Are Digital Rights Valuable? Theory and Evidence from eBook Pricing[1]

Gal Oestreicher-Singer, Arun Sundararajan

Working Paper CeDER-06-01, Center for Digital Economy Research
Leonard N. Stern School of Business, New York University

This version: January 2006

Abstract: The effective management of digital rights is the central challenge in many industries making the transition from physical to digital products. We present a new model that characterizes the value of these digital rights when products are sold both embedded in tangible physical artifacts, and as pure digital goods, and when granting rights permitted by one's digital rights management (DRM) platform may 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 its 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 a differential piracy effect that can lower a seller's pricing power. When the latter effect is sufficiently high, granting digital rights can 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 an economically significant impact on ebook prices, and establishing that the digital rights which aim to replicate physical consumption while increasing the threat of piracy are the ones that have negative impact on seller value. We also show that if the pricing of a digital good is keyed off that of an existing tangible good, optimal pricing changes for the former should be more nuanced, rather than simply mirroring changes in the price of the latter, and we discuss the effect of the technological sophistication of potential customers on optimal pricing and rights management. Our results represent new evidence of the importance of an informed and judicious choice of the different digital rights granted by a DRM platform, and provide a new framework for guiding managers in industries that are progressively being digitized.

[1]We thank Bin Gu, Ravi Mantena, Foster Provost, Jeffrey Simonoff, Yossi Spiegel, and Manuel Trajtenberg for feedback on earlier drafts and/or helpful discussions, and seminar participants at Ben-Gurion University, Haifa University, the Technion, Tel-Aviv University, the University of Connecticut, the University of Michigan and the International Conference on Information Systems for their comments. A precursor to this paper, titled "Are Digital Rights Valuable? Theory and Evidence from the eBook Industry" won the award for best overall paper at the 25th International Conference on Information Systems.

1. Introduction

This paper studies how granting digital rights influences the quality of legal and pirated digital goods, and how this affects their value to buyers accustomed to consuming these goods as physical artifacts. We present a new model of how the demand for physical, digital and pirated goods are related to pricing and rights management choices in industries being transformed by digitization, and we subsequently validate predictions from this model using data from the ebook industry.

The importance of the specific questions addressed by this paper is illustrated well by the ongoing transformation of a number of industries whose products (music, video, graphic art, magazines, newspapers, books), historically embedded in tangible artifacts, 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 tangible 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 by prior research in information systems.

However, the process of digitization in these industries has been gradual. A majority of these digital goods continue to be purchased as tangible artifacts, and their pure digital counterparts are sometimes viewed as niche products. A number of factors influence the pace of this transition. The business models of the dominant firms in some of these industries are predicated on their superior production/distribution of the tangible artifact, and these firms may thus believe that a rapid shift to digital goods could threaten their dominance. If consumers are not technologically sophisticated, they are likely to view digital products as being of lower value, and possibly as inferior substitutes for the existing physical products they are accustomed to using. Furthermore, some goods are intrinsically more easily "digitizable": they lend themselves to digitization more naturally than others, on account of the nature of the physical artifact the good is embedded in. For instance, music on a CD, while viewed as a good embedded in a physical artifact, is already digital, and thus

more easily converted into a pure digital good than, say, the contents of a physical book. Finally, and perhaps most importantly, 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 the transition towards digitizing products presents sellers with difficult technological and business decisions. These relate crucially to the strategic control of the quality of digital goods and the extent to which the presence of these digital goods leads to piracy. Both of these 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). A different range of rights is technologically feasible under each DRM platform, and the extent of such flexibility, along with the secureness or strength of the DRM system, are important considerations when choosing a platform. However, technological feasibility and 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 emusic services (Mossberg, 2003). On the other hand, granting digital rights that are too extensive raises concerns about prematurely cannibalizing established sales of physical goods[2]. More importantly, granting certain digital rights may increase the prevalence and quality of pirated substitutes for the good. 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 towards increasing the quality of legal digital goods, and towards facilitating piracy, a careful assessment of

[2]Similar issues have been recognized in the context of channel conflict caused by Web-based channels for, among other businesses, online retailing (Viswanathan, 2005) and online brokerage (Konana et al., 2000).

their relative contribution towards each is important.

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

1. The relative extent to which each digital right a seller may grant buyers contributes towards an increase in the value 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.

2. 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.

3. The technological sophistication of potential customers, which determines what fraction of them place a relatively high value on digital goods (both legal and pirated).

4. The extent to which the physical good is "digitizable", which influences how closely digital goods are assessed as substitutes for their tangible counterparts.

We capture each of these factors in our model of sellers who choose digital rights and prices for their digital goods, while simultaneously selling corresponding physical goods at pre-specified 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 the increases in value they enjoy on account of the 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. 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. However, the prevalence of piracy dampens the extent to which changes in tangible good prices are mirrored in the pricing of their digital counterparts.

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, since the potential threat of cannibalization can be effectively managed through the strategic control of price.

3. The extent to which granting digital rights affects a seller's pricing power is determined by the interplay between a direct quality effect, proportional to the increase in the quality of the legal digital good, and a differential piracy effect, proportionate to the (weighted) increase in the quality of the pirated good. If these two effects balance, then granting the right is profitable, and should increase the seller's optimal price. However, if the latter effect dominates the former substantially, granting digital rights can actually be associated with a decrease in the price of the legal digital good.

4. During early stages of the transition from tangible to digital goods, when the level of technological sophistication of a seller's consumer base is low, an increase in the level of technological sophistication may be associated with an increase in the price of digital goods relative to their tangible counterparts, although this effect is limited by the threat of piracy as a market becomes progressively more technologically sophisticated.

Our analytical results relate the prices of digital goods to the prices of their physical counterparts, the extent to which digital rights are granted to customers, the relative demand for each kind of good, and the technological sophistication of the product's target customer base. We test these results using a data set containing pricing and digital rights data for over 3,000 ebook titles, across six categories, and sold by a specialized ebook retailer. There are many reasons why the ebook industry is an excellent setting to test our theory. The digital rights associated with each ebook are determined individually by publishers rather than by a major retailer, and thus there is substantial variation in the level to which rights are granted across ebooks (the summary of variables in Table 3.1 in section 3 illustrates this well). There is also a substantial variation in prices across ebooks, and significant demand for both tangible and digital versions of a wide variety of books.

The equations we estimate are hedonic pricing indices that relate an ebook's price to the price of its tangible counterpart, the digital rights it grants, and the book's category. Our empirical findings provide strong support for the predictions of our theoretical model, showing that tangible

book prices and each of four digital rights have a significant impact on ebook prices, and that digital rights explain a significant fraction of variation in ebook prices. 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 the value a seller can derive from its legal digital goods. We also distinguish between these rights based on an intuitive categorization of them into those that aim to replicate the physical consumption experience, and those that enhance the digital consumption experience, showing that granting rights in the former set is likely to have a detrimental effect on seller value, while granting rights in the latter set is valuable.

Our paper adds to a growing literature that aims to guide managers facing piracy and rights management challenges in digital industries. Prior work has studied alternative approaches to the strategic control of digital piracy, through the strategic choice of either quality, or of pricing. Papers which model the deterrence of piracy by varying quality generally base their analysis on Mussa and Rosen's (1978) model of vertical product differentiation, and the idea is that a seller might produce a lower-priced degraded substitute for its flagship good so that piracy is either reduced or even eliminated. For instance, Poddar (2005) shows that the quality and reliability of the pirated product are among three attributes that determine whether permitting piracy is profitable or not. Alvisi and Carbonara (2002) find that in the presence of piracy and heterogenous copying costs, such differentiation arises as the optimal strategy for a monopolist to deter piracy, by diverting consumers from the pirated good to the original one. By offering legal digital products of varying quality, the monopolist can effectively discriminate between those consumers with higher and lower copying costs. Two other models that explore ideas closely related to the ones described above are those by Snir (2003), and by Wu et al. (2003). Belleflamme (2003) studies the interdependence between different producers' incentives to accommodate/deter the presence of a pirated good. The relative welfare benefits of legal and technological deterrence are studied by Png and Chen (2003) and by Takeyama (1994).

In our model, we build on the approach of many of these prior papers, preserving their notion of

the pirated good as an inferior (vertically differentiated) substitute for the legal good, but expanding on it by simultaneously modeling the demand for a tangible good horizontally differentiated from its (legal and pirated) digital counterparts. The seller chooses quality by explicitly varying the digital rights it grants buyers of its legal digital good. Further, we allow the seller to use pricing in addition to digital rights management as an instrument to deter piracy. A more involved study of the strategic choice of pricing in the presence of piracy has been done by Sundararajan (2004), who shows that while in the absence of price discrimination, a seller's optimal level of protection against piracy is at the technologically maximal level, a seller who can price discriminate always chooses a strictly lower level of protection. Since his model considers only pure digital goods, we extend its results by examining the effect of having a physical substitute in addition to a pirated version, and we provide evidence for the need to restrict certain digital rights even in the absence of price discrimination.

