with a higher number of data points. What this finding tells us is that it is sometimes worthwhile to wait for a year or two before predicting T*. However, if the cost of waiting is significant from a strategic perspective, such as a competing product entry, the manager is advised to use the predicted T* but with caution.

4.4. Alternative metrics

4.4.1. Comparing the impact of VON to speed of diffusion (q/p) on the accuracy of predicted T*

One may wonder if WOM is a necessary condition for VON to work. Van den Bulte et al. (2004) analyzed meta data sets to see if the diffusion pattern is a result of heterogeneity in the target market or social contagion. They separated out q/p and used that metric to represent the "speed" of diffusion, which they studied across various products. They found existence of both the effects, namely, heterogeneity and social contagion, although they believed that the social contagion was more a result of social cohesion rather than WOM. Since q/p is far less complicated than VON we have proposed, we tested whether the simple q/p could do the job of the more sophisticated VON in line with Occam's Razor principle of parsimony (Sober, 1994). We compared the effects of VON and q/p to check if VON offered more insights than the simpler metric. We used Equation (19) and modified it as follows compare to VON with q/p:

$$Ln(Y_{ji}) = \beta_0 + \beta_1 Ln(q_{ji}/p_{ji}) + \beta_2 Ln(L_{ji}) + \beta_3 D_j \tag{20}$$

where L_{ji} is the number of the data points of product *j* in the subset *i*, and D_j assumes value 1 if the product *j* was launched after the year 2000 and 0 otherwise. Estimating Equation (20) on the 130 data sets, we found the coefficient of *Ln(q/p)* to be insignificant, i.e., the effect of q/p was found to be insignificant. The results are reported in Table 4 (columns 4&5), with a replicate of VON-results (columns 2&3) for the ease of comparison.

A possible reason for the poor performance of the metric Ln(q/p) vis-à-vis VON is that q/p captures the speed of diffusion (i.e., a metric for V-Force) only and ignores the N-Force (i.e., Noise) in a data set while VON captures both of these forces[9].

4.4.2. Comparing the impact of VON with V-Force on the accuracy of predicted T*

Another appropriate metric we could use is the V-Force of VON. Taking into consideration equation (9) and equation (16), we work on the following function of V-Force as an alternative metric and see how it helps in assessing the accuracy of predicted T*:

$$\text{Effect of V}-\text{Force} = Ln\left(\frac{S_V(T^*)}{S_V(0)}\right) = Ln\left(\frac{(1+q/p)^2}{4(q/p)}\right)$$

Hence the effect of V-Force is a function of q/p and is highly correlated with it[10]. Like equation (20), we use the following regression to compare V-Force with VON:

$$Ln(Y_{ji}) = \beta_0 + \beta_1 Ln(V_{ji}) + \beta_2 Ln(L_{ji}) + \beta_3 D_j \tag{22}$$

where V_{ji} is the V-Force of product *j* in the subset *i*, following equation (21); L_{ji} is the number of the data points of product *j* in the subset *i*; and D_j assumes value 1 if the product *j* was launched after the year 2000 and 0 otherwise.

Table 4 (columns 6&7) reports the regression result of equation (22) and compares it with the VON results. The results show that the coefficient of *Ln(V)* is insignificant and hence the effect of V-Force doesn't have a significant effect on the accuracy of predicted T*. The result strengthened that the VON metric is required to show the accuracy of predicted T*.

4.5. Analysis with different data intervals (monthly or quarterly)

We have used annual data in our analysis and suggested that managers should wait until there were enough data points for the estimation. A natural question a manager could ask is: Why wait for years if we have access to quarterly data or monthly data? Such a data set has two differences with respect to the annual data. First, the data interval per se is small, and second, for a given period in years, the number of data points available with quarterly or monthly data will be higher. We examine the impact of smaller-interval data in this section, as it pertains to T* prediction.

Since quarterly data or monthly data are not readily available, especially for consumer durable goods like TV and room air-conditioners, we managed to obtain *quarterly* data for two hi-tech products introduced in the past 15 years, namely, iPhone and Fitbit. These products were regarded as breakthrough technologies and due to their leading positions, they were near-monopolies in their initial years. With each product, we fit the Bass Model to the quarterly data, starting from the 4th or 5th quarter, and we used our previously explained procedure to get the estimates of peak sales time (T*). We repeated the process by adding quarterly data one by one to examine the changes in those estimates and compared them with the estimates obtained using yearly data.

