Understanding New Products' Market Performance Using Google Trends

Abstract:

This paper seeks to empirically examine diffusion models and Google Trends' ability to explain and nowcast the new product growth phenomenon. In addition to the selected diffusion models and Google Trends, this study proposes a new model that incorporates the two. The empirical analysis is based on the cases of the iPhone and the iPad. The results show that the new model exhibits a better curve fit among all the studied ones. In terms of nowcasting, although the performance of the new model differs from that of Google Trends in the two cases, they both produce more accurate results than the selected diffusion models.

Keywords: big data; Google Trends; new product; diffusion; nowcasting.

1. Introduction

Firms need to bring new products to the market in order to stay ahead of the competition. At the same time, it is vital for firms to understand and predict the market performance of new products, as they are linked to a set of marketing and operational problems that ultimately affect firm profitability (Negahban et al., 2014; Qin et al., 2012).

The Bass model (Bass, 1969) captures the essence of market growth phenomena by using parameters that process intuitive interpretations in the diffusion context (Bass, 2004). It is capable of explaining and predicting empirical cases of new product diffusion, as evidenced in the prior literature. Therefore, the model and its extensions—see a review by Peres *et al.* (2010)—are widely employed and have applications in various sectors. However, as the Bass model was developed for product categories, it needs to incorporate the competition effect in order to study the market demand of individual brands. Furthermore, the market performance of a new product can be exposed to many influences, such as marketing-mix (Danaher *et al.*, 2001), technological advances (Jiang *et al.*, 2012), seasonal effect (Radas *et al.*, 1998), cross-country issues (Putsis *et al.*, 1997), and repeated purchase (Guo, 2014). This complexity can be further increased by the coupling and dynamics of those influences.

Bass (2004) emphasized the importance of model simplicity and stated that "*simple and elegant mathematical models, often referred to as 'beautiful', that match well with the phenomenon being studied will have appeal in the arena of competing ideas about the phenomenon*". However, since the diffusion influences are diverse and dynamic, it can be difficult to frame the boundary of a diffusion case, to identify the key influences in the process, and thus to structure an appropriate model for theory development and business practice. Although researchers have been trying to develop more complex models to capture more influences, the influences included in such models are still limited and exclusive, and their reported empirical improvement is usually modest—thus, the original Bass model remains the most popular in this field due to its simple structure and generalizability.

For instance, Kim *et al.* (2000) developed a diffusion model that integrates the influences of both product competition and generation, and required the estimation of 16 parameters when studying the wireless telecom service industry in Hong Kong with three product categories of only one or two generations in use; that is, the Pager (one generation), the Cellular Phone (two generations), and the Cordless Telephone 2 (one generation). Due to the model complexity, recent diffusion models that integrate the two influences (Kiesling *et al.*, 2012; Negahban *et al.*, 2014; Samuel Sale *et al.*, 2017; Stummer *et al.*, 2015; Yan *et al.*, 2011) mostly use simulations to understand the phenomena and to generate new insights, and this can limit their real-world relevance.

The shorter product life cycles of today's market also pose a significant challenge for forecasting product demand and sales (Chien *et al.*, 2010). More specifically, we now require better methods to analyse the limited information in order to inform operations (Li *et al.*, 2015; Li *et al.*, 2012). As diffusion models become increasingly sophisticated, they usually demand a large amount of data from different sources as input, and/or require advanced techniques for the estimation of massive parameters with limited data—see an example of the use of Bayesian method in Albuquerque *et al.* (2007)—which could hinder the application of the models; it is becoming the preserve of those few firms that can access both the required data and computing power. Therefore, we believe that new methods suited to understand the market performance of new products are required.

Meanwhile, big data is becoming one of the hottest business topics (Nunan, 2015; Nunan *et al.*, 2013), due to its potential to enable more advanced decision making (Kiron, 2013; Lamba *et al.*, 2017; Manyika *et al.*, 2011). In particular, a great amount of data is being generated through people's interactions with technology and is being collected by various online platforms, such as Internet search engines, Twitter, Facebook, Wikipedia, and Flickr. This collected data may provide opportunities to more accurately understand aggregated customer attitudes and behaviours towards a new product throughout its life-cycle (Schaer *et al.*, 2018). Our literature review shows that the examination and application of such data are still limited. In particular, the existing literature has been studying the use of the data in forecasting through various techniques such as regressions (Rivera, 2016; Schneider *et al.*, 2016), moving average terms (Li *et al.*, 2017) and mixed data sampling (Smith, 2016). But we still have limited knowledge of how the data sources can benefit the rich new product diffusion literature we have been accumulating.

Therefore, this study takes the lead in examining the use of diffusion models, Google Trends, and combinations thereof in explaining and nowcasting (i.e., forecasting current events for which results have not been revealed) the market performance of new products. In addition to the selected diffusion models and Google Trends, this study proposes a new model that incorporates the two. More specifically, the new model explains customer likelihood of purchase based on the conventional notion of the innovation and imitation effects drawn from the Bass framework, and Google Trends is employed to calibrate the size of the dynamic market potential in the process of market growth.

We introduce the cases of the iPhone and the iPad in our empirical analysis. These products were both subject to numerous influences through their life-cycles, such as technological upgrades, price adjustments, seasonal effects, praise and criticism from the market, and competition from other brands. The empirical results show that the new model is capable of accurately capturing the market dynamics of the products with a performance superior to that of the benchmarks. In terms of nowcasting, both Google Trends and the new model can produce more accurate results than conventional diffusion models.

The rest of the paper is structured as follows. Section 2 reviews the related literature. In Section 3, the methodology of this study is described. Then the model's performance is discussed in Section 4. The last section concludes the study.

2. Related literature

2.1. New product diffusion models

At the aggregated level, the market performance of a new product usually follows a bell-shaped curve in which the level of sales is plotted versus time (Geroski, 2000). As the most popular stream of diffusion modelling, the Bass framework (Bass, 1969) models this phenomenon by multiplying two variables: customer likelihood of purchase through time and the size of the corresponding market potential. More specifically, customer likelihood of purchase is driven by a constant diffusion driver that can be explained by the mass media effect (also known as the innovation effect) and by a dynamic driver that can be explained by the word-of-mouth effect or social contagion effect (also known as the imitation effect); the size of the corresponding market potential, on the other hand, is calculated as the overall market potential (usually assumed as a constant) minus those customers who have already purchased the product.

The literature on new product diffusion models has been continuously expanding on the basis of the Bass model (Bass, 2004) to include other diffusion influences. For instance, it is worth noting that the original Bass model was developed within an ideal environment in which only one generation of a single product exists in the system. In real situations, however, different generations of products can follow each other due to their continuous improvement, and their providers often face competitors offering the same or similar products to the market. Therefore, scholars have developed a number of models suited to illustrate the diffusion phenomena under multi-generation and/or multi-brand/category conditions (e.g., see the reviewed models in Peres *et al.* (2010)).

One issue with this line of research, however, is the complexity of the extended models. For instance, the model developed by Norton *et al.* (1987), although considered parsimonious, requires $(2 + G)$ parameters to explain the diffusion of a product with G generations. When the issues of multi-generation and multi-brand/category are considered concurrently, the modelling process can become even more complicated because both customer likelihood of purchase and market potential may change in relation to each generation and each brand. By taking a more parsimonious approach, for instance, Kim *et al.* (2000) extended the Norton-Bass model by calibrating the market potential of each generation and of each category on the basis of the market dynamics of other generations and categories. Despite this parsimony, 16 parameters still needed to be estimated when considering the case of the Wireless Telecom Service Industry in Hong Kong with three product categories: Pager (one generation), Cellular Phone (two generations) and CT2 (one generation). The increased complexity of the models also requires more data as inputs and more sophisticated parameter estimation techniques, which further limit their practical value.

In the existing literature, the Bass model is extended mainly through two approaches: the first involves modifying the market potential, while the second entails modifying customer likelihood of purchase (i.e., the innovation and imitation effects). In the field of multi-generational product diffusion, modellers usually focus on the dynamic market potential that results from the generation substitution (Jiang *et al.*, 2012). For instance, the Norton-Bass model (Norton *et al.*, 1987) envisages the later generation plundering the customer base of earlier ones when they coexist in the market; the model proposed by Mahajan *et al.* (1996) suggests that, having purchased one generation of a product, customers will become potential customers of the following ones through upgrading or leapfrogging. Meanwhile, scholars have often modified the innovation and imitation effects to explain the cross-brand/category effect. More specifically, those models consider the adoption of one brand/category to have a positive, negative, or no effect on the diffusion of other brands (Chatterjee *et al.*, 2000). In terms of the market potential of each brand/category, two competing views exist in the literature. In the former setting, different brands/categories compete for market share—e.g., Libai *et al.* (2009a)—while the latter is more likely to lead to a steady-state condition in which competing products/categories coexist in the marketplace by targeting different customer niches—e.g., Parker *et al.* (1994).

