MARCELO SALHAB BROGLIATO

ESSAYS IN
COMPUTATIONAL MANAGEMENT SCIENCE

ESSAYS IN
COMPUTATIONAL MANAGEMENT SCIENCE

MARCELO SALHAB BROGLIATO

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Part I

DIFFUSION AND DISMISSAL OF
INNOVATION: FORECASTING THE NUMBER
OF FACEBOOK'S ACTIVE USERS

INTRODUCTION

Why do entrepreneurs appear, not continuously, that is singly in every appropriately chosen interval, but in clusters? Exclusively because the appearance of one or a few entrepreneurs facilitates the appearance of others, and these the appearance of more, in ever-increasing numbers.

— *Joseph Schumpeter*

The way innovations diffuse through markets is an important and useful topic in marketing, which was made popular through Rogers' work [27] and has influenced many marketing researchers. Within innovation diffusion, Bass [2] proposed a model which forecasts how many people will have adopted new products or technologies by a given point in time. INFORMS members have voted this model as one of the Top 10 Most Influential Papers published in the 50-year history of Management Science in connection with the 50th anniversary of the journal [3].

The Bass [2] model was designed to forecast only innovation adoption, which is the first time one consumes the innovation. In the model, either the consumer has or has not consumed the innovation by a given point in time. Thus, recurrent customers are considered only once, because they have already consumed the innovation before. The model has only three parameters, which are estimated through the number of adoptors of the innovation. Though very simple, it is considered very robust.

Although it is clear that the marketing investment, the product prices, the economy itself, and many other variables affect the diffusion process, the Bass [2] model does not contemplate these variables and yet it is still able to describe the empirical adoption curve of a large number of new products and technological innovations. In order to explain the robustness of the model, Bass et al. [4] developed a general model including these variables, and they showed that this general model reduces to the Bass model as a special case. They also showed that the shape of the diffusion of innovation process is always the same, an S-curve. Norton and Bass [23] analyzed the diffusion of innovation for product substitution, explaining some unexpected changes in innovation adoption which were still unclear.

The motivation behind the present work is that there may be people who reject a particular innovation. Such people do not

recommend the innovation, on the contrary, they may publicly complain about it and bad-mouth it. This negative word-of-mouth effect of rejection has always existed [26, 6, 29, 8] but it is becoming more and more important as information can spread more easily and faster among people through the usage of new technologies [16, 15, 1]. Nowadays, before making a decision, people may search the internet for reviews and feedbacks about the innovation, and what they find affects their decision [9, 10, 13, 11, 18, 25]. A number of firms have already perceived this change caused in large part by the social media and have adapted to this new condition, like Starbucks [14], for example.

There is an extensive literature on innovation diffusion. Numerous extensions to the Bass [2] model have been proposed (for reviews, see Meade and Islam [20] and Peres et al. [24]), among which Mahajan and colleagues were the first to propose a model that includes the negative word-of-mouth [19]. The latter and other extensions that include negative word-of-mouth are much more complex than the Bass [2] model, and their parameters must be estimated using the number of people who have already adopted the innovation at a given time. A problem is that most recent innovations, like Facebook, Twitter, and Netflix do not disclose their total number of users. Actually, they deem this number confidential. They only disclose the number of active users, not including the number of users who have rejected them.

In this work, I propose an extension to the Bass model which (i) includes the negative effect of the rejections, (ii) is as simple as the Bass [2] model, and (iii) its parameters can be estimated using the number of people who have adopted and have not posteriorly rejected the innovation, i.e., the number of *active adopters*. An important difference between the number of *active adopters* and the *total adopters* is that first may decrease over time, while the latter cannot. First, I discuss the Bass [2] model and its parameters; then I describe the model extension and analyze four models of rejection; next, I detail the estimation method; after that, I estimate the models by using Facebook's active users dataset; and finally, I discuss the results and conclude.

The main contribution of this work is that the proposed extension is the first to include the effect of the rejections that can be estimated using the number of *active adopters* instead of the number of *total adopters*. The model can be applied to forecast the number of active adopters of these companies in the next quarters, which becomes an important tool for analysts, investors, and the companies themselves. As the number of active adopters is related to the market cap of these companies, these forecasts may be useful to estimate the future value

of the firms. At the time of this writing, Facebook's market cap is $222.69B, Twitter's is $23.18B, and Netflix's is $20.49B[1].