A related stream of literature has explored the benefits of piracy to a seller. A common argument is that it increases profits for a product that displays network effects, an idea first articulated by Connor and Rummelt (1991), and discussed by many others, including Bakos et el (1999) in the context of file sharing systems ("clubs") and Shy and Thisse (1999) in a duopoly model. Prasad and Mahajan (2003) and Haruvy et el. (2004) showed that piracy is not necessarily harmful to a firm seeking to launch a new product, since it establishes an initial user base, and speeds up product diffusion. The question they then address is how much piracy should be tolerated. Gu and Mahajan (2004) show that since pirated goods are found most attractive by those customers who are most price sensitive, its presence can actually reduce the intensity of price competition, by "removing" these consumers from the market, and thereby benefiting competing providers of legal digital goods. A different benefit of piracy can be that it implicitly provides free "samples" of a product to consumers who are not perfectly informed about its characteristics, thus informing them about the value of the product, and influencing legal demand (Peitz and Waelbroeck, 2003). Chellappa and Shivendu (2005) extend this idea to model sampling and strategic pricing in the presence of a pirated good. Gayer and Shy (2003) explore a related idea in the context of digital

goods enhancing the sales of tangible goods. In contrast with this line of research, we focus on the central detrimental aspect of piracy: that it generates a free substitute for legally available physical and digital goods.

Our survey of the literature indicates that the relationship between piracy, digital rights management and the strategic control of quality has not been studied empirically, and our paper thus represents the first such study. The empirical studies about digital piracy we are aware of aim primarily to estimate the effect of piracy on the demand for the legal products and on profit (for instance, Zenter, 2003; Hui and Png, 2003). An interesting prior empirical study of software piracy was by Gopal and Sanders (1998), who establish that the strength with which a government enforces its digital intellectual property laws is positively related to the robustness of the domestic software industry, an observation that is likely to generalize to other digital industries like music and video as well.

We have organized the rest of this paper as follows. Section 2 models the legal demand for a seller's tangible and digital goods, and describes the relationship between the prices of legal digital goods, their digital rights and the prices of their tangible counterparts, both in the absence of piracy and in its presence. The predictions of this model are tested in Section 3, which describes our data set, presents our empirical models, and describes the results of their estimation. Section 4 concludes with a summary of our results, some limitations of our study, and directions for future research.

2. Theory: DRM, piracy and pricing strategy

2.1. Overview of model

We model a monopoly seller who may produce two versions of a product: a physical (*tangible*) good and a *digital* good. These goods are imperfect substitutes. The quality of the tangible good s_T is exogenous, and normalized to 1. The quality s_D of the digital good is determined by the level

to which the seller grants each of a set of n digital rights $\mathbf{r} = (r_1, r_2, ..., r_n)$ to its buyers:

$$s_D = S_D(r_1, r_2, ..., r_n). \tag{2.1}$$

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 pre-specified minimum and maximum levels. For instance, if using 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, upto 35 pages every two weeks).

Granting a buyer a higher level of digital rights increases her value from the digital good, and therefore, S_D is (weakly) increasing in each of its arguments. While the set of possible values for some digital rights may be discrete, we instead assume that each r_i can vary continuously between its minimum and maximum levels. This assumption is for analytical convenience, though relaxing it is unlikely to directionally alter our results. Altering the level of digital rights within the constraints of its DRM platform imposes no direct cost on the seller, and consequently, changes to the quality s_D of the digital good have 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 of the good, there may be a *pirated* digital version available. The quality s_P of the pirated good is also determined by the level to which the seller grants the same set of n digital rights $\mathbf{r} = (r_1, r_2, ..., r_n)$ to buyers of its legal digital good:

$$s_P = S_P(r_1, r_2, ..., r_n). \tag{2.2}$$

Since an increase in the level of rights granted to legal users often facilitates the creation of

higher-quality pirated versions[3], S_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.

Consumers are heterogeneous along two dimensions. The first dimension, indexed by a *taste* variable $y \in [0,1]$ represents the consumer's relative preferences for tangible versus digital goods. Holding everything else constant, a consumer indexed by a higher value of y places a higher value on a digital good and a lower value on a physical good than a consumer indexed by a lower value of y places on the same two goods. The second dimension, indexed by a *digital type* variable $\theta \in [0,1]$ represents a consumer's preferences for digital quality, or how much value a customer ascribes to an increase in digital rights. Each customer of type θ always prefers higher digital quality (or the granting of more digital rights) to lower; 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 seller does not know the taste y or type θ of any specific consumer, but knows the distribution of y and θ in the population of potential consumers. The distribution function of y is denoted $F_h(y)$, and the distribution function of θ is denoted $F_v(\theta)$. We assume that these distributions are absolutely continuous and have corresponding density functions $f_h(\theta)$ and $f_v(\theta)$. For simplicity, we also assume that these distributions are independent.

We 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 are linearly separable in value and price, and across digital type and taste, specified using a common utility function

$$U(s,\theta,y) = u(s,\theta) - w(y),$$

where the linearly separable form is chosen for analytical convenience. The surplus a consumer gets

[3]For instance, the right to download rather than stream digital audio files increasing the threat of digital piracy over file-sharing networks, the right to print a digital book enables the creation of pirated PDF copies of the book.

from his or her purchase of the digital good is:

$$\text{Digital: } u(s_D, \theta) - w([1-y]) - p_D, \tag{2.3}$$

the surplus a consumer gets from his or her use of the free pirated good is

$$\text{Pirated: } u(s_P, \theta) - w([1-y]), \tag{2.4}$$

and the surplus a consumer gets from his or her purchase of the tangible good is

$$\text{Tangible: } u(1,1) - w(y) - p_T. \tag{2.5}$$

We use a common underlying utility function in order to maintain consistency in the variation of a customer's preferences across the three goods[4], since these goods are imperfect substitutes. Equations (2.3) and (2.4) reflect a choice of modeling a pirated good as an inferior (vertically differentiated) substitute for the legal digital good, which is consistent with a subset of the prior literature we discuss in Section 1. Equation (2.5) indicates that we model consumers with the same taste y as being *homogeneous* in their preferences for the quality of the tangible good. We do this because we wish to focus our analysis on the interplay between cannibalization of the sales of the tangible good, the threat of digital piracy and the granting of digital rights. Introducing an additional customer characteristic that models heterogeneity in preferences for tangible goods will complicate things unnecessarily. Anchoring the "common type" of customers with respect to their preferences for the quality of the tangible good at 1 simply reflects an assumption that all customers are used to using tangible goods, and this forms the benchmark for their assessment of the value of the digital good. One might think of this as benchmarking the "intrinsic quality" of the good and treating the quality of the digital and pirated goods as being measured relative to this benchmark. This is especially true for the industry (books) in which we test our theory's predictions.

Customers use exactly one version of the good, purchasing either the tangible good or the digital

[4]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/pirated goods, and the digital good is vertically differentiated from the pirated good. The subscripts of the distribution functions h (horizontal) and v (vertical) are labeled accordingly for this reason.

good, or using the free pirated good. We do not choose a specific functional form for u or w, but assume they have the following properties:

1. $\frac{du}{ds} > 0, \frac{du}{d\theta} > 0, \frac{dw}{dy} > 0.$

2. $u(s, 0) \leq w(0)$: consumers of digital type 0 derive no value from digital goods at any quality level, or some customers are just not interested in purchasing digital goods.

3. $u(1, 1) - w(1) > p_T$: the market is fully covered.

4. $\frac{d^2u}{dsd\theta} > 0$: consumers of higher digital type value the same incremental increase in quality more than consumers of lower digital type.

Assumption 1 is consistent with how we have described the variation in preferences with type θ and location[5] y. The (average) slope of $w(y)$ could indicate the extent to which the good is "digitizable", by determining how closely consumers on average assess tangible and digital goods as substitutes (for example, if $w'(y) = 0$, a tangible and digital good of equal quality are perfect substitutes). Assumption 3 posits that in the absence of digital goods, all consumers purchase the tangible good. This is towards focusing our analysis on a scenario in which consumers switch from consuming a tangible good to consuming its digital counterpart (as the industry in question transforms from being "physical" to being digital), and the associated issues of cannibalization and piracy, rather than a scenario in which the digital good expands the seller's market. We discuss the implications of relaxing this assumption in Section 4. Assumption 4 is the (standard) Spence-Mirrlees condition.

The sequence of events is as follows: the seller announces their choice of r, p_T and p_D, consumers make their choices, and each party realizes its payoffs.

[5] The results of our model are unchanged if we assume that u increases in y and alter equations (2.3) through (2.5) accordingly. We have made this choice to be consistent with how utility varies with location in standard models of horizontal differentiation with "misfit" or "transportation" costs (for instance, the Hotelling model).

2.2. Choosing digital rights in the absence of piracy

Our first set of results set a benchmark by specifying the optimal choice of digital rights and pricing in the absence of piracy. Therefore, in this subsection,

$$S_P(\mathbf{r}) = 0, \tag{2.6}$$

and customers choose between the tangible good and the digital good. The price and quality of the tangible good are exogenous, and the seller chooses the level of each of its digital rights and the price of the digital good. Given these choices, demand can be characterized as follows. For those customers with location y, define $\theta_{TD}(y, p_D)$ as the type indifferent between the tangible good and the digital good when the price of the digital good is p_D. For each y, $\theta_{TD}(y, p_D)$ is thus defined as follows:

$$u(1,1) - w(y) - p_T = u(s_D, \theta_{TD}(y, p_D)) - w([1-y]) - p_D. \tag{2.7}$$

A customer with location and type (y,θ) prefers the tangible good to the digital good if $\theta < \theta_{TD}(y, p_D)$, and prefers the digital good to the tangible good if $\theta > \theta_{TD}(y, p_D)$. Under assumptions 2 and 3, it follows that

$$u(1,1) - w(0) - p_T > u(s_D, 0) - w(1) - p_D, \tag{2.8}$$

and therefore, realized demand is as summarized in Figure 2.1.