In addition to the above examples, our review indicates that the two approaches are used (separately or collaboratively) across the literature to model various diffusion phenomena—i.e., generational diffusion with price effect (Tsai, 2013), new product diffusion with brand competition (Libai *et al.*, 2009b), new product diffusion with marketing-mix variables (Bass *et al.*, 1994), software diffusion under the influence of pirate copies (Givon *et al.*, 1995), the growth chasm in the process of diffusion (Van den Bulte *et al.*, 2007), and new product diffusion across countries (Kumar *et al.*, 2002). Therefore, we argue that the two approaches can continue guiding the development of Bass type models to explain the new product diffusion phenomena.

2.2. Online behavioural data – Google Trends

In spite of some discouraging reports (e.g., Pappas (2014)), researchers have been actively exploring the value and potential of new data sources resulting from human interactions with the Internet (Schaer *et al.*, 2018). For instance, scholars analysed the changes in Google search terms (Perlin *et al.*, 2017) and Wikipedia usage patterns (Moat *et al.*, 2013) related to finance, and found that such data could indicate the behaviour of finance market actors. Correlations are also found between influenza outbreaks and Google searches (Preis *et al.*, 2014), and between the atmospheric pressure during Hurricane Sandy and the uploaded

Flickr photos of the event (Preis *et al.*, 2013a). Furthermore, people in the public sector are beginning to pay more attention to the online platforms and to the data that were collected, for instance, during the 2011 riots in England (Panagiotopoulos *et al.*, 2014).

In regard to search traffic, which is the most widely used source of information (Schaer *et al.*, 2018), Google Trends is a data tool that reports how often a particular term has been searched, relative to the overall number of searches during the time of interest. Past studies show that Google Trends data can be applied directly to explain and forecast certain trends of interest (Jun *et al.*, 2017). Apart from the above examples involving financial predictions and influenza outbreaks, Google Trends has been recently shown to be an effective tool in explaining and predicting other phenomena such as referendum results (Mavragani *et al.*, 2016), hotel non-resident registrations (Rivera, 2016) and tourist arrivals (Bangwayo-Skeete *et al.*, 2015). Other search engines with similar functions—such as Naver (in South Korea) and Baidu (in China)—are also subjected to scrutiny in order to predict tour flows (Huang *et al.*, 2017; Kim *et al.*, 2016; Li *et al.*, 2017).

Furthermore, Google Trends has started to attract academic attention in relation to the study of the market performance of products. For instance, Vosen *et al.* (2011) found that Google Trends is a better indicator of private consumption than survey-based indicators; Jun *et al.* (2016) examined the correlation between consumer search activities and their purchase behaviours; Jun *et al.* (2017) demonstrated that Google Trends has great potential for analysing how consumers adopt new technologies and products; the work of Choi *et al.* (2012) demonstrates the ability of Google Trends in the near-term forecasting of the retail sales of a wide range of products; and Carrière-Swallow *et al.* (2013) further examined the nowcasting value of Google Trends based on the case of automobile sales in an emerging market.

However, the study of Google Trends in relation to product’s market performance is still in its early stages. Although Google Trends has shown its potential for the indication of sales trends, its validity needs to be further assessed and compared (Barreira *et al.*, 2013; Choi *et al.*, 2012; Jun *et al.*, 2017). For instance, we still have insufficient knowledge in regard to how to best use such data to explain and forecast market performance at the individual product level. Furthermore, the existing literature in this area mostly studies Google Trends (and other user/internet generated data) by means of techniques such as regressions (Rivera, 2016; Schneider *et al.*, 2016), moving average terms (Li *et al.*, 2017), mixed data sampling (Smith, 2016), and machine learning methods (Santillana *et al.*, 2015). As stated by Schaer *et al.* (2018), "*we note that the majority of the papers investigated reported positive findings for all types of user-generated data sources. The models applied most frequently are linear, in the form of an ARX model, for both nowcasting and forecasting*". In particular, the literature has made insufficient effort to compare the performance of Google Trends with new product diffusion models or to study the potential of incorporating Google Trends into those diffusion models. This is also a motivation behind this study.

3. Methodology

3.1. The model

Let us assume that a product is newly introduced into the market. If $P(t)$ indexes customer likelihood of purchase in the time period t, and $m(t)$ indicates the size of the market potential—i.e., the number of customers who are interested in buying the product in time period t—then, the sales of the new product—i.e., $S(t)$—can be explained by the product of $P(t)$ and $m(t)$.

$$S(t) = P(t)m(t) \qquad (1)$$

In the new model, $P(t)$ takes the form of Equation (2), which is derived from the original Bass model (Bass, 1969). Here, the setting of $P(t)$ can be interpreted as follows: customer likelihood of purchase is determined by customer product awareness, which is simultaneously driven by the innovation and imitation effects (advertising and word-of-mouth effects)—see the work of Norton *et al.* (1987) for a similar explanation.

$$P(t) = \frac{1 - exp(-(p+q)t)}{1 + (q/p)exp(-(p+q)t)} \qquad (2)$$

The key difference between our suggested model and prior Bass-type diffusion ones lies in the interpretation of $m(t)$. In the conventional diffusion models, the market potential is usually considered to be a constant or to be explained by a few specific factors for simplicity. However, the development of Internet technologies has enabled us to collect and record massive amounts of data on user behaviours, including those that reflect customer active interest in specific products and topics (Jun *et al.*, 2014; Preis *et al.*, 2013b), and Google Trends is one public source of such data. Hence, in this study, we attempt to let the market potential be calibrated by Google Trends.

- -
Place Table 1 about here
- -

More specifically, we consider that the overall level of customer search interest in a product can be a proportional reflection of the number of customers who are interested in it—i.e., the size of the market potential. For products involving competing generations, brands, and/or categories, we assume that these factors will only affect the changing market potential of the products themselves—i.e., $m(t)$—and not customer likelihood of purchase—i.e., $P(t)$. Then, we use Equation (3) to incorporate the modified Google Trends data into the new model. In the equation, m_{GT} indexes the size of the market potential under one unit of Google Trends, and $GT(t - t_{lag})$ represents the modified Google Trends data. Here t_{lag} indicates a possible time lag between Google Trends and the market performance of a product, as customers could take a certain amount of time to actually purchase after searching in relation to the product. When $t_{lag} = 0$, it means that the dynamics of the search behaviours will be immediately reflected in the dynamics of the market potential.

$$m(t) = m_{GT} \times GT\left(t - t_{lag}\right) \tag{3}$$

By substituting Equation (2) and Equation (3) into Equation (1), we have the new model for new product diffusion. It is worth noting that, as the Google Trends data represent the relative number of searches of a term in relation to the overall number of searches—rather than the absolute number of searches—the data can be further calibrated by multiplying the total Google searches in the corresponding time periods, in order to produce more accurate results.

Keyword search can indicate either consumer interest in a product (Du *et al.*, 2012) or the level of marketing activities (Hu *et al.*, 2014). As discussed in the previous literature review section, the Bass model is extended mainly through two approaches: to modify the likelihood of purchase (i.e., the innovation and imitation effects) and to modify the market potential. In this study, we have chosen the latter despite the fact that the former is accepted more often in the literature—e.g., see reviews of such models by Peres *et al.* (2010), Meade *et al.* (2006), and Mahajan *et al.* (2000). Note that we consider the aggregated online search behaviour (i.e., Google Trends) to be the result of the summed diffusion influences, however, we do not see the result of Google Trends as a driver for customer purchase intentions (unlike customer reviews on Amazon, eBay, iTunes, and Google Play). In other words, although, by virtue of its definition and working mechanism, Google Trends is a natural indicator for the people's aggregated interests in a product over time, right now there is less evidence that people consult Google Trends before making purchase decisions. Note that we used Google Trends data to calibrate the innovation and the imitation effects in the Bass model in our empirical analysis. The reported fit performance is less accurate (see Table 4), which confirms the validity of our approach in this study.

3.2. The benchmarking models

The new model incorporates diffusion models and Google Trends. The original Bass model and the Norton-Bass model are then introduced as representatives of the conventional diffusion models. A simple model is also introduced to demonstrate the performance of using Google Trends data alone for the product's market performance. We brief these models as follows.