We provide the results in Table 5. Looking at the results of iPhone, we see that both the quarterly data and the annual data produce very similar values for the predicted T*. This shows that the data interval may not have a major impact on the estimated value of peak sales time. For Fitbit, the predicted values for T* are also similar for both quarterly and annual data sets.[11] Given that quarterly and yearly data produce similar T* estimates, but quarterly data has higher noise, we recommend using yearly data for predicting T.[12]

[9] Since VON and V-Force are almost the same except for the Noise in the VON measure, we did additional analysis by including only those data sets that have least Noise in them to examine if they yield similar results in terms of prediction accuracy. We included VON, V-force, and data length as independent variables, and ran the regression on the reduced dataset. We found that with low noise, both the metrics were significant. These results clearly highlight the importance of incorporating market noise in assessing the prediction accuracy of T*.

[10] We thank a reviewer for suggesting this interesting metric.

[11] For iPhone and Fitbit, the actual T* values were 9 and 4, respectively. This further validates that the predicted values of T* are quite close to the actual T*.

[12] The noise levels associated with the estimates of the quarterly data are higher, perhaps due to the presence of seasonality. Sales of consumer electronic goods exhibit much more seasonality than other goods, typically in the holiday season of Quarter 4 (TraQline, 2022). More research needs to be done to explore if data interval has any major influence on T* estimates and how far a manager can ignore issues such as seasonality when using the quarterly versus yearly data. Further, hi-tech electronic goods are characterized by continual modifications, and successive generations. Hence, one might have to use a modified Bass model or specific diffusion models such as Norton and Bass (1987) and Jiang and Jain (2012).

Table 4
Effects of VON vs. q/p vs. V-Force on the Accuracy of Predicted T*.

	Metric VON (Eq.19)		Metric Speed of Diffusion (q/p) (Eq.20)		Metric V-Force (Eq.22)	
	Parameter Estimate	Approx Std Err	Parameter Estimate	Approx Std Err	Parameter Estimate	Approx Std Err
Intercept	6.5017***	(0.7188)	4.4084***	(0.6228)	4.3964***	0.5912
Data length	−1.1177***	(0.2569)	−0.7611**	(0.2597)	−0.7617**	0.2597
Dummy (Word-of-Mouse)	−0.5759**	(0.1706)	−0.4053*	(0.2029)	−0.4078*	0.2024
VON	−0.5832***	(0.1389)				
q/p			−0.0053	(0.0592)		
V-force					−0.0039	0.00607

*** $p < 0.0001$; ** $p < 0.01$; * $p < 0.05$.

Table 5
Comparing Predicted T* Using Quarterly vs Yearly Data.

iPhone			Fitbit		
Year	T* in Years (using Quarterly data)	T* in Years (using Yearly data)	Year	T* in Years (using Quarterly data)	T* in Years (using Yearly data)
2011	4.27	4.49	2016	3.25	3.11
2012	5.10	5.54	2017	3.35	3.21
2013	5.61	6.09	2018	3.73	3.64
2014	6.15	6.63	2019	4.36	4.34
2015	6.93	7.43			
2016	7.42	7.93			
2017	7.94	8.47			
2018	8.47	9.02			

4.6. Traditional Word-of-Mouth vs. Digital Word-of-Mouth (i.e., Word-of-Mouse)

Section 4.3 of our analysis indicated that products introduced after the year 2000 have a more accurate T* value, when compared to the products introduced in the pre-2000 period. To investigate this further, we divided our full dataset into two subsets based on product introduction date, one for the products introduced in the pre-2000 period and the other for the products introduced in the post-2000 period. We estimated our VON model (i.e., Equation (19)) with each subset. Since the model and the variable metrics are consistent across the two datasets, we can directly compare the VON impact through analyzing their respective coefficient estimates. Table 6 displays the results, where we see that VON is a more powerful predictor of T* accuracy for the post-2000 products (estimate = −0.8398) than for the pre-2000 products (estimate = −0.4224).Extending this finding, we can say that the word-of-mouse, where the media of info spread includes largely the Internet and social media, seems to be stronger than the traditional word-of-mouth.

