No. 037/2020/MKT

Early Prediction of Peak Sales Time in New Product Sales: Your Prediction May be Statistically Significant, But Is It Accurate?

Trichy V. Krishnan*
Great Lakes institute of management

Shenfei Fang*
Department of Marketing
Monash Business School

Dipak C. Jain
China Europe International Business School (CEIBS)

August 2020

* Corresponding authors:
1. Trichy V. Krishnan (krishnan61@gmail.com). Address: Great Lakes Institute Of Management, Prince Infocity II,I[st] Floor, 283/3, Rajiv Gandhi Salai (OMR), Kandanchavadi, Chennai – 600096, India.
2. Shenfei Fang (shanfei.feng@monash.edu). Address: Monash University, Victoria 3800, Australia.

Early prediction of peak sales time in new product sales: Your prediction may be statistically significant, but is it accurate?

Abstract

Managers dealing with new products need to forecast sales growth, especially at the time when sales reach their peak, known as the peak sales time. They use a few initial years' data to develop a forecast of the peak sales time and test for its statistical significance using traditional measures such as the t-statistic. Unfortunately, these traditional measures only indicate whether the forecast is significantly different from zero or not, not whether the forecast is *accurate*, that is, closer to the actual value. To determine the accuracy of the forecast, we develop a new metric called voice over noise (VON) that is derived from a diffusion model framework. VON is built on the premise that both the consumer voice (i.e., word-of-mouth that drives sales growth) and the market noise (i.e., distortions of the sales numbers) are important to consider in assessing the accuracy of the forecast; neither alone is sufficient. We empirically prove that a stronger VON indicates a more accurate forecast of peak sales time, which the traditional statistical measures cannot do. For our empirical test, we use 15 new products from the past 3 decades.

Keywords: forecasting, new product diffusion, VON, peak sales time, prepeak sales data

Early prediction of peak sales time in new product sales: Your prediction may be statistically significant, but is it accurate?

1. Introduction

New products create a great deal of anxiety in product managers because of the associated unpredictability concerning their commercial performance. Doubts regarding the commercial success of a new product remain even for products from relatively successful companies like Apple and Tesla. For example, when Apple announced in late 2018 that it would no longer disclose the unit sales for its iPhone, investors interpreted it as a sign that iPhone sales had "peaked" (Kovach, 2018), especially after being disappointed by flat sales figures since 2016 (Statista.com). Similarly, when Tesla reported a 31% decline in deliveries of its Model 3 in the first quarter of 2019 compared to the final quarter of 2018, people wondered whether the demand for Model 3 had already reached its "peak" in North America (O'Kane, 2019). In both cases, concerns were raised during the growth phase of the new products, and investors were worried that the peak sales (i.e., maturity) would happen sooner than they had thought. As a result, the financial market responded immediately and negatively. These two incidents underscore the importance of peak sales time. For the sake of convenience, we denote the peak sales time as simply PST.

Sales growth of a new product or service has a typical pattern shown in Figure 1. In this hypothetical case, the maximum interest to predict PST would be in years 5 through 9. PST is considered strategically important for at least two reasons:

1. In the sales evolution of a new product, the arrival of PST signifies a slowing growth rate (Kovach, 2018), which in turn affects the market valuation of the firm, especially if the new product is central to the firm's growth (e.g., Model 3 for Tesla).

2. An approaching PST informs: (a) the marketing department to prepare to change its marketing strategies to handle the maturing market, (b) the operations department to amend production plans and supply chain dynamics (Sale, Mesak, & Inman, 2015), and (c) the R&D department to develop and test the next version of the new product or a new model.

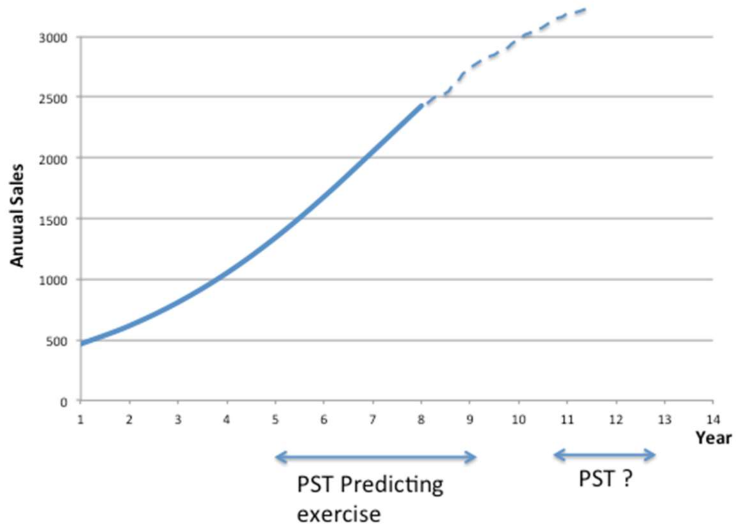

Fig. 1 Typical new product sales growth pattern

Because we only have the initial years' sales data (i.e., pre-PST data), PST prediction is naturally an extrapolation. Such an extrapolated prediction assumes that the sales pattern occurring during the available data period will continue to evolve in the same way in the future, at least until the PST occurs. Fortunately, we can assume such continuity in the case of a sales growth process of a new product for two reasons. First, the sales growth of a new product is a manifestation of the underlying adoption process that has been shown to be systematic in nature (Rogers 1962, 1995) at the individual level, primarily driven by word-of-mouth (WOM), or more precisely consumer voice, which includes the traditional WOM and the new age word-of-mouse[1]. Second, the resulting sales evolution at the aggregate level has also been shown to be systematic; in fact, the impact of consumer voice on sales growth is so systematic that Bass (1969) captured the sales growth process through an elegant mathematical function. Hence, one can safely offer an extrapolated prediction of the PST using the information contained in the few early years' data through employing models such as the Bass model (Bass, 1969).[2]