---

1 Data obtained from Yahoo! Finance website on December 1st, 2016.

THE BASS MODEL

The Bass [2] model is a simplification of the diffusion of innovation process proposed by Rogers [27]. The Rogers' classification of adopters has five classes: (i) innovators; (ii) early adopters; (iii) early majority; (iv) late majority; and (v) laggards. Bass simplified them to only two classes: innovators and imitators. The innovators are the ones who start using an innovation regardless of who else and how many other people are already using it. The imitators are the ones who concern themselves about who is using the innovation and, as long as many other people are using it, they are more inclined to adopt it. Thus, at the beginning of the diffusion of innovation, the majority of adopters are innovators. As more and more people adopt the innovation, the majority of new adopters shift to imitators.

Mathematically, the model presents itself with the following set of equations:

$$S(t) = mf(t) \tag{1}$$

$$Y(t) = mF(t) \tag{2}$$

$$\frac{f(t)}{1 - F(t)} = p + qF(t), F(0) = 0 \tag{3}$$

Both $S(t)$ and $Y(t)$ are related to the absolute amount of adopters; while $S(t)$ is the number of new adopters at time t, $Y(t)$ is the number of people who had already adopted the innovation by time t, thus $S(t) = \frac{d}{dt}Y(t)$. The model also has three positive parameters: (i) the potential market size m; (ii) the innovators parameter p; and (iii) the imitators parameter q. There are also $f(t)$ and $F(t)$ which are related to the percentage of the potential market: $f(t)$ is the percentage of the potential market which is adopting the innovation at time t, while $F(t)$ is the percentage of the potential market which had already adopted it at time t, thus $f(t) = \frac{d}{dt}F(t)$.

Therefore, $S(t)$ and $f(t)$ are related to the adoption rate at time t, while $Y(t)$ and $F(t)$ are related to the accumulated adopters at time t. In the beginning of the diffusion of innovation, there are no adopters at all, thus $Y(0) = F(0) = 0$. As the number of adopters can only increase over time, both $Y(t)$ and $F(t)$ are monotonically increasing functions, and both $S(t)$ and $f(t)$ are always greater or equal to zero.

The non-linear ordinary differential equation 3 is the main equation in the Bass model. Its left side is known as the hazard function and it expresses the probability of someone adopting the innovation, provided that he/she has not chosen to adopt it yet, i.e.,

the rate of adoption. Its right side means that this probability is at least p and increases linearly with the percentage of people who have already adopted the innovation, i.e., $F(t)$.

Solving the differential equation, Bass found a closed formula for the diffusion, which is $F(t) = \frac{1-e^{-(p+q)t}}{1+\frac{q}{p}e^{-(p+q)t}}$. This closed formula always has the famous shape of the S-curve, regardless of the values of $p > 0$ and $q > 0$. From the closed formula of $F(t)$, it is trivial to obtain the equations of $S(t)$, $Y(t)$, and $f(t)$.

The potential market size m is the unknown number of people who will have adopted the innovation after a very long time. It is not exactly the target market of the innovation, but a subset of it, as no product diffuses over its entire target market. If a company estimates and updates the model more than once for their product, the change in m is a change in the potential market and could help the company to understand whether their decisions in the meantime have increased or decreased the number of future adopters. As $\lim_{t\to\infty} Y(t) = m$, the whole potential market will have adopted the innovation at some point.

The innovator parameter p is related to the proportion of people in the potential market who adopt the innovation regardless the others. In other words, their decision to adopt is not influenced by the social system, but by other external factors. The bigger the p, the larger the number of innovators, thus the faster the diffusion at the beginning.

The imitator parameter q is related to the influence of those actually using the innovation on those who are not using it yet. This is why this parameter multiplies $F(t)$, which is the proportion of the market which had already adopted it. This influence is mainly understood as a result of the word-of-mouth recommendation. In other words, the more people use the innovation, the more other people will adopt it. The bigger the q, the larger the imitator effect, thus, the faster the diffusion.

Practitioners have been using this model to forecast future demand. First, they measure the number of adopters over time. Then, they estimate the parameters m, p, and q. Finally, they extrapolate $S(t)$ out of the measured time window and use its value as the forecast demand. They also calculate $m - Y(t)$ as a forecast of how many people have not adopted the innovation yet.

THE EXTENDED MODEL

I propose adding a new term in the differential equation 3 in order to include the effect of rejection, as in equation 4. Hereafter I will refer to: (i) the people who have adopted the innovation as *total users*; (ii) the people who have adopted the innovation and remain using it as *active users*; and (iii) the people who rejected the innovation as *inactive users*. Clearly, the function $Y(t)$ is the number of *total users*, which is equal to the sum of the number of *active users* with the number of *inactive users*.

$$\frac{f(t)}{1-F(t)} = p + qF(t) - wR(t), F(0) = 0, R(0) = 0 \qquad (4)$$

The function $R(t)$ is the percentage of accumulated *inactive users* at time t. Thus, $A(t) = F(t) - R(t)$ is the percentage of accumulated *active users* at time t. Multiplying by m, $mA(t) = Y(t) - mR(t)$ is the number of *active users* at time t. Since $A(t) \geqslant 0$, $F(t) \geqslant R(t)$, which makes sense because it is not possible to have more *inactive users* than *total users*.

