

# Diffusion Models of E-Books:

## Epidemic Model and Probit Model

Yasutaka UEDA\*

### Abstract

This paper attempts to review previous studies relating to diffusion models and determine their implication on the diffusion of e-books in Japan. The conclusions drawn from this study include findings such as that, in the case of e-books, which have both aspects of a transition from existing technology to new technology and the creation of a new technical field, they resemble a logistic model, characterized by starting out slow and a long period of time until critical mass is reached, and that such a slow start is only natural since it involves a significant change in human history, constituting a change from the physical to the non-physical for the first time in the long history of books going back more than five thousand years.

**Keywords:** Epidemic model, Probit model, Logistic model, Bass model, Network externality, Critical mass, Lock-In

## 1. Introduction

Due to the appearance of e-books, pessimistic concerns that paper books (printed books) would disappear began to be heard from persons in publishing fields troubled over the fear of books no longer selling. The appearance of e-books can be perceived as the “third revolution” following the appearance of written words (first revolution) and the invention of the printing press (second revolution) from the viewpoint of the history of books. The diffusion of e-books can be considered to essentially be equivalent to the innovation of digital technology resulting in the transformation from physical books to non-physical books. Demand for e-books differs considerably from demand presumed by a simple economic model. In the case of commodities and services discussed with an ordinary economic model, the degree of satisfaction obtained by consuming a commodity is unrelated to the behavior of other users. The effect obtained when a certain consumer dines at a restaurant is not influenced by the food consumed by the person at the next table. In contrast, demand for e-books is different. In the case of consumption of e-books, the value of an e-book for a particular person increases the greater the proliferation of e-books, or in other words, the more the same commodities and services are consumed by another person.

---

24 Nov. 2011

\* Doctor of Science (Global Information and Telecommunication)  
Professor, Edogawa University, Japan

This paper attempts to obtain suggestions for future diffusion and advancement of e-books in Japan by estimating their usage and ownership structure. Following a review of related previous studies, a description is given of the analysis framework and methods followed by a discussion of implications for the future.

Furthermore, the author visited the Gutenberg Museum located in the Mainz, Germany, the birthplace of Johannes Gutenberg, in February 2011 for the purpose of conducting research. The Gutenberg Museum is far superior to other museums in terms of being able to examine a wide range of heirlooms from the early days of printing culture. Although it can be determined from archeological artifacts that characters were inscribed on bone, bronze, porcelain and stone start-



Photo-1 Gutenberg Museum in Mainz, Germany

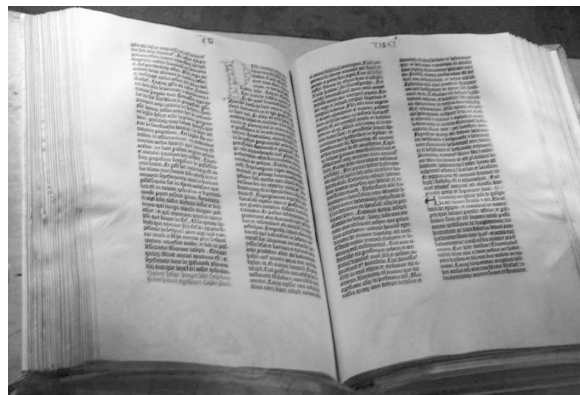


Photo-2 42-line Gutenberg Bible (1452-55)

ing around 5000 B.C. in China, and that paper was made from hemp fiber, silk threads, the bark of mulberry trees as well as components of other plants, it is quite evident that all of these constitute the inscription of characters on physical materials, making them considerably different in nature from non-physical e-books.

And the author conducted surveying e-books at local electrical appliance stores and book stores in New York, U.S.A. (October 28 to November 1, 2011).