PST, estimated through extrapolation from the pre-PST data, must be judged on two dimensions. One is the commonly-used statistical significance, which is measured by *p*-value or 95% confidence interval (CI), and the other is the accuracy. Statistical significance tells us whether the estimate is different from zero, while accuracy tells us how close the estimate is to the true value of the PST. Of course, *ex-ante,* one does not know the true value of the PST, and hence managers generally look only at statistical significance and ignore the

[1] By the new age word-of-mouse we mean all the information spread through social media, blogs, etc.
2 There is overwhelming empirical support in the literature for the Bass model, e.g., Bass (1969), Parker (1994), Bass, Krishnan, and Jain (1994), Sultan, Farley, and Lehmann (1990), and Van den Bulte and Stremersch (2004).

accuracy. In the case of PST prediction, ignoring accuracy can have strategic implications; this is explained with a practical example.

Consider the answering machine, introduced in the 1980s. The sales growth data are presented in Table 1.

Table 1
Answering Machine Sales

Year	**T**	**Sales ('000)**
1982	1	850
1983	2	2200
1984	3	3000
1985	4	4220
1986	5	6450
1987	6	8800
1988	7	11100
1989	8	12500
1990	9	13560
1991	10	15380
1992	11	14590

Now, let us assume that we are at 1988 (i.e., $t = 7$). Sales have grown from 850,000 in 1982 to 11.1 million in 1988. At $t = 7$, the product is well into the growth stage, and hence it would be useful to predict PST now. The available data set has seven data points. Fitting the Bass model on this data set, we obtained estimates of the model parameters, which were then used to conduct an extrapolated estimation of the PST with the following results: $\widehat{PST} = 7.16$ (SE: 0.91) and the 95% CI for $\widehat{PST} = (6.82, 7.49)$.

The first 7 years' data predict that peak sales would happen at year 7.16. This predicted PST is statistically significant; that is, the 95% CI of the estimate (6.82, 7.49) does not contain zero. However, the 95% CI does not include *the* actual *PST*, which is 10 (see Table 1). This means that the predicted PST in the case of answering machines is not accurate because it is significantly lower than the actual PST. The estimated PST is earlier by 2.84 years, and even the upper bound of the estimate is off by 2.5 years. If the marketing, operations, and R&D managers use such predictors to devise their respective strategies, the capital markets may unnecessarily punish the company value (e.g., Tesla).

Of course, at the time when a manager is predicting the PST, the true PST remains unknown. Can the manager at least assess how far the predicted PST is from accuracy? To help the manager in this accuracy assessment, we develop a metric to assess the PST predicted with initial years' sales data. The development of such a metric is possible because the sales growth of a new durable is a well-behaved function that is governed by the established adoption theory. We derive our proposed metric from the diffusion theory and use multiple data sets to offer empirical support on why the proposed metric would help in assessing how far away the estimated PST is from the true value. Investors and product managers can use our metric to assess the accuracy of their PST estimates and make more informed decisions on their strategies concerning the oncoming PST.

The rest of the paper is organized as follows. In section 2, we develop a new metric to assess the accuracy of the extrapolation-predicted PST and derive a functional form for it. In section 3, we empirically prove its role, whereas in section 4, we provide managerially useful guidelines derived from our empirical data. In section 5, we conclude the paper and offer managerial directions.

2. Metric to Assess the Accuracy of Predicted PST

As mentioned earlier, many researchers have shown that the sales growth pattern of a new product is driven by consumer voice. However, variables such as price, advertising, and other random disturbances, which we collectively call market noise for our purposes, will also affect the sales figures. Thus, we split the sales growth pattern into two parts: that governed by consumer voice and that by market noise. If consumer voice dominates market noise, then one would see a sales growth curve that closely follows a theoretically defined diffusion pattern, and in that case, the initial years' sales data will have clear information on the oncoming PST. If, on the other hand, market noise dominates consumer voice, we will see a jagged sales curve that will make it hard to discern a clear diffusion pattern, and in that case, the information contained in the initial years' sales data would be rather murky on the oncoming PST. In line with this reasoning, we propose voice over noise (VON) as a metric to assess whether the initial years' sales data could predict an accurate PST or not.

$$\text{Metric to Assess Accuracy of Predicted PST} = \text{VON} = \frac{Consumer\ Voice}{Market\ Noise} \tag{1}$$

We propose that with a higher VON, one can expect higher accuracy while predicting PST with initial years' sales data.

2.1 Deriving VON, the Metric to Assess Accuracy of Predicted PST

Our focus is on the rate of sales growth that occurs in the time between the launch period and PST and to what extent it is influenced by consumer voice vis-à-vis market noise. Let r denote the average periodic growth rate from t = 0 to t = PST. Then we have:

$$S_{obs}(PST) = S_{obs}(0) \cdot (1 + r)^{PST} \tag{2}$$

where $S_{obs}(t)$ is observed sales at time t. Because it is always convenient to handle growth rate on a logarithmic scale, we take the log of both sides of equation (2) and rearrange the terms to get:

$$Ln(1 + r) = \frac{1}{PST} Ln\left(\frac{S_{obs}(PST)}{S_{obs}(0)}\right) \tag{3}$$

Let us split observed sales into their two components: theoretical sales (i.e., driven purely by consumer voice) and sales influenced by market noise. Noting that the annual sales of a typical, new consumer durable grows from hundreds in the launch period to millions in the PST, it is prudent to add market noise multiplicatively to the theoretical sales (Van den Bulte & Lilien, 1997; Venkatesan, Krishnan, & Kumar, 2004; Niu, 2006). Accordingly, we have:

$$Sales_{observed}(t) = Sales(t)_{Consumer-Voice} \times \varepsilon(t) \tag{4}$$

where $Sales(t)_{Consumer-Voic}$ is the theoretical sales expected to happen at t if the product diffuses in a world of zero market noise and $\varepsilon(t)$ represents market noise, which is assumed to be log-normal with mean 1 and variance $\exp(\sigma^2) - 1$.