The negative word-of-mouth parameter w is related to how much the *inactive users* really affect the new adopters decision in the diffusion process. The bigger the w, the greater the negative influence of these *inactive users* on the new adopters. Another possible understanding is that w is related to how much the *inactive users* are bad-mouthing the innovation. The bigger the w, the more they bad-mouth the innovation.

Equation 4 could be rewritten as $\frac{f(t)}{1-F(t)} = p + qA(t) + (q-w)R(t)$. This form is useful in order to understand the impact of *active users* and *inactive users* on the rate of adoption at time t.

If $w = q$, then the rate of adoption increases linearly with the *active users*, since $\frac{f(t)}{1-F(t)} = p + qA(t)$. In other words, the positive word-of-mouth has exactly the same influence as the negative word-of-mouth on the new adopters. It also means that the rate of adoption is always greater than zero, thus the whole potential market will have adopted the innovation sooner or later.

If $w < q$, then the new adopters are more influenced by the number of *total users* than by the number of *inactive users*. It is just as if the *inactive users* do not bad-mouth the innovation so much, or at all. Again, the rate of adoption is always greater than zero, thus the whole potential market will have adopted the innovation sooner or later.

If $w > q$, then the proposed extension really differs from the Bass model. In this case, the influence of the *inactive users* is greater than

the influence of *total users*. So, if the rate of adoption were equal to zero ($w R(t) = p + q F(t)$), the innovation might not be adopted by the whole potential market, i.e., $\lim_{t \to \infty} F(t) < 1$.

The main contribution of this work is to have the choice to use the number of *active users*, which is the information that most of the companies disclose, in order to estimate the parameters of the proposed extension model. With that, both the number of *inactive users* and *total users* could be forecast. It is important to notice that the number of *active users* ($m A(t)$) could decrease over time. In fact, forecasting when this is going to happen may be crucial for corporations.

MODELS FOR R(t)

The proposed extended model already includes the effect of rejection through the $R(t)$ function and the w parameter. In order to complete the model, $R(t)$ has to be well defined. Let $r(t) = \frac{d}{dt}R(t)$ be the rate of new *inactive users* at time t. The proposed differential equations for $R(t)$ are the following:

$$\text{Model 1: } r(t) = \nu f(t) \tag{5}$$

$$\text{Model 2: } \frac{r(t)}{1 - R(t)} = \nu f(t) \tag{6}$$

$$\text{Model 3: } r(t) = \nu[F(t) - R(t)] \tag{7}$$

$$\text{Model 4: } \frac{r(t)}{1 - R(t)} = \nu[F(t) - R(t)] \tag{8}$$

These four models can be grouped into two families, one for the equations 5 and 6, and another for the equations 7 and 8. The former relates the rejection to the rate of new people adopting the innovation, as if people decide whether they will use or reject the innovation when they try it. Then, they do not change their position anymore. The latter assumption relates the rejection to the number of *active users*, as if the *active users* first adopt the innovation and then they continuously reject it.

The rejection parameter ν has a different interpretation in each family. In equations 5 and 6, it is the proportion of new adopters who will reject the innovation. In equations 7 and 8, it is the proportion of active users who are continuously rejecting the innovation.

Therefore, for all these models of rejection, the complete diffusion of innovation model has five parameters to be estimated, namely m, p, q, w, and ν.

4.1 MODEL 1

In this model, the rate of new *inactive users* at time t is proportional to the percentage of people adopting the innovation at time t, as if people decide whether they will use or reject the innovation when they are adopting it.

The differential equation 5 can be easily solved integrating both sides. Thus, $R(t) = \nu F(t)$, $A(t) = (1 - \nu)F(t)$, and $f(t)/[1 - F(t)] = p + (q - w\nu)F(t)$. As the imitator coefficient must be positive, we must have $w\nu < q$.

The solution shows that this model has exactly the same explanatory power as the Bass model, neither better nor worse. This happens because the model's solution has exactly the same equation after the linear transformation $q^* = q - wv$.

As $\lim_{t\to\infty} F(t) = 1$, thus $\lim_{t\to\infty} R(t) = v$. Hence, the proportion of *inactive users* is exactly equal to the rejection parameter.

Solving the differential equation for $F(t)$, it gets $F(t) = (1 - e^{-(p+q-wv)t})/(1 + \frac{(q-wv)}{p}e^{-(p+q-wv)t})$. Finally, as $R(t) = vF(t)$, thus $R(t) = v(1 - e^{-(p+q-wv)t})/(1 + \frac{(q-wv)}{p}e^{-(p+q-wv)t})$.

