An Evaluation of the Extended Logistic, Simple Logistic, and Gompertz Models for Forecasting Short Lifecycle Products and Services

Charles V. Trappey [a,1], Hsin-ying Wu [b]

[a]Professor (Management Science), National Chiao Tung University, Hsinchu, Taiwan.
[b]PhD student, National Chiao Tung University, Hsinchu, Taiwan.

Abstract. Many successful technology forecasting models have been developed but little research has explored the relationship between sample set size and forecast prediction accuracy. This research studies the forecast accuracy of large and small data sets using the simple logisticl, Gompertz, and the extended logistic models. The performance of the models were evaluated using the mean absolute deviation and the root mean square error. A time series dataset of four electronic products and services were used to evaluate the model performance. The result shows that the extended logistic model fits large and small datasets better than the simple logistic and Gompertz models. The findings also show that that the extended logistic model is well suited to predict market growth with limited historical data as is typically the case for short lifecycle products and services.

Keywords. Extended logistic model, Technology forecasting

1 Introduction

With the rapid introduction of new technologies, electronic products and services are often replaced within a year. The product life cycle for electronic goods, which used to be about ten years in the 1960's, fell to about 5 years in the 1980's and is now less than two years for cell phones and computer games. As product life cycles become shorter, less data becomes available for analysis. Given this market situation, it is important to use smaller data sets to forecast future trends of new electronic products and services as they are introduced.

A product life cycle is typically divided into four stages: introduction, growth, maturity and decline stage [3]. At introduction stage, the product is new to the market and the product awareness has not been built, so the feature of this stage is

[1] Professor (Management Science), National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu, Taiwan 300; Tel: +886 (3) 5727686; Fax: +886 (3) 5713796; Email: trappey@cc.nctu.edu.tw

slow sales growth. The growth stage is a period of rapid sales growth since the product is accepted widely by the market and the sales volume will boost. As the sales growth begin to slow down, the mature stage start. Therefore, the product life cycle curve can be illustrated with S-shaped curve from introduction stage to beginning of mature stage. Since lifecycle curve grow as sigmoid curve, growth curve method could be applied to forecast the future trend of products.

Growth curves are widely used in technology forecasting [4, 6-10] and are referred to as “S-shaped” curves. Technology product growth follows this curve since the initial growth is often very slow (e.g., a new product replacing a mature product), followed by rapid exponential growth when barriers to adoption fall, which then falls off as a limit to market share is approached. The limit reflects the saturation of the marketplace with the product or the replacement of the product with another. The curve also models an inflection or break point where growth ends and decline begins.

Many growth curve models have been developed to forecast the adoption rate of technology based products with the simple logistic curve and the Gompertz curve the most frequently referenced [1, 6]. However, when using these two models to forecast market share growth, care must be taken to set the upper limit of the curve correctly or the prediction will become inaccurate. Setting the upper limit to growth can difficult and ambiguous. If the product is a necessity or will likely be popular for decades, then the upper limit can be set to 100% of the market share. This means that the product will be completely replaced only after everyone in the market has purchased the product. However, when marketers consider new technology products such as a computer game or a new model cell phone, the value for the upper limit to market share growth can be difficult to estimate. That is, a computer game can be quickly replaced by another game after only reaching 10% market share.

In order to avoid the problem of estimating the market share growth capacity for the simple logistic and the Gompertz models, Meyer and Ausubel [8] proposed the extended logistic model. Under this model, the capacity (or upper limit) of the curve is not constant but is dynamic over time. Meyer and Ausubel also proposed that technology innovations do not occur evenly through time but instead appear in clusters or “innovation waves.” Thus, they formulated an extended logistics model which is a simple logistics model with a carrying capacity $K(t)$ that is itself a logistics function of time. Therefore, the saturation or ceiling value becomes dynamic and can model pulses from bi-logistic growth. Bi-logistic growth represents growth where market share seems to reach a limit but then growth is suddenly rejuvenated or reborn and begins again. Therefore, the researchers extend the constant capacity (K) of the simple logistic model to the carrying capacity ($K(t)$). This study applies this idea to the study of electronics products with $K(t)$ representing an extended logistic model.

