ARE DIGITAL RIGHTS VALUABLE? THEORY AND EVIDENCE FROM THE EBOOK INDUSTRY

Gal Oestreicher-Singer and Arun Sundararajan
Leonard N. Stern School of Business
New York University
New York, NY U.S.A.
goestrei@stern.nyu.edu **asundara@stern.nyu.edu**

Abstract

The effective management of digital rights is a crucial challenge in many industries making the transition from physical to digital products. We present an economic model that characterizes the value of digital rights when products are sold both embedded in tangible physical artifacts and as pure digital goods, and when granting digital rights may also affect the extent of digital piracy. Our model indicates that in the absence of piracy, digital rights should be unrestricted, since a seller can use their pricing strategy to optimally balance sales between physical and digital goods. However, the threat of piracy limits the extent to which digital rights should be granted: the value of digital rights is determined not only by their direct effect on the quality of legal digital goods, but by their effect on the differential quality of legal and pirated digital goods. When the latter effect is negative, granting digital rights may have a detrimental effect on value; our model indicates that this kind of effect is more likely to be observed for digital rights that aim to replicate the consumption experience of physical goods, rather than enhancing a customer's digital experience. We test the predictions of our analytical model using data from the ebook industry. Our empirical evidence supports our theoretical results, showing that four separate digital rights each have a significant impact on ebook prices, and establishing that those two that are most strongly associated with digital piracy have a negative impact on seller value. We also show, both analytically and empirically, that sellers should increase the prices of digital goods as the prices of their physical counterparts increase, but decrease them as the technological sophistication of their potential customers increases. Our results represent new evidence of the importance of an informed and judicious choice of the different digital rights permitted by one's DRM platform, and provide a framework for guiding managers in industries that are progressively being "digitized."

Keywords: Digital rights management, DRM, piracy, digital piracy, intellectual property, IP, copyright, channel conflict, cannibalization, ebook, hedonic model, hedonic price index

Introduction

This paper studies how granting digital rights influences the quality of legal and pirated digital goods, and how this affects their value to sellers accustomed to delivering these goods as physical artifacts. We present a model that links the demand for physical, digital, and pirated goods to pricing and rights management choices, and we subsequently validate the predictions of this model using data from the ebook industry.

The specific questions addressed by this paper are better understood in the broader context of the ongoing digitization of industries whose products (music, video, graphic art, magazines, newspapers, books) used to be sold embedded in tangible artifacts, but are now increasingly available as pure *digital goods*. These digital goods have a number of attractive properties for sellers. They can be produced at a small fraction of the cost of producing their physical counterparts. They can be delivered easily over the Internet, often allowing sellers to bypass costly intermediaries that limit the geographic reach of their sales. They may include electronic features that enhance their users' rendering experience. Many issues related to the economics, pricing, and delivery of these digital goods have been recognized and analyzed in the IS literature (although not the specific ones we study).

The process of digitization in these industries has been gradual. Currently, most sales continue to be in the form of tangible artifacts, and digital goods are seen as imperfect substitutes for their physical counterparts. Additionally, a target market that is more technology savvy is likely to view digital products as being of higher value, and possibly as better substitutes for the existing physical products they are accustomed to using; therefore, the technological sophistication of potential customers affects the perceived difference in value. Moreover, digital goods are subject to piracy, which is widespread in many industries, and which affects demand for both nascent digital goods, as well as the corresponding established physical goods (music being the most visible recent example, but not the only one).

The uncertain and potentially incomplete process of transition toward digitizing products presents sellers with difficult technological and business decisions. These relate most crucially to the strategic control of the quality of digital goods and the extent to which their presence leads to piracy, both of which are currently implemented through a seller's management of the *digital rights* associated with their digital goods. The technological challenges of managing digital rights are gradually being addressed by the emergence of viable industry-specific platforms for digital rights management (DRM). Technological robustness aside, it is also imperative that managers grant the different digital rights permitted by their DRM platforms in a judicious and informed way. This is because an insufficient level of digital rights can result in premature failure of a digital initiative, as illustrated by early e-music services. On the other hand, granting digital rights that are too extensive can lead to cannibalization of established sales of physical goods. Similar issues have been recognized in the context of channel conflict caused by Web-based channels for, among other businesses, online retailing (Viswanathan 2000) and online brokerage (Konana et al. 2000).

Moreover, granting certain digital rights may increase the availability of pirated substitutes. For instance, the right to download (rather than stream) digital audio files increases the desirability of purchasing legal digital music, while simultaneously increasing the threat of digital piracy over file-sharing networks. The right to print a digital book encourages ebook adoption by customers used to reading printed pages, but enables the creation of pirated PDF copies of the book. The right to a backup of a digital movie on a DVD adds value to legal downloaders, but facilitates illegal secondary sales. Since different digital rights contribute differentially toward increasing the quality of legal digital goods, and toward facilitating piracy, a careful assessment of their relative economic value is bound to be important.

To summarize, the appropriate choice of digital rights is influenced by many technological and business factors, which include

- The relative extent to which each digital right a seller may grant buyers contributes toward an increase in quality of the digital good and the corresponding increase in the prevalence of piracy. The latter is also influenced by the strength of the industry's DRM platform.

- The extent to which these increases affect the current revenues generated by the sale of physical goods, and the anticipated revenues from sales of digital goods.

- The technological sophistication of potential customers, which determines how much they value digital goods (both legal and pirated) and how closely they assess digital goods as substitutes for their tangible counterparts.