Unfortunately, it is not possible to estimate this model. The problem is that $\forall v \in \mathbb{R}^+$, $\exists \hat{v} \in \mathbb{R}^+$ such as the set of parameters (m, p, q, w, v) and $(m, p, q, wv/\hat{v}, \hat{v})$ have exactly the same residuals when estimated. That is, the model can be estimated for any value arbitrarily set for v. Intuitively, as both $F(t)$, $A(t)$, and $R(t)$ have the same shape, the parameters can be estimated with an empirical *active users* dataset and then you can slide up or down $F(t)$ just changing the values of v and w.

This result is interesting because it shows that the Bass model can already explain the diffusion of innovations which follows this model of rejection. Hence, it just confirms the robustness of the Bass model.

4.2 MODEL 2

The right side of the differential equation 6 is the rate of rejection, i.e., the probability of someone rejecting the innovation, provided that he/she has not rejected it yet. Thus, in this model, the rate of rejection is proportional to the percentage of people adopting the innovation at time t, i.e., the more people adopt the innovation, the more they reject it. But if no one is adopting, there would be no rejection also, which would hold the number of *active users* the same.

The differential equation 6 can also be solved for $R(t)$ algebraically. Using the fact that $-\frac{d}{dt}\log[1 - R(t)] = r(t)/[1 - R(t)] = vf(t)$, and integrating both sides of this equation yields:

$$-\frac{d}{dt}\int_0^t \log[1 - R(\tau)]d\tau = v\int_0^t f(\tau)d\tau \tag{9}$$

$$-\log[1 - R(t)] = vF(t) \tag{10}$$

$$1 - R(t) = e^{-vF(t)} \tag{11}$$

$$R(t) = 1 - e^{-vF(t)} \tag{12}$$

Finally, $f(t)/[1 - F(t)] = p + qF(t) - w[1 - e^{-vF(t)}]$ and $\lim_{t\to\infty} R(t) = 1 - e^{-v}$.

As $0 \leqslant F(t) \leqslant 1 \Rightarrow 0 \leqslant vF(t) \leqslant v$, we can do a good approximation of $e^{-vF(t)}$ using a Taylor series around the point $v/2$ for small values of v.

From the Taylor series centered around $v/2$, we know that $e^{-x} \approx e^{-v/2}(1 + v/2 - x)$. Thus, we have $1 - e^{-vF(t)} \approx 1 - e^{-v/2}(1 + v/2) + e^{-v/2}vF(t)$, and, finally, $f(t)/[1 - F(t)] = [p - w - we^{-v/2}(1 + v/2)] + (q - we^{-v/2}v)F(t)$.

Therefore, for small values of v, this model has approximately the same explanation power as the Bass model and we can write $f(t) = p^* + q^*F(t)$, where $p^* = p - w - we^{-v/2}(1 + v/2)$ and $q^* = q - we^{-v/2}v$.

In contrast to model 1, no parameter could be arbitrarily set in this model, thus it can be estimated using an empirical *active users* dataset.

4.3 MODEL 3

In this model, the rate of new *inactive users* increases linearly with the number of *active users*, since $A(t) = F(t) - R(t)$. Thus, while there are *active users*, a fraction v of them will be rejecting the innovation. Hence, everyone will have rejected the innovation sooner or later.

The differential equation 7 can be rewritten as the following first order linear differential equation, which has to be solved:

$$\frac{d}{dt}R(t) + vR(t) = vF(t) \tag{13}$$

The solution to this differential equation is:

$$R(t) = ve^{-vt}\int_0^t e^{v\tau}F(\tau)d\tau \tag{14}$$

$$= v\left[F(u) * e^{-vu}\right](t) \tag{15}$$

Or, as $\frac{d}{dt}\left(e^{vt}F(t)\right) = ve^{vt}F(t) + e^{vt}f(t)$, $R(t)$ can be rewritten as:

$$R(t) = F(t) - e^{-vt}\int_0^t e^{v\tau}f(\tau)d\tau \tag{16}$$

$$= F(t) - \left[f(u) * e^{-vu}\right](t) \tag{17}$$

As $F(t) = \int_0^t f(\tau)d\tau$ and $e^{v\tau} \geqslant 1$, we have that:

$$\int_0^t e^{v\tau}f(\tau)d\tau \geqslant F(t) \tag{18}$$

$$-e^{-vt}\int_0^t e^{v\tau}f(\tau)d\tau \leqslant -e^{-vt}F(t) \tag{19}$$

$$F(t) - e^{-vt}\int_0^t e^{v\tau}f(\tau)d\tau \leqslant F(t) - e^{-vt}F(t) \tag{20}$$

From equation 16:

$$R(t) \leqslant F(t) - e^{-vt} F(t) \tag{21}$$

$$R(t) \leqslant F(t)(1 - e^{-vt}) \tag{22}$$

Finally, $R(t) < F(t)$, $r(t) > 0$, and $\lim_{t \to \infty} R(t) \leqslant 1$.