5. Managerial implications of VON

The empirical results show that T*prediction is more accurate if VON is high. A key managerial question at this juncture would be: How high is considered high enough? Based on our empirical analysis of the 18 product categories, we develop three thresholds, pertaining to three levels of acceptable deviation, namely, Tight, Medium, and Slack. See Table 7.

Table 7
Threshold VON Derived Using the Empirical Data on 18 Product Categories.

	Tight	Medium	Slack
Level of acceptable deviation (α*)	8 %	10 %	12 %
Threshold VON (median)	28.7	19.6	14.3

For example, a manager might want to be extremely cautious about the prediction, which is typically 5–10 % deviation from the true value. Let's take the average and let the deviation level be 8 %, and according to our calculation this means a threshold-VON of 28.7. Suppose the new product is in its 6th year and its sales are still in the growth phase. If the manager uses the data and gets a T* of say 8 years, implying that the data point to the sales peaking in year 8. An 8 % deviation would mean that the actual peak could be within 0.6 years of the 8th year. For this to happen, the corresponding VON must be higher than 28.7. However, note that the thresholds mentioned in Table 7 were derived from the analysis carried out on all the data sets we had used. For a given product, it will be more prudent to use products of the same class.

We will explain how we evaluated the three thresholds. We use Equation (19) as the base model and keep only those variables that were significant and arrive at the following prescriptive model.

$$Ln(Y_{ji}) = \beta_0 + \beta_1 Ln(Z_{ji}) + \beta_2 Ln(L_{ji}) + \beta_3 D_j \quad (21)$$

where Y_{ji} measures the deviation in T*-prediction obtained using subset i of product j, and $Z_{ji} = (1+\alpha)VON_{ij}$ measures the VON of the corresponding product-subset as in Equation (19) but qualified by an additional factor $(1+\alpha)$, which we explain as follows. Note that, as mentioned earlier, L_{ji} is the length of the data set of the product-subset, and D_j assumes value 1 if the product was launched after the year 2000 and 0 otherwise.

For the independent variables on the right-hand side of Equation (21), we adopt the same values used in the estimation of Equation (19) along with their corresponding parameter estimates obtained from the regression results. Accordingly, we have: $\beta_0 = 6.50$, $\beta_1 = -0.58$, $\beta_2 = -1.11$, and $\beta_3 = -0.58$. What is however unknown on the right-hand side of Equation (21) is α, a parameter we have newly introduced to modify the value of VON_{ij} for all $\{i, j\}$. If α = 0, there is no modification to VON. As we start increasing the value assigned to α, the T*-difference will start decreasing monotonically because of β_1 being negative. Where do we stop? The stopping point, denoted by α*, would be a function of

Table 6
Comparison of VON Effect Between Pre-/Post- 2000 Products.

	Pre-2000			Post-2000		
	Parameter Estimate	Approx Std Err	Approx Pr > |t|	Parameter Estimate	Approx Std Err	Approx Pr > |t|
Intercept	6.1662	0.8308	<0.0001	7.8415	1.4414	<0.0001
Data length	−0.8956	0.2643	0.0011	−1.4728	0.5713	0.0133
VON	−0.4224	0.1481	0.0055	−0.8398	0.2837	0.0049

the criticality of the T* prediction. If the manager wants to be strict with the prediction, they must choose a higher α*. This will happen, for example, when the manager must make a risky investment decision. If the prediction is not going to be viewed as strict but rather loose, say, as in planning the inventory of goods, a lower threshold is advisable. Accordingly, we chose three levels of acceptable deviation (Strict, Medium, and Loose) and evaluated α* and the corresponding threshold-VON values (median). These are provided in Table 7.