Photo-3 E-Books at electrical appliance store “BESTBUY” in New York, USA

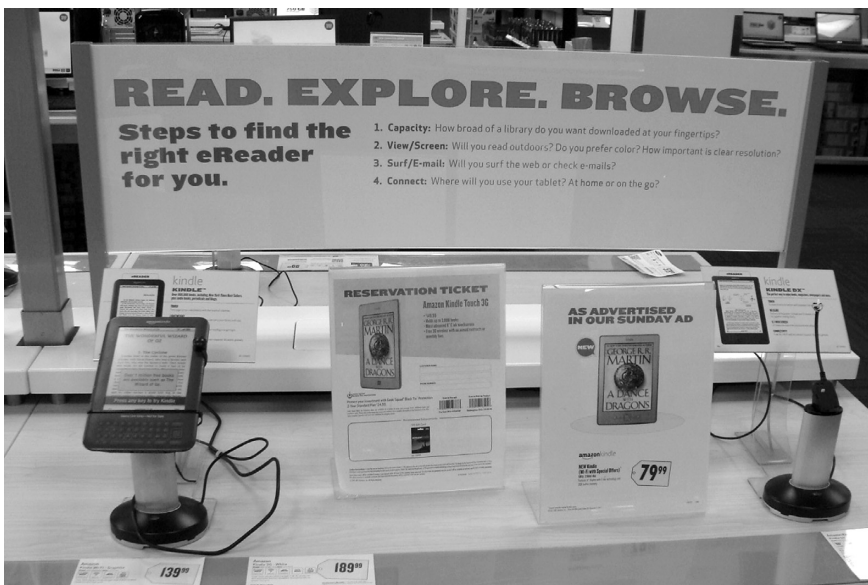


Photo-4 E-Books at electrical appliance store “BESTBUY” in New York, USA

## 2. Literature Review

Jitsuzumi (1988) conducted an analysis of the course of diffusion of mobile communications using a logistic curve. Takashima and Chou (1998) used a logistic model to conduct an empirical analysis of the number of cell phone subscribers. Kimura (2001) demonstrated that the course of diffusion of the number of cell phone subscribers can basically be effectively explained with a logistic model. Kodama (2000) confirmed that a logistic model can be applied to the diffusion of devices such as liquid crystal devices, printers and PCs. Kimura (2009) analyzed the diffusion of digital HDTV receivers using a logistic regression model. There have also been analyzes using the Bass (1969) model, which is suitable for analyzing new demand and predicting demand. Yamashita (1999) predicted the diffusion of digital televisions by referring to the course of the previous diffusion of VTR and color televisions using the Bass model, and came to the conclusion that the diffusion rate would reach 100% in 10 years. Other examples of studies using the Bass model include a prediction of the diffusion of mobile televisions (Ueda, 2003), an analysis of the diffusion of music (Asai, 2008), and an analysis of the proliferation of digital television receivers (Oniki and Honma, 2008). Although there have yet to be any studies relating to electronic books, there are numerous surveys on network externality and diffusion models. It is the intent of this paper to verify a diffusion model for e-books while referring to these surveys.

## 3. Technical Breakthrough Model and Technology Fusion Model

In the case of perceiving the appearance of e-books as a technical innovation, what then is the essential phenomenon responsible for that technical innovation? Technical innovations can broadly be divided into two types. The first type is a technical breakthrough model that brings about a technical replacement in which an existing technology becomes obsolete and is replaced with a new technology. Examples include the case of semiconductors replacing vacuum tubes or the case of CDs replacing records.

The second type of technical innovation is a technology fusion model in which existing technologies of different fields are organically fused and integrated to give birth to a new technical field. Examples of this type include the field of mechatronics resulting from the fusion of mechanical technology and electronics technology, and the field of optoelectronics, which is the fusion of electronic engineering technology and optical technology.

With respect to the relationship with demand, in contrast to the first type in the form of the technical breakthrough model essentially taking over existing demand, the second type in the form of the technology fusion model results in the creation of new demand through the creation of completely new products and services.

E-books differ from both of these types in that they constitute a type that encompasses both types. The appearance of e-books can be perceived as the “third revolution” following the appearance of written words (first revolution) and the invention of the printing press (second revolution) from the viewpoint of the history of books. The second revolution in the form of the

invention of the printing press by Gutenberg 450 years ago is a typical example of the technical breakthrough model of the first type. Following the invention of the printing press using lead typesetting by Johannes Gutenberg in the middle of the 15<sup>th</sup> century, the hand-written manuscripts prior to that time essentially disappeared. In other words, a technical replacement occurred from handwritten manuscripts and woodblock prints to typographic printing, or in other words, to the media of books. Since both the existing technology and the new technology were the same in that they both involved the “physical”, the appearance of the new technology, which was much more functional than the existing technology, led to the unavoidable demise of the old existing technology.