I did not manage to prove that $\lim_{t \to \infty} R(t) = 1$, but this result appeared in all performed simulations. If that is true, then all people will reject the innovation at some point in time - a fact that makes sense.

Unfortunately, it seems that there is no closed formula for $F(t)$. Using equation 14, the final differential equation is:

$$\frac{f(t)}{1 - F(t)} = p + qF(t) - wve^{-vt} \int_0^t e^{v\tau} F(\tau) d\tau \tag{23}$$

Or, using 16, it becomes:

$$\frac{f(t)}{1 - F(t)} = p + (q - w)F(t) + we^{-vt} \int_0^t e^{v\tau} f(\tau) d\tau \tag{24}$$

The condition $w \leqslant p + q$ is sufficient to ensure $f(t) \geqslant 0$, since $w \leqslant p + q \Rightarrow wR(t) \leqslant pR(t) + qR(t) \leqslant p + qR(t) \leqslant p + qF(t) \Rightarrow p + qF(t) - wR(t) = f(t)/[1 - F(t)] \geqslant 0 \Rightarrow f(t) \geqslant 0$. Assuming that $\lim_{t \to \infty} R(t) = 1$, then it is easy to prove that this condition is also necessary.

4.4 MODEL 4

In this model, the rate of rejection increases linearly with the number of *active users*. Thus, while there are *active users*, the rate of rejection will be greater than zero. Hence, everyone will have rejected the innovation sooner or later.

Although equation 8 is a Ricatti equation [5], none of the available techniques could solve the differential equation and it seems that there is no closed formula for $F(t)$. Hence, the equation will be analyzed through a linearization around the fixed points.

$$\begin{cases} f(F, R) = (p + qF - wR)(1 - F) \\ r(F, R) = v(F - R)(1 - R) \end{cases} \tag{25}$$

Solving the system $f(F, R) = r(F, R) = 0$, the following solutions are found:

$$u_1^* = (1, 1) \tag{26}$$

$$u_2^* = \left(\frac{p}{w - q}, \frac{p}{w - q} \right) \tag{27}$$

$$u_3^* = \left(\frac{w - p}{q}, 1 \right) \tag{28}$$

The only valid solutions are u_1^* and u_2^*. The solution u_3^* is not valid because $p + q - w > 0 \Rightarrow (w - p)/q < 1 \Rightarrow F < R$ which is not possible because it would imply a negative number of *active users*.

Finally, the linearization around u^* is:

$$\begin{bmatrix} f(R, T) - f(u^*) \\ r(R, T) - r(u^*) \end{bmatrix} = J|_{u^*} \left(\begin{bmatrix} R \\ T \end{bmatrix} - u^* \right) \tag{29}$$

$$J = \begin{bmatrix} q(1 - F) - (p + qF - wR) & -w(1 - F) \\ v(1 - R) & -v[(1 - R) + (F - R)] \end{bmatrix} \tag{30}$$

Now, let's analyze the jacobian matrices and their eigenvalue for each valid solution.

$$J|_{u_1^*} = \begin{bmatrix} -(p + q - w) & 0 \\ 0 & 0 \end{bmatrix} \tag{31}$$

As the eigenvalues of $J|_{u_1^*}$ are 0 and $-(p + q - w) < 0$, the point $u_1^* = (1, 1)$ is a sink, i.e., the neighborhood converges to u_1^* when $t \to \infty$. It may be interpreted that all users will have rejected the innovation after a long time.

ESTIMATION METHOD

The parameters of the model should be estimated using empirical data in order to check the explanation power of the model. Bass [2] used a discrete version of his differential equation with ordinary least squares. While it worked well for him, it has not in the present work. There are several well-known problems in the estimation of parameters, most of them related to approximation of derivatives and instability of the estimators. These problems have been noticed by many authors [28, 30, 32].

In the present work, the parameters of the models were estimated using a maximum likelihood function on the residuals between $A(t) = F(t) - R(t)$ from the model and the empirical values from the dataset. The residuals were assumed to be normally distributed with $\mu = 0$, which leads to the same results as the ordinary least square method. In order to calculate the residuals, $F(t)$ and $R(t)$ were calculated based on their differential equations using the 4th order Runge Kutta (RK4) numerical method [7] with $\Delta t = 0.01$.