5.1. Step-by-step recommendation for a manager

The following are recommendations for managers to ensure proper assessment of T* prediction accuracy:

Step 1. The manager should collect as many data points as possible. We recommend using yearly data but if the manager wants to use monthly or quarterly data, they should first smoothen the variation due to seasonality.
Step 2. The manager should pick a parsimonious but established growth model like the Bass model and estimate it on the data set collected (which has at least five data points) using the log–log version (see Section 4.2). If the regression doesn't converge or the estimates are found to be insignificant, then waiting is preferable to gather more data points.
Step 3. If the estimates are significant, the manager should use a Monte Carlo technique and estimate T* (Equation (14)) along with its standard deviation and VON (see Equation (16) and Section 4.2). Depending upon how strict the manager wants to be with the prediction accuracy, they can use Table 6 and decide whether to accept the predicted T* or not.
Step 4: If VON is lower than the threshold-T*, the manager can wait for one more year to collect more data and then repeat the T* prediction exercise.
Step 5. Suppose that the manager is forced to use the predicted T* even though the corresponding VON is lower than the desired threshold level. Based on the analysis on the 18 products in this research, a predicted T* having a VON lower than the threshold-VON indicates that it is likely to have around 50 % deviation with respect to the actual T*. So, the manager must be ready with a backup plan if they continue with the predicted T* and make changes in the strategies.
Step 6. The manager may wish to use another procedure (e.g., survey or analysis of an analogous product in a similar market) to corroborate the T* prediction accuracy.

5.2. An illustrative Example: Digital camera

We illustrate the step-by-step process using the digital camera data as follows.

Step 1: We took 5 subsets DC(1) through DC(5) of digital camera for predicting T*.
Step 2: The estimates obtained from the five subsets on the Bass model are presented in Table 8.

DC(1) through DC(5) in Table 8 refer to the five subsets of different data set lengths. The Bass model parameters estimated on the log–log version are given along with their standard errors in parentheses.

Step 3: Using Equation (14) and the Monte Carlo technique, the predicted T* was evaluated. Using Equation (16) and the Monte Carlo technique, V-Force, N-Force, and VON were evaluated. The noise estimate, $\widehat{\sigma_{ji}}$, is directly available from the regression output in the SAS but we used Monte Carlo method to be in line with the other two measures.
Steps 4 and 5: Not applicable to digital camera data set.
Step 6: For the two subsets DC1 and DC2 the *p*-estimate was not significant, i.e., not a significant pool of innovators, and hence we suggest that one should not use the first two subsets to predict T*. For the subsets DC3 through DC5, all the three parameters are significant and the predicted value of T* comes out to be 10.7, 10.2, and 11 years (the actual T* being 10). More importantly, for these data sets, the parameter estimates are significant and the corresponding noise levels are lower, giving a higher VON. The highest value of VON (29.8) is for T* 10.2, which is closest to 10, the actual value of T*. This example provides empirical validity to our recommended, step-by-step procedure for predicting T*.[13]

6. Contributions and directions for future research

Of all the forecasts a manager must make with respect to a new product that has found some traction in the marketplace, one of the most important would arguably be forecasting the time when the sales growth would reach the peak, denoted by T* in the diffusion literature. When a new durable is about to reach peak sales time T*, it indicates that the maturity stage is setting in. At this juncture, the firm must change its production schedule, inventory policies, and marketing strategy, and further, work on the next new product to introduce in the marketplace. Given its importance from so many angles, a manager would like to predict T* in advance so that they are able to help the firm get ready on various fronts. Although currently there are quite a few diffusion models available that the manager can use to estimate T* with the pre-peak sales data, there has been no research in the existing literature that can tell them how *accurate* such a predicted T* is. Assessing the accuracy is a challenging task because we are working with small data sets, i.e., a few early periods' data points.

Through our research, we contribute to the marketing literature by proposing a new metric, VON, and empirically demonstrate how it helps a manager assess the accuracy of predicted T*. We also provide a step-by-step procedure for a manager to adopt this new metric. The proposed metric VON is not limited to specific model specifications. We use the Bass Model only as a demonstration of the metric's usefulness.

There are several key managerial implications. Predicting T* is vital for a firm because as the firm approaches T*, it should focus on its R&D strategy, i.e., whether it should develop a new version now or later; Operations strategy, i.e., whether it should change the production rate, inventory, and supplies from vendors; and Marketing strategy, i.e., whether it should start using price as a tool to attract customers. Furthermore, products introduced in the 2000s (i.e., 2000–2010) seem to yield more accurate T* than those introduced in the 1980s and 1990s, highlighting the special role played by word-of-mouse (i.e., information spread digitally through the Internet, social media, and e-mails) in the post-2000 period in addition to traditional WOM. Lastly, data intervals do not have a major impact on the prediction accuracy, i.e., using quarterly or monthly data doesn't lead to better predictions than yearly data.