Unlike the new model, which works for both durable and non-durable products, the original Bass model is more suitable for durable ones. In the current study, note that we consider the iPhone and the iPad to be nondurable products, as they have relatively short life-cycles and generate new purchases through generational upgrade. Therefore, we refine and interpret the Bass model in the following format, in which $\frac{1-exp(-(p+q)t)}{1+(q/p)exp(-(p+q)t)}$ has the same interpretation as in the new model and m_{Bass} can be considered as the upper limit of the market potential that the sales can reach.

$$S(t) = \frac{1-exp(-(p+q)t)}{1+(q/p)exp(-(p+q)t)} m_{Bass}, \tag{4}$$

The Norton-Bass model (Norton *et al.*, 1987) works for both durable and non-durable products when $S(t)$ takes appropriate interpretations (Jiang *et al.*, 2012). To illustrate this model, the units of two product generations in use during the time period t ($S_1(t)$ and $S_2(t)$) can be explained by Equations (5) and (6),

$$S_1(t) = F_1(t)m_1\big(1 - F_2(t-\tau_2)\big) \tag{5}$$

$$S_2(t) = F_2(t-\tau_2)\Big(m_2 + F_1(t)m_1\big(1 - F_2(t-\tau_2)\big)\Big) \tag{6}$$

where τ_2 is the release time of the second generation and m_1 and m_2 represent the market potential for the two generations respectively. In both equations, $F_l(t)$ is the diffusion rate of a generation l product at time t, which takes the form shown below based on the Bass model.

In Equation (7), p and q are the coefficients for innovation and imitation, which is consistent with both the Bass model and the new model.

$$F_l(t) = \frac{1-exp\left(\left(-(p+\mathrm{q})(t-\tau_l)\right)\right)}{1+(p/\mathrm{p})exp\left(\left(-(p+\mathrm{q})(t-\tau_l)\right)\right)}, \text{ when } t > \tau_l \tag{7}$$

$$F_l(t) = 0, \text{ when } t \leq \tau_l \tag{8}$$

With the below model, we also examine the fit performance of using Google Trends data alone , where $GT(t)$ represents the Google Trends data and m_{GT} indexes the size of the market potential under one unit of Google Trends.

$$S(t) = m_{GT} \times GT(t) \tag{9}$$

3.3. Data

To examine the performance of the selected models, we employ the two cases of the iPhone and the iPad, which are both innovative Apple Inc. products. The two cases are particularly interesting as the market growth of both is subject to various influences and changes throughout their life-cycles: both products have evolved through several generations of technological development, which have repeatedly boosted their market growth; Apple adopted different advertising and pricing strategies at different stages of the market growth; Apple released the products in different countries at different times through different channels; Apple faces constant threat from competitors, such as Samsung; Apple's business model somehow differs from those of other firms in the industry; the market performance of the products is also influenced by the seasonal effect and by any praise and criticism expressed by society at large.

As the iPhone and the iPad are both well-known high-tech products, we can assume that most of their potential customers are also heavy internet and search engine users. In other words, the potential customers of the two products are more likely to Google them before making the purchase decision. Therefore, these two cases are ideal for the purpose of this study.

\- -
Place Table 2 about here
\- -

\- -
Place Figure 1 about here
\- -

We extracted the sales data of the two products from Apple’s quarterly reports, and we downloaded the monthly Google Trends data for the two terms ‘iPhone’ and ‘iPad’ directly from the Google Trends website. Bear in mind that the Google Trends data used in the new model can be calibrated by multiplying the total Google searches in the corresponding time periods. For this purpose, we obtain the yearly number of total searches made on Google. We assume that the number of Google searches increased steadily during the studied time period, which enables us to estimate the monthly number of Google searches through linear regression. Details of the data can be seen in Table 2 and Figure 1.

3.4. Parameter estimation technique

As some of the selected models require the estimation of parameters, the genetic algorithm (Venkatesan *et al.*, 2004) is introduced. Similar to other estimation techniques employed in diffusion studies—such as nonlinear least squares estimation and maximum likelihood estimation—the genetic algorithm is also included in software packages for convenient use. To estimate the model parameters, we run the genetic algorithm by minimising the below function, where $S(t)$ represents the observed data and $E(S(t))$ the data estimated by the model.

$$\sum_{t=0}^{T}\left(E(S(t))-S(t)\right)^2 \qquad (10)$$

In order to produce more accurate estimations, we used the genetic algorithm package found in MatLab with most of its default settings, with the exception of the increased population size of the estimation (set to 500) and the algorithm stopping rule (terminate if the improvement in the objective function is less than 10^{-12} for 100 consecutive generations). It should be noted that the genetic algorithm may have a tendency to produce local optima. Hence, we ran the case estimation 100 times in order to reduce the local optima rate and provide a validity check. The reported values in this study are those that produce the best fit from the 100 estimation repetitions, and the standard deviations of the 100 repeat estimates.

Note that the new model also considers the time lag between Google Trends data and the market performance of a product (i.e., between customer search behaviours and purchase behaviours). As, based on the empirical data, the time lag was uncertain, different scenarios were considered and compared. Unlike other expensive products or investments such as real estate, the studied products were consumer electronics with a relatively short life-cycle, and therefore, the average time lag between search and purchase was expected to be relatively short. We report the results of the new model with two settings: the first considered a one-time-period delay (i.e., $t_{lag} = 1$) of one-month; and the second considered no time lag (i.e., $t_{lag} = 0$). In addition, it should be noted that the employed Google Trends data are monthly, while the sales data of the products are quarterly. Hence during the parameter estimation, we first estimated the monthly sales based on the new model and the Google Trends data, then we converted the estimated data to a quarterly basis to match the observed data in the estimation function.

4. Results

4.1. Model fit performance

The graphical representation of the models' performance is plotted in Figure 2. The estimated parameters and the statistical model fit are shown in Tables 3 and 4. The reported parameters are rather stable, providing evidence for the face validity of the models. We employ sum of squared error (SSE), mean absolute error (MAE), and R squared to report the model fit, as they are widely used in the literature to assess the performance of diffusion models (Decker *et al.*, 2010; Kim *et al.*, 2000; Norton *et al.*, 1987). We would like to emphasise a few interesting issues stemming from the comparisons.

- -

Place Table 3 about here

- -

- -

Place Table 4 about here

- -

- -

Place Figure 2 about here

- -

First, as explained in the methodology section, we used two settings for t_{lag} ($t_{lag} = 1$ and $t_{lag} = 0$), which resulted in two groups of results for the new model (i.e., New Model I and New Model II). In both settings, the new model is capable of capturing the market dynamics of the products. In terms of the performance comparison between the two settings, the model with time lag tends to have better performance for the sales of the iPhone (e.g., the R^2 values of the two settings are reported as 0.9267 and 0.8256 respectively), while the model without time lag works better in explaining the sales of the iPad (e.g., the R^2 values of the two settings are 0.7763 and 0.8252). Perhaps this is because the iPad faces less competition in the tablet market (Whitney, 2014) than the iPhone does in the smartphone market, and therefore tablet customers may take less time to commit to the purchase after searching for information about the product, while smartphone customers may spend more time shopping around.

Second, both the graphical and statistical results indicate that the new model offers significantly improved model fit compared with the benchmarks in relation to both iPhone and iPad sales. Take the example of the iPhone: the reported SSE of New Model I is 2.20×10^8, compared with 6.60×10^8 for the Bass model, 4.30×10^8 for the Norton-Bass model, 4.51×10^9 for the explanatory ability of the Google Trends data, 6.42×10^8 and 7.09×10^8 for the Bass model with modified p and q by using Google Trends data; the reported MAE of the new model is 1.52×10^3, showing a visible improvement over the Bass model, the Norton-Bass model, the explanatory ability of the Google Trends data, and the Bass model with modified p and q by using Google Trends data. The adjusted R^2 of the new model is 0.9267, which is also a significant improvement compared with the other three models. In addition, the parameter estimates of the new model show that the innovation effect (p) plays an important role in driving customer likelihood of purchase (1.18×10^{-2} and 1.19×10^{-2} in the case of the iPhone and 9.75×10^{-2} and 8.57×10^{-2} in the case of the iPad). In terms of the imitation effect, it plays a moderate role in the case of the iPhone (i.e., 4.23×10^{-2} and 3.84×10^{-2}) but has limited impact in the case of the iPad (i.e., 3.07×10^{-3} and 2.03×10^{-8}).

Third, although the Bass model is capable of illustrating the overall trend of the product growth in the studied cases, for the case of the iPhone in particular, it is unable to capture the fluctuations as the new model does. Also, the Bass model cannot explain why the market performance of the iPad has been generally declining since 2014, resulting in a poor fit of performance. It should be noted that the superior performance of the new model compared with the Bass model is not based on an increased number of model parameters, as the two models both use three parameters for estimation.