The proposition of this research is that the extended logistic model with a time-varying capacity feature will be better than the simple logistic model and the Gompertz model which require a set constant capacity. Therefore, this research studies the forecast accuracy of large and small data sets using the simple logistic, Gompertz, and extended logistic models. A time-series dataset describing the

Taiwan market growth rates for two types of electronic products and two types of services were used to evaluate model performance.

2 Technological Forecasting Models

2.1 Simple Logistic Curve Model

Most biological growth follows an "S-shape" curve or logistic curve which best models growth and decline over time [8]. Since the adoption of technology and technology based products is similar to biological growth, the simple logistic model is widely used for technology forecasting.

The model of simple logistic curve is expressed as $y_T = \frac{L}{1+ae^{-bt}}$ where L is the upper bound of y_t, a describes the location of the curve, and b controls the shape of the curve. To estimate the parameters for a and b, the equation of the simple logistic model is transformed into linear function using natural logarithms. The linear model is expressed as $Y_t = \ln(y_t/L - y_t) = -\ln(a) + bt$ and the parameter a and b are then estimated using a simple linear regression.

2.2 Gompertz Model

The Gompertz model was first used to calculate mortality rates in 1825 but has since been applied to technology forecasting [5]. Although the Gompertz curve is similar to the simple logistic curve, it is not symmetrical about the infllection point which occurs at $t = (\ln(b)/k)$. Gompertz's model reaches the point of inflction early in the growth trend and is expressed as $y_t = Le^{-ae^{=bt}}$, where L is the upper bound which should be set before estimating the parameters a and b. Similar to the parameters of the simple logistic model, natural logarithms are used to transform the original Gompertz model to linear equation $Y_t = \ln(\ln(L/y_t)) = \ln(a) - bt$ and then the parameters are estimated.

2.3 Extended Logistic Model

The simple logistic model and the Gompertz model assume that the capacity of technology adoption is fixed and there is an a upper bound to growth for these models. However, the adoption of new technology is seldom constant and changes over time. As shown by Meyer & Ausubel [8], the original form of simple logistic model $\frac{df}{dt} = by\left(1 - {}^{y}\!/_{k}\right)$ is extended to $\frac{df}{dt} = by\left(1 - {}^{y}\!/_{k(t)}\right)$, where k is the upper limit of the logistic curve and $k(t)$ is the time-varying capacity and is the function which is similar to logistic curve.

This research uses the extended logistic model and assumes that $k(t) = 1 - D \times e^{-at}$, where the value of D can be any number and the value of A larger than zero. The setting of $k(t)$ comes from Chen's study [2]. This research also assumes that the capacity will fluctuate with time and the saturation level of a product could be 100% but may also just be 60% or 80% because some new products could be introduced to the market and substitute older products. Thus, a product may not achieve the 100% penetration of the market and may be replaced earlier than expected.

Finally, the extended logistic model can be expressed as $y_t = \dfrac{k(t)}{1 + C \times e^{-bt}} = \dfrac{1 - D \times e^{-at}}{1 + C \times e^{-bt}}$, where $k(t)$ is the capacity that is fluctuate with time, and a, b, C, D are the parameters computed using a nonlinear estimation method provided by a statistic software package like SYSTAT.

2.4 The Measurements of Fit and Forecast Performance

There are many statistics used to measure forecast accuracy. In this study, the Mean Absolute Deviation (MAD) and Root Mean Square Error (RMSE) are used to measure performance. The mathematical representations are shown below:

$$MAD = \frac{\sum_{i=1}^{n} \left| f_i - \hat{f}_i \right|}{n}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(f_i - \hat{f}_i \right)^2}{n}}$$

where f_i is the actual value at time t, $\hat{f}_i$ is the estimate at time t, and n is the number of observations. These measurements are based on the residuals, which represent the distance between real data and predicted data. Consequently, if the values of these measurements are small, then the fit and predicted performance is acceptable.

3 Data Collection and Data Analysis

The Taiwan market growth rate for color TVs, telephones, Asymmetric Digital Subscriber Lines (ASDL) and mobil internet subscribers were collected to test the forecast accuracy of the simple logistic model, the Gompertz model and the extended logistic model.

ADSL is a type of data communications technology which enables the use of the Internet over copper telephone lines. ADSL users must sign up for accounts so the application data models the growth rate of adoption of the technology service. Mobil internet subscribers also apply for accounts and the subscriptions for WAP, GPRS, and PHS represent this class of data service.