The demand for the tangible good is:

$$q_T(p_D) = F_H(y_D) + Q(y_D, 1, \theta_{TD}(y)), \tag{2.9}$$

where $y_D(p_D)$ is the taste value below which all consumers purchase the tangible good:

$$y_{TD}(p_D) = y : \theta_{TD}(y, p_D) = 1, \tag{2.10}$$

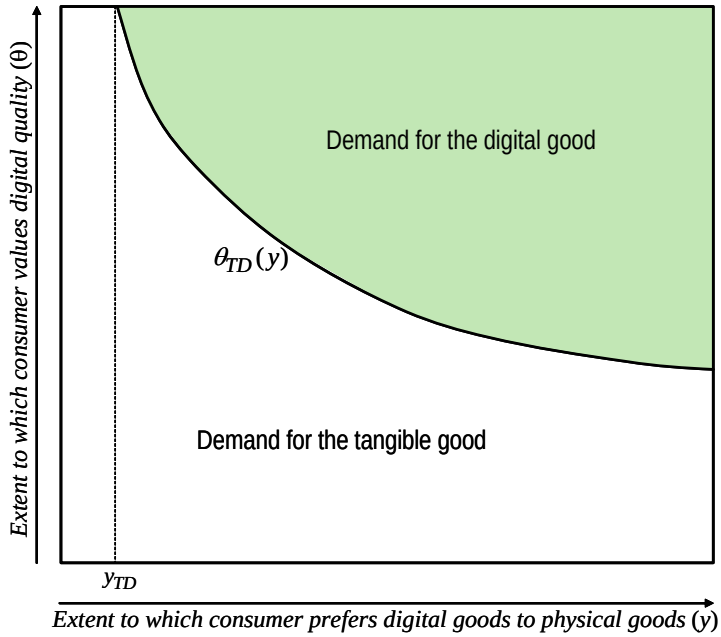


Figure 2.1: Summarizes demand for the tangible and digital good in the absence of piracy. The curve $\theta_{TD}(y)$ represents those consumers indifferent between the two goods.

and

$$Q(y_a, y_b, g(y)) \equiv \int_{y=y_a}^{y_b} \int_{\theta=0}^{g(y)} f_v(\theta) f_h(y) dy \tag{2.11}$$

is the mass of customers over the area under an arbitrary curve $\theta = g(y)$, between y_a and y_b. Correspondingly, the demand for the digital good is

$$q_D(p_D) = 1 - F_H(y_{TD}(p_D)) - Q(y_{TD}(p_D), 1, \theta_{TD}(y, p_D)). \tag{2.12}$$

Evidently, $\theta_{TD}(y)$ and y_{TD} also depends on p_T and s_D, though, to keep our equations more readable, we do not include these as arguments[6].

Based on this characterization of demand, we can derive an expression for the optimal price of the digital good, and describe corresponding optimal extent to which the seller should grant its digital rights. This analysis, which is presented in Appendix A, leads to our first two proposition:

Proposition 1. *A unit increase in the price of the tangible good leads to a corresponding unit*

[6] A parallel case involves parameter values such that $\theta_{TD}(0) < 1$. Each of its results are directionally similar to those presented in this section.

increase in the price of the digital good.

Proposition 2. *In the absence of piracy, the seller grants its customers the maximum possible level of rights permitted by its DRM platform.*

The intuition underlying the result of Proposition 2 is quite simple. In the absence of piracy, the seller can address any threat of cannibalization as effectively as is necessary through its strategic control of the price of the digital good. Its profits increase with an increase in the quality of the digital good, and therefore, it chooses digital rights to maximize the quality of its digital good. This result is especially stark because it holds when the market is fully covered; clearly, it will generalize to a model in which market expansion is possible through the introduction of the digital good, since a seller's incentives to increase s_D are stronger if it were able to expand its market in addition.

Empirically, however, we observe that sellers frequently restrict the digital rights they grant buyers. Thus, even when demand is modeled quite generally, with no restrictions on the ease with which a good can be digitized, or on the technological sophistication of the target consumer base, a model which ignores piracy yields results that seem to contradict what is observed in practice. In our next subsection, we return to the model described in section 2.1, and establish how the presence of piracy alters these predictions about pricing and rights management for digital goods.

2.3. Pricing and rights management in the presence of piracy

We now return to the complete model described in section 2.1. Granting a level of digital rights $(r_1, r_2, ..., r_n)$ to each buyer of the digital good also induces the creation of a free pirated good of quality $s_P = S_P(r_1, r_2, ..., r_n) > 0$. Buyers choose between three imperfect substitutes: the tangible good, the digital good and the pirated good. Characterizing demand completely will therefore involve analyzing three pairwise comparisons made by customers.

The first pairwise comparison, between the tangible good and the digital good, yields the functions $\theta_{TD}(y, p_D)$ and $y_{TD}(p_D)$ that are defined in (2.7) and (2.10). Next, for those customers

with taste y, define $\theta_{TP}(y)$ as the customer type that is indifferent between the tangible good and the pirated good:

$$u(1,1) - w(y) - p_T = u(s_P, \theta_{TP}(y)) - w([1-y]). \tag{2.13}$$

Notice that $\theta_{TP}(y)$ does not depend on p_D, the price of the digital good. A customer with taste and type (y,θ) prefers the tangible good to the pirated good if $\theta < \theta_{TP}(y)$, and prefers the pirated good to the tangible good if $\theta > \theta_{TP}(y)$. The taste value y_{TP} below which consumers of all types $\theta \in [0,1]$ prefer the tangible good to the pirated good is defined by:

$$y_{TP} = y : \theta_{TP}(y) = 1. \tag{2.14}$$

If $\theta_{TP}(1) > \theta_{TD}(1)$, then all relevant consumers (that is, those who will not purchase the tangible good when a digital or pirated substitute is available) prefer the digital good to the pirated good, and the analysis proceeds as in Section 2.2. Similarly, if $y_{TP} < y_{TD}(p_D)$, then all relevant consumers prefer the pirated good to the digital good, and this scenario will therefore not occur, since the seller can raise its profits by simply not offering a digital good. Therefore, the only scenario of economic relevance is when the following conditions hold:

$$\theta_{TP}(1) < \theta_{TD}(1, p_D), \tag{2.15}$$

$$y_{TP} > y_{TD}(p_D), \tag{2.16}$$

and under these conditions, there is a unique value of y at which $\theta_{TP}(y) = \theta_{TD}(y, p_D)$, which we label $y_{PD}(p_D)$.

$$y_{PD}(p_D) = y : \theta_{TP}(y) = \theta_{TD}(y, p_D) \tag{2.17}$$

Clearly,

$$y_{TD}(p_D) < y_{TP} < y_{PD}(p_D), \tag{2.18}$$

and therefore, demand of each good is as depicted in Figure 2.2.

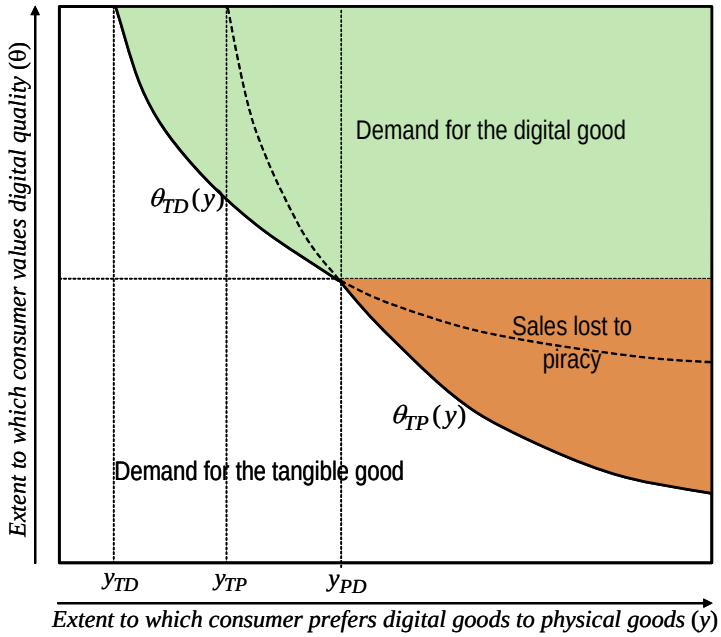


Figure 2.2: Illustrates demand for the tangible and digital goods in the presence of piracy. The curve $\theta_{TD}(y)$ defines those consumers indifferent between the tangible and digital goods, and curve $\theta_{TP}(y)$ defines those consumers indifferent between the tangible and pirated goods.

