Changing Price Elasticity of Digital Goods: Empirical Study from the E-book Industry

Completed Research Paper

Eunkyoung Lee
KAIST Business School
85 Hoegiro Dongdaemoon-gu Seoul, Korea
greenpray@business.kaist.ac.kr

Byungtae Lee
KAIST Business School
85 Hoegiro Dongdaemoon-gu Seoul, Korea
btlee@business.kaist.ac.kr

Abstract

Many industries including music and e-book industries have been making or already made the transition from physical to digital goods in a couple of decades. Since the transformation fundamentally affects the cost structure, consumers' purchasing behavior, and other critical factors in the industries, it is important to understand the changes that the transition has brought about.

This paper empirically explores the dynamics of price elasticity of e-book demand. Since the e-book market has been rapidly growing in recent years, consumers' price sensitivity and its dynamics should be investigated to develop an optimal pricing strategy.

We use massive industry sales data and find that demand of e-books is becoming price inelastic over time. This can be partially attributed to the reference price, which has been constructed by consumers through personal experiences. We also find other factors that influence the price elasticity and demand of e-books.

Keywords: Digital goods, e-books, price elasticity, dynamics of price elasticity, reference price effect

Introduction

Digital goods is a general term that is used to describe any goods that are produced, stored, delivered, and used in an electronic infrastructure such as the Internet. In a couple of decades, many industries have been making or already made the transition from physical to digital goods. Since this transformation fundamentally affects the cost structure, consumers' purchasing behavior, and other critical factors in the industries, it is very important for managers and researchers to understand the changes that the transition has brought about. The music industry is a prime example. Most people go to the online music stores like iTunes, but scarcely visit the brick-and-mortar stores in order to buy music these days. The book industry is also undergoing a digital transformation. According to an article in the Economist (2010), like many other parts of the media industry, publishing is being radically reshaped by the growth of the Internet, and now e-books threaten to undermine sales of the old-fashioned kind. It is the e-book industry that we mainly deal with in this research.

An e-book, one of the representative digital goods, has recently come into wide use with the proliferation of mobile devices including smart phones, tablet PCs, and dedicated e-book readers like Kindle. As a result, the e-book market has grown dramatically in the last few years alone. According to the American Association of Publishers, e-book sales in the U.S. reached $113 million in 2008, which was just under 0.5% of overall book sales. While the amount of e-book sales accounts for a small fraction in 2008, the market grew at a rate of 68.4% in that year. Since then, the e-book industry has maintained its rapid growth, finally becoming as large as that of paper books. Amazon.com, a leader in the bookselling industry, announced that e-book sales surpassed hardcover sales in July 2010 and paperback sales in January 2011. Moreover, in May 2011, e-books outnumbered both hardcover and paperback combined by the ratio of 105 e-books to 100 print books. Although the e-book market in Korea has yet to arrive at the maturation phase like the U.S., the situation is similar to the U.S. market a few years ago. The Korea Electric Publishing Association reported that the Korean e-book market was about $255 million in 2011, which fell below 10% of the total publishing industry at that point. The size of the Korean e-book market is still small, but it is growing at a rapid rate. It is expected to grow over time despite being in the early stage.

Although there are some paper books with no corresponding digital edition and vice versa, all contents can basically be published in both paper and digital editions. Recently, most books have actually been published in both editions in the U.S. This suggests that e-books are perfect substitutes for paper books in respect of contents. However, like other digital goods, e-books have some basic properties that differentiate them from paper books in many ways.

In terms of supply, the critical difference is the cost. Since e-books are pure digital goods, the marginal production cost is almost zero. Moreover, the retailers of e-books do not have to worry about the cost of storage and distribution. This indicates that e-books cost much less compared to paper books. Classic economic theory suggests that a lower cost leads to a reduced price in the competitive market. However, the producers and retailers cannot help but consider piracy issues and the impact of e-book price on corresponding paper book sales. Publishers are actually concerned that e-books will cannibalize existing paper book sales and that consumers will get used to low e-book prices, which may put pressure on the prices of paper books (The Economist 2010). In light of these issues, it is academically interesting and practically important to come up with an appropriate pricing strategy for e-books.

There are also differences between e-books and paper books on the demand side. For customers, e-books have both advantages and disadvantages compared to paper books. The demerits of e-books are as follows. To read e-books, consumers need to have an e-book reader device, install e-book software and payment program, and set up an e-book store account. These could be regarded as inconveniences by those who are not accustomed to using electronic devices or making online purchases. Some people might feel uncomfortable about reading electronic texts. Also, another disadvantage is the restriction on sharing contents with friends due to digital rights management (DRM) issues. On the other hand, it is a definite advantage to own and access thousands of e-books anytime with an e-book device, limited only by its memory capacity. Moreover, consumers do not need to pay shipping charges and wait for delivery when they buy e-books. They enjoy access to the e-book right after payment, and have no worries over misplacement. In consideration of all these pros and cons of e-books, consumers compare value with price to make a purchase decision.

Under these circumstances, it is important and interesting to figure out how people accept and consume e-books. Therefore, in this paper, we investigate the price elasticity of e-book demand in order to understand consumers' price sensitivity. Also, it is worthwhile to note that shortcomings of e-books are continuously being remedied by developments in technology with time. In this context, it is especially important to consider whether price elasticity has changed.

Price elasticity is the percentage change in demand caused by a percentage change in price. It has long been an interesting topic for researchers with an academic curiosity toward consumers' price sensitivity and managers seeking to detect appropriate price levels. To develop suitable pricing strategies, it is academically and practically important to understand price elasticity, a direct indicator of consumers' response to price changes. There are various factors that affect the price elasticity of demand, such as product properties, the market situation, product life cycle, the closeness of substitutes, and consumers' characteristics. In this regard, many researchers have conducted numerous econometric studies. Hoch et al. (1995) show that demographic and competitor variables influence the price elasticity of household item demand. Krishnamurthi and Raj (1991) empirically study the relationship between consumers' brand loyalty and price elasticity. Granados et al. (2012) compare price elasticity in online and offline channels. However, there is still much to be done on the dynamics of price elasticity over time. Although there have been several related studies, they are primarily concerned with price elasticity over product life cycle or brand life cycle (Simon 1979; Parker & Neelamegham 1997). This paper focuses on the dynamics of e-books' price elasticity of demand over time in the presence of reference price effects that control other factors. It is one of the first studies that empirically investigate the changing price elasticity over time by using massive industry sales data.