There are 1 data points for color TV set purchases (from 1974 to 2004) and 35 data points for telephone purchases (1964 to 2004). Twenty six quarterly data points for ADSL subscription were collected from the second quarter in 2000 to the third quarter in 2006 and 19 quarterly data points for mobile internet subscription were collected from the third quarter in 2000 to the third quarter in 2006.

Figure 1 shows the saturation rate for the four products. The growth rate for color TVs and telephones follows as S-shape curve and the curves for ADSL and mobile internet subscriptions also have basic sigmoid curve. As can be seen in Figure 1, it spent 20 years (1964-1984) for telephone getting into the mature stage whose characteristic is slower growth rate than that at the growth stage in product lifecycle. Further, color TV also spent 10 years (1974-1984) entering the mature stage. However, for ADSL and mobile internet subscriptions, it only took 5 years to begin the mature stage. Therefore, the data for color TV and telephones were categorized as long life cycle products, while the data for ADSL and mobile internet were as short life cycle products.

These four data are fitted to the models after removing the last five data points. The last 5 data points were reserved to test the prediction accuracy of the models. Further, to compare the performance of the three models using large and small data sets, the data for color TV and telephones were used to represent longer lifecycles (larger data sets) since these two products have more than 30 years historic data. Compared to color TV and telephones, the data for ADSL and mobile internet subscriptions rely on small data sets of less than 6 years and represent shorter lifecycles.

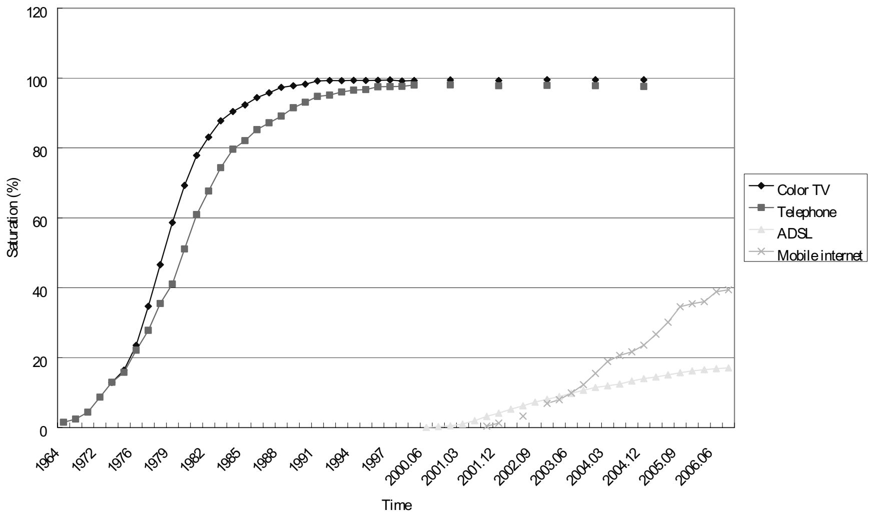

Figure 1. Market growth for Color TVs, telephones, ADSL subscription, and mobile Internet subscription

4 Analysis

The first step is to fit the data to the simple logistic, Gompertz, and the extended logistic models. After reserving the last five data points, the coefficients of the models and the statistics for MAD and RMSE are computed. The second step uses the derived models to forecast the five data points and compare the forecast with the true observations and compare.

Table 1 summarizes the fit and prediction performance of the three models. The evaluation rule is that the smaller the value for MAD and RMSE, then the better the prediction performance. Therefore, the results show that the extended logistic model has the best fit and prediction performance for both long and short data sets. Thus, the extended logistic model is suitable for predicting both long and short lifecycle products.

The simple logistic and the Gompertz model require that the values for the upper limit be set correctly, otherwise the accuracy of the models can be suffer. For the long data sets, the upper limit for color TVs and telephones is easy to see and is set at 100%. However, for the short data sets, the capacity for ADSL and mobile Internet service is difficult to determine. Thus, for the simple logistic and the Gompertz model, different upper limits are set for the short data sets. The upper limit for the saturation rate of ADSL is set to 100%, 50% or 30% whereas and the upper bound for mobile Internet is set to either 100% or 50%. As shown in Table 1, the extended logistic model with time-varying capacity yields the best fit and prediction performance.