Each of these factors is explicitly captured in our model. Specifically, we model sellers who choose digital rights and prices for their digital goods, while simultaneously selling the corresponding physical goods at prespecified price levels. The extent to which each digital right is granted affects the quality of the digital good, and may affect the quality of a (free) pirated substitute. Customers vary in how much they value increases in quality induced by this granting of digital rights. Additionally, they vary in their tastes for physical and digital goods, captured by the difference in value they ascribe to a digital good and physical good of equal quality. This difference varies across customers, and the extent of variation (or heterogeneity) is an inverse measure of the technological sophistication of potential customers in an industry. We characterize the appropriate pricing and digital rights choices of a seller, both in the absence and presence of piracy. The results of our analytical model indicate, among other things, that

1. The price of a digital good is increasing in the price of its corresponding physical counterpart, and decreasing in the extent of technological sophistication of the target customer base.

2. In the absence of a threat of piracy, the price of a digital good is increasing in the level of each associated digital right, and a seller should always choose to grant the highest level of rights permitted by its DRM platform, despite a potential threat of cannibalization.

3. If granting digital rights leads to piracy, the profitability of these rights to a seller depends both on their *direct effect* on the quality of the legal digital version, and (more crucially) on their *differential effect* on the quality of the legal and the pirated versions. This quality difference interacts with the price of the physical good and the technological sophistication of

customers in its impact on value; a seller may, therefore, wish to restrict certain digital rights even if granting them increases the value of its digital good.

Our analytical results suggest empirical equations that relate the prices of digital goods to the prices of their physical counterparts, the extent to which digital rights are granted to customers, and the technological sophistication of the product's target customer base. We test these results by estimating two sets of empirical equations, using a data set comprising of 1,788 ebooks across three categories sold by a specialized ebook retailer. The first is a hedonic pricing index that relates ebook price to the level of digital rights granted by each ebook. This model is estimated on the entire data set, and indicates the magnitude of the value each digital right provides an ebook seller through its effect on ebook prices. The second model includes the price of the physical book and the technological sophistication of the customer base for different categories of books, and is estimated on a pilot set of 300 randomly selected ebooks for which we have also collected physical book price data. Both models strongly support the predictions of our theoretical model, showing that physical book prices, technological sophistication, and each of four digital rights have a significant impact on seller value. More importantly, our results indicate 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, while other digital rights result in a net increase in a seller's value.

Our paper adds to a growing literature that aims to guide managers facing piracy and rights management challenges in digital industries. Salient work on this topic has studied the relationship between piracy deterrence and network effects (Connor and Rummelt 1991), the relative welfare benefits of legal and technological deterrence (Png and Chen 2003), controlling piracy through strategic pricing and rights management (Sundararajan 2004) or vertical differentiation (Snir 2003, Wu et al. 2003), and the relationship between piracy enforcement incentives and domestic software production (Gopal and Sanders 1998). Space constraints preclude describing a more detailed survey, which indicates that the relationship between rights management, piracy, and the strategic control of quality has not been studied empirically.

We have organized the rest of this paper as follows. The next section presents our analytical economic model which characterizes demand for the seller's goods and the sales lost to piracy, and derives the relationship between the prices of digital goods, their digital rights and the other factors described above, both in the absence of piracy and in its presence. The predictions of this model are tested in the subsequent section, which describes our data set, presents our empirical models, and describes the results of their estimation. The final section concludes with a summary of our results, some limitations of our study, and our ongoing work in addressing these limitations.

Choosing Digital Rights and Pricing Strategy: An Analytical Model

This section briefly describes our economic model, and presents the analytical results and predictions that form the basis for our subsequent empirical study.

Overview

Our model is of a monopoly seller who produces two versions of a product: a physical (tangible) good and a digital good. The tangible good is of constant and exogenous quality s_T. The quality s_D of the digital good is determined by the level to which the seller grants different digital rights (r_1,r_2,…r_n) to buyers of the digital good:

$$s_D = f_D(r_1, r_2, r_3, ..., r_n).$$

Examples of *digital rights* for different digital products include the right to print an ebook, the right to backup a downloaded video file, the right to play an MP3 file on a portable device, and the right to lend an ebook. The range of feasible values for each digital right r_i is determined by an exogenously specified DRM platform used by the seller, and each right, therefore, has prespecified minimum and maximum levels. For instance, under Adobe's DRM platform for ebooks, a seller can vary the right a buyer has to print pages of an ebook by allowing unlimited printing (the highest level), allowing no printing whatsoever (the lowest level), or specifying a fixed number of pages that can be printed during each time period (an intermediate level; for example, up to 35 pages every two weeks). Granting a buyer a higher level of digital rights increases her value from the digital good and, therefore, f_D is (weakly) increasing in each of its arguments. Altering the level of digital rights within the constraints of its DRM platform imposes no direct cost on the seller and, consequently, varying the quality s_D of the digital good has no associated variable cost. The physical good has a constant variable cost c per unit sold; the digital good has a constant variable cost of zero.

In addition to the legal physical and digital versions, there may be a *pirated* digital version of the good available. The quality s_P of the pirated good is also determined by the level to which the seller grants different digital rights (r_1,r_2,…r_n) to buyers of the digital good:

$$s_P = f_P(r_1, r_2, r_3, ..., r_n)$$

Since an increase in the level of rights granted to legal users often facilitates the creation of higher-quality pirated versions, f_P is (weakly) increasing in each of its arguments. The price of the tangible good is denoted p_T, and the price of the digital good is denoted p_D; the pirated good is free.