There are a few areas for further research. First, our finding on the role of word-of-mouse effect on the products introduced in the post-2000 era is worth further investigation, i.e., further study on how social networks and viral diffusion are contributing to the diffusion process. It would be interesting to know how today's fast and mobile consumer voice affects T* prediction. Second, noting that market noise has been found to play a highly important role in prediction, future researchers can further investigate the noise component and advise managers on remedial actions to be taken to reduce the noise. Third,

[13] One may wonder how useful it is to get an accurate prediction of T* when we are at the peak sales. Note that peak is not recognized until we go well past the peak, and hence even knowing that we are near the peak sales time is of great use for a manager.

Table 8
Estimates Obtained from the 5 Subsets of Digital Camera (DC).

Subset (Sample Element)	Length (No. of Data Points)	*p*-est. (SD)	*q*-est. (SD)	*m*-est. '000 (SD)	Predicted T* (years)	V-Force est.	N-Force (σ) est.	VON est.
DC(1)	6	0.0010 (0.0057)	0.5005 (0.0879)	44,309 (237,133)	10.1	4.0378	0.2264	17.8363
DC(2)	7	0.0009 (0.0024)	0.4976 (0.0536)	56,765 (167,050)	11.5	4.2821	0.1700	25.183
DC(3)	8	0.0021 (0.0008)	0.5213 (0.0371)	21,891 (10,603)	10.7	3.9198	0.1491	26.2877
DC(4)	9	0.0019 (0.0004)	0.5285 (0.0267)	19,446 (3,964)	10.2	3.3703	0.1131	29.8021
DC(5)	10	0.0019 (0.0003)	0.5072 (0.0231)	25,591 (4,535)	11	3.230	0.1158	29.5698

Numbers in parentheses are standard errors of the corresponding estimates.

VON may be used to examine the growth models between different generations of products as well as those in different countries that launch new products at different times. Fourth, the definition of peak-sales time can be looked at from a different perspective. One way would be to measure continually the change in sales rate and have a threshold to define the peak-sales time as something like "if change in sales rate is lower than certain value", then we declare that peak sales is reached. Such definitions would be hard to formulate mathematically but they would be more pragmatic and worth pursuing.

CRediT authorship contribution statement

Trichy V. Krishnan: Methodology, Formal analysis, Data curation, Conceptualization, Writing - original draft, Writing - review & editing. **Shanfei Feng:** Methodology, Formal analysis, Data curation, Writing - original draft, Writing - review & editing. **Dipak C. Jain:** Writing – review & editing, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank China Europe International Business School (CEIBS) for the financial support provided through a research grant. We also thank the editor and anonymous reviewers for their valuable feedback, which greatly contributed to enhancing the quality of this paper.

