Is the Copyright Monopoly a Best-Selling Fiction?

Stan J. Liebowitz

School of Management

University of Texas at Dallas

June 2008

Abstract:

Copyright is provided as an incentive for authors to create new works. This putative benefit, however, is also thought to come at the expense of a putative monopoly deadweight loss since copyright appears to provide a monopoly to the copyright owner. This paper attempts to determine the impact of copyright on the prices of books and thus whether and to what extent copyright confers a monopoly pricing component. The empirical findings indicate that the price of copyrighted works may not be higher than the prices of non-copyrighted peers and if they do have higher prices the differences are quite small. An explanation for no price differential is based on uniform prices that are found in many entertainment industries even when variations in demand seem somewhat predictable. If a small price increase is accepted as more likely interpretation of results I calculate a range of deadweight losses based upon reasonable market assumptions and find the size of the deadweight loss is very small compared to industry revenue. Under either scenario, the underconsumption problem that plays such a prominent role in standard copyright analysis is seen to be either irrelevant or miniscule compared to the deadweight losses from a leading alternative system for distributing creative works.

Economic analyses of copyright most often focus on balancing what is known as the underproduction/underconsumption tradeoff, or in somewhat clearer nomenclature, the access/incentive tradeoff. Intellectual property laws are thought to increase incentives for creation by reducing competition in selling the particular intellectual property, allowing the copyright owner greater profits with which to pay for the costs of creating the work. On the other hand, forbidding someone else from selling an exact copy of the work is thought likely to raise the price of the work and therefore decrease consumer access. These models are so well-known that their implications, such as the welfare loss due to underconsumption of individual titles, are generally taken for granted. The internal logic of these models is unassailable.

This paper examines the empirical foundations for the belief that the copyright monopoly is actually an economic monopoly in the sense of creating a deadweight loss in the consumption of the work. In recent years the putative welfare losses from copyright have taken on increased importance due to the rise of the Internet and difficulties with traditional business models for copyright-based industries.

A great deal of discussion, presently and in the past, has focused on the proper duration and scope of intellectual property laws. After the American copyright law was extended in the late 1990s there were heated discussions, lawsuits, amicus briefs by economists, and numerous writings on both sides of the issue.[1] Lawsuits against various forms of copying, such as Napster, Aimster, Grokster, and various other file-sharing programs, have brought attention and fanned disputes about the proper boundaries of copyright protection. As the prerecorded music industry continues to shrink, new business models are being suggested that depend on alternatives to copyright law.[2] Patent law has had some of the same debate, although there are no rock stars to liven up that discussion.[3] Nevertheless, the underlying economic model has not changed.

Alternatives to traditional copyright have been well-known to economists and lawyers for several centuries. Prizes and awards can be given either privately or by the government; government could provide revenues to creators of intellectual products and then provide those products freely to the market; being first in a market might provide sufficient revenue that copyright is unnecessary; becoming famous might provide sufficient reward and revenue from ancillary markets; [4] a compulsory license might be granted allowing

---

[1] For example, see Akerlof et al. (2003) and Liebowitz and Margolis (2005).

[2] See Fisher (2004) for the most detailed proposal out of many put forward by law school ademics. The Songwriters Association of Canada has adopted a scaled-down version of Fisher's basic idea. See http://arstechnica.com/news.ars/post/20071205-canadian-songwriters-propose-monetizing-p2p-in-canada.html .

[3] See Kremer (1998) for one notable example.

[4] Sometimes it is suggested that other accoutrements engendered by the creation of successful works—fame, for example—provide sufficient incentive to generate this creative activity without any direct financial payments. Perhaps, the theory goes, some musicians could make sufficient revenues from performing, that revenues from CDs would be irrelevant; but there are few or no instances of successful musical groups providing free recorded music in order to increase concert revenues. Nor do we see athletes or actors willing to work for free so they can cash out in other markets, such as endorsements.

any publisher to print a book so long as they pay a fixed rate to the author, and so on.[5] A related but different claim that has achieved some recent prominence is that creators receive almost none of the royalties generated by their creations, negating the incentive effect.

Virtually all economic discussion of intellectual property law and its alternatives depend on the size of the incentive/access tradeoff. How much of a deadweight loss do intellectual property laws create on the access side of the market? How much incentive do they provide, and how is that incentive transformed into actual creation?

Alas, little is known about the particulars of this tradeoff.

Economists are rightly interested in maximizing social welfare. Central to the economics of this debate is the concept that intellectual property laws create monopolies, real economic monopolies. If there were no monopolies then there would be no access or underconsumption issues to worry about. Without an economic monopoly, the standard model says, there would be no gains to creators. Without a monopoly, the standard model says, there would be no copyright-induced inefficiency.[6]

The purpose of this paper is to examine the existence and size of the monopoly price differential in the case of books. I endeavor to infer the extent of monopoly power by measuring the price increase, if any, caused by copyright. Amazingly, this has, to my knowledge, never been previously examined. I find, perhaps surprisingly, that the monopoly deadweight loss is zero or close to it. The results are, of course, limited to this particular market under investigation, which is books, although it seems likely that any such results may apply to other copyright markets with similar characteristics.

I. The Nature of the Copyright "Monopoly"

It is an economic truism that prices in monopoly markets are higher than prices in competitive markets. These higher prices lead to a decline in consumption and a resulting deadweight loss because consumption is reduced to a level whereby the last unit consumed is worth more than the cost of producing it and additional units would share this attribute but are not brought to market.

Economists, along with almost everyone else, tend to equate intellectual property protection with monopoly. Unlike patents, however, which protect an inventor against later independently created but similar inventions, copyright, for all practical purposes, merely protects a work from unauthorized versions of itself. Copyright allows the creator, or anyone to whom he has assigned his copyright, to be the exclusive agent capable of making reproductions of the work (which has been extended to various forms of 'reproduction' including public performance, electronic transmissions to the public and

[5] See Plant (1934) for an early economic analysis of many of these points.