Customers are heterogeneous along two dimensions.[1] The first dimension, indexed by a *type* variable $\theta \in [0,1]$, represents their preferences for *digital quality*, or equivalently, how much a customer values her digital rights. Each customer of type θ prefers higher quality to lower quality; however, all else being equal, a customer with a higher θ always values a digital good of fixed quality more than a customer with a lower θ. The second dimension, indexed by a *location* variable $y \in [0,1]$ represents the customer's *taste*, or more specifically, her relative preferences for tangible versus digital goods. Holding quality constant, and for a fixed value of θ, a customer indexed by a higher value of y places a higher value on the digital good and a lower value on the physical good than a customer indexed by a lower value of y, or customers increasingly prefer the digital good to the tangible good as y increases. The indexes θ and y are independently and uniformly distributed, and this assumption is made for analytical simplicity. A generalization is discussed briefly in the final section of this paper.

For clarity, we subsequently refer to the physical version as the *tangible good* (whose associated variables have the subscript *T*), the legal digital version as the *digital good* (whose associated variables have the subscript *D*), and the pirated digital version as the *pirated good* (whose associated variables have the subscript *P*). Customer preferences for the tangible good, the digital good, and the pirated good are modeled using the following utility functions:

Tangible good:	$s_T - ty - p_T$
Digital good:	$\theta s_D - t(1-y) - p_D$
Pirated good:	$\theta s_P - t(1-y)$

This simply means, for instance, that the value a customer of type θ and location y gets from buying a digital good of quality s_D at a price p_D is $[\theta s_D - t(1-y) - p_D]$. The parameter t indicates the diversity of tastes across potential customers. Since the range of y is normalized to [0,1], a higher value of t indicates that, on average, customers *vary more* in their relative preferences for the tangible and digital goods. We, therefore, use this as a (inverse) measure of the *technological sophistication* of the target customer set. All customers are accustomed to using the physical good; the more technologically sophisticated they are on average, the less likely they are to view the digital good as being "different" on average, and this corresponds to t having a lower value.[2] Customers use exactly one version of the good, purchasing either the tangible good or the digital good, or using the free pirated good.

[1]For those familiar with models of imperfect competition, our model includes aspects of both horizontal and vertical differentiation. The tangible good is horizontally differentiated from both the digital and pirated goods, and the digital good is vertically differentiated from the pirated good.

[2]A complementary interpretation of t is as a product-category specific measure of how much the digital good can capture the attributes of the physical artifact that embodied the good in the past. This is a valid interpretation as well; however, it is not pertinent to our paper, since our empirical model compares products of a specific kind (books), rather than different product categories.

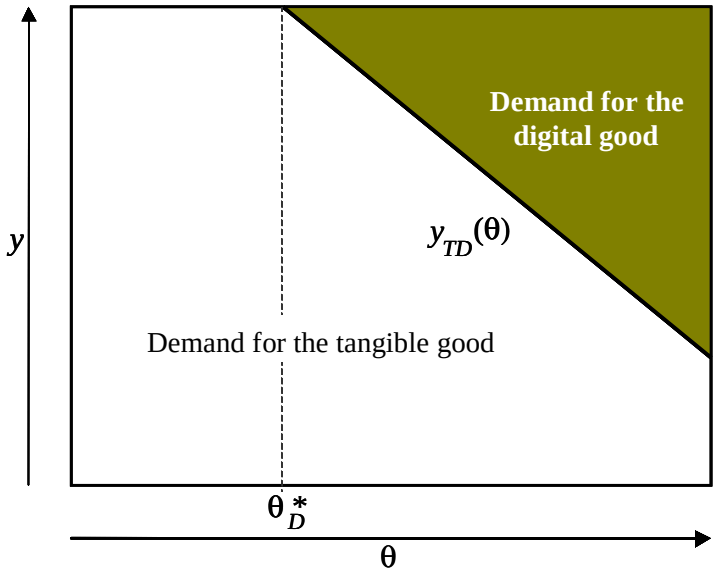


Figure 1. Summary of Demand Across Different Customer Types and Locations
Each customer (θ, y) in the shaded green region prefers the digital good to the tangible good. Since θ and y are both uniformly distributed, demand for each good is simply the area of the corresponding region.

Choosing Digital Rights in the Absence of Piracy

Our first set of results specifies the optimal choice of digital rights and pricing in the *absence* of piracy. Therefore, in this subsection,

$$f_P(r_1, r_2, r_3, ..., r_n) = 0,$$

and customers choose between the tangible good and the digital good. The seller chooses the price p_D and the quality s_D of the digital good, given fixed levels of s_T and p_T. A choice of s_D indirectly implies a corresponding choice of digital rights $(r_1, r_2, \ldots r_n)$, one of perhaps many combinations of rights that yield the same quality level s_D.

Given the prices p_T and p_D and quality levels s_T and s_D, demand is characterized as follows. For the set of customers of type θ, let the location of the customer indifferent between the tangible good and the digital good be $y_{TD}(\theta)$. This implies that

$$s_T - ty_{TD}(\theta) - p_T = \theta s_D - t[1 - y_{TD}(\theta)] - p_D, \tag{1}$$

and consequently

$$y_{TD}(\theta) = \frac{(s_T - \theta s_D) - (p_T - p_D)}{2t} + \frac{1}{2}. \tag{2}$$

Therefore, a customer indexed by (θ, y) prefers the tangible good to the digital good if $y < y_{TD}(\theta)$, and prefers the digital good if $y > y_{TD}(\theta)$. If p_T and s_T are specified to ensure that all target customers buy one or the other (that is, if all customers would purchase the tangible good if it was the only version available), then the demand realized is as summarized in Figure 1. The demand for each good (tangible, digital) is the area of its corresponding portion. The value θ_D^* is the lowest type of customer who might purchase the digital good: that is, $y_{TD}(\theta_D^*) = 1$.