However, the appearance of e-books differs from this second revolution. This is because it constitutes a transition from physical books to non-physical books. Following the invention of typographic printing using lead typesetting by Gutenberg, people have perceived books to be that in which multiple pieces of paper containing printed characters and pictures are tightly bound and provided with somewhat hard front and back covers. However, e-books cannot be considered to be physical since they are the result of storing text by a binary series of 0s and 1s and displaying on a computer screen in the form of flashing points of light. Since they represent a change from physical books to non-physical books, this does not constitute a complete replacement of printed books with e-books. Since there both things that can be done and cannot be done by printed books (ordinary books), there are also things that can and cannot be done by e-books. In order to read an electronic book, it is necessary to take the trouble to provide certain tools, such as an e-book reader. Since the e-book reader is a computer, it is also necessary to continuously provide a power supply. Thus, an e-book cannot simply be put away in a back pocket in the manner of a paperback. In other words, the disappearance of printed books will not occur in the future.

In addition, e-books also have the potential to create numerous forms of new demand unable to be realized with printed books. Archives can be constructed offering vast amounts of digital data and search technologies in the manner of Google’s “Book Research”. Books can also be converted to multimedia by integrating written text with images, movies, voice and music. E-books are also able to overcome hardships confronting the physically challenged or elderly by utilizing voice delivery systems and enlarged displays. Automated translation services are also expected to be available between countries and regions using different languages. Since data can be shared easily, non-commercial publication becomes easier. There is also the potential for numerous forms of distribution systems to be attempted, and the creation of new screening and selection standards not simply based on whether or not a book sells as in the past.

In this manner, e-books have both the aspects of replacement and creation in that in addition to replacing demand for some technologies, they also create new technical fields, thereby contributing to an increase in overall book demand.

## 4. Epidemic Model and Probit Model

Stoneman (1983) classified a non-linear model used to explain technology and other diffusion phenomena into two types consisting of an epidemic model and a probit model.

### 4.1. Epidemic Model

The epidemic model is a model that developed from a model representing a population increase, the course of transmission of an infectious disease and the like, and is typically the most frequently used non-linear model for extremely effectively explaining the diffusion of technologies, products and services. Those functions that represent the epidemic model consist of a logistic curve and a double exponential function (log-log model).

In a logistic regression model, the probability of the use of an e-book is used as the response variable, and can be explained with such factors as personal attribute factors, the ownership and utilization status of electronic book readers, or the diffusion status of e-books. The occurrence probability  $p(x)$  of an arbitrary  $x$  in the logistic regression model is represented by the following formula (1):

$$p(x) = \frac{\exp(Z)}{1 + \exp(Z)} = \frac{1}{1 + \exp(-Z)} \quad (1)$$

wherein,  $Z$  represents the linear composite variable of several explanatory variables.

The function formula obtained by linear transformation of formula (1) is as indicated below:

$$\log \frac{p(x)}{1-p(x)} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_r X_r \quad (2)$$

wherein, the right side of the formula is the linear regression model, and regression analysis can be carried out by using the product of logit transformation of  $p(x)$  for the response variable, and  $x_1, x_2, \dots, x_r$  as explanatory variables.

Here,  $M_t$  represents the market size at time  $t$ , and transforming formula (1) by using  $N_t$  as the latent market size at time  $t$  yields the following formula (3):

$$M_t = \frac{N_t}{1 + \alpha \cdot \exp[-\beta \cdot t]} \quad (3)$$

wherein,  $t$  is an explanatory variable representing time,  $M_t$  represents the diffusion rate at time  $t$  (cumulative number of subscribers),  $N_t$  is a constant that represents the upper limit of the diffusion rate at time  $t$ ,  $\beta$  represents the diffusion speed, and  $N_t$  represents the upper limit of latent diffusion and does not change. According to Rodgers (1995), with respect to a logistic model, information relating to the use of a technology is conveyed by persons already using that technology to persons not yet using that technology, and persons not yet using that technology make a judgment as to whether or not they will accept that technology. Consequently, as the

evaluation of that technology by existing users improves, the total number of users newly using that technology increases.

On the other hand, the log-log model is one of the main models used in the fields of medicine and biology.

#### 4.2. Probit Model

The probit model is a model used to explain diffusion phenomena from an economic viewpoint. Functions used to represent the probit model consist of a standard normal cumulative distribution function and a logarithmic normal cumulative distribution function.