We are provided with e-book sales data in both years of 2011 and 2012 from Kyobobook[1], one of the largest book retailers in Korea. Since we have actual sales data, we can calculate a more accurate estimate of price elasticity than past studies that derived sales from sales rank (e.g. Chevalier & Goolsbee 2003; Brynjolfsson et al. 2003; Chose et al. 2006).

The main purpose of this paper is to explore the dynamic price elasticity of digital goods with actual sales data from the stand point of reference price effects. Our empirical results contribute academically to understanding consumers' perceptions of e-books and price sensitivity of demand. Also, it would be practically useful for managers to establish pricing strategies in new digital good industries including e-books.

The rest of this paper is organized as follows. First, we present the related theoretical and empirical literature on price elasticity and the e-book market. Second, we show the research hypotheses with theoretical background and logical reasoning, and then describe the data. Third, we develop and analyze the empirical model and discuss the results. Finally, we conclude the paper by mentioning limitations and future research.

Literature Review

This paper is related to two primary streams of previous literature. The first stream of relevant work is research on price elasticity. Over the last few decades, researchers have extensively studied price elasticity in many ways. Some of them study the determinants of price elasticity. Hoch et al. (1995) relate store-level price sensitivity to consumer demographic characteristics and the nature of local market area competition. They find that the former is much more influential than the latter. According to them, factors that affect consumers' price elasticity include family size, education level, size and price of homes, and race. Krishnamurthi and Raj (1991) empirically study the relationship between consumer brand loyalty and price elasticity. They use the conditional logit formulation to model choices among the brands. Their results show that loyal consumers are less price sensitive than non-loyals in the choice decision but more price sensitive in the quantity decision. Before their study, the results between loyalty and price elasticity have been mixed. They suggest that conflicting results are caused by modeling different stages of decision process or unit of

---

[1] We had an interview with a manager of the Digital Contents Business Team in Kyobobook. He provided practical information on the Korean e-book market, and granted access to data. We take this opportunity to express our gratitude.

data, and decompose the decision stages to better understand the elasticity components and unify findings from previous research. Granados et al. (2012) analyze the difference in price elasticity across sales channels. They use a data set with information for millions of airline tickets sold in the U.S. market, and employ the log-linear demand model to test their hypotheses. Mixon and Ressler (2000) demonstrate that the relative age of music recordings is negatively related to the price elasticity of demand. There are several studies about the dynamics of price elasticity. Simon (1979) empirically reveals typical changes in price elasticity over the brand life cycle. He shows that the magnitude of price elasticity decreases in the introduction and growth stage, reaches its minimum in the maturity stage, and increases during the decline stage. Parker and Neelamegham (1997) extend the work of Simon (1989). They study price elasticity dynamics over the product life cycle, and show consistent results with Simon (1989), which means that total sales price elasticity declines in absolute value but then increases over the product life cycle. Fibich et al. (2005) analytically study the effect of the reference price on the price elasticity of demand. The definition of reference price is the price consumers have in mind, in other words, the consumers' anticipated price (Winer 1986). It is an internal standard against which observed prices are compared (Kalyanaram & Winer 1995). According to the model by Fibich et al. (2005), reference price in the current stage is the weighted average of reference price and shelf price in the previous stage. Their results suggest that the reference price can have a considerable effect on price elasticity.

The second stream of literature relevant to our study is research on e-books. Despite the rapid growth in the e-book market, researchers have paid less attention to the theoretical analysis and empirical study of this market. However, studies on e-books have recently started in earnest. Vasileiou et al. (2009) provide an overview of the overall e-book marketplace by using data from e-book publishers and e-book aggregators. Oestreicher-Singer and Sundararajan (2010) present a model dealing with the value of the digital rights and pricing for digital goods that are also available as a tangible version. The interesting result in their study is that specific digital rights are associated with a significant increase in the threat of piracy and a corresponding reduction in a seller's pricing power. Predictions are derived from the analytical model, and then proven using data from the e-book industry. Jiang and Katsamakas (2010) examine the effects of e-book entry on the book market by developing a game theory model. They show that e-book entry leads to intricate price adjustments, as well as share and profit changes. Also, they find that the advent of an e-book seller does not necessarily expand the total readership.

Previous studies about the e-book industry give us an insight into the factors that influence consumers' price sensitivity of the e-book demand, and research on price elasticity provides a background for our empirical model. Complementing these studies, we empirically explore the dynamics of price elasticity of e-book demand.

Development of Hypotheses

Research Context

An e-book is a book in digital form, and it is produced on, published through, and readable on electronic devices such as computers, dedicated e-book readers, and smart phones. Books are historically embedded in paper, but now increasingly available as pure digital goods due to the development of the Internet and technologies. Although theories abound about when the first e-book appeared, the leading theory is that it was invented in 1971 with the first steps of Project Gutenberg, a digital library for books from public domain. However, it is only in relatively recent years that the process of digitization in the book industry has radically accelerated. Only a few years ago, e-book sales accounted for a very small portion of overall book sales, but now, the e-book market in the U.S. has become larger than the paper book market. Likewise, the Korean e-book market is rapidly growing.

We had an interview with the managers in Kyobobook, the largest book retailer in Korea, and they provided e-book sales data for this research. According to the interviewees, Kyobobook has tried to increase the amount of e-book contents by persuading publishers who have been reluctant to produce digital editions, and their efforts are getting results. E-book sales in 2012 have increased by about forty percent over the

previous year, reaching a total of 15 billion won. This is still a very small fraction compared to the overall book sales of 580 billion won, which means that the Korean e-book market is in the initial stage[2].