Our first proposition describes the seller's optimal choice of price for the digital good at a quality level s_D implied by a fixed choice of digital rights.

Proposition 1. *The optimal price p_D of the digital good is specified by*

$$p_D = p_T + \frac{s_D - s_T + t}{3} - \frac{2c}{3}.$$

Therefore, the price of the digital good is

(1) increasing in the price p_T of the tangible good
(2) increasing in the quality of the digital good s_D and, consequently, increasing in the level of each digital right r_i granted to its buyers
(3) decreasing in the level of technological sophistication of its potential customers (since it is increasing in t)

Based on the choice of price p_D implied by Proposition 1 for each level of quality s_D, the optimal choice of digital rights is described in Proposition 2.

Proposition 2. *In the absence of piracy, the seller should grant buyers the maximum level of digital rights permitted by its DRM platform.*

Mathematically, Proposition 2 follows from the fact that profits as a function of s_D are strictly convex in s_D, and this function is maximized at its upper endpoint. Intuitively, the seller chooses the highest level of quality because any threat of cannibalization can be controlled equally effectively at this level of quality simply by varying pricing appropriately, and the highest level of quality provides the seller with maximum pricing power.

Pricing and Digital Rights in the Presence of Piracy

We now return to the complete model described earlier in the overview. Granting a level (r_1,r_2,…r_n) of digital rights to buyers of the digital good also results in the creation of a free pirated good of quality $s_D = f_D(r_1, r_2, r_3, ..., r_n) \geq 0$. Buyers choose between three substitutable versions—the tangible good, the digital good and the pirated good. To characterize demand completely, we need to consider the three pair-wise comparisons made by potential customers. First, for any customer location *y*, if the customer type indifferent between the pirated good and the digital good is θ^*_{PD}, then

$$\theta^*_{PD} s_P - t(1-y) = \theta^*_{PD} s_D - t(1-y) - p_D, \tag{3}$$

and therefore

$$\theta^*_{PD} = \frac{p_D}{s_D - s_P}. \tag{4}$$

Next, for the set of customers of type θ, let the location of the customer indifferent between the tangible good and the pirated good be $y_{TP}(\theta)$. This implies that

$$s_T - t y_{TP}(\theta) - p_T = \theta s_P - t[1 - y_{TP}(\theta)], \tag{5}$$

and consequently

$$y_{TP}(\theta) = \frac{(s_T - \theta s_P) - p_T}{2t} + \frac{1}{2}. \tag{6}$$

Therefore, a customer indexed by (θ, y) prefers the tangible good to the pirated good if $y < y_{TP}(\theta)$, and prefers the pirated good if $y > y_{TP}(\theta)$. Finally, for the set of customers of type θ, the location of the customer indifferent between the tangible good and the digital good is $y_{TD}(\theta)$ as defined in equation (2).

Equations (2) and (6) determine the lowest customer types θ_P^* and θ_D^* who might choose the pirated and digital goods respectively. Since the slope of $y_{TP}(\theta)$ is higher (less negative) than the slope of $y_{TD}(\theta)$, it follows that if $\theta_P^* > \theta_D^*$, there is no demand for the pirated good. Similarly, if $\theta_{PD}^* \geq 1$, there is no demand for the pirated good. In both of these cases, the results of the previous subsection describe the seller's demand.

Furthermore, it follows from (2), (4) and (6) that

$$\theta_P^* < \theta_D^* \Rightarrow \theta_D^* < \theta_{PD}^*,$$

and since θ_{PD}^* is defined as the type indifferent between the pirated good and digital,

$$y_{TP}(\theta_{PD}^*) = y_{TD}(\theta_{PD}^*).$$

The demand for the tangible good and digital good, and the sales lost to piracy can therefore be completely characterized now, and this characterization is summarized in Figure 2. The demand for each good (tangible, digital, pirated) is the area of its corresponding portion.

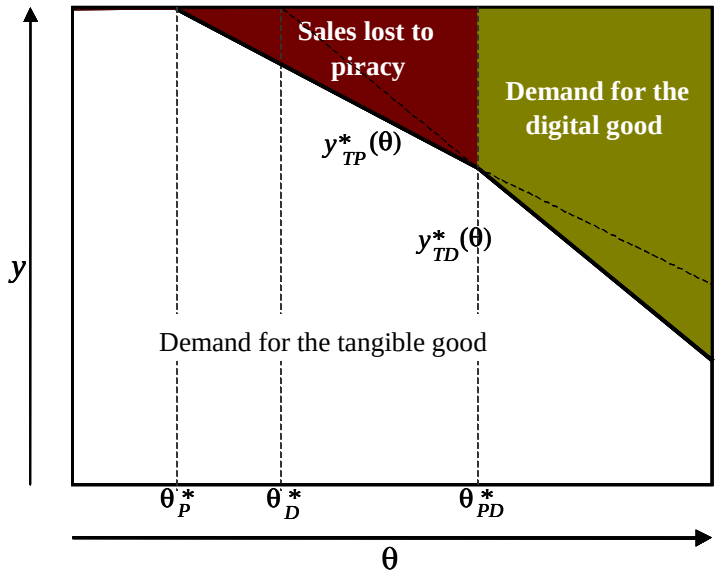


Figure 2. Summary of Demand Across Different Customer Types and Locations

Each customer (θ, y) in the shaded green (lighter) region prefers the digital good to the other two, and each customer in the shaded red (darker) region prefers the pirated version. In the figure

$$\theta_P^* = \frac{s_T - p_T - t}{s_P}, \theta_D^* = \frac{s_T - p_T + p_D - t}{s_D}, \text{ and } \theta_{PD}^* = \frac{p_D}{s_D - s_P}$$