Fourth, the Norton-Bass model plots the market growth of the two products through a stepped line—i.e., after a new generation is released the sales quickly reach a peak and stabilise thereafter. By closely reviewing the sales figures, we see that the growth curves of the two products usually experience several declines—rather than a stabilisation—during the life-cycle of each generation. Without the further introduction of specific factors—e.g., the forward-looking effect found in Shi *et al.* (2014)—the Norton-Bass model is only capable of providing explanations of the process in such a manner. It should also be noted that, in the original Norton-Bass model and its later applications, although the sales of each generation can decline due to the release of a newer one, the overall sales of the product category do not decline. Therefore, the overall market potential constantly increases and user dis-adoption is not considered. In our estimation, the growth in market potential of different generations is allowed to be negative, which is of particular relevance for the case of the iPad, as its market performance has been declining since 2014 ($m_6 = -1.35 \times 10^3$ and $m_8 = -1.29 \times 10^3$). In addition, the Norton-Bass model requires 12 and 11 estimated parameters for the cases of the iPhone and the iPad respectively, due to the many generations. The large number of model parameters poses difficulties in model estimation, especially when the available data are limited.

Last but not least, the direct use of Google Trends for product sales does not produce very accurate results. In particular, it fails to explain why the iPhone still experiences high market growth in the later stages of the studied period, when the Google Trends data stop increasing. This issue can be explained by the new model: although the market potential (Google Trends) does not grow, the sales may still increase as more of the potential customers commit to purchase due to their increased understanding of the product—i.e., $P(t)$ in the new model. In addition, our results indicate a potential time lag between the Google Trends data and the sales of the iPhone, which also cannot be captured by the direct use of the Google Trends data.

4.2. Nowcasting performance

The ability of forecasting is highly valued in terms of operations and market planning (Marmier *et al.*, 2010; Mascle *et al.*, 2014). Therefore, a key use of new product diffusion models is to estimate the model parameters based on available knowledge and/or data, and then to predict the future market growth (Bass, 2004). Similarly, the forecasting ability of Google Trends has started to receive increasing attention in the recent literature (Carrière-Swallow *et al.*, 2013; Perlin *et al.*, 2017).

It should be noted that conducting forecasting with Google Trends requires the corresponding future Google Trends data, which are not usually available or reliable. The same issue applies to the new model, which also requires Google Trends data as input. Nevertheless, Google Trends can offer great value in terms of nowcasting (Barreira *et al.*, 2013; Carrière-Swallow *et al.*, 2013; Preis *et al.*, 2014; Smith, 2016) because Google constantly updates the Google Trends data at short time intervals. In particular, the delay in the release of sales data for a product presents a limitation for decision-makers by restricting their ability to accurately assess current conditions. This issue makes nowcasting, or the prediction of the present, an important practice.

To examine the nowcasting performance of the selected approach, we followed Decker *et al.* (2010): we first divided the data set into calibration and nowcasting periods, and then we used the calibration period data to estimate the model parameters (if any) so as to predict the sales in the nowcasting periods. The parameter estimation technique used here is consistent with the one employed previously in the model fit analysis. Based on the model fit results, we used New Model I for iPhone and New Model II for iPad to achieve better nowcasting performance.

Apart from the new model and the Bass model, we also introduce the Time Series Modeler of IBM SPSS 24 to perform the nowcasting task. This procedure can automatically determine the best-fitting ARIMA or exponential smoothing model, and it can determine whether to consider the Google Trends data (as an independent variable) based on the significance of its statistical relationship with the dependent series (the sales data). The Norton-Bass model was excluded for being incapable of performing the nowcasting task as requested by our empirical setting. More specifically, as in the case of the iPad, some of the calibration periods do not cover any data drawn from the last generation, the Norton-Bass model cannot estimate the market potential of the latter (i.e., m_l) in order to perform the nowcasting analysis. We also introduce the naïve method and drift method as benchmarks, as simple methods often can beat complex models in terms of forecasting.

We conducted five sets of experiments—using the suggested model, the Bass model, the SPSS Time Series Modeler, the naïve method and the drift method—over different forecasting periods for one, two, and three data points ahead. The comparative results will show how the effective use of cumulative knowledge and of the available data about the situation improves the nowcasting performance (Armstrong *et al.*, 2015). Akin to the selection of the measures for model fit, we introduce the measures of SSE, MAE, and mean absolute percentage error (MAPE), which are widely used in the field, to demonstrate and compare the models' nowcasting performance.

\- -

Place Table 5 about here

\- -

Table 5 reports the results, showing that the new model performs better in estimating the market dynamics of the iPad in general: the MAPE results of the new model for the three nowcasting periods are 12.14%, 13.61%, and 11.92%, respectively, which are better than those provided by the Bass model (27.72%, 48.60%, and 37.93%) by the SPSS Time Series Modeler (14.98%, 8.97%, and 18.73), by naïve method (21.90%, 30.56%, and 17.30%), and by drfit method (20.07%, 35.98%, and 16.38%). The SPSS Time Series Modeler produces more accurate results in the case of the iPhone for all three nowcasting periods. The Bass model exhibits low performance for all three nowcasting periods in both cases.

5. Conclusions

Today's new products are mostly exposed to a large number of dynamic influences that can be either expected or unexpected by the market. The identification of the key diffusion influences and the corresponding modelling of the diffusion phenomena to inform market planning and operations are difficult. Although the incorporation of more influences into the diffusion model can help to increase its performance, it will correspondingly increase its complexity, and thus increase the difficulties linked to its application. Therefore, we believe that modelling new product diffusion requires a parsimonious, accurate, and generalised approach.

In previous studies, diffusion modellers had primarily focussed on developing ever more sophisticated models, which we see as becoming increasingly difficult and impractical in the complex and dynamic market. Further, increased model complexity makes model applications difficult. For instance, in order to incorporate one additional diffusion influence (i.e., the generational effect), the Bass-Norton model has to introduce seven more parameters to the iPhone case of this study and eight more parameters to the iPad one compared with the original Bass model and the new model. On the other hand, big data has become a hot topic. Our literature review shows that, although a few scholars have pioneered the match between some online data and various social and business phenomena, including new product performance, their studies are not linked to the rich diffusion literature.

Therefore, this study examines the use of diffusion models, of the Google Trends data, and of a combination of the two to explain and nowcast the market performance of new products. We summarise the key results of this study and their implications in Table 6. To sum up, the empirical results show that the combination of diffusion models and Google Trends can more accurately explain those fluctuations in the growth curve that are not captured by the conventional diffusion models or by the Google Trends data alone. Therefore, our study will serve as a useful reference in this field. We hope that, by illustrating the potential of using one behavioural data set to develop a better understanding of the market performance of a new product, this study may inspire future researchers to utilise data of a similar type in diffusion research.

- -
Place Table 6 about here
- -

In developing the new model, we propose and validate the use of customer search interests as a direct indicator of the dynamic market potential of a product. We believe that the new model has great value for understanding and explaining the diffusion phenomenon especially in today's business environment, in which many diffusion influences coexist and interact and traditional models are not suitable as they explain the phenomenon based on a few pre-assumed influences. The new model provides an applicable way to account for the complex diffusion influences, and it decreases complexity without losing accuracy. In the cases of the iPhone and the iPad, market growth can be influenced by a combination of numerous factors that cannot be identified and measured individually so as to understand the overall trend. By incorporating Google Trends data, the new model can provide a more accurate understanding and estimation of the phenomena because all these influences—which affect customer search behaviours—are reflected in the search results. However, it should be noted that our aim was not to identify the most accurate model for nowcasting sales but, rather, to evaluate the extent to which the popularity of Google search queries can complement the diffusion models. Future research can develop alternative approaches of integrating Google Trends data into diffusion models, and compare their performance.

In today's complex and dynamic market, business practice constantly demands better tools to predict the sales of new products. In particular, the issue of nowcasting is of greater importance in today's business management, where decision-makers need to observe trends before the actual data are released. Our comparison shows that, although the nowcasting performance of the new model and of the Google Trends data alone (using the SPSS Time Series Modeler) differs in different cases, both models produce more accurate results than the Bass diffusion one. The implementation of the two approaches is easy to follow, and the required Google Trends data are publicly accessible and updated in real time. Moreover, the Google Trends data and the new model can be useful for those firms that need to be constantly alerted to the sales of competitors in order to calibrate their own strategies. This is because the Google Trends data can be subcategorised into different regions of interest, and a domestic supplier can thus use the approaches to analyse the competitive sales of a multinational supplier within its region.

References

Albuquerque P, Bronnenberg BJ, Corbett CJ. 2007. A Spatiotemporal Analysis of the Global Diffusion of ISO 9000 and ISO 14000 Certification. *Management Science* **53**(3): 451-468.

Armstrong JS, Green KC, Graefe A. 2015. Golden rule of forecasting: Be conservative. *Journal of Business Research*: in Press.

Bangwayo-Skeete PF, Skeete RW. 2015. Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tourism Management* **46**: 454-464.

Barreira N, Godinho P, Melo P. 2013. Nowcasting unemployment rate and new car sales in south-western Europe with Google Trends. *Netnomics* **14**(3): 129-165.

Bass FM. 1969. A New Product Growth for Model Consumer Durables. *Management Science* **15**(5): 215-227.