[6] Later in the paper I will propose circumstances whereby the copyright monopoly, in a partial sense at least, may increase efficiency by lowering costs and may be able to provide the author with payments for creation without the creation of an underconsumption deadweight loss. The empirical results will lend credence to the view.

so forth) but it does not provide any protection from independently created competing works no matter how similar they may be.

Although there is little doubt that copyright, by definition, grants the copyright owner a monopoly on making copies of the owner's particular work, Kitch (2000) has reminded us that this does not mean that there is necessarily any economic monopoly or any deadweight loss. After all, every firm has a monopoly on their own product. Farmer Smith has a monopoly on Farmer Smith's wheat. The problem (for Farmer Smith) is that farmer Smith's wheat is just like every other farmer's wheat, so that his wheat growing doesn't provide any economic monopoly.

The heart of the matter, from an economic perspective, is whether there are close substitutes for the nominal monopoly.[7] Tiger Woods has a monopoly of sorts in the activity of playing golf. Other golfers are not identical, nor as good, as he is. There is free entry into the market for professional golfers and there is plenty of competition, but we would not say that the result looks like the idealized wheat market found in economic textbooks where all farmers earn a zero economic returns. It is another question entirely whether the government should break some of Tiger Woods fingers to increase competition and promote economic efficiency.

It will be useful, at this juncture, to distinguish between the two separate components of this industry—creating and publishing. These are really two separate markets, authoring and printing and either could be monopolistic or competitive. If publishing were a complete monopoly, as was the case when the British government provided such a license to the Stationers Company, there would be no need for copyright. By contrast, it is easy to imagine an industry where publishers wield no monopoly power. In such a world we might think of book publishing as being perfectly competitive, with any economic profits being competed away. This would mean that authors would reap any above normal profits from the sales of their books and indeed, the entire industry rent would go to those authors with hard to replace talent. Nevertheless, as easy to imagine as it might be, I do not believe that this is how most copyright industries are normally portrayed. Instead, a more common view, a seemingly very common view, is that copyright industries consist of large powerful corporations that usurp most of the industry rents, leaving the creative artists with little to show for their efforts. [8]

The act of creating is often treated as being very competitive since there are many individuals attempting to become published authors. Even if we limit ourselves to those

---

[7] Of course, different books (titles) are not exactly like one another. It has been suggested that models of monopolistic competition are appropriate since there is free entry but the items in question are not identical to one another, which I believe is a correct point. See Yoo (2004) for an insightful discussion of applying monopolistic competition to models of copyright. When I use the term 'competition' I could as easily use the term 'monopolistic competition' since free entry and zero profit are endemic to both.

[8] Most horror stories about rapacious corporations relate to the recording industry although the book industry shares many of the same characteristics. See Ku (2002) starting on page 306 and Connolly and Krueger (2006) who claim that recording artists do not generate much if any revenue from record contracts. Although this may be true for the many works which achieve almost no market success, it is clearly false for the set of more successful works, as can be verified by looking at the income of J. K. Rowling or Paul McCartney.

individuals who have secured contracts with professional publishers and agents, the great majority of such authors are not likely to wield much if any economic monopoly. Nevertheless, it seems clear that there is a small minority of artists, often household names, who wield clout by virtue of their difficult-to-imitate talent and this talent provides them monopoly power.

Regardless of where the locus of monopoly might be housed, economists have been unanimous in their depiction of a monopoly deadweight loss on the consumption side of this market.[9] The deadweight loss involves consumers who would like to consume the work but are restricted from doing so by the supra competitive price in the market. Some economists and legal scholars also focus on individuals interested in reusing portions of the work in their own creative endeavors as opposed to pure consumption. Our focus will remain on the deadweight loss from reduced consumption.

With a focus on this putative deadweight loss, the key becomes how much consumption might be reduced by the existence of copyright in the book market. Although we cannot measure reduced consumption, we can measure the cause of the reduced consumption, i.e. the increase in price that causes consumption to decline below the optimal level.

II. The Structure of the Publishing Industry

Book publishing is the oldest of the industries relying on copyright. After Gutenberg invented the printing press in the mid 1400s, various government-granted monopolies on printing existed in various European countries, such as the monopoly given to the Stationer's Company in England, chartered in 1557. The Act of Queen Anne in 1709, which gave the copyright to the author for a term of 14 year, is considered the first modern style copyright law. Current copyright law in most industrialized countries provides a duration equal to the author's life plus 70 years.

Book publishing encompasses several major business categories. According to Greco (2005) trade books (fiction and nonfiction) compose the largest component, with the category often separated into adult and juvenile. The next largest segment consists of professional and scholarly publications. Educational books represent two more large categories, with the first consisting of books intended for students in grades through high school and the other part specializing in books intended for college and university students. Finally we find mass market paperbacks and bookclubs which are treated as separate categories, less because of content then for their different distribution and pricing schemes.

The barriers to entry in the book publishing industry appear to be relatively low and falling. Greco (2005) reports that RR Bowker, publisher of the Books In Print, follows 53,000 American publishers although the Census Bureau, using a more stringent

[9] Landes and Posner (2001) have a more nuanced view than most analysts, suggesting that that there are many costs in using the copyright system which reduce the optimal duration and scope of copyright independent of the deadweight loss. Still, the deadweight loss from monopoly plays an important role in their analysis.

definition of "publishing establishments" reports less than one tenth of this number. The relatively recent innovation of printing on demand allows small start ups to produce books with relatively small initial investments. In spite of the large number of industry sellers, it would be wrong to describe the industry as atomistic since, according to Greco (2000), the top 20 publishers accounted for over 70% of the consumer book revenues in the mid 1990s and the four largest firms accounted for approximately 38% of the revenues. Even so, this level of concentration is not terribly high relative to many other industries and is non inconsistent with the possibility of competitive behavior.