Our next proposition describes the seller's optimal choice of price for the digital good at quality level s_D and s_P implied by a fixed choice of digital rights. The exact expression for p_D is available on request. Based on a simple approximation of this expression,

Proposition 3. *The seller's optimal price p_D of the digital good is*

(1) increasing in the price p_T of the tangible good
(2) increasing in the quality of the digital good s_D
(2) increasing in the difference in quality $(s_D - s_P)$ between the digital good and the pirated good
(4) decreasing in the level of technological sophistication of its potential customers

The expression for p_D depends directly on the quality s_D of the digital good (the *direct effect* of increasing quality that increases price), as well as on the difference $(s_D - s_P)$ between the quality of the digital and pirated good (the *differential effect* of increasing relative quality). An equal increase in s_D and s_P does not change the differential effect, but has a positive direct effect, and increases p_D. Therefore, at a level of r_i for which

$$\frac{df_D}{dr_i} = \frac{df_P}{dr_i},$$

an increase in the digital right r_i increases the seller's pricing power, and increases the price of the digital good. In contrast, at a level of r_i for which $\frac{df_P}{dr_i}$ is sufficiently higher than $\frac{df_D}{dr_i}$, the negative differential effect of a decrease in $(s_D - s_P)$ will dominate the positive direct effect of an increase in s_D, and increasing r_i will reduce a seller's pricing power.

Note that in Figure 2, those customers who purchase the digital good are in the right *upper* portion of the customer space: they place a high value on digital rights, but also have a strong preference for digital goods over physical goods. This suggests that digital rights that enhance the *digital experience* of the good are likely to have a more positive impact on seller value than those digital rights that allow buyers of digital goods to *replicate* the consumption experience they associate with a *physical good*. For example, the right to play one's purchased digital music files on multiple authorized rendering devices (enhancing the digital experience) is likely to have a higher positive effect on pricing than the right to burn these music files on a CD (replicating the physical good experience). Similarly, the ability to copy portions of the text of an ebook and use them in other electronic documents is a right tailored toward improving the digital experience, and is appropriate for those target customers who buy the ebook; on the other hand, the right to print pages of an ebook permits an inferior replication of the physical book experience, which these customers do not value especially highly in the first place.

Of course, the effect that an increase in each desirable digital right has on the quality of the pirated good is critical (as discussed after Proposition 3). We return to this issue in the next section.

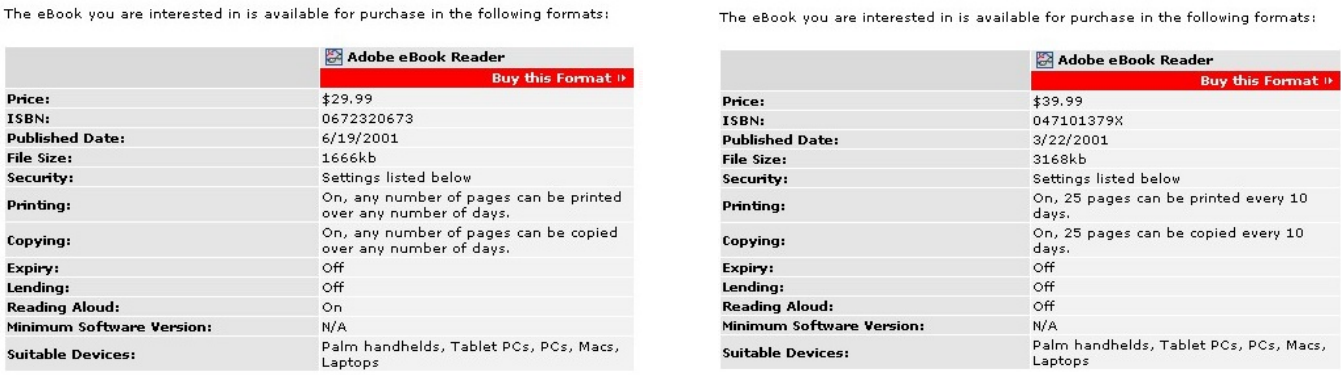
How Valuable Are Digital Rights? Evidence from the eBook Industry

This section describes our data, our empirical models, and the results of their estimation, and discusses the relationship of these results to the predictions of our analytical model.

Data

Over the last year, we have collected the prices and digital rights of over 10,000 ebooks sold by a specialty Web-based ebook retailer. The results presented are based on our data set from April 2004. The data collection was done by comprehensively mining the site of this seller, using a Perl-based spider written by one of the paper's authors specifically for this project.

The books sold by this retailer are categorized based on subject. We have focused our analysis on three categories that we believed *ex ante* would have target customer sets with different levels of technological sophistication: *computers*, *philosophy*, and *science fiction*. The smallest of these categories had 596 ebooks: to ensure a data set that was balanced across categories, we chose a random sample of 596 data points from the other two categories.