The demand for the tangible good is:

$$q_T(p_D) = F_h(y_{TD}(p_D)) + Q(y_{TD}(p_D), y_{PD}(p_D), \theta_{TD}(y, p_D)) + Q(y_{PD}(p_D), 1, \theta_{TP}(y)), \qquad (2.19)$$

where the function Q was defined in (2.11), the demand for the digital good is:

$$\begin{aligned} q_D(p_D) &= F_h(y_{PD}(p_D)) - F_h(y_{TD}(p_D)) - Q(y_{TD}(p_D), y_{PD}(p_D), \theta_{TD}(y, p_D)) \qquad (2.20) \\ &\quad + [1 - F_h(y_{PD}(p_D)][1 - F_v(\theta_{TD}(y_{PD}(p_D), p_D))] \end{aligned}$$

and the demand for the pirated good is

$$q_P(p_D) = [1 - F_h(y_{PD}(p_D))][F_v(\theta_{TD}(y_{PD}(p_D), p_D))] - Q(y_{PD}(p_D), 1, \theta_{TP}(y)) \qquad (2.21)$$

This is summarized in Figure 2.2. Based on this characterization of demand, the optimal price for the digital good can be characterized. This characterization leads to our next proposition:

Proposition 3. *The optimal price of the digital good is of the form:*

$$p_D^* = [p_T - c]\delta(p_D^*) + \frac{q_D(p_D^*)}{-\left[\dfrac{dq_D(p_D^*)}{dp_D}\right]}, \tag{2.22}$$

where $\delta(p_D^*) < 1$. *Therefore, a change in the price of the tangible good causes a less than proportionate change in the price of the digital good*

The intuition for the result of Proposition 3 is illustrated in Figure 2.3. An increase in the price of the digital good shifts the curve $\theta_{TD}(y, p_D)$ upwards, reflecting an increase in the fraction of consumers who now prefer the tangible good to the digital good. The relative fractions of consumers who prefer the tangible good over the pirated good are not affected, since the curve $\theta_{TP}(y)$ does not shift. There are corresponding shifts in y_{TD} (to the right) and y_{PD} (to the left). The upward shift of $\theta_{TD}(y, p_D)$ therefore reduces demand for the digital good in two ways. First, it causes a shift in demand from the digital good to the tangible good, illustrated by the green (left) shaded area. In addition, it causes a shift in demand from the digital good to the pirated good, as illustrated by the red (right) shaded area. The discount factor $\delta(p_D^*)$, the ratio of the change in demand for the tangible good caused by this shift to the change in demand for the digital good, is therefore always less than 1 in the presence of piracy.

Proposition 3 thus indicates that a seller who offers both digital and tangible goods must approach pricing changes for the former with care. Changes in the price of the digital good must be more nuanced, rather than simply mirroring changes in the price of the tangible good. We return to this observation when discussing our empirical findings in Section 4.

Next, consider the effect of a change in any right r_i on the price of the digital good. Notice that since

$$\frac{dp_D^*}{dr_i} = \frac{dp_D^*}{ds_D}\frac{dS_D}{dr_i} + \frac{dp_D^*}{ds_P}\frac{dS_P}{dr_i}, \tag{2.23}$$

it follows that

$$\frac{dp_D^*}{dr_i} \propto \frac{dS_D}{dr_i} - \left[\frac{-dp_D^*/ds_P}{dp_D^*/ds_D}\right]\frac{dS_P}{dr_i}. \tag{2.24}$$

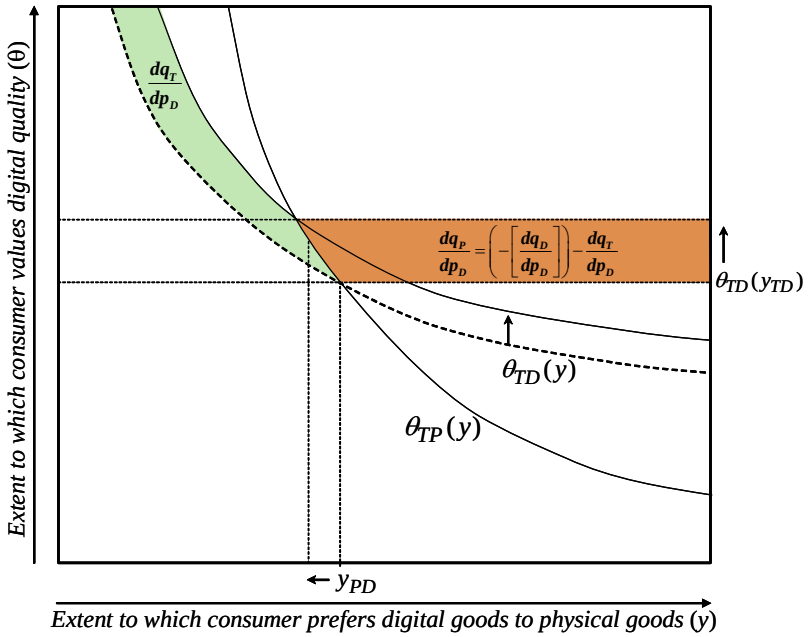


Figure 2.3: Illustrates the shift in demand from a small increase in the price p_D of the digital good. The upward shift of the curve $\theta_{TD}(y)$ reduces demand for the digital good in two ways. Part of this demand shift is captured by the tangible good (the green shaded area on the left) while the rest is lost to piracy (the red shaded area on the right). Notice that the curve $\theta_{TP}(y)$ remains unchanged by an increase in p_D, since the relative surplus from the tangible and pirated goods do not change.

(2.24) reveals that there are two separate effects that increasing the extent to which a digital right is granted has on the price of the digital good. The first is a *direct quality effect* (represented by the term $\frac{dS_D}{dr_i}$) which is positive, and proportionate to the extent to which granting the additional rights will increase the quality of the legal digital good. The second is a *differential piracy effect* (represented by the term $\left[\frac{-dp_D^*/ds_P}{dp_D^*/ds_D}\right]\frac{dS_P}{dr_i}$) which is negative, and proportionate to the extent to which granting the additional rights will increase the quality of the pirated good. Since an increase in the quality of the digital good has a positive impact on price, while an increase in the quality of the pirated good has a negative impact on price, (2.24) indicates that an increase in the extent to which a digital right is granted does not always increase the price of the digital good, or the seller's total profits, for that matter. This is in contrast with the result of Proposition 2, and an (unstated) related result, which indicates that in the absence of piracy, increasing the level to which a right is granted always increases p_D.

Furthermore, since the pirated good is an (imperfect) substitute for the legal digital good, a

unit increase in the quality of the legal digital good is likely to cause a higher increase in its price than a corresponding unit decrease in the quality of the pirated good, or, it is likely that

$$\frac{dp_D^*}{ds_D} > -\left(\frac{dp_D^*}{ds_P}\right) \tag{2.25}$$

Together, (2.24) and (2.25) suggest that if granting a digital right causes an equal increase in the quality of the digital good and the quality of the pirated good, or a higher increase in the quality of the digital good than the quality of the pirated good, granting this right will be associated with an increase in the price of the digital good, since the direct quality effect dominates the differential piracy effect. In contrast, if the quality of the pirated good increases substantially more than the quality of the legal digital good on account of granting a specific right, or more precisely, if

$$\frac{ds_P}{dr_i} > \left[\frac{-dp_D^*/ds_D}{dp_D^*/ds_P}\right]\frac{ds_D}{dr_i},$$

then the differential piracy effect dominates the direct quality effect, and granting of this digital right will be associated with a *decrease* in the price of the digital good. Again, we return to this observation when interpreting our empirical results in Section 4, towards discussing the kinds of rights that are likely to fall into each category.

Finally, consider the effect of a change in the shape of the distribution of y. Suppose we index the distribution of y by a parameter α such that $F_h(y;\alpha_1)$ first-order stochastically dominates $F_h(y;\alpha_2)$ if $\alpha_1 > \alpha_2$, or

$$\alpha_1 > \alpha_2 \Rightarrow F_h(y;\alpha_2) > F_h(y;\alpha_1) \text{ for all } 0 < y < 1.$$

One might consider a market corresponding to a larger α more *technologically sophisticated* than a market with a lower value of α, since there is a higher fraction of consumers with values of y closer to 1 (recall that a consumer with a higher value of y places relatively more value on the digital good). This is illustrated in Figure 4.

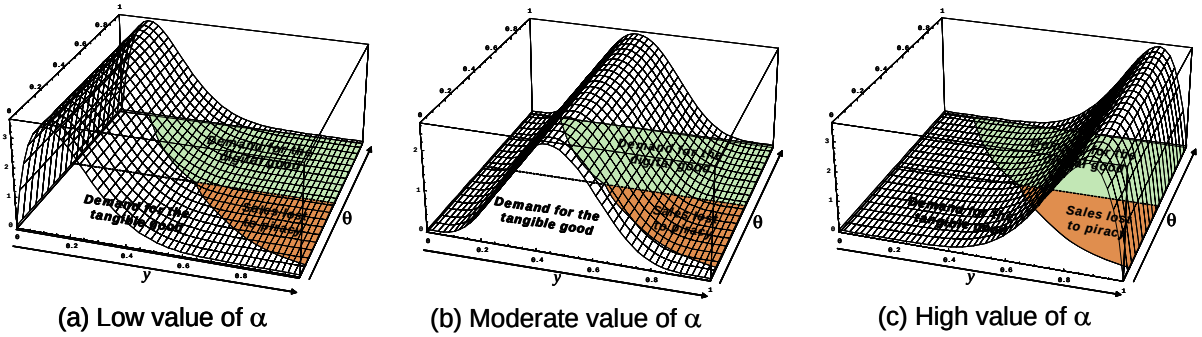


Figure 2.4: Illustrates how the mass of consumers shifts over the parameter space as a market becomes more technologically sophisticated. The two horizontal axes represent consumer taste (y) and digital type (θ), and the vertical axis represents the distribution of taste $f_h(y;\alpha)$, with a higher value of α parametrizing a distributional shift to the right.