The market for the books appears to allow free entry both in the authorship of books and in printing. It is estimated that more than 150,000 new titles are produced in the US alone every year. When there is free entry into a market we do not expect to find monopoly profits except for those cases where some firms have an unusually valuable attribute that cannot be easily imitated, if at all. In the case of authoring, such monopoly power may survive free entry if the author has unusual talent. For this reason, it is unlikely that the copyright "monopoly" leads to any economically meaningful monopoly except in those rare cases where competing works cannot match the appeal of an unusually creative effort; in other words, a Tiger Woods of literature. The cause of the monopoly in this case is the confluence of the copyright law and the inherent talent of the author.

III. The Methodology

One fairly straightforward methodology for measuring the impact of copyright on price, and the one adopted here, is to compare prices for copyrighted and non-copyrighted books that are otherwise similar. The concept itself is simple, but its implementation is not.

How does one find sets of similar books where some are under copyright and others are not? I decided to use current books, or more accurately since I have dawdled in writing this paper, almost current books, because the current market is the one we are most familiar with and the Internet allows some quick spot checking.

To start, books that have lost their copyright tend to be rather old since copyright length since 1978 has been set to exist for any title for 50 years after the author has died (70 years with the 1998 'Sonny Bono' revision to the copyright law). Books that have lost their copyright are, therefore, quite old. Prior to the 1976 copyright law (which took affect in 1978), copyright in the United States lasted for an initial term of 28 years followed by another term of 28 years if the copyright owner renewed the copyright.[10]

Examinations by Landes and Posner (2003) and Rappaport (1998) revealed that only a small percentage of copyrighted books were renewed when the original 28 years term had ended. That would mean that many books written prior to the late 1940s would have had

[10] The original initial and renewal term was 14 years prior to 1831 when the initial term was increased to 28 years. In 1909 the renewal term was also increased to 28 year. The term of copyright registration was increased to 47 years in 1962 and 67 years in 1998. Although copyright registration (and renewal) are no longer necessary, there are potential benefits to registering copyright, such as the ability to ask for statutory damages and reimbursement of attorney fees in lawsuits.

their copyright expire. But those books had their copyright lapse because 28 years after initial publication they no longer appeared to have a future market value that was sufficient to cover the small (1$) cost of copyright renewal during this period of analysis.[11]

To find old book titles that still have new copies being sold today and which therefore have measurable market prices, we need to examine books that were unusually successful upon initial publication. Liebowitz and Margolis (2005) mention that bestsellers tend to have very long durations in the market and that is where I looked. Liebowitz and Margolis also mention that bestsellers made up a surprisingly large share (60% in 1986) of the total hardcover trade market and this was confirmed by Sorensen (2007) who found that hardcover fiction titles that made the NY Times Best-Seller list were responsible for 84% of total hardcover fiction sales in 2002.

I focus in this paper on trade books, which includes mass market paperbacks. The particular trade books that we focus on are best sellers, since these books have the greatest likelihood of remaining in print. The process consisted in having research assistants examine best seller lists in long-ago decades (from 1895 to 1940) in order to generate a list of books. Once this list of titles was created, data were added to each title. The key variable indicating whether copyright still existed on the book was culled in a two stage process.

Due to the nature of copyright legislation, books written more than 56 years prior to the longer term of copyright that began in 1978 will have lost their copyright. Even with a sturdy market, the second 28 year term that began with the renewal of copyright would have expired before 1978. The period around 1922 is the cutoff between books that are still under copyright and those that are not. This cutoff provides the first stage of determining copyright.

On the other hand, some books may not have had their copyright renewed after the first 28 year period. Therefore we examined each book copyrighted after 1922 to see whether the copyright was renewed 28 years, plus or minus 1 year, after initial publication. This required considerable effort since at the time there was no way to check on copyright renewals without going to a library site containing reference books listing all copyright renewals and checking each title by hand.

Besides the copyright variable, we collected from *Books in Print* the list price of the book, the number of pages, the type of binding, the publisher, and the ISBN number and other special characteristics such as whether the book had large type, was illustrated, was a special edition, was a textbook, was for juveniles, and so forth. From Amazon's web site we tried to find the selling price on Amazon and the category of book content. Not all books had complete data for every variable.

As a comparison, best sellers from the 2002-2004 period were also examined to see how new best-sellers might differ from old best sellers.

[11] As reported in Landes and Posner (2003) p 212.

IV. Data Issues and Book Pirates

Books in Print (BIP) has both a continuously updated electronic database, and a yearly hardcopy database. The hardcopy database, which in 2004-5 contained five volumes listing the "titles" of books in print, consisted of 14,484 large pages, each with four columns of very small text. Nevertheless, the electronic database is bigger and more complete.

Even so, I chose to use the hardcopy database despite the fact that it was more laborious to generate the numbers by hand that way. There were several reasons I made this decision.

First, and most importantly, by using the 2004-2005 physical copies of *BIP* book, the database was fixed in time and could be double checked in the future. Since the electronic copy of the database is continuously changing, it would not be possible to double check the data in the future should the need arise.

Second, the electronic version contained entries of books that are not really books published in the normal sense. For example, the electronic version contained data from on-demand publishers such as IndyPublish and Books on Demand which do not show up in the physical database. There are many other publishers which provide highly questionable data. For example, Classic Books, which had hundreds of entries in the original data set, turns out to have the same prices (depending on binding) for every book. And Classic Books claimed to have a version of many of the bestsellers in the sample, even some that appeared to still be under copyright.