Bass FM. 2004. Comments on "A New Product Growth for Model Consumer Durables": The Bass Model. *Management Science* **50**(12): 1833-1840.

Bass FM, Krishnan TV, Jain DC. 1994. Why the Bass Model Fits without Decision Variables. *Marketing Science* **13**(3): 203-223.

Carrière-Swallow Y, Labbé F. 2013. Nowcasting with Google Trends in an Emerging Market. *Journal of Forecasting* **32**(4): 289-298.

Chatterjee R, Eliashberg J, Rao VR. 2000. Dynamic Models Incorporating Competition. In *New Product Diffusion Models*. Mahajan V, Muller E, Wind Y (eds.), Kluwer Academic Publishers: New York.

Chien C-F, Chen Y-J, Peng J-T. 2010. Manufacturing intelligence for semiconductor demand forecast based on technology diffusion and product life cycle. *International Journal of Production Economics* **128**(2): 496-509.

Choi H, Varian HAL. 2012. Predicting the Present with Google Trends. *Economic Record* **88**: 2-9.

Danaher PJ, Hardie BGS, Putsis WP, Jr. 2001. Marketing-Mix Variables and the Diffusion of Successive Generations of a Technological Innovation. *Journal of Marketing Research* **38**(4): 501-514.

Decker R, Gnibba-Yukawa K. 2010. Sales Forecasting in High-Technology Markets: A Utility-Based Approach. *Journal of Product Innovation Management* **27**(1): 115-129.

Du RY, Kamakura WA. 2012. Quantitative Trendspotting. *Journal of Marketing Research* **49**(4): 514-536.

Geroski PA. 2000. Models of technology diffusion. *Research Policy* **29**(4-5): 603-625.

Givon M, Mahajan V, Muller E. 1995. Software Piracy: Estimation of Lost Sales and the Impact on Software Diffusion. *The Journal of Marketing* **59**(1): 29-37.

Guo X. 2014. A novel Bass-type model for product life cycle quantification using aggregate market data. *International Journal of Production Economics* **158**: 208-216.

Hu Y, Du RY, Damangir S. 2014. Decomposing the Impact of Advertising: Augmenting Sales with Online Search Data. *Journal of Marketing Research* **51**(3): 300-319.

Huang X, Zhang L, Ding Y. 2017. The Baidu Index: Uses in predicting tourism flows –A case study of the Forbidden City. *Tourism Management* **58**: 301-306.

Jiang Z, Jain DC. 2012. A Generalized Norton–Bass Model for Multigeneration Diffusion. *Management Science* **58**(10): 1887-1897.

Jun S-P, Park D-H. 2016. Consumer information search behavior and purchasing decisions: Empirical evidence from Korea. *Technological Forecasting and Social Change* **107**: 97-111.

Jun S-P, Park D-H, Yeom J. 2014. The possibility of using search traffic information to explore consumer product attitudes and forecast consumer preference. *Technological Forecasting and Social Change* **86**: 237-253.

Jun S-P, Sung T-E, Park H-W. 2017. Forecasting by analogy using the web search traffic. *Technological Forecasting and Social Change* **115**: 37-51.

Kiesling E, Günther M, Stummer C, Wakolbinger L. 2012. Agent-based simulation of innovation diffusion: a review. *Central European Journal of Operations Research* **20**(2): 183-230.

Kim N, Chang DR, Shocker AD. 2000. Modeling Intercategory and Generational Dynamics for a Growing Information Technology Industry. *Management Science* **46**(4): 496-512.

Kim S, Shin DH. 2016. Forecasting short-term air passenger demand using big data from search engine queries. *Automation in Construction* **70**: 98-108.

Kiron D. 2013. Organizational Alignment is Key to Big Data Success. In *MIT Slow Management Review.*

Kumar V, Krishnan TV. 2002. Multinational Diffusion Models: An Alternative Framework. *Marketing Science* **21**(3): 318-330.

Lamba K, Singh SP. 2017. Big data in operations and supply chain management: current trends and future perspectives. *Production Planning & Control* **28**(11-12): 877-890.

Li D-C, Chen W-C, Chang C-J, Chen C-C, Wen IH. 2015. Practical information diffusion techniques to accelerate new product pilot runs. *International Journal of Production Research* **53**(17): 5310-5319.

Li H, Amini M. 2012. A hybrid optimisation approach to configure a supply chain for new product diffusion: a case study of multiple-sourcing strategy. *International Journal of Production Research* **50**(11): 3152-3171.

Li X, Pan B, Law R, Huang X. 2017. Forecasting tourism demand with composite search index. *Tourism Management* **59**: 57-66.

Libai B, Muller E, Peres R. 2009a. The Diffusion of Services. *Journal of Marketing Research* **46**(2): 163-175.

Libai B, Muller E, Peres R. 2009b. The Role of Within-Brand and Cross-Brand Communications in Competitive Growth. *Journal of Marketing* **73**(3): 19-34.

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.

Mahajan V, Muller E, Wind Y. 2000. *New-Product Diffusion Models*. Kluwer Academic Press: Boston & Dordrecht.

Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, Byers AH. 2011. Big data: The next frontier for innovation, competition, and productivity, McKinsey Global Institute.

Marmier F, Cheikhrouhou N. 2010. Structuring and integrating human knowledge in demand forecasting: a judgemental adjustment approach. *Production Planning & Control* **21**(4): 399-412.

Mascle C, Gosse J. 2014. Inventory management maximization based on sales forecast: case study. *Production Planning & Control* **25**(12): 1039-1057.

Mavragani A, Tsagarakis KP. 2016. YES or NO: Predicting the 2015 GReferendum results using Google Trends. *Technological Forecasting and Social Change* **109**: 1-5.

Meade N, Islam T. 2006. Modelling and forecasting the diffusion of innovation - A 25-year review. *International Journal of Forecasting* **22**(3): 519-545.

Moat HS, Curme C, Avakian A, Kenett DY, Stanley HE, Preis T. 2013. Quantifying Wikipedia Usage Patterns Before Stock Market Moves. *Sci. Rep.* **3**.

Negahban A, Yilmaz L, Nall T. 2014. Managing production level in new product diffusion: an agent-based simulation approach. *International Journal of Production Research* **52**(17): 4950-4966.

Norton JA, Bass FM. 1987. A Diffusion Theory Model of Adoption and Substitution for Successive Generations of High-Technology Products. *Management Science* **33**(9): 1069-1086.

Nunan D. 2015. Addressing the market research skills gap. *International Journal of Market Research* **57**(2): 177-178.

Nunan D, Di Domenico M. 2013. Market research and the ethics of big data. *International Journal of Market Research* **55**(4): 2-13.

Panagiotopoulos P, Bigdeli AZ, Sams S. 2014. Citizen–government collaboration on social media: The case of Twitter in the 2011 riots in England. *Government Information Quarterly* **31**(3): 349-357.

Pappas S. 2014. Data Fail! How Google Flu Trends Fell Way Short.

Parker P, Gatignon H. 1994. Specifying competitive effects in diffusion models: An empirical analysis. *International Journal of Research in Marketing* **11**(1): 17-39.

Peres R, Muller E, Mahajan V. 2010. Innovation diffusion and new product growth models: A critical review and research directions. *International Journal of Research in Marketing* **27**(2): 91-106.

Perlin MS, Caldeira JF, Santos AAP, Pontuschka M. 2017. Can We Predict the Financial Markets Based on Google's Search Queries? *Journal of Forecasting* **36**(4): 454-467.

Preis T, Moat HS. 2014. *Adaptive nowcasting of influenza outbreaks using Google searches*.

Preis T, Moat HS, Bishop SR, Treleaven P, Stanley HE. 2013a. Quantifying the Digital Traces of Hurricane Sandy on Flickr. *Sci. Rep.* **3**.

Preis T, Moat HS, Stanley HE. 2013b. Quantifying Trading Behavior in Financial Markets Using Google Trends. *Sci. Rep.* **3**.

Putsis WP, Balasubramanian S, Kaplan EH, Sen SK. 1997. Mixing Behavior in Cross-Country Diffusion. *Marketing Science* **16**(4): 354-369.

Qin R, Nembhard DA. 2012. Demand modeling of stochastic product diffusion over the life cycle. *International Journal of Production Economics* **137**(2): 201-210.

Radas S, Shugan SM. 1998. Seasonal Marketing and Timing New Product Introductions. *Journal of Marketing Research* **35**(3): 296-315.

Rivera R. 2016. A dynamic linear model to forecast hotel registrations in Puerto Rico using Google Trends data. *Tourism Management* **57**: 12-20.

Samuel Sale R, Mesak HI, Inman RA. 2017. A dynamic marketing-operations interface model of new product updates. *European Journal of Operational Research* **257**(1): 233-242.