It therefore seems natural to conclude that an increase in the technological sophistication of a market would result in a higher demand for digital goods. However, the effect of such an increase (in technological sophistication) on price is not straightforward to infer. This is because the same shift that causes an increase in the potential demand for the digital good would also cause an increase in the potential demand for the pirated good. Qualitatively, if an increase in α causes the distribution of y to move from being very right-skewed to being somewhat balanced (for instance, like a move from (a) to (b) in Figure 4), one might expect this shift to cause an increase in the price of the digital good: the primary shift in consumer mass is towards a region in which there is some potential demand for the digital good, but little potential demand for the pirated good[7]. On the other hand, if an increase in α causes the distribution of y to become very left-skewed (for instance, like a move from (b) to (c) in Figure 4), the effect on p_D is unclear. While the seller might increase the price of the digital good to benefit from the higher potential demand for that digital good, the seller's ability to increase price is limited by the larger potential demand for the pirated good, and in fact, the seller may need to lower the price of its digital good in order to ameliorate some of the loss in demand for the tangible good caused by piracy. We discuss this further in Section 4 when interpreting the variation in the price of digital goods across categories.

[7]While the positions of the curves $\theta_{TD}(y)$ and $\theta_{TP}(y)$ are unaffected by changes in $f_h(y)$, the resulting price changes may shift the curves. The figure thus illustrates what would motivate such a price change, prior to the change in price.

3. Are digital rights valuable? Evidence from the ebook industry

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

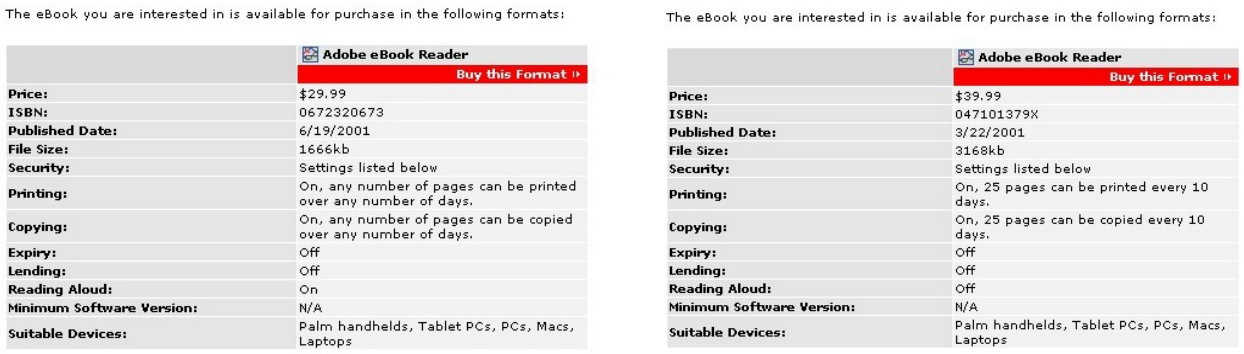
3.1. Data

We have collected the prices, digital rights and category affiliation of over 30,000 ebooks sold by a specialty Web-based ebook retailer. We have also collected data about the prices of the tangible paperback versions of a subset of these ebooks from a leading online seller of physical books. Combining the two data sets created a set of over 4,000 book titles, for which both an electronic version and a paperback version exists. The results presented are based on our data set from January 2005.

We gathered data about the prices of the tangible counterparts of our ebooks, because 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. We conjecture that information about this intrinsic quality (we use the term "quality" as it relates to a buyer's willingness to pay, rather than an assessment of the book's literary merits) is likely to be contained in the price of the tangible book, which, according to our model, influences the price of the digital good. 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.[8]

The ebooks in our data set 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: print, copy, expiry, lend and read aloud. None of the books in our data set have restricted the "expiry" right. The "lend" and "read aloud" rights have binary settings (on

[8] We also restrict ourselves to books with paperback versions for this reason, to avoid a higher price for a tangible book being on account of it being a hardcover (or some other superior) version, rather than paperback version.



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 3.1: Illustrates the digital rights granted for two ebooks in our sample. 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 $PrintAll = 0, CopyAll = 0, Read = 0, Lend = 0$.

or off). The rights "print" and "copy" 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 3.1. 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 them 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 ($CopyPartial = 1$). The variables associated with each of our data points are summarized on Table 1, along with some descriptive statistics of the data set, and the correlation matrix[9] is summarized in Table 3.1A.

The ebook retailer categorizes its books by subject. We focus our analysis on six categories that we believed ex ante would have target customer sets with different levels of technological sophistication: Computers, Fiction, Children's and Young Adults Fiction, Biography and Autobiography,

[9]We also computed the variance inflation factors for our data, which range from 1.2 to 4.7, thus confirming that our data does not display higher-order multicollinearity.

Variable	**Brief Description**	**Data points**	**Range**	**Mean**	**SD**
EPrice	ebook price	3105	1.98-165	14.68	18.01
Log*[EPrice]*	Log of ebook price	3105	0.68-5.1	2.27	0.82
TPrice	Tangible book price	3105	1.99-387	25.77	23.32
Log*[TPrice]*	Log of tangible book price	3105	0.68-5.13	2.95	0.78
Variable	**Brief Description**	**Data points**	**Range**	**% non-zero**	
CopyAll	Allows unlimited copying of text	3105	{0,1}	49%	
CopyPartial	Allows limited copying of text	3105	{0,1}	16%	
PrintAll	Allows unlimited printing of pages	3105	{0,1}	16%	
PrintPartial	Allows limited printing of pages	3105	{0,1}	14%	
Lend	Allows lending	3105	{0,1}	5.4%	
Read	Has the 'Read Aloud' feature	3105	{0,1}	39.19%	
K_1	Business category	3105	{0,1}	15%	
K_2	History category	3105	{0,1}	5%	
K_3	Fiction category	3105	{0,1}	57%	
K_4	Young category	3105	{0,1}	6%	
K_5	Biography category	3105	{0,1}	5%	

Table 3.1: Summary of variables and some descriptive statistics

Business and Economics, and History. The smallest of these categories has 164 pairs of ebooks and paperback books. Our final data set has a total of 3105 observations, where each observation corresponds to a unique title.

The first stage of our analysis investigates how the price of the ebook is related to the price of its tangible counterpart. Since we believe that a substantial part of the intrinsic quality of a book is captured in the price of the tangible version, a significant relationship between the price of the ebook and the price of the tangible book seems natural. We therefore estimate the following equation:

$$log[EPrice] = \alpha + \beta_1 \log[TPrice], \tag{3.1}$$

where we use the log transformation to account for non-constant variance. The results of this estimation are presented in Table 3.2, and confirm the expected relationship. The value of the regression coefficient β_1 (0.60) suggests partial support for Proposition 3, since it indicates that when the prices of the digital good and the tangible book are comparable, a unit increase in the price of the tangible good causes a less than unit increase in the price of the digital good. Also,

Variable	Estimated Value (SE)
α	0.51^{***} (0.05)
β_1	0.6^{***} (0.02)
$R^2 = 32.2\%, F = 1463.15$	*** significant with $p \leq 0.001$

Table 3.2: The relationship between tangible and digital book prices

the relative low R^2 indicates that there are factors beyond the price of the tangible book and the intrinsic quality it represents that influence the price of the ebook.

3.2. A hedonic price index for digital rights

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

$$\begin{aligned} log[EPrice] &= \alpha + \beta_1 \log[TPrice] + \beta_2 CopyAll + \beta_3 CopyPartial \\ &\quad + \beta_4 PrintAll + \beta_5 PrintPartial + \beta_6 Lend + \beta_7 Read + \sum_{i=1}^{5} \gamma_i K_i \end{aligned} \tag{3.2}$$

Our model's results from Section 3 predict that a digital right whose positive direct quality effect balances or dominates the differential piracy effect will have a positive coefficient β_i associated with it. On the other hand, a right for which the differential piracy effect is substantially higher than the positive direct quality effect (that is, a digital right which is not assessed as being especially valuable by customers of the digital good, but whose granting substantially increases the quality and availability of pirated versions of the good) will have a negative coefficient β_i associated with them.