5 Discussion and Conclusion

This study compares the fit and prediction performance of the simple logistic, Gompertz, and the extended logistic models for four electronic products. Since the simple logistic and Gompertz curves require the correct setting of upper limits for accurate market growth rate predictions, these two models may not be suitable for short life cycle products with limited data. Therefore, to solve this problem, the extended logistic model proposed by Meyer and Ausubel was used in this study. Since the extended logistic model estimates the time-varying capacity from the data, it tends to perform better for long data sets and for short data sets. Besides, the extended logistic model is also suitable for predicting both long and short lifecycle products.

There are limitations to the use of the extended logistic model. For example, Meade & Islam [6] use telephone data from Sweden to compare the simple logistic, extended logistic, and the local logistic models. They concluded that the extended logistic model had the worst performance. However, the growth curve of their data set was resembled a linear curve, not like sigmoid curve. Since the capacity of extended logistic model is time-varying and is a logistics function of time, it is not suitable for the data with linear curve. Therefore, we suggest that if forecasters wish to apply the extended logistic model, then they should confirm that the data is not linear. This research is concerned with accurate market predictions for short life cycle, small data set problems. For this study, we conclude that the extended logistic model is most suitable and research is planned to conduct a more rigorous and systematic testing of the model as proposed by Meade and Islam [8].

Table 1. Performance measures for the extended logistic, Gompertz, and the simple logistic models

Data length	Model	Statistic	Fit			Forecast		
			Extended logistic	Gompertz	Simple logistic	Extended logistic	Gompertz	Simple logistic
Long data	Color TV	MAD	0.0053	0.0297	0.0361	0.0025	0.0042	0.0045
		RMSE	0.0071	0.0502	0.0551	0.0026	0.0043	0.0046
	Phone	MAD	0.0063	0.0290	0.0323	0.0049	0.0125	0.0171
		RMSE	0.0083	0.0440	0.0453	0.0057	0.0130	0.0174
Short data	ADSL 100%	MAD	0.0036	0.0154	0.0329	0.0140	0.1037	0.2846
		RMSE	0.0043	0.0228	0.0483	0.0147	0.1060	0.2933
	ADSL 50%	MAD	0.0036	0.0140	0.0274	0.0140	0.0671	0.1688
		RMSE	0.0043	0.0170	0.0372	0.0147	0.0688	0.1716
	ADSL 30%	MAD	0.0036	0.0102	0.0212	0.0140	0.0345	0.0833
		RMSE	0.0043	0.0119	0.0264	0.0147	0.0352	0.0839
	Mobile Internet 100%	MAD	0.0057	0.0134	0.0337	0.0117	0.0627	0.2397
		RMSE	0.0073	0.0156	0.0439	0.0152	0.0699	0.2503
	Mobile Internet 50%	MAD	0.0057	0.0059	0.0217	0.0117	0.0146	0.0503
		RMSE	0.0073	0.0076	0.0248	0.0152	0.0166	0.0518

5 Reference

[1] Bengisu M, Nekhili R. Forecasting emerging technologies with the aid of science and technology databases. Technological Forecasting and Social Change 2006;73:835-844.

[2] Chen C-P. The Test of Technological Forecasting Models: Comparison between Extended Logistic Model, Fisher-Pry Model, and Gompertz Model. Master thesis, National Chiao Tung University, 2005.

[3] Kotlor P. Marketing management. 11th ed. 2003, New Jersey: Prentice Hall.

[4] Levary RR, Han D. Choosing a technological forecasting method. Industrial Management 1995;37:14.

[5] Martino JP. Technological Forecasting for Decision Making. 3rd ed. 1993, New York: McGraw-Hill.

[6] Meade N, Islam T. Forecasting with growth curves: An empirical comparison. International Journal of Forecasting 1995;11:199-215.

[7] Meade N, Islam T. Technological forecasting--model selection, model stability, and combining models. Management Science 1998;44:1115.

[8] Meyer PS, Ausubel JH. Carrying Capacity A Model with Logistically Varying Limits. Technological Forecasting and Social Change 1999;61:209-214.
[9] Meyer PS, Yung JW, Ausubel JH. A Primer on Logistic Growth and Substitution. Technological Forecasting and Social Change 1999;61: 247-271.
[10] Rai LP, Kumar N. Development and application of mathematical models for technology substitution. PRANJANA 2003;6: 49-60.