Applying this split to the two time periods of interest—that is, the launch period and PST—we get:

$$S_{obs}(0) = S(0) \cdot \varepsilon(0)$$

$$S_{obs}(PST) = S(PST) \cdot \varepsilon(PST) \tag{5}$$

where $S_{obs}(0)$ and $S_{obs}(PST)$ are, respectively, the observed launch period sales and peak sales, $S(0)$ and $S(PST)$ are the respective theoretical launch period sales and peak sales, and $\varepsilon(.)$ is the market noise represented by a log-normal term with mean 1 and variance $\exp(\sigma^2) - 1$, which is assumed to be independent-and-identically-distributed (IID) across observations. In other words, both $\varepsilon(0)$ and $\varepsilon(PST)$ come from the same log-normal distribution.

Taking the log on both sides of both the expressions in equation (5), we get:

$$Ln[S_{obs}(0)] = Ln[S(0)] + Ln[\varepsilon(0)]$$

$$Ln[S_{obs}(PST)] = Ln[S(PST)] + Ln[\varepsilon(PST)] \tag{6}$$

In equation (6), $Ln[\varepsilon(.)]$ follows the normal distribution because $\varepsilon(.)$ has been assumed to follow a log-normal distribution. Hence we have:

$$Ln[S_{obs}(0)] = Ln[S(0)] + \varepsilon_N(0) \tag{7}$$

$$Ln[S_{obs}(PST)] = Ln[S(PST)] + \varepsilon_N(PST) \tag{8}$$

where $\varepsilon_N(.)$ is normal error and is IID across *t*. The log-normal distribution with mean 1 and variance $\exp(\sigma^2) - 1$ becomes the normal distribution with mean $(-\sigma^2/2)$ and variance σ^2.

Subtracting the launch period sales (equation 7) from the peak sales (equation 8), we get:

$$Ln\left(\frac{S_{obs}(PST)}{S_{obs}(0)}\right) = Ln\left(\frac{S(PST)}{s(0)}\right) + \varepsilon_N(0, 2\sigma^2) \tag{9}$$

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

Substituting equation (9) in growth rate equation (3), we have:

$$Ln(1+r) = \frac{1}{PST} Ln\left(\frac{S_{obs}(PST)}{S_{obs}(0)}\right) = \frac{1}{PST} Ln\left(\frac{S(PST)}{S(0)}\right) + \frac{1}{PST}\varepsilon_N(0, 2\sigma^2) \quad (10)$$

in which the observed growth rate in the time period {0, PST} is split into its two components, namely, consumer voice and market noise-induced growth rates.

We are interested in modeling the relative strength of consumer voice over market noise. Noting that the standard deviation of the error term (i.e., the second term on the right-hand side of equation 10) is $\frac{\sqrt{2}}{PST}\sigma$, the VON is obtained by taking the ratio between the first term and the square root of the second term as follows:

$$VON = \frac{Voice}{Noise} = \frac{\frac{1}{PST} Ln\left(\frac{S(PST)}{S(0)}\right)}{\frac{\sqrt{2}}{PST}\sigma} = \frac{1}{\sqrt{2}\,\sigma} Ln\left(\frac{S(PST)}{S(0)}\right) \quad (11)$$

The two theoretical sales figures, $S(0)$ and $S(PST)$ in equation (11), can be evaluated with any good diffusion model whose parameters could be estimated with a given set of initial years' sales data. Each model will give a unique set of sales figures, although the resulting estimate of noise (σ) will be commensurate with those figures. Thus, VON measured from the initial years' sales data is model specific. For reasons mentioned in the introduction, we choose the parsimonious and empirically proven Bass model, which states that:

$$S(t) = [M - CS(t)][p + q\frac{CS(t)}{M}] \quad (12)$$

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 $S(t) = \frac{dCS(t)}{dt}$, Bass (1969) reformulated the model into a differential equation, the solution of which gives the following sales function:

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

The sales function, S(*t*), is expressed as a function of t, the only independent variable in the model.[3] Our goal is to use the sales equation and derive a mathematical function for $\frac{Sales(PST)}{Sales(0)}$ found in equation (11). Sales(0) can be obtained from expression (13) directly by letting *t* be zero. To obtain Sales(PST), expression (13) must

---

[3] This makes the model parsimonious, which is highly important for prediction purposes, especially when we consider extrapolation-based prediction as we do with PST. In spite of just one independent variable, the Bass model has performed remarkably well in the empirical tests on many product categories across several countries and decades.

be differentiated with respect to t and set to zero to attain PST, which, when inserted back into the equation, will give us the peak sales, S(PST). It can be shown that (Bass, 1969):

$S(0) = pm$, and $S(PST) = \frac{m(p+q)^2}{4q}$

Inserting these two theoretical sales functions in the VON equation (11) and rearranging terms, we get:

$$VON = \frac{1}{\sqrt{2}\,\sigma} Ln\left(\frac{S(PST)}{S(0)}\right) = \frac{1}{\sqrt{2}\sigma} Ln\left(\frac{(1+q/p)^2}{4(q/p)}\right) \quad (14)$$

VON as expressed in equation (14) is the metric we propose to assess the accuracy of the PST predicted with initial years' sales data.[4] This represents the relative strength of consumer voice over market noise in shaping the sales growth of a new product. If the initial years' sales growth is influenced more by consumer voice than by market noise, we expect to have a higher VON (expression 14). We had already conjectured that sales growth influenced more by consumer voice indicates a theoretically tight predictable diffusion pattern. Hence, putting these two inferences together, we claim that if the initial years' sales data of a new product contain a high VON, then a more accurate prediction of PST is possible.