Figure 1 shows the number of e-books sold by Kyobobook a month from January 2011 to December 2012. On the horizontal axis, each number indicates a month: 1 for January 2011, 2 for February 2011, in the same manner, 13 for January 2012, and 24 for December 2012. This graph confirms that the sales of e-books have generally increased during the past two years.
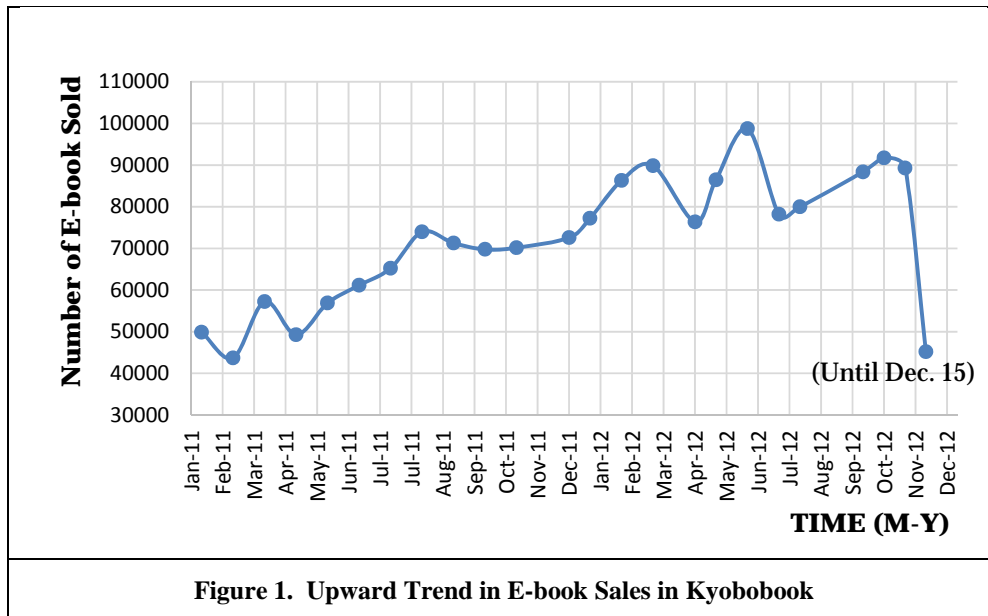

Figure 1. Upward Trend in E-book Sales in Kyobobook

As of December 15, 2012, Kyobobook has sold 111,361 titles of e-books. Consumers can purchase an e-book in Kyobobook through their smart phones, dedicated e-book devices, tablet PCs, and personal computers with network connections. Each e-book has a list price and belongs to one of 27 genres. Some e-books are only for adults, so the system prohibits the sale of adult books to minors. We describe the details of products and sales in the next section dealing with the data.

Research Hypotheses

Reference price is defined as the price consumers have in mind. It is usually constructed by consumers through personal experiences such as purchasing, observing, or being exposed to intentional and unintentional price information. According to Fibich et al. (2005), differences between the reference price and the shelf price affect the demand and price elasticity. If the reference price is higher than the shelf price, consumers are likely to sense a gain that will increase demand. On the contrary, if the shelf price is higher than the reference price, consumers are likely to sense a loss that will decrease demand. Kalyanaram and Winer (1995) find that consumers are more sensitive to losses than gains. In other words, the amount of demand that consumers reduce when they sense a loss is larger than that of demand increased by consumers when sensing a gain. Fibich et al. (2005) express the reference price using the following formulation:

$r_n = \eta r_{n-1} + (1 - \eta)p_{n-1},$

---

[2] Kyobobook is preparing to release a new e-book reader device and an entirely new e-book service in late February 2013. The new service called 'Sam' is a kind of a flat sum system. A certain number of e-books would be offered to consumers at a flat rate of a corresponding monthly fee. Kyobobook expects that this new service will increase the sales in 2013 by 40 billion won. Since this new service has not been released, the present paper deals with the general sales system, in which e-book purchases are made when a consumer pays the price of an e-book.

where r_n and p_n are reference and shelf prices at the nth buy, respectively, and η is a discrete memory parameter depending on the product category.

Since the Korean e-book market is in the initial stage, there are many customers who have never purchased an e-book. In this case, it is reasonable to assume that the first reference price would be influenced by the price of similar products in consumers' minds. It is easy to think of a paper book as a similar product to an e-book, but the survey[3] conducted by Kyobobook in late 2011 suggests that this is not true. The truth is that many consumers perceive an e-book not as a paper book but as a file that they can download for free through the Internet, just like music files or movie files[4]. It means that the reference price of e-books for consumers without any previous purchase experience is nearly zero. Because consumers are more sensitive to losses as mentioned above, we expect the price elasticity, hereafter referred to in absolute value, to be high under the circumstance that the reference price is almost zero. However, more and more people are observing that the price is not zero, and more willing to purchase e-books as time goes on. Also, the value of e-books in the minds of consumers could be enhanced by changes to external factors such as developments in technology, widespread use of mobile devices, and an increase in the amount of e-book contents. These indicate that the reference price of e-book is rising in the initial stage. This change in the reference price over time is expected to influence the price elasticity of e-book demand. Therefore, we suggest the following hypothesis.

HYPOTHESIS 1 (H1).   In the initial stage, e-book demand is becoming less price elastic over time.

The demand and price elasticity of e-books are also affected by product properties. In this study, we consider three main properties of e-books: whether an e-book has a corresponding paper book, whether an e-book is intended for adult readers only, and whether an e-book is one of the best books of the year 2011 or 2012 in overall book sales.

First, the existence of substitutes plays a critical role in determining consumers' price elasticity. Generally, if there are close substitutes for a product, the demand of the product becomes more price elastic. Parker and Neelamegham (1997) interpret the advent of competitive substitutes as entering a decline phase in the product life cycle, and show that it leads to an increase in price elasticity. For an e-book, its corresponding paper book is a very close substitute. However, we do not expect the demand of e-books with corresponding paper books to be more price elastic than those existing only in electronic form. It would be just the opposite. The reason for this thought is also explained by the reference price. In Korea's book industry, the average price level of e-books is about 60% of the average price of paper books. The e-book price is less than the price of its corresponding paper book in almost all cases. According to Smith and Nagle (1995), adding a premium product to the product line enhances consumers' perceptions of lower-priced products. In our context, it implies that the price of a paper book might be, or at least affects the reference price of its corresponding e-book. Therefore, the existence of a paper version would cause the reference price of an e-book to be high. Even though the survey conducted two years ago suggest that consumers perceive an e-book as a downloadable file available free of charge, the situation is changing quickly. There is convincing evidence in Amazon.com that the e-book market has entered the maturation phase. The prices of e-books, i.e. Kindle editions, have become almost the same with, or for some, even higher than the prices of corresponding printed editions in Amazon.com. For this reason, we build the following hypothesis.