Santillana M, Nguyen AT, Dredze M, Paul MJ, Nsoesie EO, Brownstein JS. 2015. Combining Search, Social Media, and Traditional Data Sources to Improve Influenza Surveillance. *PLOS Computational Biology* **11**(10): e1004513.

Schaer O, Kourentzes N, Fildes R. 2018. Demand forecasting with user-generated online information. *International Journal of Forecasting*.

Schneider MJ, Gupta S. 2016. Forecasting sales of new and existing products using consumer reviews: A random projections approach. *International Journal of Forecasting* **32**(2): 243-256.

Shi X, Fernandes K, Chumnumpan P. 2014. Diffusion of multi-generational high-technology products. *Technovation* **34**(3): 162-176.

Smith P. 2016. Google's MIDAS Touch: Predicting UK Unemployment with Internet Search Data. *Journal of Forecasting* **35**(3): 263-284.

Stummer C, Kiesling E, Günther M, Vetschera R. 2015. Innovation diffusion of repeat purchase products in a competitive market: An agent-based simulation approach. *European Journal of Operational Research* **245**(1): 157-167.

Tsai B-H. 2013. Modeling diffusion of multi-generational LCD TVs while considering generation-specific price effects and consumer behaviors. *Technovation* **33**(10–11): 345-354.

Van den Bulte C, Joshi YV. 2007. New Product Diffusion with Influentials and Imitators. *Marketing Science* **26**(3): 400-421.

Venkatesan R, Krishnan TV, 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.

Vosen S, Schmidt T. 2011. Forecasting private consumption: survey-based indicators vs. Google trends. *Journal of Forecasting* **30**(6): 565-578.

Whitney L. 2014. iPad sheds more market share but still dominates. In *CNET*.

Yan H-S, Ma K-P. 2011. Competitive diffusion process of repurchased products in knowledgeable manufacturing. *European Journal of Operational Research* **208**(3): 243-252.

Appendix

Below, to demonstrate the performance of the selected models, we present two more cases drawn from the video game industry. The first dataset is related to the weekly sales of Nintendo Switch from its launch to March 2018 (54 data points). The second dataset, which includes 20 data points, pertains to the weekly sales of a video game (i.e., Call for Duty: Black Ops II) from its launch to March 2013.

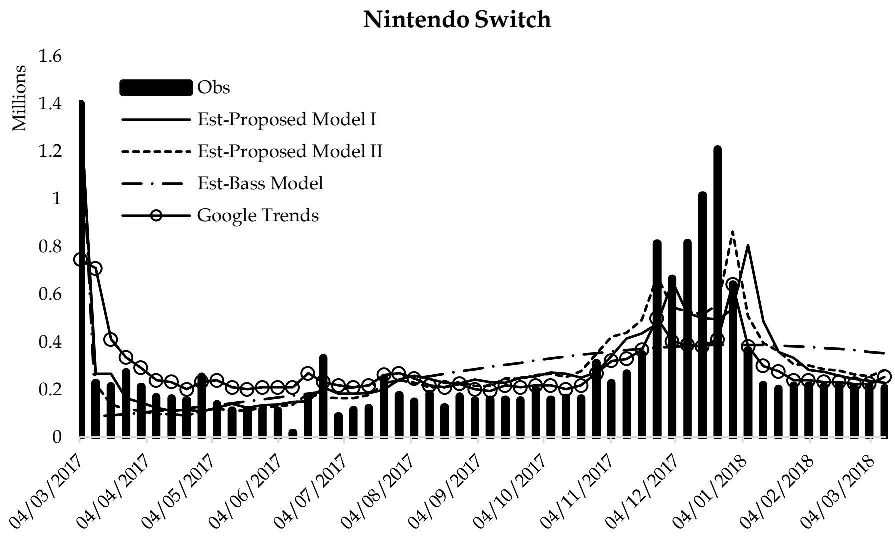
As discussed in the paper, the iPhone and the iPad have relatively short life-cycles, and their market growth is continuously boosted by the introduction of new generations. Conversely, as Nintendo Switch and the studied video game are durable products with rather long life-cycles, most customers will be one time buyers. Therefore, we take the density function of the Bass model, instead of Equation (2), to explain customer likelihood of purchase in the time period t:

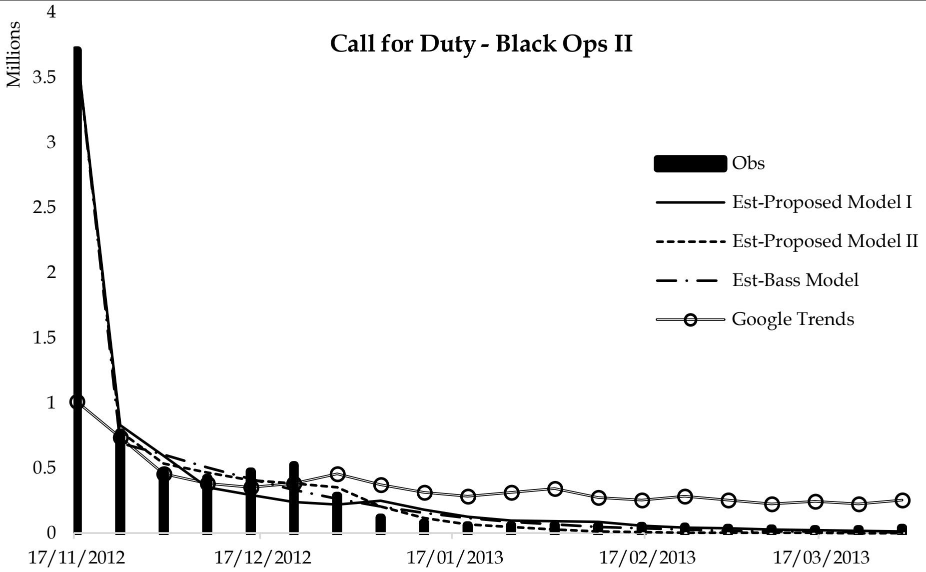
$$P(t) = \frac{((p+q)^2/p)exp(-(p+q)t)}{(1+(q/p)exp(-(p+q)t))^2} \tag{11}$$

The same parameter estimation technique is applied to the model and the benchmarks.

The results (see figure and table below) show that the incorporation of the Bass Model and the Google Trends data can better explain the sales performance of both products both graphically and statistically.

Model Fitting Performance





	Nintendo Switch			Call For Duty II		
	SSE	*MAE*	*R Squared*	*SSE*	*MAE*	*R Squared*
New Model I	1.54×10^{12}	1.08×10^{5}	0.6267	1.77×10^{11}	6.19×10^{4}	0.9861
New Model II	1.25×10^{12}	1.04×10^{5}	0.6981	5.83×10^{10}	4.39×10^{4}	0.9958
Bass Model	2.28×10^{12}	1.46×10^{5}	0.4457	9.19×10^{10}	4.96×10^{4}	0.9928
Google Trends	1.89×10^{12}	1.09×10^{5}	0.5454	7.16×10^{11}	1.71×10^{5}	0.9748

Table 1: Summary of the Notations Used in This Study

Notation	Interpretation
$S(t)$	Sales of the new product in the time period t
$P(t)$	Customer likelihood of purchase in the time period t
$m(t)$	Number of customers who are interested in purchasing the product in the time period t
p	Coefficient for the innovation effect
q	Coefficient for the imitation effect
m_{GT}	The market potential in relation to one unit of Google Trends (new model)
m_{Bass}	The market potential in relation to one unit of Google Trends (Bass model)
m_l	The market potential of general l of the product (Norton-Bass model);
$GT(t)$	Google trends data in the time period t;
t_{lag}	The time delay between Google Trends and the market performance of the product;

Table 2: Summary of the Data Used in the Empirical Analysis

Product	Generation	Time Period in the Market	Sales Data Used	Google Trends Data Used
iPhone	*iPhone*	29 Jun.2007 – 11 Jul. 2008	Quarterly based sales data for the iPhone from Apple's quarterly reports; 126 data points in total, from 3rd Quarter of 2007 to 2nd Quarter of 2017.	Monthly based data from Jan. 2007 to Jun. 2017; 126 data points in total.
	iPhone 3G	11 Jul. 2008 – 7 Jun. 2010		
	iPhone 3GS	19 Jun. 2009 – 12 Sep. 2012		
	iPhone 4	24 Jun. 2010 – 10 Sep. 2013		
	iPhone 4S	14 Oct. 2011 – 9 Sep. 2014		
	iPhone 5	21 Sep. 2012 – 10 Sep. 2013		
	iPhone 5S & 5C	20 Sep. 2013 – 21 Mar. 2016		
	iPhone 6 & Plus	19 Sep. 2014 – 7 Sep. 2016		
	iPhone 6S & Plus	25 Sep. 2015 – Now		
	iPhone 7 & Plus	16 Sept. 2016 - Now		
iPad	*iPad 1*	3 Apr. 2010 – 2 Mar. 2011	Quarterly based sales data for the iPad from Apple's quarterly reports; 29 data points in total, from 2nd Quarter of 2010 to 2nd Quarter of 2017.	Monthly based data from Jan 2010 to Jun. 2017; 90 data points in total.
	iPad 2	11 Mar. 2011 – 18 Mar. 2014		
	iPad 3	16 Mar. 2012 – 23 Oct. 2013		
	iPad 4 & Mini	2 Nov. 2012 – 16 Oct. 2014		
	iPad Air	1 Nov. 2013 – 21 Mar. 2016		
	iPad Air 2 & Mini 2	22 Oct. 2014 – 21 Mar. 2017		
	iPad Pro	11 Nov. 2015 – 5 Jun. 2017		
	iPad Pro (9.7inch)	31 Mar. 2016 – 5 Jun. 2017		
	iPad (2017)	24 Mar. 2017 - Now		

Number of Searches on Google

Yearly based; available data: 1998, 2000, and from 2007 to 2013.