The empirical data format was (t_i, x_i), where t_i is the time and x_i is the value, and the measurements were not equally spaced over time. The following Log Likelihood equation was used:

$$\text{residual}(t_i, x_i \mid m, p, q, w, \nu) = x_i - m \cdot [F(t_i \mid p, q, w) - R(t_i \mid \nu)]$$

(32)

$$\text{Log Likelihood}(\vec{t}, \vec{x} \mid m, p, q, w, \nu) = -\sum_{i=1}^{N} [\text{residual}(t_i, x_i \mid m, p, q, w, \nu)]^2$$

(33)

During the evaluation of $F(t)$ and $R(t)$ using the Runge Kutta numerical method, sometimes the exact value of t_i was not reached because of the chosen Δt. In these cases, the value of x_i was calculated using a linear approximation with the nearest points. Let $(\hat{t}_k, \hat{x}_k)$ and $(\hat{t}_{k+1}, \hat{x}_{k+1})$ be values calculated from the Runge Kutta method, such that $\hat{t}_k < t_i < \hat{t}_{k+1}$. Then, the calculated value at t_i was $\hat{x}_k + \left(\frac{\hat{x}_{k+1} - \hat{x}_k}{\hat{t}_{k+1} - \hat{t}_k}\right) \cdot (t_i - \hat{t}_k)$.

Since there is no explicit solution for the parameters $m, p, q, w,$ and ν which maximize the LogLikelihood function, the parameters were estimated using the Truncated Newton Constrained (TNC) method [12, 22, 21] from the SciPy Python Library [17]. The constraints were $m > 0$, $p > 0$, $q > 0$, $w \geqslant 0$, and $\nu \geqslant 0$ for all models. The initial guess for the TNC method was the same for all models estimation.

Sometimes, the method did not converge and another initial guess had to be used.

Even though it has not been done yet, the confidence interval, estimator average and estimator variance for each parameter will be calculated using the bootstrap method.

PRELIMINARY RESULTS

The models' parameters were estimated using the number of Facebook's *active users* from December 2004 to March 2013 [31]. In the dataset, x_i was the number of Facebook's *active users* at time t_i. The dataset had 23 non-equally time spaced measures. The users who have accessed Facebook at least once in each month were counted in the number of *active users* in that month.

The model 1 has not been estimated because it is not possible to estimate it.

The estimated models can be seen at figures 1, 2, and 3. These figures also have forecasts for the number of *active users*, *inactive users*, and *total users* of Facebook for the next 4 years (from $t = 100$ to $t = 140$).

In spite of the favorable goodness of fit using model 2 (see figure 1), the model does not seem to provide a plausible forecast, because it would mean that Facebook is reaching a stable number of *active users* and the rejections are near the end.

Models 3 and 4 have very similar outcomes (see figures 2 and 3). Their Bayesian Information Criterion (BIC) are also close, but model 4 has a better fit with the data. Their forecast makes more sense than the forecast of model 2. It predicts that Facebook is very close to the peak of *active users* and, in approximately 3 years, it is going to decline. It is also interesting to notice that, according to these outcomes, Facebook may not reach its total potential market.

The difference between the outcomes of model 2 and models 3 and 4 could be explained by the fact that they have different rationales behind their models of rejection. While differential equation of the model 2 uses the rate of new *total users*, the differential equations of the models 2 and 3 use the proportion of *active users*.
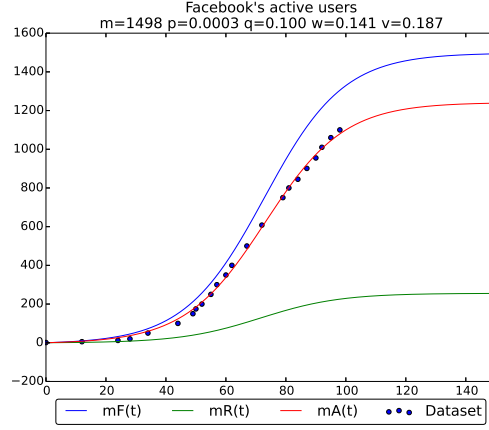
Figure 1: Fit of Model 2 with Facebook's active users dataset. $mF(t)$ is the total users, $mR(t)$ is the inactive users, and $mA(t)$ is the active users. The unit of these functions are thousands of people. The parameters are $m = 1,497.50$, $p = 0.000331$, $q = 0.100088$, $w = 0.140595$, and $v = 0.187188$. The goodness of fit are $R^2 = 99.84\%$ and BIC=10,566.52.