HYPOTHESIS 2-a (H2-a).   The demand for e-books that only exist in electronic form is more price elastic than the demand for e-books with corresponding paper books.

Second, the e-book demand would be affected by whether it is only for adults or for everyone. From July to September 2012, three of the top 15 best-selling e-books in Kyobobook were adult e-books. One of the

---

[3] We could not obtain the raw data of this survey because of security concerns. Kyobobook is internally reviewing our request to access the survey data. However, during the interview, the managers provided a statement of the results in the context of the reference price.

[4] Although illegal file sharing is prohibited by law, there are still people who share files through peer to peer networks and other means.

bestsellers was Fifty Shades of Grey, which is the first to sell more than one million copies for Kindles. Yes24, the second largest e-book retailer in Korea, also announced during the same period that six of the top 10 best-selling e-books were for adults. This is quite impressive since the number of adult e-books occupies a very small portion of overall e-books[5]. A manager of Yes24 said that the reason behind high sales of adult e-books is that readers do not have to be self-conscious of others noticing the book cover. In the context of adult goods, consumers feel more comfortable when their usage remains confidential (Edelman 2009). According to Wondracek et al. (2010), the online adult industry is among the most profitable business branches on the Internet. This would also be applicable to e-books for adults. Therefore, even though the number of potential consumers of adult e-books is definitely less than that of others[6], we construct the following hypothesis.

HYPOTHESIS 2-b (H2-b).   The demand for adult-only e-books is greater than the demand for e-books for all ages.

Last, the quality of contents would apparently have a positive effect on the demand of e-books. To assess the quality, we investigate whether an e-book belongs to the list of best-selling books among all books including e-books and paper books. At the end of the year, Kyobobook announces the top 200 titles of books in that year. We combine the 200 best books of 2011 and 200 best books of 2012 in our list. With this data, we suggest the following hypothesis.

HYPOTHESIS 2-c (H2-c).   The demand for e-books belonging to the list of best books is greater than the demand for others.

Price elasticity can be influenced by consumers' characteristics as well as product properties. We predict that the price sensitivity of consumers who buy a number of e-books is different from that of other consumers. For simplicity, consumers who purchase more e-books than the average reading volume of Korean readers are called heavy users, and other consumers as non-heavy users hereafter. According to Kim and Rossi (1994) studying the relationship purchase volume and price sensitivity, consumers with high purchase frequency or high purchase volume are far more price sensitive than consumers with low frequency or low volume of purchase. In their study, a standard multinomial logit model is used to analyze the choices of households among competing brands of canned tuna fish. Kim and Rossi (1994) explain that since consumers with higher volume have a better sense of price distribution, they are more price elastic. This theoretic point of view cannot be applied to our context. Unlike cannot tuna fish, e-books in Korea are relatively new products, thus making it difficult to observe price dispersion. In contrast to Kim and Rossi (1994), Parker and Neelamegham (1997) show that price elasticity is mostly inelastic as repeat purchases increase. They consider this result as a part of price elasticity dynamics over the product life cycle. Since we have defined a heavy user as a consumer who repeatedly buys e-books, we predict that heavy users would be less price sensitive than non-heavy users. This conjecture is consistent with Krishnamurthi and Raj (1991), who empirically explore the relationship between loyalty and price elasticity in purchase behavior. Therefore, we suggest the following hypothesis.

HYPOTHESIS 3 (H3).   The heavy user demand is less price elastic than the non-heavy user demand.

An overview of the suggested hypotheses is presented in Table 1.

---

[5] The number of adult e-books is only 285 in our data. It corresponds to about 0.002% of all e-books.

[6] The system prohibits non-adults from purchasing adult e-books.

Table 1. Hypotheses		
Context	Index	Hypothesis
Dynamics	H1	In the initial stage, e-book demand is becoming less price elastic over time.
Product properties	H2-a	The demand for e-books that only exist in electronic form is more price elastic than the demand for e-books with corresponding paper books.
	H2-b	The demand for adult-only e-books is greater than the demand for e-books for all ages.
	H2-c	The demand for e-books belonging to the list of best books is greater than the demand for others.
Consumers' characteristic	H3	The heavy user demand is less price elastic than the non-heavy user demand.

Data

Data Collection

For this study, we use Kyobobook's e-book sales data from January 1, 2011 to December 15, 2012. The data set provided by Kyobobook contains all purchase orders from consumers during this period. Each order gives information on the book sold and when and to whom, which means that our data set covers the ISBN of the book sold, the date purchased, and a customer number of the purchaser for each purchase order. Customers who have previously bought an e-book in Kyobobook are distinguished using an assigned customer number in this data set. All demographic characteristics are excluded due to the company's privacy agreement. We also receive e-book data containing the ISBN, title, list price, author, publisher, genre, publication date of the corresponding paper book, if any, registration date of the e-book, the number of pages of the e-book, whether it is for adults, whether it has a corresponding paper book, whether it is one of best books in 2011 or 2012, and whether it is one of the bestselling e-books in each month during the period. There are a total of over 1.7 million purchase orders in the data set.

Key Variables and descriptive Statistics

To test the hypotheses constructed above, we aggregate the purchase orders in two ways. One is to investigate the dynamics of price elasticity of e-book demand over time, and the other is to verify the difference of price elasticity between heavy users and non-heavy users.