3. Empirical Analysis on How VON Can Help Assess PST Prediction Accuracy

3.1 Meta Dataset for Analysis

We use pre-PST sales data for our empirical analysis. Bearing in mind that such data could have different cutoff years before PST, our first step is to build a metadata set for empirical analysis. Consider a new product, whose PST is well beyond 5 years. At the end of the 5th year after introduction, the manager can use the 5 years' sales data to estimate the parameters of the Bass model and use those estimates to predict PST. Then, at the end of the 6th year, the manager can use the 6 years' sales data to obtain another prediction for PST. The manager can keep predicting a PST every year until he or she actually goes past the PST. Assuming this product hits peak sales in year 9, the manager can potentially predict PST every year from years 5 to 9 for a total of five PST predictions, each with a longer data set than its predecessor. Thus, for a researcher, every product actually gives multiple data sets from the initial years' sales figures, depending upon how many initial years are to be considered.

We have sales data on 15 different product categories for our empirical test. The categories are listed in column 2 of Table 1 and range from cordless phones of the 1980s to the internet of the 2000s.

---

[4] Researchers in engineering and image processing use a similar metric to assess the efficacy of a signal called signal-to-noise ratio. See http://en.wikipedia.org/wiki/Signal-to-noise_ratio for more details.

Table 2
Products Used in the Empirical Analysis

Sl #	Product	First-use Decade	Energy Saver/ Entertainment	Observed PST (Years After Introduction)	Potential Number of Data Sets (*PST*–5+1)	Number of Usable Data Sets
COL 1	COL 2	COL 3	COL 4	COL 5	COL 6	COL 7
1	Answering machine	1980	Saver	9	5	4
2	VCR	1980	Entertainment	20	16	13
3	Desktop PC	1990	Saver	12	8	7
4	Fax machine	1990	Saver	11	7	7
5	Cordless phone	1980	Saver	14	10	10
6	DVD player	2000	Entertainment	9	5	5
7	Digital camera	2000	Entertainment	10	6	5
8	CD player	1990	Entertainment	16	12	11
9	Cell phone	1990	Saver	15	11	8
10	Digital TV	2000	Entertainment	13	9	9
11	Cell-Finland	1980	Saver	12	8	8
12	Internet-China	2000	Saver	12	8	8
13	Cellphone-China	1990	Saver	14	10	10
14	Printer	1990	Saver	13	9	4
15	Cellphone-India	2000	Saver	10	6	6

For every product j mentioned in Table 2, we focused on the early years' data points up to the PST of product j, PST_j (see column 5) and formed the following subsets as described earlier. The first subset had sales data points from $t = 1$ to $t = 5$ years (i.e., the first five sales data points), the second had sales data points from $t = 1$ to $t = 6$ years, the third from $t = 1$ to $t = 7$ years, and so on, until the last with sales data points from $t = 1$ to $t = PST_j$. Thus, for product j with PST at PST_j we have ($PST_j - 5 + 1$) subsets (see column 6). For example, consider the DVD player in row 6. The observed PST is year 9, and so the DVD player will give us (9-5+1) = 5 subsets. The potential number of subsets contributed by each product is given in column 6 of Table 1.

Each subset belonging to a product is a sample element in our analysis. Consider the DVD player again, which has contributed five subsets. Each of these five subsets is a sample element in our analysis. Thus, we collected 130 subsets (the sum of column 6 in Table 1) in all, thus leading to 130 sample elements. Each sample element, which is a data set of sales growth, was further scrutinized to determine whether it had useful information about the sales growth of the concerned product. We fitted the Bass model on each sample element for this screening, and out of the 130 subsets (or sample elements), 15 had problems with convergence, yielded estimates with high standard errors,[5] suffered from takeoff phenomenon (Golder & Tellis, 1997),[6] or were considered outliers. These 15 sample elements were removed, yielding a sample size of 115 (i.e., 130-15). Column 7 of Table 2 gives the number of usable sample elements (i.e., subsets) each product yielded. We call this the **meta set**.

3.2 Estimating PST, Noise, and VON on the Meta Set

Having formed the meta set, we now characterize each of its elements in three dimensions, namely, the PST it predicts, the noise it has, and the value of VON it holds.

Consider product *j* whose PST is PST_j. We will get (PST_j−5+1) number of subsets. The Bass model was estimated on each subset using the log-log version of the Srinivasan and Mason (1986) method, which sets up the estimation equation (without the error term) 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 cumulative density function at time *u*, which is simply equation (12) where *m* was taken to be 1. Whereas Srinivasan and Mason (1986) included normal error additively, we include log-normal error multiplicatively in the above expression. Taking the log on both sides, we get:

$$Ln(Sales_j(t)) = Ln(m_{ji}) + Ln(F_{ji}(t) - F_{ji}(t-1)) + \varepsilon \sim N(0, \sigma^2) \tag{15}$$

The multiplicative inclusion of error confirms our operationalization of the focal variable VON (equation 14), which was derived assuming a multiplicative error structure (equation 4) in the Bass model. We used the nonlinear least squares (NLS) procedure in SAS for the estimation.[7]

For each subset or sample element in the meta set, 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, i.e., the unexplained variation (i.e., regression residuals) in the subset,

---

[5] Note that we use the closed-form expression of the Bass model. Although closed-form expression is better than the differential-form expression for forecasting purposes, it does not build the impact of random shocks that might distort the diffusion pattern at discrete points. Such shocks, if present in an early years' sales data set, are likely to yield statistically insignificant estimates.

[6] Three of the 15 products 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 break point (Golder & Tellis, 1997). Including the pre-takeoff data points will distort the estimation of the growth element and hence we removed those points from further analysis. These products were: the answering machine, DVD players, and the digital camera.