The results of this estimation are presented in Table 3.3. The signs of the coefficients are quite striking. They indicate that three of the four digital rights – copying, reading aloud and lending – are associated with a significant increase in ebook prices. However, a fourth right – printing – is associated with a significant (and substantial) *decrease* in the price of the digital version of the

Variable	Coefficient	**Estimated Value (SE)**	**Corresponding value of e^{β_i}**
Constant	α	2.235 ** (0.006)	9.267
Log[*TPrice*]	β_1	0.274 ** (0.012)	–
CopyAll	β_2	0.052 * (0.023)	1.053
CopyPartial	β_3	0.405 ** (0.042)	1.499
PrintAll	β_4	−0.338 ** (0.035)	0.713
PrintPartial	β_5	−0.394 ** (0.045)	0.674
Lend	β_6	0.132 ** (0.046)	1.141
Read	β_7	0.408 ** (0.028)	1.504
K_1 *(Business)*	γ_1	−0.376 ** (0.035)	0.687
K_2 *(History)*	γ_2	−0.808 ** (0.049)	0.446
K_3 *(Fiction)*	γ_3	−1.189 ** (0.037)	0.305
K_4 *(Young)*	γ_4	−1.485 ** (0.049)	0.227
K_5 *(Biography)*	γ_5	−0.859 ** (0.049)	0.424
$R^2 = 66.26\%, F = 507.1$		* significant with $p \leq 0.05$ ** significant with $p \leq 0.01$ *** significant with $p \leq 0.001$	

Table 3.3: The effect of digital rights on ebook prices

book.[10]

In the context of our 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. Recall from our model that the set of customers who purchase ebooks are those

[10]Since we estimate a semilog equation, and the rights and category are binary variables, the value of e^{β_i} is of economic significance. For instance, when *CopyPartial* = 1, the value of the price of the ebook changes by a (multiplicative) factor of =1.5, or according to our estimated price equation, ebook prices are 50% higher. When *PrintParial* = 1, the value of eBook changes by a (multiplicative) factor of 0.67, or according to our estimated price equation, ebook prices are 33 percent lower. We do not focus on the magnitude of these changes, but on their predicted direction.

whose preferences are such that they favor digital goods over physical goods, and thus, the ability to create printed copies of their ebooks is unlikely to have much of a positive direct effect on ebook quality. This right aims to embed an aspect of the physical consumption experience into the digital good, rather than enhancing digital quality. Furthermore, 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 identical to the original ebook in quality). This results in a negative differential piracy effect, which our results show dominates any positive direct quality effect[11].

In contrast, copying rights enhance the digital experience associated with purchasing an ebook, 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 thus suggest that the direct quality effect that copying rights have on ebook quality dominate the piracy effect. This is natural, since copying is a right that enhances the digital consumption experience of the ebook, thereby increasing the digital quality valued by the customers for the good, rather than aiming to embed a characteristic of a tangible good into its digital version.

Moreover, the extent to which the former effect dominates the latter is stronger for partial copying rights. This can attributed to the fact that the positive direct effect is likely to be similar to both cases – the ability to copy relevant parts from a book to one's document. However, the ability to "copy all" will have a stronger negative differential piracy effect, since it does facilitates the creation of complete pirated ebooks, albeit of fairly low quality.

As currently implemented, the read-aloud right does not make pirating an ebook any easier or

[11] An alternative (and potentially intuitive) explanation for this result might be that granting each right does increases price, but less expensive books simply grant all rights, thereby creating the impression that the granting of the right is associated with a decrease in price. However, this is inconsistent with our empirical results, because our estimation shows that some rights are associated with an increase in price and some rights are associated with a decrease in price. We have also compared the distributions of the tangible book prices with printing rights and without, using the Kolmogorov-Smirnov test, which indicates that their distributions do not vary significantly.

more difficult. It is implemented using software that renders text into a computerized voice, and while it may permit the creation of MP3-based "books on tape," they are likely to be of fairly low quality. One can create higher-quality pirated versions by buying the book-on-tape (on which the audio rendering is of far higher quality) and digitizing its content. Similarly, the coefficients of our variable corresponding to lending rights suggest that while the right to lend may lead to a negative differential piracy effect, it also has a strong direct effect on quality that makes it overall value have a positive effect on price.

Our estimates of the variation in ebook price across categories have an interesting interpretation in light of our qualitative analysis of technological sophistication in section 2. Recall that if one considers the extent to which the distribution of taste values y is right-skewed as a measure of the technological sophistication of a market, an increase in this level may be associated with an increase in the price of digital goods relative to their tangible counterparts if technological sophistication is low. If one orders the categories in decreasing order of their estimated category coefficients (or in decreasing order of their ebook prices after controlling for tangible book price and the effects of digital rights) – Computers; Business and Economics; History; Biography and Autobiography; Fiction; and Children's and Young Adult Fiction – this resembles a partial ordering of the technological sophistication of these categories. At the very least, it seems reasonable to assert that the category with the highest adjusted ebook prices – Computers – is also the one with the most technological sophisticated market. The coefficients of the associated categorical variables are consistent with overall preferences of consumers for books being skewed towards preferring tangible books at this point in time. We do not have data that can validate that this effect is in fact due to these categories having customers who vary in their technological sophistication (for instance, from a survey of readers), although this represents an interesting direction for future research. The explanatory power of our model is fairly high (the R^2 of 66.2% is more than double that of the model that does not include digital rights and categories), thereby indicating that the extent to which digital rights are granted explains a substantial portion of the variation in ebook prices.

We have found these predicted signs of our estimated coefficients to be robust across a number

of other estimations. We estimate how $Log[EPrice]$ varies with changes in $Log[TPrice]$ and digital rights (that is, without the category variables) and find that all but one of our coefficients are statistically significant, the signs of these coefficients for each digital right are the same as reported in Table 3.3, and the R^2 of 50% suggests that digital rights by themselves do increase the explanatory power of the model[12]. We have also estimated the equation (3.2) using data sets collected at other points in time, finding that the signs of each coefficient are preserved.

In the final stage of our analysis, we investigate the interaction between our categories and our digital rights, towards understanding whether granting different digital rights led to different changes in ebook prices across categories. We do so by estimating the following model:

$$Log[EPrice] = \alpha + \beta_1 Log[TPrice] + \sum_{i=2}^{7} \beta_i R_i + \sum_{i=1}^{5} \gamma_i K_i + \sum_{i=2}^{7} \sum_{j=1}^{5} \phi_{ij} R_i K_j, \qquad (3.3)$$

where we label our six digital rights variables (*CopyAll, Copy Partial,...*) R_2 through R_7 for expository convenience. Our estimates are summarized in Table 3.4.

About half of our interaction coefficients are significant at the 5% level. The main effects remained consistent with our original model, thus strengthening our findings, although the coefficients of the variables associated with CopyAll are not significant. A few additional insights also emerge from this estimation. For example, granting printing rights has the most detrimental effect on ebook price for the Computer category, perhaps reflecting a higher propensity of consumers in this category to pirate such books if it is easy to do so. Granting these rights continues to have a negative effect on ebook price across five of our six categories. The exception was the Fiction category, for which granting these rights results in a positive impact on ebook price. In contrast, this was also the category for which Lend and CopyPartial rights have a negative effect on ebook price, and the former is probably a consequence of books in this category being read once, and unlike other categories, having little "reference" value. Furthermore, the ReadAloud right has a significant and high positive effect on price for the Computers and Business categories, perhaps reflecting the

[12] In a model that estimates how $Log[EPrice]$ varies with changes in just the digital rights, two of our coefficients are not statistically significant, their signs are the same as reported in Table 3.3, and the R^2 is 37%.

	No interaction	*Business*	*History*	*Fiction*	*Young*	*Biography*
No interaction		−0.21 (0.15)	−0.41* (0.18)	−0.94*** (0.14)	−1.15*** (0.21)	−0.57*** (0.17)
PrintAll	−0.70 *** (0.11)	0.17 (0.14)	−0.09 (0.21)	0.77*** (0.13)	0.51 * (0.21)	0.20 (0.22)
PrintPartial	−0.82 *** (0.14)	0.56** (0.18)	−0.03 (0.18)	1.21 ** (0.22)	0.80 * (0.34)	0.26 (0.20)
CopyAll	−0.02 (0.12)	−0.16 (0.14)	−0.03 (0.18)	0.06 (0.12)	0.07 (0.20)	0.10 (0.17)
CopyPartial	0.29* (0.12)	0.01 (0.15)	0.38** (0.18)	−0.44 * (0.2)	0.03 (0.33)	0.31 (0.19)
Lend	0.30 * (0.10)	−0.34* (0.13)	0.68* (0.38)	−0.54 ** (0.18)	−0.31 (0.20)	0.43 (0.31)
Read	1.05 *** (0.15)	−0.28 (0.17)	−0.54 ** (0.19)	−0.86 *** (0.16)	−0.92 *** (0.22)	−0.81 *** (0.20)
Log[TPrice]	0.24 *** (0.01)					
Constant	2.07 *** (0.15)					
$R^2 = 68.6\%$	* significant with $p \leq 0.05$ ** significant with $p \leq 0.01$ *** significant with $p \leq 0.001$					

Table 3.4: How digital rights and categories interact. The coefficients in the first column (labeled "No interaction") correspond to the estimates of α and the β_i coefficients, in the first row correspond to the estimates of the γ_i coefficients, while the other entries are estimates of the ϕ_{ij} coefficients. For example, the entry in the row labeled PrintPartial and the column labeled Business corresponds to the coefficient estimated for the variable PrintPartial*Business.

positive direct effect granting this kind of read-aloud right has on quality for these category (the fact that the voice is digital and lacks intonation may not matter as much for computer or business books as it does for books in the Fiction).