We start to aggregate the orders by month, and the ISBN to calculate how many units of an e-book were sold for a month. Since the purpose is to observe the changing price elasticity over time, we consider one month as a time variable. In other words, the value of a time variable is 1 for the records aggregated in January 2011, 2 for the records aggregated in February 2011, and so on. Of course, the value is 13 for the records aggregated in January 2012, and the value is 24 for the records aggregated in December 2012. We expect that the larger the value of a time variable, the less price elastic the demand of e-books. The dependent variable in this empirical study is the sales volume of an e-book. The independent variables include the list price of an e-book, a time variable as explained above, and dummy variables on whether it has a corresponding paper book, whether it is only for adults, whether it is one of the best e-books in the previous month, whether it, in itself or its corresponding paper book, is one of the best books of 2011 or 2012, and whether it belongs to a certain genre. We also use the variables about publication date, registration date, and the number of pages in our analysis. Key variables are summarized in Table 2. Table 3 represents the descriptive statistics of sales variable and price variable. Table 4 shows the correlation matrix between variables.

Table 2. Description of Key Variables for Model 1[7]		
Type	Variable	Description
Dependent	$SALES_{it}$	Sales volume of the e-book i during the time t
Independent	$PRICE_i$	List price of the e-book i
	$TIME_t$	Number of months that has passed from the start of 2011
	SUB_i	Dummy variable for existence of the paper version of the e-book i
	$ADULT_i$	Dummy variable for whether the e-book i is only for adults
	B_PM_i	Dummy variable for whether the e-book i is one of the best e-books in time $t-1$
	B_11_i	Dummy variable for whether the e-book i, in itself or its corresponding paper book, is one of the best books of 2011
	B_12_i	Dummy variable for whether the e-book i, in itself or its corresponding paper book, is one of the best books of 2012
	$GENRE_{ij}$	Dummy variable for whether the e-book i, belongs to the genre j ($j = 0, 1, \cdots, 26$)[8]
	SNC_PUB_i	Number of months/years that has passed from the date when the paper version of the e-book i was published
	SNC_REG_i	Number of months/years that has passed from the date when the e-book i was published
	$PAGE_i$	Number of pages of the e-book i

Table 3. Descriptive Statistics of monthly *SALES* and *PRICE* for Model 1					
Variable	Observation	Mean	S.D.	Minimum	Maximum
SALES	296,489	5.53	26.25	1	1,800
PRICE	296,489	5,087.97	5,798.21	0	440,000

Table 4. Correlation[9] Matrix of Key Variables for Model 1											
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]
[1] *SALES*	1.00										
[2] *PRICE*	-0.01	1.00									
[3] *TIME*	-0.01	-0.01	1.00								
[4] *SUB*	-0.01	-0.03	-0.01	1.00							

---

[7] The definition and explanation of Model 1 and Model 2, as used in the titles of tables, are in the next section.

[8] Every e-book belongs to one of 27 genres classified by Kyobobook. The 27 genres are as follows: Education, family/living/cook, health/medical book, business/economics, language, summary books, comic books, magazine, society/politics/law, novel, dictionary/reference book, poetry/essay, children's book, travel/hobby, history/custom/mythology, books in serials, arts/culture, foreign book, kids' book, liberal arts, self-improvement, science/engineering, genre fiction, religion/spirituality, teens, computer/Internet, and unclassified book. This also applies to Table 5.

[9] Correlations of genre dummy variables are excluded in the table for brevity.

[5] *ADULT*	0.06	-0.03	0.01	0.00	1.00					
[6] *B_PM*	-0.01	0.00	-0.01	0.00	-0.00	1.00				
[7] *B_11*	0.08	0.02	0.01	0.00	-0.01	0.00	1.00			
[8] *B_12*	0.19	0.03	0.03	-0.02	-0.01	-0.00	0.23	1.00		
[9] *SNC_PUB*	-0.12	0.04	-0.32	0.08	-0.05	0.02	-0.01	-0.05	1.00		
[10] *SNC_REG*	-0.11	0.05	-0.44	0.02	-0.02	0.02	-0.01	-0.05	0.77	1.00	
[11] *PAGE*	0.02	0.32	-0.06	0.02	0.05	-0.00	0.02	0.01	0.15	0.16	1.00

We now prepare to test the hypothesis about the difference of price elasticity between heavy users and non-heavy users. According to Statistics Korea, from July 15, 2010 to July 14, 2011, the ratio of people aged 13 and above reading books is 61.8%, and the average number of books read by the same group is 20.8. Since our data contains purchase orders for about two years, we define heavy users as customers who have bought more than 42 e-books. Consumers who are not heavy users are naturally grouped as non-heavy users. We first divide the orders into two parts by customer type, and then aggregate the orders in each part by ISBN. The dependent variable is the sales volume of an e-book for the whole period. The independent variables are similar to the above case. Key variables are summarized in Table 5. Table 6 represents the descriptive statistics of sales variable and price variable. Table 7 and 8 show the correlations between variables for heavy users and non-heavy users, respectively.

Table 5. Description of Key Variables for Model 2		
Type	**Variable**	**Description**
Dependent	$SALES_i$	Sales volume of the e-book *i*
Independent	$PRICE_i$	List price of the e-book *i*
	SUB_i	Dummy variable for existence of the paper version of the e-book *i*
	$ADULT_i$	Dummy variable for whether the e-book *i* is only for adults
	B_11_i	Dummy variable for whether the e-book *i* is one of the best books of 2011
	B_12_i	Dummy variable for whether the e-book *i* is one of the best books of 2012
	$GENRE_{ij}$	Dummy variable for whether the e-book *i* belongs to the genre *j* $(j = 0, 1, 2, \cdots, 26)$
	SNC_PUB_i	Number of months/years that has passed from the date when the paper version of the e-book *i* was published
	SNC_REG_i	Number of months/years that has passed from the date when the e-book *i* was published
	$PAGE_i$	Number of pages of the e-book *i*

Table 6. Descriptive Statistics of monthly *SALES* and *PRICE* for Model 2						
	Variable	Observation	Mean	S.D.	Minimum	Maximum
For heavy users	*SALES*	35,995	19.25	60.97	1	1,504
	PRICE	35,995	4,988.62	5,988.88	0	240,000
For non-heavy users	*SALES*	45,820	20.34	124.52	1	9,986
	PRICE	45,820	5,696.48	9,648.66	0	440,000