[7] We use the log-log version of the estimation equation of Srinivasan and Mason (1986). As mentioned in the introduction, 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.

2. $\widehat{VON}_{J\iota}$ (equation 14), the consumer voice over market noise, and

3. $\widehat{PST}_{J\iota}$, the PST as predicted by the subset's estimates (equation 13).[8]

Note that while the actual PST, i.e., PST_j, is the same for all the subsets pertaining to product j, the predicted PST, i.e., $\widehat{PST}_{J\iota}$, may be different with the different subsets.

To demonstrate our estimation process, consider the estimates obtained from running the five subsets of the DVD player on the Bass model, which are presented in Table 3. DVD(1) through DVD(5) in Table 3 refer to the five subsets of different data set lengths (see columns 1 and 2 of Table 3). The Bass model parameters estimated on the log-log version are given in columns 4, 5, and 6, along with their standard errors in parentheses. Using equation (13) and the bootstrapping technique, the predicted PST and its standard error are evaluated (see column 7 of Table 3, where the corresponding standard errors are reported in the parentheses). Using equation (14) and the bootstrapping technique, VON was evaluated (see column 9 of Table 3).

Table 3
Estimates Obtained From the 5 Subsets of DVD Players

Subset (Sample Element)	Length (# of Data Points)	Actual Peak Sales Time t_j*	p* est. $\widehat{p_{J\iota}}$ (SD)	q* est. $\widehat{q_{J\iota}}$ (SD)	m* est. $\widehat{m_{J\iota}}$ (SD)	t* est. $\widehat{PST_{J\iota}}$ (SD)	Noise est. $\widehat{\sigma_{J\iota}}$	VON est. $\widehat{VON}_{J\iota}$
COL 1	COL 2	COL 3	COL 4*	COL 5*	COL 6*	COL 7*	COL 8	COL 9
DVD(1)	5	9	0.0060 (0.0029)	1.0955 (0.1373)	6131 (1103)	4.8517 (0.8107)	0.3003	14.516
DVD(2)	6	9	0.0061 (0.0024)	0.9296 (0.1222)	9086 (1713)	5.4928 (0.8448)	0.3255	11.113
DVD(3)	7	9	0.0062 (0.0020)	0.8676 (0.0877)	10472 (1236)	5.7394 (0.6806)	0.3076	10.845
DVD(4)	8	9	0.0064 (0.0022)	0.7525 (0.0907)	13241 (1817)	6.3703 (0.844)	0.3719	8.127
DVD(5)	9	9	0.0063 (.0022)	0.6269 (0.0974)	18147 (3493)	7.4283 (1.2042)	0.4413	6.281

* Numbers in the parentheses are standard errors of the corresponding estimates

3.3 Estimating the Impact of VON on the Accuracy of Predicted PST

[8] Because direct estimation of predicted PST, noise, and VON from the $\{\widehat{p_{J\iota}}, \widehat{q_{J\iota}}, \widehat{m_{J\iota}}\}$ estimates will not give us the distributional properties of these three metrics, we use the bootstrapping method with 1,000 iterations to obtain the variances of these three estimates.

Having characterized each element, where each element is a data set of the meta set in terms of (a) what PST it predicts and (b) VON, our objective is to assess whether VON influences the deviation of predicted PST from the actual PST. Accordingly, we use the following regression on the meta set to estimate this influence.

$$Ln(PST\ Dif) = \beta_0 + \beta_1 \, \mathrm{Ln}\left(\widehat{VON}_{ji}\right), \tag{16}$$

where $PST\ Dif = \left|\widehat{PST}_{ji} - PST_j\right| * 100/PST_j$, and j pertains to product j and i to a subset in it.

We expect β_1 to be negative, indicating that a larger VON would ensure a more accurate PST prediction. This is because a large VON of a given data set signifies that the data collectively capture the underlying consumer voice so substantially that they overcome the negative influence of noise.

We now consider three additional explanatory factors that are likely to influence the observed variation in the predicted PST accuracy in the 115 sample elements of the meta set.

3.3.1 Data set length

The first factor we include is the number of data points in each subset used in the estimation. 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 & Rao, 1990; Mahajan & Sharma, 1986; Parker, 1994; Srinivasan & Mason, 1986; Van den Bulte & Lilien, 1997). Noting that we are dealing with small data sets, even additions of one data point make a significant impact on the estimation and subsequent prediction because the information it brings will be of substantial value to the small base. 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. Extending this argument, we conjecture that a longer subset (i.e., one with a higher number of data points) will result in better estimation of consumer voice and hence would lead to less deviation in the prediction of PST. Hence, we expect the impact of data set length on the PST deviation to be negative.

Next we consider two socioeconomic factors derived from Rogers (1995, p.206): "the nature of communication channels" and "the nature of the social system."

3.3.2 Changes in social communication channel over time (word-of-mouse effect)

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 came in when the social media and social websites started going mainstream in the 2000s. The products introduced in the 2000s may experience stronger consumer-voice, which would affect the accuracy of PST prediction. The 15 products in our data come from three different decades (see Table 3). In Column 3, we provide the average PST of the products pertaining to each decade and in Column 4 the mean percentage deviation of predicted PST for all the products belonging to that decade. These were estimated from the 115 subsets of the products as described earlier. The products introduced in the 2000s seem to reach the peak-sales much quicker (10.8 years as against 13 to 14 years for the products from 1980s and 1990s). Further, the percentage deviation in predicted PST is in slightly lower for the products introduced in the 2000s. Accordingly, we will include in equation (16) a word-of-mouse' effect, a dummy for products from 2000 to signal the extra influence they get from social media, blogs, and other new age communication channels. We expect the 2000s' products to have a negative effect on the PST deviation.