References

Bass, F. M. (1969). A new product growth model for consumer durables. *Management Science, 15*(5), 215–227.
Bass, F. M., Krishnan, T. V., & Jain, D. C. (1994). Why the Bass Model Fits without Decision Variables. *Marketing Science, 13*(3), 203–223.
Bass, F.M., Jain, D. C., & Krishnan, T.V. (2000). Modelling the Marketing-Mix Influence in New-Product Diffusion. New-Product Diffusion Models.
Base, F.M., 2004. Comments on "a new product growth for model consumer durables the bass model". *Management science* 50, 12_supplement, 1833-1840.
Bersch, C. V., Akkerman, R., & Kolisch, R. (2021). Strategic planning of new product introductions: Integrated planning of products and modules in the automotive industry. *Omega, 105*, Article 102515.
Bewley, R., & Griffiths, W. E. (2003). The penetration of CDs in the sound recording market: Issues in specification, model selection and forecasting. *International Journal of Forecasting, 19*(1), 111–121.
Chandrasekaran, D. & Tellis, G.J. (2018). A summary and review of new product diffusion models and key findings. In *Handbook of Research on New Product Development*. Edward Elgar Publishing.
Golder, P. N., & Tellis, G. J. (1997). Will It Ever Fly? Modeling the Takeoff of Really New Consumer Durables. *Marketing Science, 16*(3), 256–270.
Hardie, B. G., Fader, P. S., & Wisniewski, M. (1998). An empirical comparison of new product trial forecasting models. *Journal of Forecasting, 17*(3–4), 209–229.
Ho, T. H., Savin, S., & Terwiesch, C. (2002). Managing demand and sales dynamics in new product diffusion under supply constraint. *Management Science, 48*(2), 187–206.
Jain, D. C., & Rao, R. C. (1990). Effect of price on the demand for durables: Modeling, estimation, and findings. *Journal of Business & Economic Statistics, 8*(2), 163–170.
Jain, D. C., & Rao, R. C. (2020). Timing for the revised edition of a new textbook. Working paper, China Europe International Business School (CEIBS), 1-5.
Jiang, Z., Qu, X., & Jain, D. C. (2019). Optimal Market Entry Timing for Successive Generations of Technological Innovations. *MIS Quarterly, 43*(3), 787–806.
Kalish, S., & Lilien, G. L. (1986). A market entry timing model for new technologies. *Management Science, 32*(2), 194–205.
Kotler, P., & Keller, K. L. (2016). Marketing Management 15th Global edition (Global). *Harlow: Pearson Education Limited.*
Krankel, R. M., Duenyas, I., & Kapuscinski, R. (2006). Timing successive product introductions with demand diffusion and stochastic technology improvement. *Manufacturing & Service Operations Management, 8*(2), 119–135.
Lattin, J. M., & Roberts, J. H. (1989). *The role of individual-level risk-adjusted utility in the diffusion of innovation*. Working Paper, Graduate School of Business, Stanford University.
Lenk, P. J., & Rao, A. G. (1990). New Models from Old: Forecasting Product Adoption by Hierarchical Bayes Procedures. *Marketing Science, 9*(1), 42–53.
Mahajan, V., & Shama, S. (1986). A simple algebraic estimation procedure for innovation diffusion models of new product acceptance. *Technological Forecasting and Social Change, 30*(4), 331–346.
Mahajan, V., & Muller, E. (1996). Timing, diffusion, and substitution of successive generations of technological innovations: The IBM mainframe case. *Technological Forecasting and Social Change, 51*(2), 109–132.
Mansfield, E. (1961). Technical change and the rate of imitation. *Econometrica: Journal of the Econometric Society*, 741–766.
Meade, N., & Islam, T. (1995). Forecasting with growth curves: An empirical comparison. *International Journal of Forecasting, 11*(2), 199–215.
Meade, N., & Islam, T. (1998). Technological forecasting—Model selection, model stability, and combining models. *Management science, 44*(8), 1115–1130.
Meade, N., & Islam, T. (2006). Modelling and forecasting the diffusion of innovation–A 25-year review. *International Journal of Forecasting, 22*(3), 519–545.
Moorthy, K. S., & Png, I. P. (1992). Market segmentation, cannibalization, and the timing of product introductions. *Management Science, 38*(3), 345–359.
Negahban, A., & Smith, J. S. (2018). Optimal production-sales policies and entry time for successive generations of new products. *International Journal of Production Economics, 199*, 220–232.
Niu, S.-C. (2006). A Piecewise-Diffusion Model of New-Product Demands. *Operations Research, 54*(4), 678–695.
Norton, J. A., & Bass, F. M. (1987). A diffusion theory model of adoption and substitution for successive generations of high-technology products. *Management science, 33*(9), 1069–1086.
Parker, P. M. (1994). Aggregate diffusion forecasting models in marketing: A critical review. *International Journal of Foresting, 10*(2), 353–380.
Prasad, A., Bronnenberg, B., & Mahajan, V. (2004). Product entry timing in dual distribution channels: The case of the movie industry. *Review of Marketing Science, 2*, 83–102.
Qin, R., & Nembhard, D. A. (2012). Demand modeling of stochastic product diffusion over the life cycle. *International Journal of Production Economics, 137*(2), 201–210.
Ramírez-Hassan, A., & Montoya-Blandón, S. (2020). Forecasting from others' experience: Bayesian estimation of the generalized Bass model. *International Journal of Forecasting, 36*(2), 442–465.
Rogers, E.M. (1962, 1995). *Diffusion of Innovations*. New York: Free Press.
Schwarz, J. A., & Tan, B. (2021). Optimal sales and production rollover strategies under capacity constraints. *European Journal of Operational Research, 294*(2), 507–524.
Sober, E. (1994). Let's Razor Occam's Razor. In Knowles, Dudley (ed.). Explanation and Its Limits. Cambridge University Press. 73–93.
Srinivasan, V., & Mason, C. H. (1986). Nonlinear least squares estimation of new product diffusion models. *Marketing Science, 5*(2), 169–178.
Tian, Y., Huffman, G. J., Adler, R. F., Tang, L., Sapiano, M., Maggioni, V., & Wu, H. (2013). Modeling errors in daily precipitation measurements: Additive or Multiplicative? *Geophysical Research Letters, 40*(10), 2060–2065.
Toubia, O., Goldenberg, J., & Garcia, R. (2014). Improving penetration forecasts using social interactions data. *Management Science, 60*(12), 3049–3066.
TraQline.com (2022) Spotting Seasonal Sales Trends in Consumer Data . Retrieved September 9 2022 from https://www.traqline.com/newsroom/blog/spotting-seasonal-sales-trends-in-consumer-data/.
Trusov, M., Rand, W., & Joshi, Y. V. (2013). Improving prelaunch diffusion forecasts: Using synthetic networks as simulated priors. *Journal of Marketing Research, 50*(6), 675–690.