Table 7. Correlation Matrix of Key Variables for Heavy Users for Model 2								
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
[1] *SALES*	1.000								
[2] *PRICE*	-0.069	1.000							
[3] *SUB*	-0.012	-0.045	1.000						
[4] *ADULT*	0.244	-0.023	0.002	1.000					
[5] *B_11*	0.042	0.021	0.001	-0.004	1.000				
[6] *B_12*	0.067	0.035	-0.021	-0.004	0.310	1.000			
[7] *SNC_PUB*	-0.128	0.064	0.072	-0.057	-0.004	-0.043	1.000		
[8] *SNC_REG*	-0.081	0.075	0.031	-0.028	-0.005	-0.042	0.760	1.000	
[9] *PAGE*	0.072	0.374	0.032	0.029	0.008	0.003	0.137	0.166	1.000

Table 8. Correlation Matrix of Key Variables for Non-Heavy Users for Model 2								
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
[1] *SALES*	1.000								
[2] *PRICE*	-0.005	1.000							
[3] *SUBS*	-0.013	-0.027	1.000						
[4] *ADULT*	0.037	-0.017	0.002	1.000					
[5] *B_11*	0.207	0.007	0.001	-0.003	1.000				
[6] *B_12*	0.290	0.013	-0.021	-0.003	0.332	1.000			
[7] *SNC_PUB*	-0.045	0.031	0.062	-0.055	-0.015	-0.045	1.000		
[8] *SNC_REG*	-0.020	0.040	0.030	-0.034	-0.018	-0.048	0.789	1.000	
[9] *PAGE*	0.025	0.275	0.028	0.030	0.006	0.002	0.169	0.165	1.000

Empirical Model and Results

The Log-Linear Demand Model

To estimate price elasticity over time econometrically, we employ the log-linear model in this empirical study. Since the definition of price elasticity is the percentage change in demand D because of a percentage change in price P, or $\eta = \delta D / \delta P \cdot P / D = \ln D / \ln P$, the power of *PRICE* in the log-linear model gives price elasticity. This model has been extensively used by researchers in studies of elasticity (Granados et al. 2012; Krishnamurthi & Raj 1991; Hughes et al. 2006). The log-linear specification is multiplicative as follows[10]:

$$SALES = e^{\alpha} \cdot PRICE^{\beta_1 + \beta_2 TIME + \beta_3 SUB} \cdot TIME^{\beta_4} \cdot SUB^{\beta_5} \cdot ADULT^{\beta_6} \cdot B_{PM}{}^{\beta_7} \cdot B_{11}{}^{\beta_8} \cdot B_{12}{}^{\beta_9}$$

$$\cdot \prod_j GENRE_j^{\sigma_j} \cdot e^{\varepsilon}$$

This model is used in line with Granados et al. (2008). They break the power of price into the base elasticity and its difference of channels. In this study, we break it into the base, its difference over time, and its difference from the existence of a substitute. The log transformation of the above equation is as follows:

$$\ln SALES = \alpha + \beta_1 \ln PRICE + \beta_2 TIME \ln PRICE + \beta_3 SUB \ln PRICE + \beta_4 \ln TIME + \beta_5 \ln SUB +$$
$$\beta_6 \ln ADULT + \beta_7 \ln B_{PM} + \beta_8 \ln B_{11} + \beta_9 \ln B_{12} + \sum_j \sigma_j \ln GENRE_j + \varepsilon$$

---

[10] For the dummy variable, *SUB*, we take 0 or 1 in the power of price term, and take 1 or e in the natural logarithm term to make this equation reasonable without loss of generality.

This is hereafter referred to as Model 1, which is used to verify the changing price elasticity over time.

The other model is similarly constructed to investigate the difference of price elasticity between heavy users and non-heavy users. We call it Model 2. It is as follows:

$$\ln SALES = \alpha + \beta_1 \ln PRICE + \beta_2 SUB \ln PRICE + \beta_3 \ln SUB + \beta_4 \ln ADULT + \beta_5 \ln B\_11 + \beta_6 \ln B\_12$$
$$+ \sum_j \sigma_j \ln GENRE_j + \varepsilon$$

The data to be used in this model are not aggregated by month, so this model does not contain any time variable like *TIME* and *B_PM*. We will run this demand model with a data set for heavy users, and repeat the same with a data set for non-heavy users. In this model, β_1 is the price elasticity of e-books that do not have a corresponding paper book, and the parameter β_2 represents the difference between the price elasticity of e-books with corresponding paper books and e-books without. In other words, the price elasticity of e-books with corresponding paper books is $\beta_1 + \beta_2$.

Since the models used in this study are multivariate regression models, multicollinearity among the independent variables should be checked (Greene 2008). We calculate the variance inflation factor, VIF, the indicator for multicollinearity. In our data set, the sizes of VIFs for all variables are less than 10. It implies that the models do not have the problems.

Endogeneity Validation

As we showed in the previous section, the demand of e-books are influenced by various factors including e-book prices, many observable e-book features, an unobservable e-book quality, and so on. However, the e-book price itself is affected by the unobservable e-book quality as well. For example, the popularity of the writer could influence both the price and demand of the e-book. Also, the policies or strategies that publishing companies used could have an effect on both of them. This brings the endogenous price problem, which means that the estimate of e-book price may be inconsistent and biased.

According to Berry et al. (1995), micro-data does not by itself solve the problem of unobserved product characteristics that are correlated with prices, but it allows one to introduce product-specific instruments to control for unobserved attributes. To address the potential endogeneity problem, we take a few instruments, and employ three-stage least squares (3SLS) approach which combines two-stage least squares (2SLS) and seemingly unrelated regression (SUR) methods. This approach is used to derive the parameters of the full system since endogenous variables in some equations of the model are used as explanatory ones in other equations (Kuruzovich et al. 2008). It is widely used for handling the endogeneity problem and contemporaneous cross-equation correlation between error terms.

In our model, 3SLS is operated as follows. The endogenous variable, *PRICE*, is regressed against the predetermined exogenous variables of the entire system in the first stage. The exogenous variables are *SUB*, *ADULT*, *SNC_PUB*, *SNC_REG*, and *PAGE* in this case. In the next stage, the entire system is estimated with the predicted value of *PRICE* by using the 2SLS mechanism. Then the variance-covariance matrix of the system is constructed from the residuals obtained from the second stage, and the parameters of the entire system are estimated by using feasible generalized least square (FGLS) in the third stage.