The eBook you are interested in is available for purchase in the following formats:

	Adobe eBook Reader
	Buy this Format
Price:	$29.99
ISBN:	0672320673
Published Date:	6/19/2001
File Size:	1666kb
Security:	Settings listed below
Printing:	On, any number of pages can be printed over any number of days.
Copying:	On, any number of pages can be copied over any number of days.
Expiry:	Off
Lending:	Off
Reading Aloud:	On
Minimum Software Version:	N/A
Suitable Devices:	Palm handhelds, Tablet PCs, PCs, Macs, Laptops

The eBook you are interested in is available for purchase in the following formats:

	Adobe eBook Reader
	Buy this Format
Price:	$39.99
ISBN:	047101379X
Published Date:	3/22/2001
File Size:	3168kb
Security:	Settings listed below
Printing:	On, 25 pages can be printed every 10 days.
Copying:	On, 25 pages can be copied every 10 days.
Expiry:	Off
Lending:	Off
Reading Aloud:	Off
Minimum Software Version:	N/A
Suitable Devices:	Palm handhelds, Tablet PCs, PCs, Macs, Laptops

Figure 3. Illustrates Digital Rights Granted for Two eBooks in Our Data Set

Under our definition of variables associated with digital rights, the first ebook would have *PrintAll = 1, CopyAll = 1, Read = 1* and *PrintPartial = 0, CopyPartial = 0, Lend = 0.* The second ebook would have *PrintPartial = 1, CopyPartial = 1,* and *PrintAll = 0, CopyAll = 0, Read = 0, Lend = 0.*

These ebooks are offered in up to four formats (Adobe eBook, Microsoft Reader, Palm Reader, and MobiPocket Reader). Every ebook is available in the Adobe eBook format, which was also the format whose DRM platform offers the widest range of digital rights; therefore, we choose to focus on this format. In addition to its price and category, each ebook has five digital rights associated with it: printing, copying, expiry, lending and reading aloud. None of the books in our data set have restricted the right expiry. The rights lending and reading aloud have binary settings (on or off). The rights printing and copying have three kinds of settings: unrestricted (print as many pages as one wants as frequently as needed, copy as much text of the book as one wants), none (no printing allowed, no copying allowed), and partial. There is a range of different levels of rights granted under partial, along two dimensions: the number of pages, and the frequency (for example, copy up to 25 pages every 10 days, print up to 35 pages every 7 days). The digital rights for two ebooks as depicted on the retailer's site are displayed in Figure 3.

We have not comprehensively analyzed all the details associated with partial rights across our data set. However, all the books we examined that had partial print or copying rights specified it based on *both* a fixed set of pages and a fixed frequency. For simplicity, therefore, we treat all books with partial printing rights as having the same printing rights (*PrintPartial = 1*), and all books with partial copying rights as having the same copying rights[3] (*CopyPartial = 1*). We have also collected the tangible book prices for a small random sample of ebooks; this data is used in the analysis of the (later) subsection on tangible book prices, and is described there. The variables associated with each of our data points are summarized on Table 1, along with some descriptive statistics of the data set.

Table 1. Summary of Variables and Some Descriptive Statistics

Variable	Brief Description	Data Points	Range	Mean	SD
EPrice	Ebook price	1788	1.49 - 214.5	17.23	22.11
Log[*EPrice*]	Log of ebook price	1788	0.39 - 5.36	2.393	0.899
Variable	**Brief Description**	**Data Points**	**Range**	**% Non-Zero**	
CopyAll	Allows unlimited copying of text	1788	{0,1}	9%	
CopyPartial	Allows limited copying of text	1788	{0,1}	24%	
PrintAll	Allows unlimited printing of pages	1788	{0,1}	6%	
PrintPartial	Allows limited printing of pages	1788	{0,1}	7%	
Lend	Allows lending	1788	{0,1}	20%	
Read	Has the read aloud feature	1788	{0,1}	57%	
CatSciFi	Science fiction category	1788	{0,1}	33%	
CatPhilosophy	Philosophy category	1788	{0,1}	33%	

[3]Therefore, segmenting into "restricted by # of pages" and "restricted by time" did not work.

A Simple Hedonic Price Index for Digital Rights

Our first model estimates the value of each digital right to an ebook seller using a simple hedonic pricing equation:

$$\log[EPrice] = \alpha + \beta_1 CopyAll + \beta_2 CopyPartial + \beta_3 PrintAll + \beta_4 PrintPartial + \beta_5 Lend + \beta_6 Read$$

Our analytical results from the previous section predict that each digital right 1 has a positive direct effect on digital good quality and a positive differential effect (or one for which the positive direct effect dominates the negative differential effect) will have a positive associated β_i. Correspondingly, those rights for which the negative differential effect dominates the positive direct effect (that is, those digital rights which are not assessed as being especially valuable by customers of the digital good, but which substantially increase the quality and availability of pirated substitutes) will have a negative β_i.