Van den Bulte, C., & Lilien, G. L. (1997). Bias and systematic change in the parameter estimates of macro-level diffusion models. *Marketing Science, 16*(4), 338–353.
Van den Bulte, C., Lilien, G. L., & Stremersch, S. (2004). Social Contagion and Income Heterogeneity in New Product Diffusion: A Meta-Analytic Test. *Marketing Science, 23*(4), 530–544.
Venkatesan, R., Krishnan, T. V., & Kumar, V. (2004). Evolutionary estimation of macro-level diffusion models using genetic algorithms: An alternative to nonlinear least squares. *Marketing Science, 23*(3), 451–464.
Wilson, L. O., & Norton, J. A. (1989). Optimal entry timing for a product line extension. *Marketing Science, 8*(1), 1–17.
Xiao, Y., & Han, J. (2016). Forecasting new product diffusion with agent-based models. *Technological Forecasting and Social Change, 105*, 167–178.
Young, P. (1993). Technological growth curves: A competition of forecasting models. *Technological Forecasting and Social Change, 44*(4), 375–389.
Yusuf, A. S., & Busalim, A. H. (2018). Influence of e-WOM engagement on consumer purchase intention in social commerce. *Journal of Services Marketing*.
Jiang, Z., & Jain, D. C. (2012). A Generalized Norton-Bass Model for Multigeneration Diffusion. *Management Science, 58*(10), 1887–1897.

Trichy V. Krishnan is a Research Advisor to Rajagiri Business School and Rajagiri College of Social Sciences, India. He has published extensively in leading academic journals, including Marketing Science, Management Science, Journal of Marketing Research and International Journal of Research in Marketing, and has written many award-winning case studies as well. He was previously employed with Centre for Economic Research, Netherlands, Rice University, Houston, and NUS Business School, Singapore.

Shanfei Feng is a lecturer in the Department of Marketing at Monash Business School of Monash University in Australia. Her research interests include Business-to-Business services, customer relationship, and new product diffusion. She has published in journals including International Journal of Research in Marketing, Journal of Service Research, International Journal of Foresting and Journal of Business and Industrial Marketing.

Dipak C. Jain is a globally recognized marketing and innovation expert whose insights have inspired a generation of business leaders to pursue success with significance. He's currently the Co-Director of CEIBS Social Security and Ageing Finance Institute and Distinguished Professor of Marketing at China Europe International Business School (CEIBS). His areas of research expertise include the marketing of high-tech products; market segmentation and competitive market structure analysis; cross-cultural issues in global product diffusion; new product innovation; and forecasting models. He has published more than 75 articles in leading academic journals.