The *PRICE* model in this procedure is as follows.

$$\ln PRICE = \alpha + \beta_1 \ln SUB + \beta_2 \ln ADULT + \beta_3 SNC\_PUB + \beta_4 SNC\_REG + \beta_5 PAGE + \varepsilon$$

Results

Table 9 shows the results for the 3SLS approach in Model 1.

Table 9. Results of Model 1[11]		
	Coefficient	S.D.
ln*SALES*		
ln*PRICE*	-108.990***	7.423
*TIME·*ln*PRICE*	0.003***	0.000
*SUB·*ln*PRICE*	108.813***	7.417
ln*TIME*	-0.241***	0.026
ln*SUB*	-972.834***	66.229
ln*ADULT*	1.404***	0.038
ln*B_PM*	-0.534***	0.120
ln*B_11*	0.816***	0.079
ln*B_12*	4.110***	0.132
GENRE	(included)	
Constant	69.081***	4.491
ln*PRICE*		
ln*SUB*	0.588***	0.044
ln*ADULT*	-0.317***	0.010
SNC_PUB	0.008***	0.000
SNC_REG	-0.004***	0.000
PAGE	0.002***	0.000
Constant	7.045***	0.044
Observations (N)	296,489	
Note: *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively		

The coefficient of *TIME· lnPRICE* is 0.003, and this is highly significant. It implies that the price elasticity is -108.987 (=-108.990+0.003×1) in January 2011, and -108.918 (=-108.990+0.003×24) in December 2012 for e-books with no corresponding paper books[12]. In other words, the price elasticity increases by 0.003 each month. This means that the e-book demand is getting less price elastic over time. Therefore, H1 is accepted. This is consistent with other researchers' conclusions. Parker and Neelamegham (1997) suggest that the elasticity is generally highest during the earliest phases of the life cycle, and Simon (1979) also shows that the magnitude of price elasticity decreases in the introduction and growth stage, and reaches its minimum at the maturity stage. The result about H1 conforms to these previous studies. Similarly to H1, H2-a is also accepted because the coefficient of *SUB· lnPRICE* is significantly positive. Since the coefficients of *lnADULT*, *lnB_11*, and *lnB_12* have positive values and are all significant, H2-b and H2-c are also accepted. One unexpected finding is that the coefficient of *lnB_PM* is significantly negative. It indicates that e-books with higher sales in the previous month would sell less this month. Although this model cannot diagnose the exact reason, this may be because of the limited number of e-book consumers.

---

[11] The coefficient and standard deviation of dummy variables for genres are excluded in the table for brevity.

[12] Price elasticity in the Model 1 is $\beta_1 + \beta_2 \cdot TIME + \beta_3$. Since $\beta_3 = 0$ for e-books existing only in electronic form, the price elasticity in January 2011 (*TIME* = 1) is $\beta_1 + \beta_2 \times 1$, and the price elasticity in December 2012 (*TIME* = 24) is $\beta_1 + \beta_2 \times 24$.

Table 10 represents the results for 3SLS approach in Model 2.

Table 10. Results of Model 2				
	For heavy users		For non-heavy users	
	Coefficient	S.D.	Coefficient	S.D.
ln*SALES*				
ln*PRICE*	-461.434***	83.780	-729.345***	143.019
SUB·ln*PRICE*	460.678***	83.764	728.607***	143.007
ln*SUB*	-4375.97***	795.987	-7049.784***	1383.818
ln*ADULT*	2.690***	0.585	1.789***	0.766
ln*B_11*	0.037	1.175	0.128	1.444
ln*B_12*	7.652***	1.432	9.507***	1.825
GENRE	(included)		(included)	
Constant	184.134***	32.254	7056.323***	1383.924
ln*PRICE*				
ln*SUB*	2.473***	0.113	1.602***	0.124
ln*ADULT*	-0.221***	0.038	-0.349***	0.040
SNC_PUB	0.009***	0.001	0.006***	0.000
SNC_REG	-0.003***	0.001	-0.001*	0.001
PAGE	0.002***	0.000	0.002***	0.000
Constant	5.212***	0.113	6.142***	0.123
Observation (N)	35,995		45,820	
Note: *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively				

The results represent the difference of price elasticity between heavy users and non-heavy users. For the e-books with corresponding paper books, the price elasticity is -0.756 (=-461.434+460.678) for heavy users, and -0.738(=-729.345+728.607) for non-heavy users. On the contrary to what we expect in H3, the price elasticity of non-heavy users is less in absolute value than price elasticity of heavy users. However, in the case of the absence of the corresponding paper book, the heavy user demand is less price elastic than the non-heavy user demand as H3 explains. In other words, the price elasticity of heavy users is highly affected by the existence of the paper version. They are more price elastic in e-books than non-heavy users when paper book versions are available, but they are less price elastic when there is no substitute for the e-books. It could be interpreted as heavy users are more accustomed to reading paper books than e-books when both version are available. The reason behind this could be because the Korean e-book market is in the initial stage. It would be interesting to test H3 with data in the U.S. e-book market, which has entered the maturation phase.

In the results about Model 2, the coefficients of ln*B_11* are not statistically significant unlike the coefficients of ln*B_12*. It means that H2-c is partially supported in this model.

To sum up, the empirical results highly support H1, H2-a, and H2-b, partially support H2-c, and do not support H3.

Conclusion

This paper empirically investigated the changing price elasticity of e-book demand in Korea over time. We employed the log-linear demand model to identify price elasticity, and used massive industry sales data from January 1, 2011 to December 15, 2012 provided by Kyobobook, the largest book retailer in Korea. To solve the price endogeneity problem, 3SLS approach is operated. By analyzing the model and data, we found strong evidence of changes in price elasticity over time, which means that the demand of e-books is getting less price elastic as time goes on. This could be explained by the reference price effect. Also, we identified other factors that influence price elasticity and demand. The results of this study can be utilized to establish a pricing strategy for e-books. For example, publishers could set higher price for e-books having corresponding paper books than those without. This is because, according to our results, the existence of a corresponding paper book reduces consumers' price elasticity.