This brings us to a potentially more troubling problem. It is apparently very easy to put out a business shingle stating "publisher" these days. Most of the publishers on the list are minor publishers. Some of them may fall into the 'fly-by-night' category. A little web searching revealed stories claiming that some of these publishers did not pay for rights to works still under copyright.[12] Such accusations of piracy were made against Kessinger, Amereon and others. Obviously, if publishers of works classified as copyrighted are not paying royalties, thus having lower costs than other publishers of copyright works, the estimation of the impact of copyright would be negatively biased if their data were included.

The hardcopy *BIP* database did not, in many cases, contain complete book information for the books from these publishers, thus seemingly providing one method to reduce the impact of such publishers.

The penalty from using the hardcopy BIP, besides the costs of turning pages by hand and reading very small print, was a reduction in data. The original electronic BIP based data set consisted of 855 observations based on 358 unique titles. The data set from the

[12]This story about Kessinger by the unhappy son of the copyright owner makes the point very well: http://www.targetmarketingmag.com/bcs/enews/fullStory.bsp?sid=32029&var=story as do stories about Amereon being sued for infringement by Piers Anthony and the wife of Robert Heinlein. Publishers such as Kessinger state on their website that they care very much about copyright and that they want to be notified if one of their published books is thought to infringe copyright. One cynic referred to the approach as "catch us if you can."

hardcopy of *BIP* consisted of 419 observations and 191 titles. The title with the greatest number of variants (16) was Edith Wharton's *The Age of Innocence*.

To these old titles were added 82 recent best-selling titles and 153 binding-variants of these best-sellers which were all under copyright.

V. Some Initial Empirics

One question is whether books of a particular type tend to have the same prices. In related markets, for example, CDs tend to have similar list prices, movie admission prices in a multiplex do not depend on the particular feature being shown, iTunes charges the same price for all songs, and so forth. If market sellers are not price searchers then the idea of profit maximization by individual monopolists would seem to be one that would not be apposite.

Figure 1: List Price of Recent Bestsellers

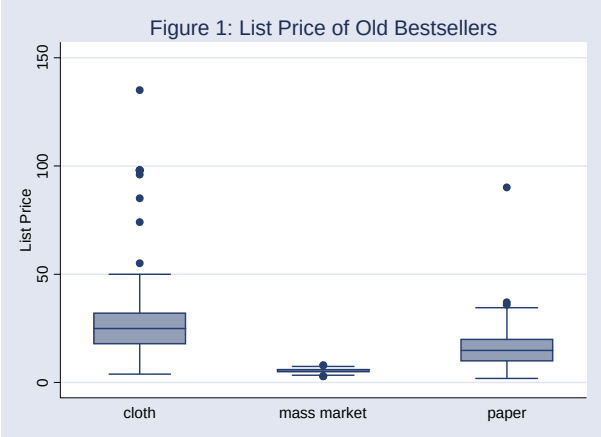
The Box Plots in Figure 1 indicate that for the set of recent best-sellers, prices are tightly grouped, especially for mass market paperbacks (24 out of 27 are at exactly $7.99, with the other 3 at $7.50). Even with almost no variation, the correlation between pages and price is .46. The prices of other paperbacks are also tightly concentrated around $15 although there is somewhat greater variation, due in part to a greater variety of books in the "paperback" classification.[13] The correlation between price and pages is -.02. Hardcover books (cloth) have a somewhat greater price variation than the other two categories but are fairly tightly grouped in the $24-$28 dollar range.[14] The correlation between price and number of pages is .55 for this group.

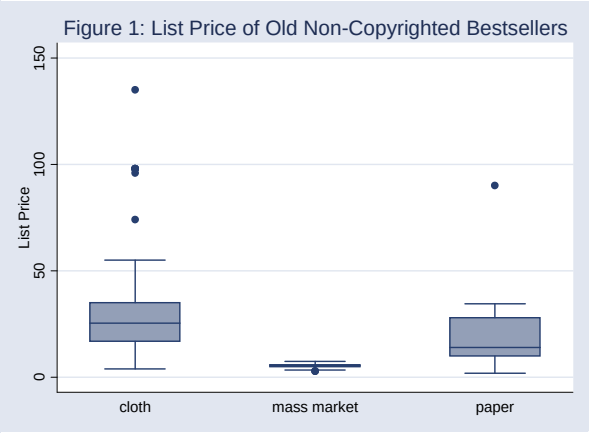
Older books had considerably greater variation in their prices. This was more pronounced for the non-copyrighted old books than for the copyrighted old books. This is for the simple reason that non-copyrighted titles can have many publishers and some publishers

[13] The clear outlier at $30 is a book that according to BIP is part of a role-playing Star-Wars game. Books with large print, which tend to have slightly higher prices, were removed.

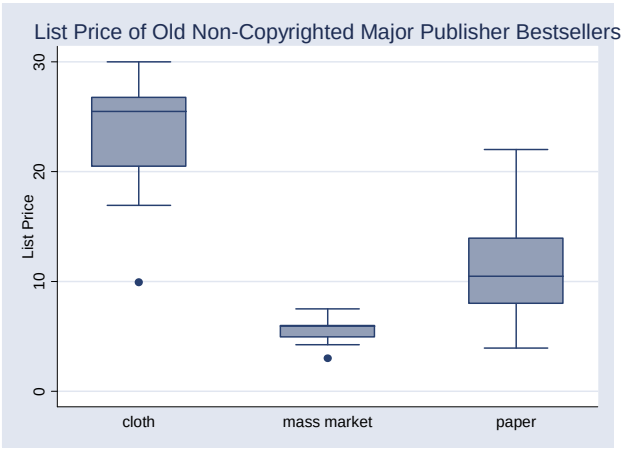
[14] The outlier at $10 is a Christmas book of only 48 pages.

of non-copyrighted books seem to be specialized into producing high priced or low priced books.[15]





In order to weaken the impact of some of these questionable publishers, Figure X provides a box plot for non-copyrighted old best-sellers from major publishers, as described in section A below. It is clear that the major publishers do not have the extreme prices of the minor publishers. Although the pricing is not as tight as for recent bestsellers, it is considerably tighter than when all publishers are included.