The intuitive explanations for the magnitude and signs of these coefficients might be of independent interest for the publishing industry, but from our point of view, they merely serve to further validate the link between the effects our theoretical model ascribes to changes in digital rights, and the presence of estimated coefficients that reflect these effects in our data.

4. Conclusions and ongoing work

We have presented a model characterizing the choice of digital rights and pricing for digital goods offered by a firm who also sells a physical version of the digital good, the granting of digital rights for which may also lead to an increase in digital piracy. The predictions of this model are validated by our empirical results, and suggest important new guidelines for managers in industries that are progressively being digitized. As the pace of industry transformation by information technology accelerates over the coming years, managing such transitions will become central to continued business success across a wider variety of industries, thereby increasing the value of IS research that contributes to our understanding of this transformational power of IT (Agrawal and Lucas, 2005), and of how the Internet affects markets and industry structure (Ellison and Ellison, 2005). We hope our study makes such a contribution.

Summarizing our key results:

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

2. When granting digital rights in the presence of digital piracy, the value of these rights is governed by two opposing effects: *a direct quality effect*, or how much the granting of the right increases willingness to pay for the legal digital goods, and a *differential piracy effect,* that measures how much it reduces the relative surplus of consuming legal tangible and digital goods through its inducing digital piracy.

– When the direct quality effect balances the differential piracy effect, then granting the right is beneficial to the seller. It raises prices and profits from the sales of digital goods. This is more likely to be associated with rights that enhance the digital consumption experience of the good, rather than replicating the physical consumption experience, since consumers who purchase the digital good belong to the segment who value digital quality. If they did value the physical consumption experience, these consumers would have purchased the tangible good in the first place.

– When the differential quality effect associated with the right exceeds the direct quality effect significantly, granting the right may affect the seller's pricing power adversely, since consumers who prefer digital goods over tangible goods need to be offered sufficient surplus to ensure that they do not simply resort to using a pirated copy of the good. We conjecture that this kind of right is generally likely to be one that attempts to *replicate* the physical consumption experience digitally, for two reasons. First, such rights do not contribute to those aspects of consumption that the segment who purchases digital goods values, and the direct quality effect is therefore likely to be low. Furthermore, rights that attempt to replicate the physical consumption experience digitally typically involves providing an opportunity to render the digital good into a *tangible* artifact (onto a CD, onto paper), and facilitating such rendering is likely to be associated with an increase in the quality and distribution of pirated copies, since it involves relaxing, to some extent, technological protection that prevent copying.

3. In testing our theory's predictions using pricing and digital rights data for over 3000 ebooks, we find that each of four separate digital rights have both a statistically and economically significant impact on ebook prices, after controlling for tangible book price levels and category. Our results also show that a right which aims to replicate physical consumption characteristics – printing – is the one with a negative impact on ebook prices, while rights enhancing the consumption experience by exploiting the fact that the book is now a digital good – copying and reading aloud via digital audio – have a positive impact on prices. This provides empirical support for the discussion in (2) above.

4. The variation in ebook prices we observe across categories suggests that an increase in technological sophistication of consumers is associated with an increase in the price of legal digital goods. Such a possibility is discussed qualitatively by our theoretical model as being associated with an industry in an early stage of transformation from tangible artifacts to digital goods, which seems consistent with the book publishing industry. Technological sophistication is a fairly broad term, however, and there are many aspects to it that our model and data have not captured. Studying this relationship more closely is an interesting direction for future research, since it may

prove important to choosing the timing of the transition, to the extent that sellers in an industry have control over it.

There are many other directions for future research suggested by our paper. We assume that tangible book prices are exogenous, which, apart from being simpler, reflects a model consistent with the industry our data is from. Our conversations with publishers have suggested they still price the tangible book independently, and then key the ebook price off this. However, a model of the simultaneous choice of tangible and digital pricing in the presence of piracy would be a useful extension, especially as digital goods gain prominence in more industries. While most digital goods in other industries do not currently have the rich variation in rights across products that made our empirical study of ebook pricing viable, they are likely to in the near future, and may present the opportunity for studies similar to ours across other industries like music and broadcast video. There are also indications that new forms of pricing are forthcoming in the ebook industry, most notably pay-per-use pricing from Amazon.com and Google. As this industry matures, it would be interesting to examine how our empirical findings evolve, and this is a direction of research we hope to pursue in the future.

5. References

1. Agrawal, R. and Lucas, H., 2005. The Information Systems Crisis: Focusing on High-Visibility and High-Impact Research. *MIS Quarterly* 29, 381-398.

2. Bhattacharjee, S., Gopal, R. and Sanders, G., 2003. Digital Music And Online Sharing: Software Piracy 2.0? *Communications of the ACM* 46 (7), 107-111.

3. Bakos, Y., Brynjolfsson, E. Lichtman., D., 1999. Shared Information Goods. *Journal of Law and Economics* 42, 117–156.

4. Belleflamme, P., 2002. Pricing Information Goods in the Presence of Copying. Mimeo, Queen Mary University of London.

5. Berndt, E., 1990. *The Practice of Econometrics.* Addison-Wesley Press

6. Chellappa, R., and Shivendu, S., 2003. Managing Piracy: Pricing and Sampling Strategies for Digital Experience Goods in Vertically Segmented Markets. *Information Systems Research* (forthcoming).

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

8. Ellison, G. and Ellison, S., 2005. Lessons about Markets from the Internet. *Journal of Economic Perspectives* 19 (2), 139-158.

9. Haruvy, E., Mahajan, V., and Prasad, A., 2004. The Effect of Piracy on the Market Penetration of Subscription Software. *Journal of Business* 77, S81-S108.

10. Gayer, A., and Shy, O., 2003. Internet and Peer-to-Peer Distributions in Markets for Digital Products. *Economics Letters* 81, 51-57.

11. Gu, B., and Mahajan, V., 2004. The Benefits of Piracy: A Competitive Perspective. *Sixteenth Workshop on Information Systems and Economics.*

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

13. Hui, K., Png., I., 2003, Piracy and the Legitimate Demand for Recorded Music. *Contributions to Economic Analysis and Policy* 2(1).

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

15. Mossberg, W. 2003. Apple's new service beats illegal free sites. The Wall Street J. (April 30) D1.

16. Mussa, M., and Rosen, S., 1978. Monopoly and Product Quality. *Journal of Economic Theory* 18, 301-317.

17. Peitz, M., and Waelbroeck, P., 2003. Piracy of Digital Products: A Critical Review of the Economics Literature. CESifo Working Paper Series No. 1071

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

19. Poddar, S., 2004. Software Piracy: When Protection is Optimal to the Software Developer. Working Paper 0312, National University of Singapore.

20. Prasad, A., and Mahajan, V., 2003. How Many Pirates should a Software Firm Tolerate? An Analysis of Piracy Protection on the Diffusion of Software. *International Journal of Research in Marketing* 20, 337-353.

21. Shy, O., Thisse., J. F., 1999. A Strategic Approach to Software Piracy. *Journal of Economics and Management Strategy* 8, 163–190

22. Snir, E. M., 2003. The Record Industry in an Era of File Sharing: Lessons from Vertical Differentiation. In *Proceedings of the 24th International Conference on Information Systems*, Seattle, Washington, USA..

23. Sundararajan, A., 2004. Managing Digital Piracy: Pricing and Protection. *Information Systems Research* 15, 287-304.

24. Takeyama, L. N., 1994. The Welfare Implications of Unauthorized Reproduction of Intellectual Property in the Presence of Network Externalities. *Journal of Industrial Economics* 42, 155-166.

25. Viswanathan, S., 2005. Competing Across Technology-Differentiated Channels: The Impact of Network Externalities and Switching Costs. *Management Science* 51, 483-496.

26. Zentner, A., 2003. Measuring the Effect of Music Downloads on Music Sales. http://home.uchicago.edu/~alezentn/musicindustrynew.

A. Appendix: Proofs

The proofs in this appendix use the following additional notation:

$$\Delta Q(y_a, y_b, g(y)) \equiv \int_{y=y_a}^{y_b} \left[f_v(g(y)) f_h(y) \frac{dg(y)}{dp_D} \right] dy. \tag{A.1}$$

The proof of Proposition 1 follows from the following lemma:

Lemma 1. *The optimal price of the digital good, p_D^*, is uniquely defined by the following equation:*

$$p_D^* = [p_T - c] + \frac{q_D(p_D^*)}{-\left[\dfrac{dq_D(p_D^*)}{dp_D}\right]}. \tag{A.2}$$

where:

$$\frac{q_D(p_D^*)}{-\left[\dfrac{dq_D(p_D^*)}{dp_D}\right]} = \frac{1 - F_H(y_D(p_D^*)) - Q(y_D(p_D^*), 1, \theta_{TD}(y, p_D^*))}{\Delta Q(y_D(p_D^*), 1, \theta_{TD}(y, p_D^*))} \tag{A.3}$$

is the ratio of demand for the digital good to the slope of the demand for the tangible good, evaluated at the optimal price.