The situation in which a different consumer has an effect on the decision made by a certain consumer as to whether or not to consume an e-book is not assumed in the case of ordinary commodities and services. If the benefit of a certain consumer expected to be obtained by subscribing to a network outweighs the cost to be borne as a result thereof, it is logical for that consumer to subscribe, while in the case the benefit does not outweigh the cost, the consumer elects not to subscribe. The “network effect”, in which the convenience of another consumer is improved by the subscription to a network by a certain consumer, is referred to as “network externality” in the sense that the decision of a certain consumer has an effect on another consumer without being mediated by a market mechanism.

Since the number of users of a technology is small during the early stages of diffusion of that technology, the effect of network externality acts in a direction that delays the decision to use that technology by latent users, and since a large number of users are already using that technology as proliferation advances, the effect acts in a direction that promotes the decision to use a latent market that is newly beginning to be used.

The standard normal cumulative distribution function is in the form of an S-curve in which, although the initial diffusion speed is extremely slow, it rapidly begins to diffuse after having reached a critical mass. The function formulas obtained by linear transformation are as indicated below.

$$Z_t = \beta_0 + \beta_1 \cdot t \quad (4)$$

$$Z_t = \frac{(t - \mu d)}{\partial d} \quad (5)$$

The logarithmic normal cumulative distribution function depicts a curve that demonstrates a sudden rise followed by quickly reaching a limit. The function formulas obtained by linear transformation are as indicated below.

$$Z_t = \beta_0 + \beta_1 \cdot \log t \quad (6)$$

$$Z_t = \frac{(\log t - \mu d)}{\partial d} \quad (7)$$

## 5. Bass Model

The Bass model (1969) is used as a model that is suited for predicting the diffusion and demand of new demand for durable consumer goods, and has been previously applied to demand analyses of durable consumer goods, including television receivers and other home appliances. Bass (1969) divided the number of consumers of goods into innovators, which decide on a purchase without being affected by the number of existing purchasers, and imitators, which decide on a purchase while being affected by existing purchasers, and formulated a probability  $p(t)$  of a non-purchaser prior to a time  $t$  making a purchase at that time  $t$  in the form of the formula indicated below:

$$p(t) = \frac{f(t)}{1-F(t)} = P + qF(t) \quad (8)$$

wherein,  $p$  represents an innovation coefficient,  $q$  represents an imitation coefficient,  $f(t)$  represents a probability function at the time of purchase, and  $F(t)$  represents the probability distribution function thereof. At each time, the number of innovators accounts for a fixed proportion ( $p$ ) of the number of non-purchasers, and the number of imitators is proportional to the number of existing purchasers. In this model, an amount of time is required until the quantity of goods sold reaches a peak, and demand is gradually stimulated by existing purchasers through word-of-mouth.

## 6. Conclusions and future development

In the case of a land-based telephone, for example, calls cannot be made unless the other person has a telephone. Consequently, the number of persons who had previously subscribed to telephone service is a major factor in determining the number of subscribers at that time. In contrast, in the case of e-books, it is not necessary for friends or acquaintances to be reading an e-book. Even if they are reading printed books, the person in question is able to freely enjoy reading e-books. In addition to the psychological factor of the proliferation of e-books among others, one of the main factors that determine whether or not one uses e-books is the evaluation of the convenience of electronic books by others already using e-books. In addition, there is probably also the sense of superiority of being a person who has subscribed to a new service faster than others as well as the sense of satisfaction derived from owning an electronic book representing the latest technology. Moreover, reduced hardware (e-book reader) costs, reduced software (e-book) costs and intensified publicity and advertising activities stemming from increased competition probably also contribute to deciding whether or not to use e-books.

However, the increase in latent market size triggered by the creation of new demand is analyzed as depicting an extremely gradual logistic curve. In the logistic model, although the startup speed is somewhat fast, the time to reach critical mass is long, thereby resulting in a gradual slope for the curve at the time of startup.



For authors, publishing houses, printing companies, libraries and the like, e-books are frequently perceived as a threat by virtue of depriving them of a previously acquired interest. However, within the long history of books going back more than 5,000 years, the method by which books are read has been divided into two facets consisting of physical books (printed books) and non-physical books (e-books) for the first time in history. During the course of time going back more than 5,000 years, although the products of inscribing, writing or printing characters and letters that have been organized into sentences to a certain degree on both clay plates, paper made from bamboo parchment (bamboo writing strips) or animal skins and present-day paper, and winding into scrolls or binding into publications, have been defined as books, since e-books constitute a significant historical change from physical books to non-physical books, it is only natural that they demonstrate a gradual start, thus eliminating the need for any sense of panic on the part of related industries.