In Table 1 you will find summary statistics for the data. First note that the period of time runs from 1895 to 1940 for the old books and 2002 to 2004 for the new books. The average book contains approximately 400 pages, with recent books being slightly longer. Book lengths run from forty pages (children's books) to over a thousand pages. Most of the books are fiction and only a quarter of the old books are from major publishers. On average, old books appear to be slightly more expensive. Slightly more than one quarter of the old books are still under copyright. Most of the old books are classified as "trade paper" but most of the new books are hardcover (cloth).

[15] An extreme case of this is Best Books which, according to BIP, charged $98 for each of their 5 cloth titles. Whether this means that the entries in BIP are not informative, as appears to be the case for the very high fixed-price Classic Books (which is responsible for the $90 observation for paperbacks) or whether the company just has a fixed price policy, is not clear.

Table 1									
Variable	Obs	Mean	Min	Max		Obs	Mean	Min	Max
	----------Old Books----------					----------Recent Books----------			
# of pages	419	371.72	40	1108		153	428.47	48	864
listprice	419	26.10	2	135		153	19.77	6.99	32.95
copyright	419	0.27	0	1		153	1.00	1	1
fiction	419	0.77	0	1		153	0.54	0	1
majorpub	419	0.25	0	1		153	1.00	0	1
cloth	419	0.20	0	1		153	0.53	0	1
massm	419	0.08	0	1		153	0.18	0	1
paper	419	0.55	0	1		153	0.29	0	1
year	419	1917	1895	1940		153	2003	2002	2004

Information about the leading publishers for old titles is shown in Table 2. There we show the number of books and the percentage of books that are copyrighted for each of the top ten publishers ranked by number of books published. The shift to the hardcopy BIP reduces the number of small and suspect publishers such as Kessinger and Buccaneer. Nevertheless, there are still some suspect publishers (e.g., Amereon) playing a role and we will address that later.[16]

Table 2: Old Titles by Data Source and Leading Publishers					
---------Hardcover BIP----------			----------Electronic BIP----------		
Publisher	# Books	% Copyrighted	Publisher	# Books	% Copyrighted
Classic Books	52	1.89%	IndyPublish.com	80	0.00%
Penguin Books	17	39.29%	Classic Books	67	4.48%
Dover Publications	13	0.00%	Kessinger Publishi	67	23.88%
Simon & Schuster	10	52.38%	Penguin Books	39	33.33%
Reprint Services Co	10	16.67%	Simon & Schuster	32	50.00%
North Books	8	0.00%	Buccaneer Books	24	54.17%
Thorndike Press	7	41.67%	1st World Publishi	18	0.00%
Harper Collins	6	45.45%	Fredonia Books	15	6.67%
Random House	6	0.00%	Wildside Press	15	0.00%
Amereon Limited	5	28.57%	Amereon Limited	14	42.86%

VI. Regression Results

The first and simplest examination of the relationship between the price of books and copyright is to run a regression with price as the dependent variable and using variables that are supposed to explain the price as the covariates.[17] Publishers of books under

[16] Thorndike Press specializes in large print books and does this for both copyrighted (where it appears to have permission of the copyright owner) and non-copyrighted books.

[17] I obtained data on both the list price and in some instances the selling price on Amazon, but had far more observations using list price. There are three reasons to prefer list price. Since the correlation between the

copyright are normally under contract to pay the author whereas publishers of books not under copyright are not. The form of the price variable that seems to make the most sense, under these conditions, is to take natural log. There are two reasons for this. First, the royalty payments to authors are generally calculated as a percentage of revenues. The price differential is thus not a fixed amount but a (relatively) fixed percentage.[18] Second, any monopoly power is likely to be best expressed as a percentage of price.[19]

Table 3: Regression on Ln(Price); All Publishers, Clustered								
	Old	Old	Full	Full	Cheap, Old	Cheap, Old	Dear, Old	Dear Old
Copyright	-0.1163	-0.0231	-0.0946	0.0114	0.2430	0.1363	-0.3474	-0.1964
	(0.88)	(0.31)	(0.71)	(0.16)	(3.61)***	(2.37)**	(2.52)**	(1.83)*
No. of Pages	0.0010	0.0008	0.0009	0.0006	0.0003	0.0006	0.0006	0.0005
	(5.41)***	(4.48)***	(5.81)***	(5.06)***	(1.37)	(1.78)*	(2.03)**	(2.25)**
Recent			0.0355	0.0631				
			(0.50)	(0.74)				
Fiction	-0.1539	-0.1531	-0.1242	-0.1244	-0.1064	-0.2870	-0.1463	-0.0809
	(1.78)*	(2.59)**	(1.97)*	(2.73)***	(1.73)*	(2.82)***	-1.23	-0.94
Mass Market	-0.9297	-0.5772	-0.8106	-0.5830	-0.5478	-0.4963	0.0000	0.0000
	(6.38)***	(9.22)***	(7.41)***	(13.46)***	(7.76)***	(4.72)***	(.)	(.)
Hardcover	0.692	0.717	0.634	0.676	0.152	0.133	0.397	0.990
	(5.48)***	(5.79)***	(6.92)***	(10.17)***	-1.22	-0.71	(2.83)***	(5.83)***
Library binding	1.2655	1.0637	1.2547	1.0379	0.0044	0.0000	0.8121	1.2134
	(7.33)***	(4.03)***	(7.16)***	(4.03)***	(0.06)	(.)	(3.48)***	(10.70)***
Perfect binding	0.5093	0.0000	0.5118	0.0000	0.0000	0.0000	-0.1345	0.0000
	(3.11)***	(.)	(3.29)***	(.)	(.)	(.)	(3.30)***	(.)
Childrens	-0.6014	-0.3491	-0.6139	-0.3483	-0.5488	-0.5732	-0.2261	-0.0041
	(4.84)***	(2.96)***	(4.92)***	(3.19)***	(6.15)***	(4.85)***	(1.99)*	(0.06)
Large Print	-0.2431	-0.0345	-0.1637	0.0179	0.1259	-0.0055	-0.5562	-0.0300
	(1.70)*	(0.43)	(1.67)*	(0.25)	(0.83)	(0.34)	(4.27)***	(0.35)
Constant	2.469	2.489	2.487	2.476	2.226	2.318	3.247	2.809
	(13.54)***	(34.05)***	(16.35)***	(41.20)***	(14.34)***	(12.79)***	(23.51)***	(21.44)***
Publisher Dummie	No	Yes	No	Yes	No	Yes	No	Yes
Observations	419	419	572	572	174	174	167	167
Adjusted R-square	0.59	0.88	0.61	0.90	0.46	0.78	0.41	0.89
* significant at 10%; ** significant at 5%; *** significant at 1%; high-priced are >23 and low-priced are <15								