Table 3: Parameter Estimates

iPhone				**iPad**			
New Model I [a]		*New Model II* [b]		*New Model I*		*New Model II*	
p	1.18×10^{-2} (6.74×10^{-6})	p	1.19×10^{-2} (6.95×10^{-6})	p	9.75×10^{-2} (2.88×10^{-3})	p	8.57×10^{-2} (9.90×10^{-7})
q	4.23×10^{-2} (2.02×10^{-5})	q	3.84×10^{-2} (2.17×10^{-5})	q	3.07×10^{-3} (6.42×10^{-4})	q	2.03×10^{-8} (1.37×10^{-6})
m_{GT}	1.05×10^{4} (8.80×10^{-2})	m_{GT}	1.05×10^{4} (1.42×10^{-1})	m_{GT}	4.61×10^{3} (2.72×10^{1})	m_{GT}	4.49×10^{3} (6.89×10^{-3})
Bass Model		*Google Trends*		*Bass Model*		*Google Trends*	
p	1.42×10^{-3} (2.05×10^{-4})	m_{GT}	1.01×10^{4} (4.22×10^{-5})	p	2.07×10^{-2} (2.32×10^{-6})	m_{GT}	3.99×10^{3} (3.48×10^{-5})
q	5.98×10^{-2} (3.11×10^{-3})			q	5.84×10^{-2} (1.49×10^{-5})		
m_{Bass}	1.88×10^{4} (2.47×10^{2})			m_{Bass}	7.37×10^{3} (2.07×10^{-3})		
Norton-Bass Model				*Norton-Bass Model*			
p	1.00×10^{0} (1.41×10^{-8})	q	1.00×10^{0} (1.30×10^{-8})	p	8.26×10^{-1} (2.45×10^{-2})	q	9.10×10^{-1} (2.33×10^{-2})
m_1	5.41×10^{2} (5.41×10^{-5})	m_2	1.10×10^{3} (9.15×10^{-5})	m_1	1.54×10^{3} (4.63×10^{0})	m_2	2.41×10^{3} (1.33×10^{1})
m_3	9.53×10^{2} (8.32×10^{-5})	m_4	3.19×10^{3} (6.49×10^{-5})	m_3	1.56×10^{3} (1.71×10^{1})	m_4	2.88×10^{2} (1.28×10^{1})
m_5	4.67×10^{3} (5.78×10^{-5})	m_6	2.65×10^{-3} (6.65×10^{-5})	m_5	7.70×10^{-2} (5.38×10^{-2})	m_6	-1.35×10^{3} (3.17×10^{1})
m_7	1.66×10^{3} (1.87×10^{-5})	m_8	2.82×10^{3} (9.11×10^{2})	m_7	1.87×10^{2} (9.01×10^{1})	m_8	-1.29×10^{3} (8.16×10^{1})
m_9	2.97×10^{3} (8.65×10^{2})	m_{10}	1.35×10^{3} (8.84×10^{2})	m_9	3.91×10^{2} (3.02×10^{1})		

The values in brackets are the standard deviation of the 100 repeats;
The unit for market potentials is 10^3*;*
[a]: $t_{lag} = 1$;
[b]: $t_{lag} = 0$.

Table 4: Model Fit Performance

	iPhone			iPad		
	SSE	*MAE*	*R Squared*	*SSE*	*MAE*	*R Squared*
New Model I	2.20×10^8	1.52×10^3	0.9267	1.92×10^8	2.32×10^3	0.7763
New Model II	7.24×10^8	2.38×10^3	0.8256	1.37×10^8	1.81×10^3	0.8252
Bass Model	6.60×10^8	2.23×10^3	0.8328	4.75×10^8	3.11×10^3	0.3967
Norton-Bass Model	4.30×10^8	1.84×10^3	0.8518	3.15×10^8	2.58×10^3	0.5978
Google Trends [c]	4.51×10^9	7.75×10^3	0.4950	6.11×10^8	3.76×10^3	0.3162
Bass Model – p [d]	6.42×10^8	2.27×10^3	0.8336	4.74×10^8	3.10×10^3	0.3975
Bass Model – q [e]	7.09×10^8	2.47×10^3	0.8326	4.67×10^8	3.05×10^3	0.4050

The unit for SSE results is 10^6*;*
The unit for MAE results is 10^3*;*
[c]*: The explanation ability of the Google Trends data alone;*
[d e]*: We also used the Google Trends data to calibrate parameters* p *and* q*, and reported the performance.*

Table 5: Nowcasting Performance

Model	Nowcasting Period	iPhone				iPad			
		SSE	*MAE*	*MAPE*	*Ranking*	*SSE*	*MAE*	*MAPE*	*Ranking*
*New Model**	1 Quarter	1.70×10^8	1.30×10^4	31.78%	4	1.92×10^6	1.39×10^3	12.14%	1
	2 Quarters	2.30×10^8	1.04×10^4	23.48%	2	3.67×10^6	1.35×10^3	13.61%	2
	3 Quarters	2.32×10^8	8.38×10^3	15.79%	2	4.93×10^6	1.27×10^3	11.92%	1
Bass Model	1 Quarter	2.86×10^8	1.69×10^4	41.24%	5	1.00×10^7	3.17×10^3	27.72%	5
	2 Quarters	4.20×10^8	1.36×10^4	31.10%	3	4.76×10^7	4.72×10^3	48.60%	5
	3 Quarters	7.80×10^8	1.38×10^4	23.71%	5	5.32×10^7	3.85×10^3	37.93%	5
SPSS Time Series Modeler	1 Quarter	3.66×10^6	1.91×10^3	4.66%	1	2.93×10^6	1.71×10^3	14.98%	2
	2 Quarters	1.36×10^8	8.14×10^3	17.61%	1	1.94×10^6	9.37×10^2	8.97%	1
	3 Quarters	2.58×10^8	6.68×10^3	10.11%	1	1.50×10^7	2.12×10^3	18.73%	3
Naïve Method	1 Quarter	9.48×10^7	9.74×10^3	23.73%	2	6.26×10^6	2.50×10^3	21.90%	4
	2 Quarters	2.15×10^9	3.24×10^4	72.53%	4	2.00×10^7	2.91×10^3	30.56%	3
	3 Quarters	1.12×10^9	1.42×10^4	21.05%	3	1.93×10^7	2.11×10^3	17.30%	4
Drift Method	1 Quarter	1.22×10^8	1.10×10^4	26.90%	3	5.26×10^6	2.29×10^3	20.07%	3
	2 Quarters	2.59×10^9	3.55×10^4	79.64%	5	2.64×10^7	3.47×10^3	35.98%	4
	3 Quarters	1.07×10^9	1.42×10^4	21.90%	4	1.55×10^7	1.95×10^3	16.38%	3

The unit for SSE results is 10^6*;*
The unit for MAE results is 10^3*;*
*: *Based on the results of previous analysis, New Model I was used for the iPhone and New Model II was used for the iPad.*