Figure 2: Fit of Model 3 with Facebook's active users dataset. $mF(t)$ is the total users, $mR(t)$ is the inactive users, and $mA(t)$ is the active users. The unit of these functions are thousands of people. The parameters are $m = 1,967.64$, $p = 0.000184$, $q = 0.097867$, $w = 0.330511$, and $v = 0.006912$. The goodness of fit are $R^2 = 99.83\%$ and BIC=11,485.68
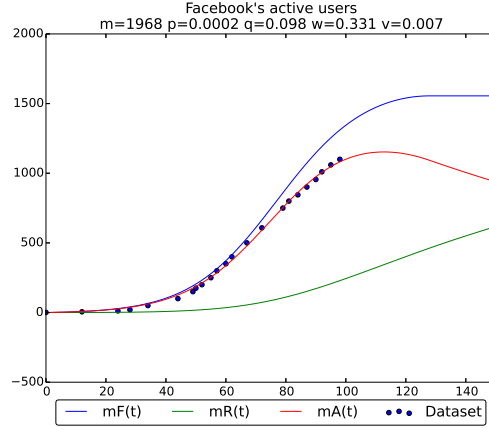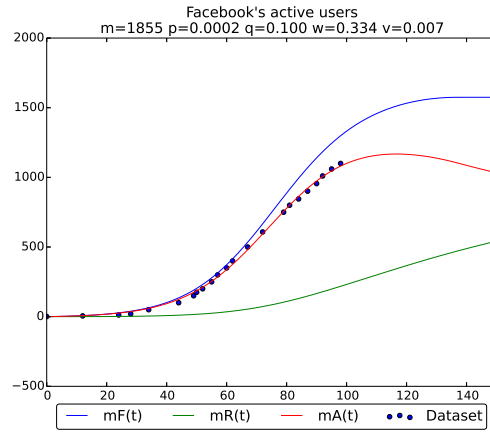
Figure 3: Fit of Model 4 with Facebook's active users dataset. $mF(t)$ is the total users, $mR(t)$ is the inactive users, and $mA(t)$ is the active users. The unit of these functions are thousands of people. The parameters are $m = 1,854.85$, $p = 0.000183$, $q = 0.099738$, $w = 0.334454$, and $v = 0.007007$. The goodness of fit are $R^2 = 99.84\%$ and BIC=10,724.55

CONCLUSION

The main contribution of this work is the parameter estimation through the empirical number of *active users* dataset, forecasting the number of *total users*, the number of *active users*, and the number of *inactive users*.

If the adopters who have rejected the innovation follow the equations of model 1 and 2, then the proposed extended model is transformed into the Bass model through a linear transformation of the parameters. This confirms the Bass model robustness.

Model 2 does not seem to be a good model of rejection, since the number of *active users* never decreases which does not seem to be plausible.

Models 3 and 4 had very similar results when fitting the Facebook dataset. The lack of analytical solutions for them, however, is a barrier to better understand their behavior, and to know whether they will always have similar outcomes or they will diverge depending on the data. Model 3 seems to be more analytically manageable.

It is important to notice that, in models 3 and 4, the innovation may not be adopted by the whole potential market, but it would in the Bass model. Whether it will be adopted by the whole potential market or not depends on the parameters w and v. For instance, it seems that $mF(t)$ is not converging to m at figures 2 and 3. It is a major difference between the proposed extended model and the Bass model.

As this is a working paper, it is also intended to include the analyses of other datasets, like either Twitter's number of *active users*, or WhatsApp's, or Netflix's, or Reddit's, or Dropbox's, or Waze's, etc. This would enhance the proposed model power of forecasting.

It is also intended to run a backtest with the available data. First, the parameters of the model are estimated using a subset of the dataset. Then, through extrapolation, the number of *active users* is forecast. Finally, it is compared to this part of the dataset - which must not have been used in the estimation.

The main limitation of this work is that it has no theory to support which of the models of rejection best fit with empirical data. Although there is an extensive literature on negative word-of-mouth, this literature does not predict which model would be the best. But, if any of these companies discloses the number of *total users* and *active users*, the estimation method could be adapted to estimate the parameters using both pieces of information at the same time, which

would make possible to verify which of the models of rejection provides the best fit.

Future work could explore other models of rejection and also other estimation methods, like nonlinear least squares [30] and Kalman filter [32].

BIBLIOGRAPHY

[1] Ana Babic, Francesca Sotgiu, Kristine de Valck, and Tammo HA Bijmolt. The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. *Journal of Marketing Research*, 2015.

[2] Frank M Bass. A new product growth for model consumer durables. *Marketing science*, 15(5):215–227, 1969.

[3] Frank M Bass. Comments on "a new product growth for model consumer durables the bass model". *Management science*, 50 (12_supplement):1833–1840, 2004.

[4] Frank M Bass, Trichy V Krishnan, and Dipak C Jain. Why the bass model fits without decision variables. *Marketing science*, 13 (3):203–223, 1994.