We run a simple OLS regression with and without publisher dummies and using the Stata cluster command to correct for heteroskedasticity across publishers. First we run these regressions for only the set of old books and then for both old and new books. Then we

list prices and the Amazon transaction prices was very high (.95) it probably doesn't make much difference which is used. Amazon appears to base discounts on popularity of the book and large price changes can result when a book loses popularity and is repriced, possibly providing misleading results. Third, one can easily go back to 'Books in Print' to get the list prices for previous years. One cannot easily retrieve historic Amazon data in a useful fashion.

[18] Royalty rates, which can differ, often increases as the quantity of sales increases although for bestsellers the rate is presumably at the top step for the great majority of copies sold.

[19] For example, the Lerner measure of monopoly power is expressed as a percentage of price.

run regressions for subsets of the observations based on the price of the books. The results are found in Table 3.[20]

The copyright variable does not provide a consistent result across these specifications. In the regression using only old books, the copyright coefficient indicates, surprisingly, that copyright *lowers* price by 11.6% although the inclusion of publisher dummy variables weakens the coefficient and in neither case is the impact statistically significant. Similar but slightly weaker results occur when recent best-sellers are included in the sample, even though these best-sellers are the books most likely to contain real monopoly elements. Recent best-sellers have small positive coefficients but this result is statistically very weak.

All the regressions show that the number of pages is positively related to price, that cloth, library, and perfect bindings all raise the price relative to the excluded paperback bindings, and that mass market books have lower prices. All of these findings are consistent with expectations. Works of fiction and children's books appear to have lower prices, which is also not surprising. There appears to be weak evidence that large print books are lower in price than regular books, which we know is not true since large print books are routinely slightly more expensive than the same title with regular print, but since large print books also have more pages the negative coefficient may indicate that large print books have prices that are lower than expected given their extra pages.

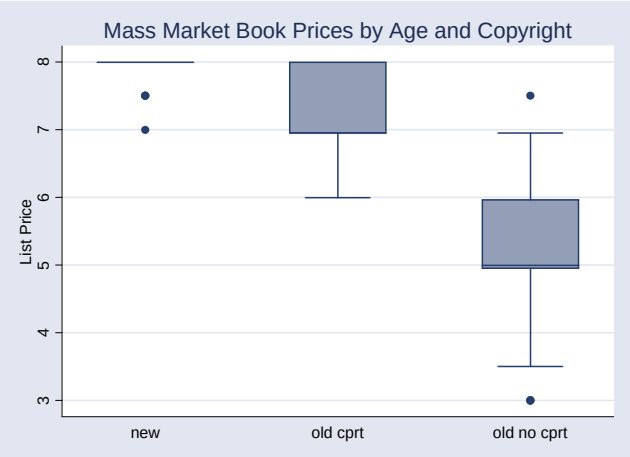
The second half of the table includes regressions for high and low price books. I decided to separate the books by price because, contrary to the overall regression results, copyrighted mass market books appear to have considerably higher prices than non-copyrighted books as seen in Figure X. Figure X shows non-copyrighted mass market books, the right hand box plot, clearly have lower prices than new mass market books (which are all copyrighted with a almost uniform price $7.99) or old copyrighted books.

Given this, it is reasonable to expect that the impact of copyright for mass market books, and perhaps all inexpensive books, should be positive. This was confirmed by quantile regressions (not shown) and also by separating books into a high-priced group (>$23) and a low-priced group (<$15).[21] For low-priced books the copyright variable is positive and

[20] The overall regressions were also run trying to control for the impact of outliers, using quantile (median) regressions and also Stata's built-in RREG routine (which attempts to weaken the influence of outliers by eliminating observations with large Cook's D and decreasing the weighting of observations with large residuals). The results were largely the same as presented in Table 3 except for the full sample with dummy variables where the copyright coefficient in the quantile regression indicated a 19% increase in book prices and the RREG regression found a 7% increase. This result appears to be due to new bestsellers since the coefficient was negative in each case when only old books were included in the regression, which is the more precise test of the impact of copyright per se.