Table 6: Summary of Results & Implications

	Considered Diffusion Influences	Required Data for Estimation	Model Fit Performance	Nowcasting Performance	Key Implications
New Model	• Innovation effect • Imitations effect • Various other influences through people' searching behaviour	• Historical sales data • Google Trends data	• Best fit performance in both cases • Able to capture the fluctuations in the curves	• 2nd best nowcasting performance in the iPhone case • Best in the iPad case	• The model is parsimonious and generalised • It takes various diffusion influences into consideration • Although the model's estimation requires additional Google Trends data, the data is easy to obtain • The estimated parameters can provide additional managerial and operational implications • It provides superior fit and nowcasting performance
Bass Model	• Innovation effect • Imitations effect	• Historical sales data	• 3rd best in both cases • Unable to capture the fluctuations in the curves	• Lowest nowcasting performance in both cases	• The model is parsimonious and generalised • Only two diffusion influences are considered in the model • The estimated parameters can provide additional managerial implications • Although it can explain/predict the general growth trend of new products, it cannot capture the fluctuations in the their growth curves
Norton-Bass Model	• Innovation effect • Imitations effect • Generational effect	• Historical sales data	• 2nd best in both cases • Able to capture some of the fluctuations in the curves	• Not reported	• It considers one diffusion influence more than the Bass model; i.e., the generational effect • It requires a large number of parameters to be estimated • The increased model complexity results in a better fit performance than the Bass model; however, it fails to perform nowcasting analysis in the cases of this study
Google Trends	• No diffusion influences are considered	• Historical sales data • Google Trends data	• Lowest fit performance in both cases • Poor match to the fluctuations in the curves	• Best in the iPhone case • 2nd best in the iPad case	• This approach does not reflect the market context of new product growth • The use of Google Trends data alone produces the lowest fit performance compared with the other models • Combining with SPSS Time Series Modeler, it shows superior nowcasting performance

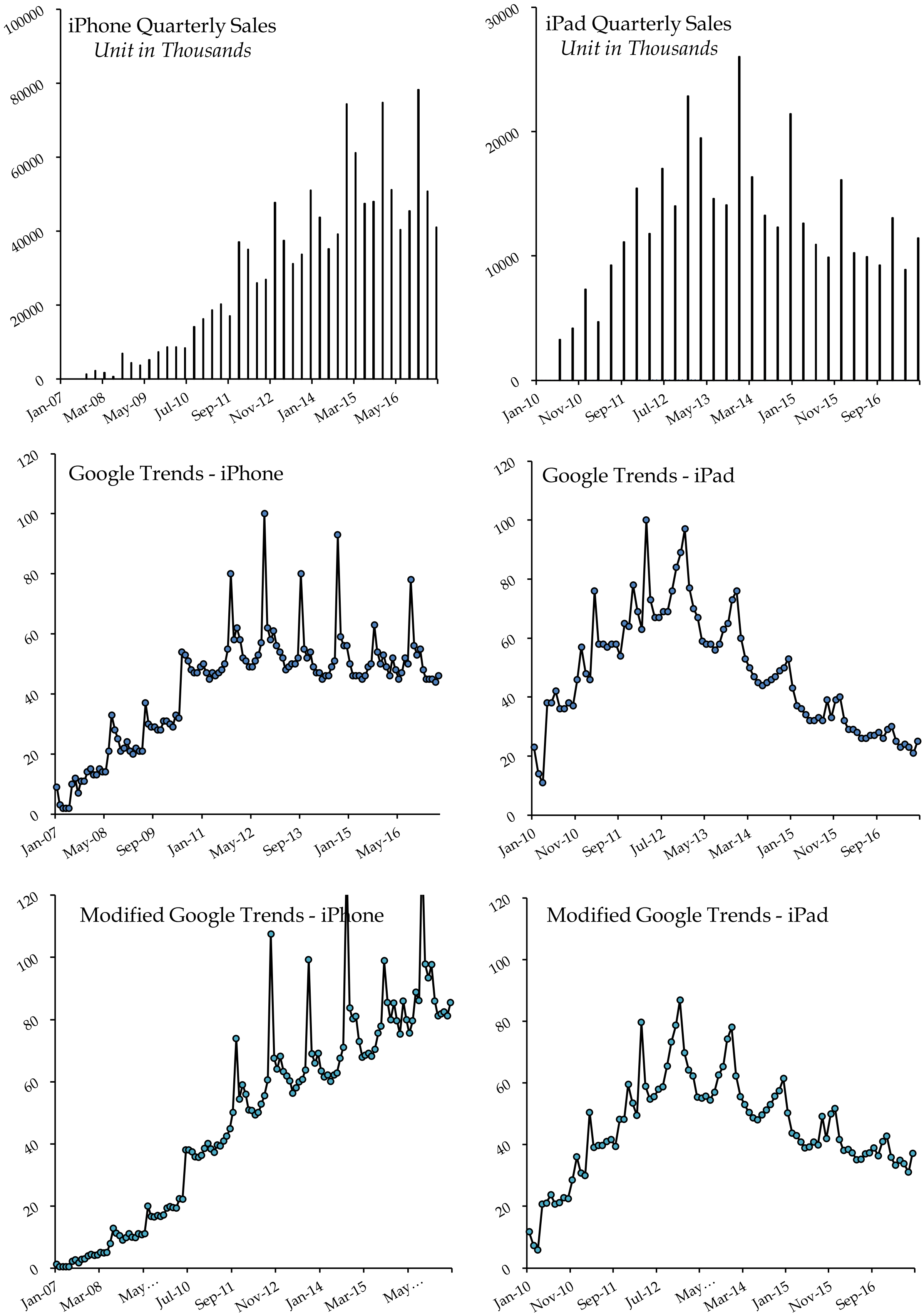


Figure 1: Sales Data and Google Trends Data: iPhone & iPad

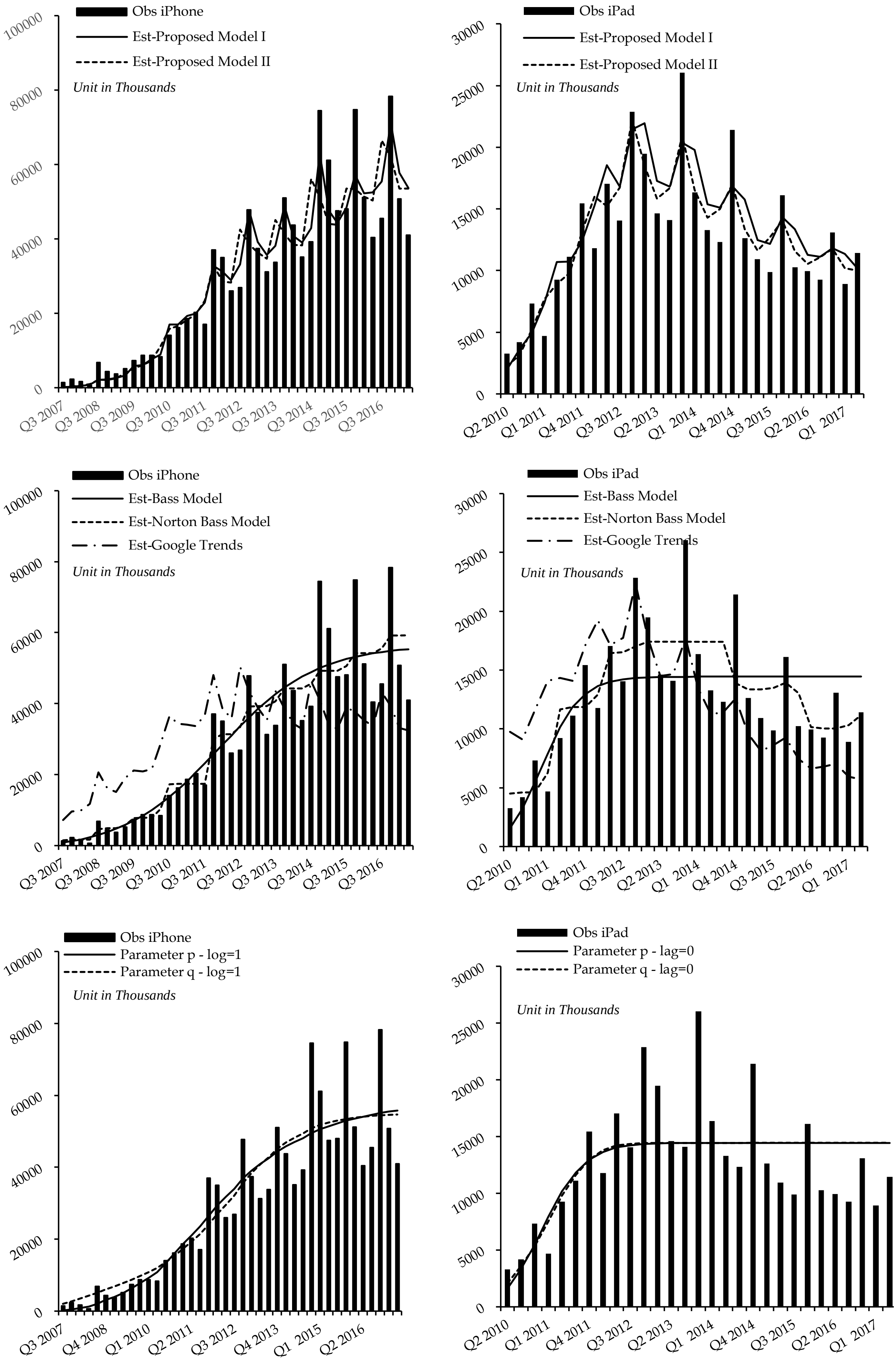


Figure 2: Results of Model Fit (Unit in Thousands)