[5] Sergio Bittanti, Alan J Laub, and Jan C Willems. *The Riccati Equation*. Springer-Verlag New York, Inc., 1991.

[6] Paula Fitzgerald Bone. Word-of-mouth effects on short-term and long-term product judgments. *Journal of business research*, 32(3): 213–223, 1995.

[7] John Charles Butcher. *The numerical analysis of ordinary differential equations: Runge-Kutta and general linear methods*. Wiley-Interscience, 1987.

[8] Francis A Buttle. Word of mouth: understanding and managing referral marketing. *Journal of strategic marketing*, 6(3):241–254, 1998.

[9] Pei-Yu Chen, Shin-yi Wu, and Jungsun Yoon. The impact of online recommendations and consumer feedback on sales. *ICIS 2004 Proceedings*, page 58, 2004.

[10] Chrysanthos Dellarocas. The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management science*, 49(10):1407–1424, 2003.

[11] Chrysanthos Dellarocas, Xiaoquan Michael Zhang, and Neveen F Awad. Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive marketing*, 21(4):23–45, 2007.

[12] Ron S Dembo and Trond Steihaug. Truncated-newtono algorithms for large-scale unconstrained optimization. *Mathematical Programming*, 26(2):190–212, 1983.

[13] Wenjing Duan, Bin Gu, and Andrew B Whinston. Do online reviews matter?—an empirical investigation of panel data. *Decision Support Systems*, 45(4):1007–1016, 2008.

[14] John Gallaugher and Sam Ransbotham. Social media and customer dialog management at starbucks. *MIS Quarterly Executive*, 9(4):197–212, 2010.

[15] David Godes and Dina Mayzlin. Using online conversations to study word-of-mouth communication. *Marketing science*, 23(4): 545–560, 2004.

[16] Bernard J Jansen, Mimi Zhang, Kate Sobel, and Abdur Chowdury. Twitter power: Tweets as electronic word of mouth. *Journal of the American society for information science and technology*, 60(11): 2169–2188, 2009.

[17] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–. URL http://www.scipy.org/. [Online; accessed 2014-12-16].

[18] Mira Lee and Seounmi Youn. Electronic word of mouth (ewom) how ewom platforms influence consumer product judgement. *International Journal of Advertising*, 28(3):473–499, 2009.

[19] Vijay Mahajan, Eitan Muller, and Roger A Kerin. Introduction strategy for new products with positive and negative word-of-mouth. *Management Science*, 30(12):1389–1404, 1984.

[20] Nigel Meade and Towhidul Islam. Modelling and forecasting the diffusion of innovation–a 25-year review. *International Journal of forecasting*, 22(3):519–545, 2006.

[21] Stephen G Nash. Newton-type minimization via the lanczos method. *SIAM Journal on Numerical Analysis*, 21(4):770–788, 1984.

[22] Stephen G Nash. A survey of truncated-newton methods. *Journal of Computational and Applied Mathematics*, 124(1):45–59, 2000.

[23] John A Norton and Frank M Bass. A diffusion theory model of adoption and substitution for successive generations of high-technology products. *Management science*, 33(9):1069–1086, 1987.

[24] Renana Peres, Eitan Muller, and Vijay Mahajan. Innovation diffusion and new product growth models: A critical review and research directions. *International Journal of Research in Marketing*, 27(2):91–106, 2010.

[25] Jürgen Pfeffer, T Zorbach, and KM Carley. Understanding online firestorms: Negative word-of-mouth dynamics in social media networks. *Journal of Marketing Communications*, 20(1-2):117–128, 2014.

[26] Marsha L Richins. Negative word-of-mouth by dissatisfied consumers: A pilot study. *The journal of marketing*, pages 68–78, 1983.

[27] Everett M Rogers. *Diffusion of innovations*. The Free Press, New York, 1st edition, 1962.

[28] David C Schmittlein and Vijay Mahajan. Maximum likelihood estimation for an innovation diffusion model of new product acceptance. *Marketing science*, 1(1):57–78, 1982.

[29] Robert E Smith and Christine A Vogt. The effects of integrating advertising and negative word-of-mouth communications on message processing and response. *Journal of Consumer Psychology*, 4(2):133–151, 1995.

[30] V Srinivasan and Charlotte H Mason. Technical note-nonlinear least squares estimation of new product diffusion models. *Marketing science*, 5(2):169–178, 1986.

[31] The Associated Press. Number of active users at facebook over the years, May 2013. URL http://news.yahoo.com/number-active-users-facebook-over-230449748.html.

[32] Jinhong Xie, X Michael Song, Marvin Sirbu, and Qiong Wang. Kalman filter estimation of new product diffusion models. *Journal of Marketing Research*, pages 378–393, 1997.