[21] The quantile regressions did not always converge and did not allow the clustering correction for heteroskedasticity, thus it seemed preferable to show the OLS regressions with cutoffs. I use $23 and $15 as cutoffs, but they are representative of a fairly smooth change that occurs in the copyright coefficient as you move from low-priced books to high-priced books. Nevertheless, the quantile regressions at .2 and .8 (available upon request) provided very similar results for the copyright coefficients.

significant, although less so when publisher dummy variables are included.[22] But for high priced books, copyright is negative and significant, indicating that copyright *lowers* the price. What these results would mean, if accurate, is that copyright appears to lower the price of expensive hardcover books but raise the price of inexpensive mass market and regular paperbacks. We will return to this finding later.



Nevertheless, there is little in these regressions to support a view that copyright leads to higher prices for books in an overall sense.

A. *Problems with the publishers*

I have already discussed some of the problems with the small publishers. Since every publisher claims to be legitimate, it is unclear that we could solve the potential problem of publishers violating copyright by merely removing the known offenders. Even legitimate small publishers can cause problems if they cater to specialized markets outside the mainstream, or if their data in BIP is not kept up to date.

One potential method for removing suspect publishers is to limit the sample to only those publishers that are well established and who therefore are unlikely to intentionally violate copyright. The cost of doing this is that we will have a smaller sample size.

The algorithm I used to restrict publishers was fairly simple. Any publisher of a then-current best-seller in the 2002-2004 period was deemed to be major. Any publisher that was a household name, or at least was one in my household, was deemed a major publisher. Any publisher that came up first in a Google search under its name and had a group of associated URLs underneath the main description was treated as a major publisher. University presses, no matter how major, were not treated as major publishers because they tend to have very different types of books. All other presses were deemed minor and excluded from the analysis below.

[22] Both median regressions and Stata's built-in RREG routine indicated that when outliers were deemphasized, the copyright coefficient was considerably smaller (approximately half) than the 24.3% and 13.6% values shown in Table 3.

This left us with 40 publishers (some might be different imprints from the same owner), 286 books and 143 titles.

Table 4:Major Publishers, clustered								
	Old	Old	Full	Full	Cheap, Old	Cheap, Old	Dear, Old	Dear, Old
Copyright	0.1550	0.0007	0.1845	0.0705	0.2340	0.0443	-0.0675	-0.1061
	(1.24)	(0.01)	(1.55)	(0.93)	(3.78)***	(0.48)	(1.02)	(2.32)**
No. of Pages	0.0008	0.0009	0.0005	0.0006	0.0006	0.0008	0.0006	0.0009
	(2.99)***	(3.32)***	(3.23)***	(3.77)***	(1.88)*	(2.81)***	(2.05)*	(3.16)***
Recent			0.063448	0.000447				
			(1.06)	(0.01)				
Fiction	-0.0245	-0.2664	-0.0395	-0.1605	-0.2050	-0.4439	0.0412	-0.0715
	(0.22)	(3.78)***	(0.76)	(2.75)***	(1.14)	(1.65)	(0.48)	(1.78)*
Mass Market	-0.6842	-0.5799	-0.6660	-0.5851	-0.0793	-0.1668	0.0000	0.0000
	(6.27)***	(7.32)***	(10.06)***	(12.12)***	(0.62)	(8.71)***	(.)	(.)
Hardcover	0.7356	0.7244	0.6140	0.6684	0.2129	0.2681	0.4926	0.5259
	(5.52)***	(4.64)***	(9.46)***	(9.43)***	(1.98)*	-1.6	(5.30)***	(5.08)***
Library binding	1.0846	0.5049	1.0525	0.4787	0.0000	0.0000	0.6704	0.4477
	(7.46)***	(12.80)***	(8.11)***	(22.86)***	(.)	(.)	(6.42)***	(12.59)***
Children's	-0.5478	-0.4698	-0.5764	-0.4874	-0.3631	-0.3845	0.0917	0.1647
	(5.27)***	(4.44)***	(5.24)***	(5.51)***	(3.61)***	(6.90)***	(1.05)	(2.12)**
Large Print	-0.0629	0.0076	0.0225	0.0416	-0.0402	0.0205	0.0000	0.0000
	(0.74)	(0.53)	(0.30)	(0.47)	(0.40)	(2.96)***	(.)	(.)
Constant	2.122	2.290	2.249	2.354	1.751	1.986	2.534	2.479
	(8.66)***	(18.48)***	(14.50)***	(31.68)***	(14.46)***	(11.20)***	(16.31)***	(15.85)***
Publisher Dummi	No	Yes	No	Yes	No	Yes	No	Yes
Observations	140	140	286	286	55	55	54	54
adj R-squared	0.65	0.82	0.76	0.84	0.46	0.84	0.65	0.89

t statistics in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; expensive>13 for all books, >9 for old books; ^ no clustering

I reran the regressions with this reduced sample and the results are shown in Table 4. The first four regressions are the same as those in Table 3. One immediate difference is that copyright is now always positive. The impact appears economically strong, between 15% and 20% when there are no publisher dummy variables, but it is not statistically significant. When publisher dummy variables are include, however, the result is small (~7%) when new books are included in the sample and virtually nonexistent when only old books are used.

The next four regressions reprise examining the difference of copyright's impact between high price books (dear) and low price books (cheap). Because the distribution of book prices for major publishers differs from that of all publishers (it has a considerably lower mean, 12.7 versus 26.1, and median, 11 versus 18.95) the cutoff prices to compare low-price and high-priced books needed to be lowered. The overall result of a negative impact of copyright for high priced books and a positive impact for low price books still basically holds, but it is weaker than before. If publisher dummy variables are not included, the copyright coefficient for inexpensive books is large and statistically significant but including the publisher dummy variables weakens this result tremendously although copyright's impact is still moderately bifurcated.