Although this research sheds light on the dynamics of price elasticity of e-book demand, it has some limitations. First, demographic factors were not used in this study, because we were not provided with such data due to the privacy agreement of Kyobobook. According to Hoch et al. (1995), consumer demographic variables such as race, education level, family size and so on are critical determinants of price elasticity of household item demand. We expect these factors to be also significant in our context. In addition, we did not take into account the channel that consumers used for their purchases. The data provided by Kyobobook did not contain any information on channels. Similar to the study on difference between channels by Granados et al. (2012), it would be an interesting topic to investigate the channel effect in the e-book industry. For further research, it would also be necessary to consider digital rights management (DRM). Since an e-book is a pure digital good, e-book sellers should have control over piracy issues. This may influence the consumers' price elasticity, and consequently affect the pricing strategy of e-books.

It is difficult but nevertheless important to determine an appropriate price in rapidly growing or changing industries like the e-book market. Since price elasticity directly indicates consumers' response to price change, researchers have also expressed a keen interest in the subject. In this paper, we empirically confirmed that the price elasticity of e-book demand has been slightly but significantly changing. Our implications on pricing strategies serve as valuable contributions to the academia, where there has been little empirical evidence despite the importance, as well as practical fields.

There are several worthwhile issues that we leave for further research. For example, it would be interesting to compare the price elasticities of paper books and e-books after the e-book market is mature enough such as the U.S. market. It might show the future of the book market, and even other industries where digital goods exists. Also, certain data limitations mentioned above would suggest expanding empirical research.

References

Berry, S., Levinsohn, J., and Pakes, A. 1995. "Automobile Prices in market Equilibrium," *The Econometric Society* (63:4), pp. 841-890.

Brynjolfsson, E., Hu, Y., and Smith, M.D. 2003. "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers," *Management Science* (49:11), pp. 1580-1596.

Chevalier, J., and Goolsbee, A. 2003. "Measuring Prices and Price Competition Online: Amazon.com and BarnesandNobles.com," *Quantitative Marketing and Economics* (1:2), pp. 203-222.

Edelman, B. 2010. "Markets: Red Light States: Who Buys Online Adult Entertainment?" *The Journal of Economic Perspectives* (23:1), pp. 209-220.

Fibich, G., Gavious A., and Lowengart, O. 2005. "The Dynamics of Price Elasticity of Demand in the Presence of Reference Price Effects," *Journal of the Academy of Marketing Science* (33:1), pp. 66-78.

Ghose, A., Smith, M.D., and Telang, R. 2006. "Internet Exchanges for Used Books: An Empirical Analysis of Product Cannibalization and Welfare Impact," *Information Systems Research* (17:1), pp. 3-19.

Granados, N., Gupta, A., and Kauffman, R.J. 2008. "Designing Online Selling Mechanisms: Transparency Levels and Prices," *Decision Support Systems* (45:4), pp. 729-745.

Granados, N., Gupta, A., and Kauffman, R.J. 2012. "Online and Offline Demand and Price Elasticities: Evidence from the Air Travel Industry," *Information Systems Research* (23:1), pp. 164-181.

Greene, W. H. (2008). Econometric Analysis, 6th ed. Prentice-Hall: NJ.

Hoch, S.J., Kim, B., Montgomery, A.L., and Rossi, P.E. 1995. "Determinants of Store-Level Price Elasticity," *Journal of Marketing Research* (32:1), pp. 17-29.

Hughes, J.E., Knittel, C.R., and Sperling D. 2008. "Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand," *The Energy Journal* (29:1), pp. 113-134.

Jiang, Y., and Katsamakas, E. 2010. "Impact of E-book Technology: Ownership and Market Asymmetries in Digital Transformation," *Electronic Commerce Research and Applications* (9:5), pp. 386-399.

Kalyanaram, G., and Winer, R.S. 1995. "Empirical Generalizations from Reference Price Research," *Marketing Science* (14:3), pp. 161-169.

Kim, B., and Rossi, P.E. 1994. "Purchase Frequency, Sample Selection, and Price Sensitivity: The Heavy-User Bias," *Marketing Letter* (5:1), pp. 57-67.

Krishnamurthi, L., and Raj, S.P. 1991. "An Empirical Analysis of the Relationship between Brand Loyalty and Consumer Price Elasticity," *Marketing Science* (10:2), pp. 172-183.

Kuruzovich, J., Viswanathna, S., Agarwal, R., Gosain, S., and Weitzman, S. 2008. "Marketspace or Marketplace? Online Information Search and Channel Outcomes in Auto Retailing," *Information Systems Research* (19:2), pp. 182-201.

Mixon, F.G., and Ressler R.W. 2000. "A Note on Elasticity and Price Dispersions in the Music Recording Industry," *Review of Industrial Organization* (17:4), pp. 465-470.

Oestreicher-Singer, G., and Sundararaja, A. 2010. "Are Digital Rights Valuable? Theory and Evidence from eBook Pricing," *CeDER Working Paper* No. 06-01.

Parket, P.M., and Neelamegham, R. 1997. "Price Elasticity Dynamics over the Product Life Cycle: A Study of Consumer Durables," *Marketing Letters* (8:2), pp. 205-216.

Simon, H. 1979. "Dynamics of Price Elasticity and Brand Life Cycles: An Empirical Study," *Journal of Marketing Research* (16:4), pp. 439-452.

Smith, G.E., and Nagle, T.T. 1995. "Frames of Reference and Buyers' Perception of Price and Value," *California Management Review* (38:1), pp. 98-116.

The Economist. April 3, 2010. E-publish or perish. 65.

Vasileiou, M., Hartley, R., and Rowley, J. 2009. "An Overview of the E-book Marketplace," *Online Information Review* (33:1), pp. 990-1029.

Winer, R.S. 1986. "A Reference Price Model of Brand Choice for Frequently Purchased Products," *Journal of Consumer Research* (13:2), pp. 250-256.

Wondracek, G., Holz, T., Platzer, C., Kirda, E., and Kruegel, C. 2010. "Is the Internet for Porn? An Insight into the Online Adult Industry," *In Proceeding of 9th WEIS*.