The results of this estimation are presented in Table 2. The signs of the coefficients are quite striking. They indicate that granting two of the four digital rights—copying and reading aloud—result in significant *increases* in ebook prices. However, granting the other two rights—printing and lending—result in significant (and substantial) *decreases* in prices. Moreover, this direction of change persists as one increases the level of printing and copying rights granted to buyers.[4]

In the context of our analytical model's predictions, these results have a straightforward interpretation. The negative effect of increasing printing rights on prices indicates that the right to print ebooks increases the value of pirated substitutes for books substantially more than it increases the value of the digital good. As discussed in earlier, given that the set of customers who purchase ebooks are those whose tastes favor digital goods over physical goods, the ability to create printed copies of their ebooks is unlikely to have much of a positive direct effect on ebook quality. On the other hand, the *PrintAll* right facilitates the creation of near-perfect copies of many ebooks (although Adobe's own PDF distiller does not permit this, other free PDF-creation software like Win2PDF allow one to print an entire ebook through their distiller and create an unprotected PDF file which is almost iden-

Table 2. The Effect of Digital Rights on eBook Prices

Variable	Coefficient (β_i)	Corresponding value of e^{β_i}
Constant	2.314 *** (0.031)	10.10
CopyAll	0.560 *** (0.201)	1.75
CopyPartial	0.219 ** (0.101)	1.24
PrintAll	–0.636 *** (0.188)	0.53
PrintPartial	–0.208 * (0.115)	0.81
Lend	–1.11 *** (0.105)	0.33
Read	0.461 *** (0.048)	1.59

R^2 = 14.1%, F = 49.8

**significant with p < 0.1*
***significant with p < 0.05*
****significant with p < 0.01*

[4]Since we estimate a semilog equation, the value of e^{β_i} is what has economic significance. For instance, when *CopyPartial = 1,* the value of *EBook* changes by a (multiplicative) factor of $e^{\beta_2} = 1.24,$ or according to our estimated price equation, ebook prices are 24% higher. When *PrintPartial = 1,* the value of *EBook* changes by a (multiplicative) factor of $e^{\beta_4} = 0.81,$ or according to the estimated price equation, ebook prices are 19 percent lower. We do not focus on the magnitude of these changes, but on their predicted direction.

Table 3. Summary of Variables and Some Descriptive Statistics for Our Sample of 300 eBooks

Variable	Brief Description	Data Points	Range	Mean	SD
EPrice	ebook price	300	2.5 – 137.5	23.72	18.83
TPrice	Tangible book price (paperback)	300	1.75 – 108	25.89	15.32
CatSciFi	Science fiction category	300	{0,1}		
CatPhilosophy	Philosophy category	300	{0,1}		

tical to the original ebook in quality). This results in a negative differential effect, which our results show dominates the positive direct effect. Similarly, our results about lending rights suggest that they do not increase ebook quality significantly, but create high-quality free substitutes (ebooks one's friends have purchased, for instance, that one can borrow).

In contrast, copying rights enhance the ebook digital experience, and there is alignment between the preferences of ebook buyers and the value delivered by this right. This may be especially true for reference books and textbooks. While copying may facilitate piracy, current copying rights are restricted only to text, and not to figures or images; moreover, pirated versions created by copying text lose the typesetting and layout of the original. Our results indicate that the positive direct effect that copying rights have on ebook quality dominate any differential effect. As currently implemented, the read-aloud right does not make pirating an ebook any easier or more difficult (it may permit the creation of low-quality MP3-based "books on tape," but one can create higher-quality pirated versions by buying the book-on-tape and digitizing its content) and, therefore, is likely to have both a positive direct effect and a positive differential effect.

Tangible Book Prices and Technological Sophistication

The previous subsection indicates that while each digital right has a significant effect on the price of the ebook, the explanatory power of all the digital rights in determining this price is low (R^2 = 14 percent). This is not surprising, since a substantial part of the intrinsic quality of a book is associated with its content, the popularity of its author, and so on, independent of whether its format is tangible or digital. This intrinsic quality information is likely to be "contained" in the price of the tangible book, and our analytical results do indicate that p_T influences the price of the digital good. Moreover, the categories we chose (computers, philosophy, and science fiction) were ones we anticipated having customers with different levels of technological sophistication, which our model suggests will impact ebook pricing.

Our second model, therefore, estimates the effect of the price of the tangible book and the category of the book on the price of the ebook:

$$EPrice = \alpha + \beta_1 TPrice + \beta_2 CatSciFi + \beta_3 CatPhilosophy.$$

The data set we use is the random sample of 300 ebooks (100 per category) for which we have collected tangible book prices, and which is summarized in Table 3. For consistency, we collected the prices of the (non-mass market) paperback copy of each of these books.[5] To insure against variation in tangible book prices due to seasonal discounting or sales at the time of collection, we collected the *list price* of the tangible book. Therefore, the coefficient β_i also contains information about the average discount offered on the list price of the tangible books.

The results of this estimation are summarized in Table 4, and show that both tangible book prices and category have a significant effect on ebook prices. The model's estimates of the effect of technological sophistication are rather interesting. Interpreted in the context of our model, these results suggest that the science fiction category has more technologically sophisticated customers (and, therefore, a lower value of *t* and, consequently, lower ebook prices) than the computers and philosophy categories, and the philosophy category has less technologically sophisticated customers than the other two. We do not have data that can validate

[5]Some of the books in our sample did not have paperback versions. Our method of selecting the random sample was as follows: we selected a random sample of 200 books from each category. For each category, we then randomly selected books from this set of 200, one at a time. If the book had the appropriate paperback version available, we collected its price. If not, we discarded the book from our set. This proceeded until we had 100 books in each category with their corresponding tangible book prices.

Table 4. The Effect of Physical Book Prices and Technological Sophistication on eBook Prices

Variable	Coefficient (β_i)
Constant	4.716 ** (1.842)
TPrice	0.905 *** (0.062)
CatSciFi	–3.743 *** (1.723)
CatPhilosophy	4.746 ** (1.927)

R^2 = 62.8%, F = 169.5 ***significant with p < 0.05*
****significant with p < 0.01*

that this effect is in fact due to these categories having customers who vary in this way (for instance, from a survey of readers), but the empirical results seem to support our *ex ante* intuition. In the context of our analytical model's predictions (under which the coefficient β_1 should be 1), the value of the coefficient of the tangible book price may be due to tangible books being sold at a discount off their list price. Additionally, the model has fairly high explanatory power (R^2 = 62.8 percent), which is consistent with our conjecture that the hidden information about the intrinsic quality of the ebook is contained in the tangible book price.