Table 4
Word-of-mouse effect

Decade	# of Products	Average Actual PST	Estimated Mean *PST Diff.*
1980	4	13.8	26.94%
1990	6	13.0	27.29%
2000	5	10.8	25.68%

3.3.3 Product type effect

Following Rogers' (1995) argument that the nature of social systems affects the new product adoption rate, it is reasonable to think that products meant to save time and effort for consumers would face a different adoption pattern than those designed to bring entertainment value (Horsky, 1990). In our data set, out of the 15 products, 10 are energy-saving and five are entertainment providers. When analyzed in isolation (as reported in Table 5), they however do not seem to show any differences in the speed of adoption (i.e., both have the same PST) or differences with respect to the accuracy of predicted PST (see Table 5). We will however let the data reveal the impact statistically.

Table 5
Product type effect

Type	# of products	Average Actual peak-sales time	Estimated *PST Dif*
Energy Saving	10	10	25.92 %
Entertainment	5	10.46	27.57 %

Equation (16) is now expanded to include the three additional explanatory factors:

$$Ln(PST\ Dif) = \beta_0 + \beta_1 \mathrm{Ln}\left(\widehat{VON}_{ji}\right) + \beta_2\ Ln\left(DataLength_{ji}\right) + \beta_3\ Dum\left(ProductType_j\right) + \beta_4\ Dum(WordOfMouse_j) \tag{17}$$

where $PST\ Dif = \left|\widehat{PST}_{ji} - PST_j\right| * 100/PST_j$, $\widehat{VON}_{ji}$ = Value of VON estimated from fitting Bass model to subset i of product j, $DataLength_{ji}$ = Number of data points in subset i of product j,

$Dum(ProductType_j)$ = 0 if product j is an energy saver and 1 if it is an entertainment provider, and $Dum(WordOfMouse_j)$ = 1 if product j belongs to decade 2000–2010, and 0 otherwise.

3.3.4. Regression results on meta-set

The actual PST of the 15 products averaged 12.4 years. The average subset length in the meta set was 9.64. The average deviation of predicted PST from the actual PST across all 115 observations was estimated to be 26.93%,[9] and the average *VON* estimated across all the observations was found to be 14.24. The results of regressing equation (17) on the meta set of 115 observations are provided in Table 6.

Table 6
Influence of VON on predicted PST's deviation or lack of accuracy

Parameter	Variable-concerned	Estimate	Std. Error	p-value	Sig or Not Sig
β_0	Intercept	6.9780	0.8192	<.0001	Sig
β_1	VON	-0.6062	0.1478	<.0001	Sig
β_2	Data set length (# of data points)	-1.0454	0.2693	0.0002	Sig
β_3	Product type (entertainment vs. energy saver)	-0.0205	0.1732	0.9062	Not Sig
β_4	Word-of-mouse	-0.5357	0.1937	0.0067	Sig

VON has a negative and significant impact on the PST deviation (i.e., the lack of accuracy of forecast) even in the presence of three other independent variables. This implies that a higher VON would significantly indicate a higher accuracy of the predicted PST. Put differently, if a new product enjoys a clearer consumer voice—that is, one stronger than the market noise—the manager would be able to not only understand its growth pattern but also attain a more accurate prediction of the PST. The *relative* strength of voice with respect to noise is what gives the proposed metric the power to assess the accuracy of extrapolated prediction of PST.[10]

[9] The genetic algorithm-based estimation procedure used by Venkatesan et al. (2004) reported deviations of estimated PST forecasts in absolute numbers for seven products: clothes' dryers, room air conditioners, color TVs, ultrasound, mammograms, foreign language, and accelerated program. When we matched these reported deviations with the actual PST information of these products as we have in our data, we found that the average percentage deviation of forecasts of their study is 23.72%. Although we do not use these products in our analysis, we ran a separate analysis on these data sets and calculated the deviations of predicted PST (i.e., using log-log method in NLS setting) and found the average to be 22.17%. This shows that the log-log version in NLS estimation seems to perform with almost the same effectiveness as the genetic algorithm method, with the added advantage of being much easier to use, especially for a manager.

[10] As mentioned earlier, Van den Bulte and Stremersch (2004) had used, in a different context, the metric q/p to represent the shape of the diffusion curve. In line with the Occam's Razor principle of parsimony (Sober, 1994), we tested whether the simple q/p could do the job of the more complicated VON. We took equation (14) and therein replaced VON with q/p and repeated the empirical analysis on the 115 early years' data sets gathered from 15 new products. We found that q/p

What makes this finding even more interesting is that products of the 2000s appear to enjoy the benefits of word-of-mouse, which not only makes their sales grow at a quicker pace (see Table 4) but also enables a more accurate PST prediction (i.e., coefficient β_4 in Table 6 is negative and significant). The accuracy of PST prediction is not different between product types, whether they are energy savers or entertainment providers.

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 PST 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 2 before predicting PST. 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 PST with caution.

4. Directions for Use of VON

The empirical results show that PST prediction is more accurate if the sales growth pattern exhibits a substantially high VON. A key managerial question at this juncture would be: How high is considered high enough? Is there a threshold VON that a manager could use as a benchmark to decide whether a particular observed VON is high enough? Based on our empirical analysis of the 15 product categories, we develop such a threshold VON. We use equation (17) as a predictive model, keeping only those variables found to be significant. In equation (18) below,

$$Predicted\ (PST\ Dif)_{ij} =$$

$$\exp\left[\beta_0 + \beta_1\ Ln((1+C)VON)_{ij} + \beta_2\ Ln(DataLength)_{ij}\ + \beta_5\ (WordOfMouse)_{ij}\right] * \frac{100}{PST} \qquad (18)$$

where subscripts i and j refer to the subset and the product, respectively. For the independent variables on the right-hand side, we use the values used in the estimation of equation (15), along with their corresponding parameter estimates. Accordingly, we have: $\beta_0 = 6.98$, $\beta_1 = -0.61$, $\beta_2 = -1.05$, and $\beta_5 = 0.54$. What is however unknown on the right-hand side of equation (18) is C, a parameter we have newly introduced to modify the value of the VON_{ij} for all $\{i, j\}$. If C = 0, there is no modification to VON. As we start increasing the value assigned to C, the predicted PST difference will start decreasing monotonically thanks to β_1 being negative. We are interested in finding out at what critical value of C—which we may call C*—the PST difference would become insignificant—that is, when the predicted PST would be statistically close to the actual PST. The more accurate a manager wants the predicted PST to be, the higher the C*. If a manager wants