Proof. First, differentiating both sides of (2.9) and (2.12) with respect to p_D and using (A.1) yields:

$$\frac{dq_T(p_D^*)}{dp_D} = \Delta Q(y_{TD}(p_D^*), 1, \theta_{TD}(y, p_D)) \tag{A.4}$$

$$\frac{dq_D(p_D^*)}{dp_D} = -\Delta Q(y_{TD}(p_D^*), 1, \theta_{TD}(y, p_D)) \tag{A.5}$$

The monopolist chooses p_D^* to solve:

$$\max_{p_D}[p_T - c]q_T(p_D) + p_D q_D(p_D). \tag{A.6}$$

The first-order condition for (A.6) is:

$$[p_T - c]\frac{dq_T(p_D)}{dp_D} + p_D\frac{dq_D(p_D)}{dp_D} + q_D(p_D) = 0, \tag{A.7}$$

which in conjunction with (A.4), (A.5) and (2.12) yields equations (A.2) and (A.3). ■

Proof of Proposition 1

If the demand for the digital good is positive, then

$$[1 - F_H(y_D(p_D^*))] > Q(y_D(p_D^*), 1, \theta_{TD}(y, p_D^*)),$$

and thus the expression in (A.3) is strictly positive.

Proof of Proposition 2

For a fixed level of digital quality s_D, denote the optimal price of the digital good as $p_D^*(s_D)$, and the corresponding demand for the digital good as $q_D(p_D^*(s_D), s_D)$ The seller's problem of choosing digital rights is therefore equivalent to determining the level s_D that maximizes profits:

$$\max_{s_D}\Pi(s_D) \equiv [p_T - c][1 - q_D(p_D^*(s_D), s_D)] + [q_D(p_D^*(s_D), s_D)], \tag{A.8}$$

and then choosing a combination of rights that implements this level of quality. However, notice that using the envelope theorem, it follows that:

$$\frac{d\Pi(s_D)}{ds_D} = [p_D^*(s_D) - (p_T - c)]\frac{\partial q_D(p_D^*(s_D), s_D)}{\partial s_D}. \tag{A.9}$$

From (A.2) the first term on the RHS is strictly positive. Furthermore, differentiating both sides of (2.7) with respect to s_D and rearranging yields:

$$\frac{d\theta_{TD}(y, p_D)}{ds_D} = \frac{u_1(s_D, \theta_{TD}(y, p_D))}{-u_2(s_D, \theta_{TD}(y, p_D))} < 0 \tag{A.10}$$

for all y and p_D, since $u_1(s,\theta) > 0$ and $u_2(s,\theta) > 0$. (A.10) in turn implies that

$$\frac{\partial q_D(p_D^*(s_D), s_D)}{\partial s_D} > 0, \tag{A.11}$$

which, in conjunction with (A.9) implies that the seller's profits are strictly increasing in s_D. Since

$$s_D = S_D(r_1, r_2, ..., r_n), \tag{A.12}$$

and S_D is strictly increasing in each of its arguments, the result follows.

Our next lemma characterizes the optimal price of the digital good in the presence of piracy.

Lemma 2. *The optimal price p_D of the digital good is:*

$$p_D^* = [p_T - c]\delta(p_D^*) + \frac{q_D(p_D^*)}{-\left[\dfrac{dq_D(p_D^*)}{dp_D}\right]}, \tag{A.13}$$

where:

$$\delta(p_D^*) = \frac{\Delta Q(y_{TD}(p_D^*), y_{PD}(p_D^*), \theta_{TD}(y, p_D^*))}{\Delta Q(y_{TD}(p_D^*), y_{PD}(p_D^*), \theta_{TD}(y, p_D^*)) - \left[\frac{d\theta_{TP}(y_{PD}(p_D^*))}{dy}\frac{dy_{PD}(p_D^*)}{dp_D}\right][[1 - F_h(y_{PD}(p_D^*))]f_v(y_{PD}(p_D^*))]} \tag{A.14}$$

is the discount on the digital good induced by piracy, and:

$$\frac{q_D(p_D^*)}{-\left[\dfrac{dq_D(p_D^*)}{dp_D}\right]} = \frac{1 - F_h(y_D(p_D^*)) - Q(y_D(p_D^*), y_{PD}(p_D^*), \theta_{TD}(y, p_D^*))}{\Delta Q(y_D(p_D^*), y_{PD}(p_D^*), \theta_{TD}(y, p_D^*)) - \left[\frac{d\theta_{TP}(y_{PD}(p_D^*))}{dy}\frac{dy_{PD}(p_D^*)}{dp_D}\right][[1 - F_h(y_{PD}(p_D^*))]f_v(y_{PD}(p_D^*))]} \tag{A.15}$$

is the ratio of demand for the digital good to the slope of the demand for the tangible good, evaluated at the optimal price.

Proof. Differentiating both sides of (2.19) and (2.20) with respect to p_D, cancelling out common terms, and simplifying using (A.1) yields:

$$\frac{dq_T(p_D^*)}{dp_D} = \Delta Q(y_{TD}(p_D^*), y_{PD}(p_D^*), \theta_{TD}(y, p_D)) \tag{A.16}$$

$$\begin{aligned}\frac{dq_D(p_D^*)}{dp_D} &= -\Delta Q(y_{TD}(p_D^*), y_{PD}(p_D^*), \theta_{TD}(y, p_D^*)) \\ &\quad + \left[\frac{d\theta_{TP}(y_{PD}(p_D^*))}{dy}\frac{dy_{PD}(p_D^*)}{dp_D}\right][[1 - F_h(y_{PD}(p_D^*))]f_v(y_{PD}(p_D^*))]\end{aligned} \tag{A.17}$$

The monopolist chooses p_D^* to solve:

$$\max_{p_D}[p_T - c]q_T(p_D) + p_D q_D(p_D). \tag{A.18}$$

The first-order condition for (A.18) is:

$$[p_T - c]\frac{dq_T(p_D)}{dp_D} + p_D\frac{dq_D(p_D)}{dp_D} + q_D(p_D) = 0, \tag{A.19}$$

or

$$p_D = \frac{[p_T - c]\frac{dq_T(p_D)}{dp_D}}{-\left[\frac{dq_D(p_D^*)}{dp_D}\right]} + \frac{q_D(p_D^*)}{-\left[\frac{dq_D(p_D^*)}{dp_D}\right]}$$

which in conjunction with (A.16) and (A.17) yields equations (A.14) and (A.15). ■

Proof of Proposition 3

By the definition of $y_{PD}(p_D)$,

$$\theta_{TP}(y_{PD}(p_D), p_D) = \theta_{TD}(y_{PD}(p_D), p_D) \tag{A.20}$$

Differentiating both sides of (A.20) with respect to p_D and rearranging yields:

$$\frac{dy_{PD}(p_D)}{dp_D} = \frac{\frac{d\theta_{TD}(y_{PD}(p_D),p_D)}{dp_D}}{\left(\frac{d\theta_{TP}(y_{PD}(p_D),p_D)}{dy} - \frac{d\theta_{TD}(y_{PD}(p_D),p_D)}{dy}\right)} \tag{A.21}$$

Now, differentiating both sides of (2.7) with respect to p_D and rearranging yields

$$\frac{d\theta_{TD}(y, p_D)}{dp_D} = \frac{1}{u_2(s_D, \theta_{TD}(y, p_D))} > 0, \tag{A.22}$$

since $u_2(s, \theta) > 0$. Next, (A.20) in conjunction with the fact that $u_{12}(s, \theta) > 0$ implies that

$$u_2(s_P, \theta_{TP}(y, p_D)) < u_2(s_D, \theta_{TD}(y, p_D)). \tag{A.23}$$

Finally, differentiating both sides of (2.7) and (2.13) with respect to y and rearranging yields:

$$\frac{d\theta_{TD}(y, p_D)}{dy} = \frac{w_1(y) + w_1(1-y)}{-u_2(s_D, \theta_{TD}(y, p_D))}, \tag{A.24}$$

$$\frac{d\theta_{TP}(y, p_D)}{dy} = \frac{w_1(y) + w_1(1-y)}{-u_2(s_P, \theta_{TP}(y, p_D))}, \tag{A.25}$$

which, using (A.23) and the fact that $w_1(y) > 0$ implies that

$$\left[\frac{d\theta_{TP}(y_{PD}(p_D), p_D)}{dy} - \frac{d\theta_{TD}(y_{PD}(p_D), p_D)}{dy}\right] < 0. \tag{A.26}$$

(A.21), (A.22) and (A.26) imply that

$$\frac{dy_{PD}(p_D)}{dp_D} < 0., \tag{A.27}$$

(A.22) and (A.27) therefore imply that

$$-\left[\frac{d\theta_{TP}(y_{PD}(p_D^*))}{dy}\frac{dy_{PD}(p_D^*)}{dp_D}\right][[1 - F_h(y_{PD}(p_D^*))]f_v(y_{PD}(p_D^*))] > 0 \tag{A.28}$$

so long as there is non-zero demand for the pirated good. Also, from (A.1),

$$\Delta Q(y_{TD}(p_D^*), y_{PD}(p_D^*), \theta_{TD}(y, p_D^*)) > 0.$$

Inspecting the definition of $\delta^*(p_D)$ in (A.14), this implies that the denominator of the RHS of (A.14) is strictly greater than its numerator, which completes the proof.