Conclusions and Ongoing Work

We have presented a model characterizing the choice of digital rights and pricing for digital goods offered by a firm that also sells a physical version of the good, and when some digital rights may lead to increased piracy. The predictions of this model are supported by our empirical results, and suggest important new guidelines for managers in industries that are progressively being digitized. Summarizing our key results,

1. In the absence of piracy, digital rights are always valuable through their direct effect on increasing the quality of digital goods. Any concerns relating to the cannibalization of the sales of physical goods can be effectively addressed by a strategic choice of pricing.

2. Granting certain digital rights may increase the quality (and prevalence) of pirated digital goods. A crucial driver of the value of these rights in the presence of piracy is their differential effect on quality: how much they change the difference in quality between the legal and pirated digital goods

    - When the direct effect balances or dominates the differential effect, digital rights have a positive effect on a seller's value from them.

    - If the differential effect is negative and dominates the direct effect, granting digital rights can be detrimental to the value a seller derives from them. This is especially likely for digital rights that attempt to replicate the consumption experience provided by physical goods, rather than enhancing a customer's digital experience.

3. Empirical evidence from the ebook industry suggests that four different digital rights have a significant impact on ebook prices. There is strong negative differential effect associated with specific digital rights, which are typically those rights that aim to replicate the experience of the physical good rather than enhancing the digital experience; granting these rights results in a significant (and often substantial) lowering of ebook prices.

4. Sellers should increase the prices of digital goods as the prices of their physical counterparts increase, but decrease them as the technological sophistication of their potential customers increases. This is suggested by our economic model, and supported by some evidence from our data set.

A limitation of this model is its treatment of the quality and price of the physical good as being exogenously specified. This restriction is not for tractability or ease of analysis: it is a reflection of what we believe is the current approach taken by publishers toward the pricing and rights management of their ebooks; they price the corresponding physical books independently, using their standard pricing methods. Subsequently, the choice of ebook price and digital rights is made (typically by a separate ebook division) while treating the price of the physical book as fixed, and as containing all information relevant to the intrinsic quality of the book. An extension that considers a more general model (similar to Weber 2002) while admitting digital piracy would guide managers at a later stage of digital industry evolution. A natural related extension would study multiple versions of the digital good; there is no evidence of this yet in the ebook industry, but one would expect more active efforts toward targeted differentiated rights management and pricing (Clemons and Weber 1994) as these digital industries mature.

An extension toward further generality would admit arbitrary distributions of customer types θ and locations y. This extension could enrich our characterization of technological sophistication as reducing the fraction of customers whose tastes indicate a preference for the physical good, and increasing the fraction of customers who place a higher value on digital quality.

We propose to carry out more extensive empirical validation of our predictions. A first step toward this would automate the collection of physical book prices; a limitation of automated collection, however, is that the range of books at our specialty retailer is far wider than the range of ebooks offered by leading Web retailers. We are also exploring sources of data that can enhance our specification of the extent of technological sophistication for buyers of different categories of books.

References

Clemons, E., and Weber, B. "Segmentation, Differentiation, and Flexible Pricing: Experiences with Information Technology and Segment-Tailored Strategies." *Journal of Management Information Systems* (11:2), 1994, pp. 9-36.

Connors, K., and Rummelt, R. "Software Piracy: An Analysis of Protection Strategies," *Management Science* (37), 1991, pp. 125-139.

Gopal, R., and Sanders, L. "International Software Piracy: Analysis of Key Issues and Impacts," *Information Systems Research* (9), 1998, pp. 380-397.

Konana, P., Menon, N., and Balasubramanian, S. "The Implications of Online Investing," *Communications of the ACM* (43:1), 2000, pp. 34-41.

Png, I., and Chen, Y. "Information Goods Pricing and Copyright Enforcement: Welfare Analysis," *Information Systems Research* (14), 2003, pp. 107-123.

Snir, E. "The Record Industry in an Era of File Sharing: Lessons from Vertical Differentiation," in *Proceedings of the 24th International Conference on Information Systems*, S. T. March, A. Massey, and J. I. DeGross (Eds.), Seattle, WA, 2003, pp. 72-84.

Sundararajan, A., "Managing Digital Piracy: Pricing and Protection,"*Information Systems Research* (15), 2004, forthcoming.

Viswanathan, S. "Competition Across Channels: Do Electronic Markets Complement or Cannibalize Traditional Retailers," in *Proceedings of the 21st International Conference on Information Systems*, W. J. Orlikowski, S. Ang, P. Weill, H. C. Krcmar, and J. I. DeGross (Eds.), Brisbane, Australia, 2000, pp. 513-519.

Weber, T. "Mixed Differentiation of Information Goods under Incomplete Information," in *Proceedings of the 23rd International Conference on Information Systems*, L. Applegate, R. D. Galliers, and J. I. DeGross (Eds.), Barcelona, 2002, pp. 81-94.

Wu, S., Chen, P., and Anandalingam, G. "Fighting Information Good Piracy with Versioning," in *Proceedings of the 24th International Conference on Information Systems*, S. T. March, A. Massey, and J. I. DeGross (Eds.), Seattle, WA, 2003, pp. 617-629.