In addition, many of the current e-books are in the form of applications, and when in this form of being downloaded as files, enable e-books to only be read with specific hardware. Consequently, when the quantity of software (e-books) able to be used only with specific hardware (e-book readers) increases, the market superiority of that specific hardware (e-book reader) will be locked in. The current situation surrounding e-books is such that the software (e-books) that can be read with hardware is limited, and has the risk of being locked in. If the platform for reading e-books develops in the manner of a “cloud service” that enables all electronic books to be read regardless of the hardware, this “lock-in” will not occur, and since the expected benefits to consumers obtained by subscribing to a network will improve, the diffusion speed of e-books overall is expected to be accelerated. Since there are no limitations on terminals used to store and read data on the Internet, the resulting image is that of converting a bookshelf to a cloud. As long as the terminal is provided with a browser, books can be read that have been purchased over the Internet from anywhere in the world.

Finally, a discussion is provided of some of the issues raised by this paper. It is desirable to conduct an empirical analysis based on quantitative data relating to the U.S. and Japanese markets in order to conduct a more precise analysis. The conducting of such in-depth empirical research will be an issue for the future.

## 7. Acknowledgements

Financial support by Edogawa University to this study is gratefully acknowledged. Note that the author is responsible for any errors in this paper.

## References

- [ 1 ] Asai S. [2007], “Information Transmission by Mass Media and Word-of-Mouth Effects as Seen During the Course of Proliferation of Music,” *Journal of Information and Communication Research* (2007), pp. 27–38
- [ 2 ] Bass, F. M. [1969], “A New Product Growth Model for Consumer Durables,” *Managing Science*, 15 (5), pp. 215–227

- [ 3 ] Jitsuzumi T. [1998], "Analysis of the Current State of the Mobile Communication Industry," *Institute for Posts and Telecommunications Policy (IPTP) Monthly Report* No. 122, 1998. 11, pp. 4-23
- [ 4 ] Kimura M. [2001], "Quantitative Analysis of Creation of New Demand by Digital Fusion," *Journal of Information and Communication Research* (2001), pp. 89-102
- [ 5 ] Kimura M. [2004], "Comparative Analysis of Cell Phone Proliferation in Three Northeast Asia Countries," *Journal of Information and Communication Research* (2004), pp. 27-38
- [ 6 ] Kimura M. [2009], "Exploring the Determinants of Digital HDTV Adoption," *Journal of Public Utility Economics*, Volume 61, No. 2, November 2009, pp. 45-54
- [ 7 ] Kodama F. [2000], "Innovation Management in The Emerging IT Environments," POM Facing The New Millennium (selected papers from The First World Conference on Production and Operation Management POM Sevilla 200), 2000
- [ 8 ] Li Zeng, Shishikura M., Kasuga N. [2008], "An Analysis of Demand for New Devices in the Transition from Terrestrial Broadcasting to Digital — The Case of Digital Televisions —", *Journal of Information and Communication Research* No. 87 (vol. 26 No. 2) 2008, pp. 67-76
- [ 9 ] Oniki. Homma [2008], "An Economic Analysis of the choice at a Time of Analog Wave Suspension for Digitalization of Terrestrial Broadcasting," JSICR Annual conference (2008. 6. 15)
- [10] Rogers, M. R. [1995], "Diffusion of Innovations" 4th ed., New York, The Free Press, 1995
- [11] Sharif, M. N., Ramanathan, K. [1981], "Binomial Innovation Diffusion Models with Dynamic Potential Adopter Population," *Technological Forecasting and Social Change* 20, 1981, pp. 63-87
- [12] Stoneman, P. [1983], "The Economic Analysis Technological Change," Oxford University Press, 1983
- [13] Takashima • Chou [1998], "Effects of Mobile Communication Rates on Proliferation Saturation Level," *Journal of Information and Communication Research* (1998), pp. 11-21
- [14] Ueda Y. [2003], "Expansion of the Possibilities of Television Advertising Effects by Mobile Televisions," *Journal of Advertising* (No. 44) (2003. 8), pp. 101-112
- [15] Yamashita T. [1999], "Study on the Course of Proliferation of Digital Television Receivers," *Journal of Information and Communication Research* (1999), pp. 34-45