---

alone is not sufficient for assessing the accuracy of PST forecast and that by augmenting it with noise, we get a better metric (i.e., VON). Results can be obtained from the authors.

to be extremely confident in his forecast, we have to choose a very high C*. This will happen, for example, when the manager has to make a risky investment decision that depends critically on the PST. We chose three levels of confidence and evaluated C* and the corresponding threshold-VON values. These are provided in Table 7 below.

Table 7
Threshold VON derived using the empirical data on 15 product categories[11]

Level of Confidence Needs to be Placed on the Prediction	Low	Medium	High
Threshold VON	25	43	70

For example, if the manager wants to be extremely cautious with the prediction, he or she can choose a high confidence level—that is, a threshold-VON of 70. This implies that if the manager predicts PST using the $\{p, q\}$ estimates that also yield a VON (evaluated as explained in section 3) lower than 70, then he or she should not rely on the predicted PST.[12]

4.1 Recommendation for a manager

The following are recommendations for managers to ensure accurate PST predictions:

Step 1. The manager should collect as many data points as possible and make sure that the known intra-period noises such as seasonality and takeoff periods are removed.

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 3.2). If the estimates are found to be insignificant, the manager should abandon the prediction exercise and wait to gather more data points.

Step 3. If the estimates are significant, the manager should use a bootstrapping technique and estimate PST (equation 13) along with its standard deviation and the VON (see equation 14 and section 3.2). If the predicted PST turns out to be insignificant (i.e., not significantly different from zero), the manager should abandon the prediction exercise. If it is significant, the manager should look at VON. Depending upon how cautious the manager wants to be with respect to using the predicted PST, the manager can use Table 7 and choose to use

[11] Note that "no difference between the forecast and actual PST" is what we want to show beyond reasonable doubt, and hence using a significant level of, say, 10% is scientifically more conservative than using a significant level of 1%. Accordingly, a higher significance level would give a more conservative C*—that is, a higher C* and a higher threshold-VON. Using our empirical data, we found that the threshold-VON of 70 corresponds to a 10% significance level, and the threshold-VON of 25 corresponds to a 1% significance level. We used the estimated distribution properties of predicted PST and derived the z-statistic to arrive at the various critical values.
[12] We also carried out a simulation to evaluate a threshold VON; it turned out to be 13. Details can be sent on request.

or ignore the predicted PST. If the manager is forced to use the predicted PST against the recommendation from Table 7, he or she should proceed with caution.

Step 4. The manager may wish to use another procedure (e.g., survey or analysis of an analogous product in a similar market) to corroborate the PST prediction accuracy.

5. Conclusion and Directions for Future Research

New product managers handling new products whose sales have yet to reach their peak would more likely want to use the scarce initial years' data points and estimate a diffusion model to predict when the peak sales would happen. That prediction would help managers conduct a better production planning and marketing strategy and also possibly help them find an optimal time to launch the next version of the product if it happens to be a multigenerational product like a smartphone. However, the predicted PST may be statistically significant (i.e., different from zero) but not accurate (i.e., different from the true value of PST).

Based on the literature of new product diffusion, we split the sales growth pattern into two parts: the part governed by consumer voice and that by market noise. The accuracy in PST prediction will be low if there is lot of market noise in the data for a given level of consumer voice, and this is what we show in this research using a new metric called VON. We gathered 115 initial years' data sets from 15 products introduced in the 1980–2010 period and showed that VON is a good indicator of whether predicted PST can be relied upon with respect to its accuracy.[13] Specifically, if a new product enjoys a clearer VON—that is, consumer voice is stronger than the market noise—the manager would be able to attain a more accurate prediction of the PST. We further ran a simulation exercise to garner threshold values of VON for different confidence levels.

Besides the impact of VON, we found that when trying to predict PST, it is always better to wait as long as possible and gather more sales data points. Another finding, perhaps more interesting to managers, is that the products introduced in the 2000s (i.e., 2000–2010) seem to yield more accurate PST than those introduced in the 1980s and 1990s, indicating the special role played by word-of-mouse (i.e. information spread through internet, social media and emails) in the post-2000 period.

This paper contributes to management research in the following ways:

---

[13] In a separate analysis, we added 12 more products to the set of 15 from the pre1980s era such as color TVs and room air-conditioners, typically found in many growth model papers. Running the analysis on these 27 products, we obtained very similar results that indicated the robust impact of VON and data set length on the accuracy of PST forecast and one-step-ahead forecasts. Details can be sent upon request.

1. We have introduced a new metric, VON, that helps a manager assess the accuracy of PST prediction. The accuracy is different from the statistical significance of a forecast because the latter indicates whether the forecast is different from zero, while the former indicates whether the forecast is closer to the actual value.
2. We have introduced a new estimation method for the growth models, namely the log-log version in NLS. The log-log formulation enables the error term to be proportional to the actual sales number because the error is included multiplicatively. In the traditional methods, researchers use additive error. Multiplicative error is more natural because the sales number in a typical diffusion data set (i.e., of a consumer durable) varies over a huge range, from a few hundred in the initial stages to millions at the PST. The log-log version is easy to run and performs well, as discussed in section 3.3.4 and footnote 9. It can be used for estimation of diffusion models in place of the traditionally-used NLS procedure.
3. The distinction between statistical significance of the forecast and forecast accuracy does not get its due attention in marketing literature, and by bringing it to the forefront through our metric, VON, we are able to offer strategic contribution to diffusion research.
4. Our proposed metric, VON, captures the <u>relative</u> strength of diffusion force (i.e., consumer voice) over market noise and adds to the research of Van den Bulte and Stremersch (2004), who used a rather straightforward measure (q/p) to represent diffusion force, although in a different context.

There are a few areas one can pursue 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 investigating further because one may be able to discern why market noise seems to play a smaller role vis-à-vis consumer voice; perhaps social networks and viral diffusion are contributing to this phenomenon. It would be interesting to know how today's fast and mobile consumer voice affects PST prediction. Second, noting that market noise has been found to play a highly important role in prediction, future researchers can look into diagnosing the noise component and advising managers on remedial actions to be taken to reduce the noise. Third, one can consider the competitor's effect. For instance, consider the fierce competition between Apple and Samsung on smartphones. Consumers always compare the latest models between the two brands (even before the products' launch), and the companies advertise the differences in an attempt to influence consumers' perceptions. We may expect that consumer voices on competitors' products could well affect each product's growth curve; the link can now be examined through VON. Fourth, because consumer voice and market noise could both have a time lag effect, 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.

REFERENCES

Bass, F. M. 1969. A new product growth model for consumer durables. *Management Science*, 15(5), 215−227.

Bass, F. M., T. V. Krishnan, and D. C. Jain. 1994. Why the Bass Model Fits without Decision Variables. *Marketing Science*, 13(3): 203-223.

Bass, F. M., K. Gordon, T. L. Ferguson, M L. Githens. 2001. DIRECTV: Forecasting Diffusion of a New technology prior to Product Launch. *Interfaces* 31(3, Part 2 of 2), S82-S93.

Golder, P. N., and G. J. Tellis. 1997. Will It Ever Fly? Modeling the Takeoff of Really New Consumer Durables. *Marketing Science*, 16(3): 256-270.

Ho, T.-H., S. Savin, and C. Terwiesch. 2002. Managing Demand and Sales Dynamics in New Product Diffusion Under Supply Constraint. *Management Science*, 48(2), 187-206.

Horsky, D. 1990. A diffusion model incorporating product benefits, price, income and information. *Marketing Science*, 9(4), 342-365.

Hu, Y., and C. van den Bulte. 2014. Nonmonotonic Status Effects in New Product Adoption. *Marketing Science*, online print.

Jain, D. C., and R. C. Rao. 1990. Effect of price on the demand for durables: Modeling, estimation, and findings. *Journal of Business & Economic Statistics*,8(2), 163-170.

Kovach, S. 2018. The iPhone has reached its peak, and now it's time to change how you think about Apple. Retrieved May 21, 2019 from https://www.cnbc.com/2018/11/02/the-iphone-has-peaked.html

Lenk, P. J., and A. G. Rao. 1990. New Models from Old: Forecasting Product Adoption by Hierarchical Bayes Procedures. *Marketing Science*, 9(1): 42-53.

Mahajan, V. and S. Shama. 1986. A simple algebraic estimation procedure for innovation diffusion models of new product acceptance. *Technological Forecasting and Social Change*, 30(4), 331-346.

Niu, S.-C. 2006. A Piecewise-Diffusion Model of New-Product Demands. *Operations Research*, 54(4): 678-695.

O'Kane, S. 2019. Tesla is selling fewer cars, but says it has 'sufficient' cash. Retrieved May 21, 2019 from https://www.theverge.com/2019/4/3/18291922/tesla-sales-decline-q1-2019-deliveries-production

Parker, P.M. 1994. Aggregate diffusion forecasting models in marketing: A critical review. *International Journal of Foresting*, 10(2), 353-380.

Rogers, E.M. 1962. 1995. *Diffusion of Innovations*. New York: Free Press.

Schmittlein, D. C., and V. Mahajan. 1982. Maximum Likelihood Estimation for an Innovation Diffusion Model of New Product Acceptance. *Marketing Science*, 1 (1), 57-78.

Silverman, G. 2011 *The Secrets of Word-of-Mouth marketing: How to Trigger Exponential Sales Through Runaway Word of Mouth*. 2nd ed. AMACOM, New York, NY.

Sober, E. 1994. Let's Razor Occam's Razor. In Knowles, Dudley (ed.). *Explanation and Its Limits*. Cambridge University Press. pp. 73–93.

Srinivasan, V., and C. H. Mason. 1986. Nonlinear least squares estimation of new product diffusion models. *Marketing Science*, 5(2), 169-178.

Statista.com 2019. Unit sales of the Apple iPhone worldwide from 2007 to 2018 (in millions). Retrieved May 21, 2019 from https://www.statista.com/statistics/276306/global-apple-iphone-sales-since-fiscal-year-2007/

Sultan, F., J. U. Farley, and D. R. Lehmann. 1990. A meta-analysis of applications of diffusion models. *Journal of Marketing Research*, 27(1), 70-77.

Van den Bulte, C., and G. L. Lilien. 1997. Bias and systematic change in the parameter estimates of macro-level diffusion models. *Marketing Science*, 16(4), 338−353.

_______________, and S. Stremersch 2004. Social Contagion and Income Heterogeneity in New Product Diffusion: A Meta-Analytic Test. *Marketing Science*, 23(4), 530-544.

Venkatesan, R., T. V. Krishnan, and V. Kumar. 2004. Evolutionary estimation of macro-level diffusion models using genetic algorithms: An alternative to nonlinear least squares. *Marketing Science*, 23(3), 451-464.