Because the inclusion of publisher dummy variables makes a large difference in the measured copyright coefficient it is important to determine whether the regression is more informative with or without publisher dummies. The determination depends on

whether publishers tend to specialize in publishing either copyrighted or non-copyrighted works and also whether they tend to specialize in publishing books in particular pricing ranges. For example, if some publishers of copyrighted books also happened to specialize in high priced books (e.g., particularly expensive bindings and paper) independent of copyright, the inclusion of the publisher dummy would correctly lower the measured impact of copyright. On the other hand, if publishers who specialized in copyrighted books charged higher prices only because the books were copyrighted, inclusion of the publisher dummies might pick up some of the copyright impact and bias downward the measured impact of copyright.

The correct specification would depend on whether the publishers of copyrighted works tended to choose high prices because their works were copyrighted, requiring royalty payments and possibly exhibiting some monopoly power, or whether they focused on high price books for some other reason. With no *a priori* reason to believe that publishers of copyrighted works also, by chance, specialize in expensive books, we would favor the results including the publisher dummies.

Clearly, the copyright coefficients are empirically more similar (and closer to zero) when publisher dummies are included, perhaps indicating that publishers do have different pricing strategies and that these strategies are also related to copyright. I make one additional attempt to further refine the data so as to hopefully weaken any price or copyright specialization by publishers that might muddy the measured impact of copyright.

B. *Further Winnowing of Publishers*

My approach to weaken any misleading relationship between prices and copyright due to publisher specialization is to remove specialized publishers. To that end I restrict the major publishers from the previous section to only those publishers who publish both copyrighted and non-copyrighted works. As long as publisher choose similar quality bindings for both copyrighted and non-copyrighted books for any given binding type, inclusion of the publisher dummy should not bias the copyright variable.

Using a rule that the share of copyrighted books in the portfolio of a publisher's books can be no more than 80% nor less than 20%, we are left with only nine publishers, 135 books, and 70 titles.

The results are found in Table 5.

This final winnowing down seems to have been something of a success. The publisher dummies no longer have as dramatic an impact on the copyright variable although they continue to lower slightly the impact of the copyright coefficient. Additional good news, as far as providing a consistent story, is that the coefficients for the copyright variable are within a fairly narrow range running from -2% to +7%, although none of these coefficients is significantly different than zero. These results indicate that copyright does not have a substantial positive impact, or even a small one that we can state with a high degree of confidence.

Table 5: Optimal Publishers, clustered								
	Old	Old	Full	Full	Cheap, Old	Cheap, Old	Dear, Old	Dear, Old
Copyright	0.0602	-0.0200	0.0717	0.0505	0.1707	0.2209	-0.0779	-0.1127
	(0.84)	(0.26)	(1.03)	(0.78)	(3.42)**	(11.87)***	(1.64)	(2.65)**
No. of Pages	0.0009	0.0008	0.0007	0.0007	0.0001	0.0002	0.0009	0.0010
	(3.96)***	(3.22)***	(3.65)***	(3.51)***	(0.95)	(1.13)	(4.40)***	(3.20)***
Recent			-0.0108	-0.0345				
			(0.17)	(0.42)				
Fiction	-0.1968	-0.2936	-0.2057	-0.2070	-0.0134	0.0051	-0.0005	-0.0752
	(2.51)**	(3.11)***	(4.12)***	(3.20)***	(0.22)	(0.07)	(0.01)	(1.73)*
Mass Market	-0.5934	-0.5993	-0.5827	-0.5572	-0.1137	-0.0277		
	(9.60)***	(6.25)***	(11.38)***	(8.15)***	(1.24)	(2.03)*		
Hardcover	0.7752	0.7833	0.6696	0.6895	0.1087		0.5399	0.5291
	(5.90)***	(8.54)***	(8.63)***	(11.51)***	(1.34)		(4.93)***	(4.17)***
Children's	-0.4011	-0.4933	-0.4062	-0.4523	-0.4138	-0.3153	0.1443	0.1792
	(3.09)***	(4.79)***	(3.04)***	(5.23)***	(3.85)***	(9.61)***	(2.29)*	(2.20)**
Large Print			0.0993	0.1329				
			(0.95)	(1.18)				
Constant	2.2825	2.4263	2.3653	2.3897	1.9151	1.7771	2.3923	2.4004
	(21.07)***	(17.30)***	(30.29)***	(25.66)***	(13.01)***	(63.05)***	(20.76)***	(14.66)***
Publisher D	No	Yes	No	Yes	No	Yes	No	Yes
Observations	90	90	135	135	34	34	36	36
Adjusted R-squa	0.68	0.72	0.78	0.8	0.63	0.71	0.73	0.74
* significant at 10%; ** significant at 5%; *** significant at 1%,								

The bifurcated regressions continue to indicate a division in the impact of copyright based on the price of books, with low priced books apparently having their prices raised by approximately 20% under copyright and high priced books have their prices lowered by about 10%. Unlike the regressions for all major publishers, the copyright coefficient for optimal publishers was strengthened by the inclusion of publisher dummy variables.

VII. What might these results mean?

What are we to make of these empirical findings? First, we would be hard pressed to conclude that copyright leads to higher prices, or at least much higher prices. The full sample generally indicated zero or negative impacts of copyright on price. Because this sample was tainted with publishers likely to bias the copyright variable downward, we endeavored to conduct two refinements to the data, with the first refinement nested in the second.

First, when the most obscure publishers were removed, we found positive (but statistically insignificant) coefficients for copyright, perhaps as much as the 15% found in Table 4. But the inclusion of publisher dummy variables lowered the measured impact of copyright to 7% or less, and these results were not statistically significant. The inclusion of publisher dummies clearly was important. We then attempted to judge whether the publisher dummy variables belonged in the regression.

In order to help determine whether publishers might be specializing in books that are either copyrighted or not copyrighted while at the same time specializing in pricing either

high or low such that the publisher dummy might be picking up some of the copyright impact, we further restricted the sample to include only major publishers who published both copyrighted and non-copyrighted works.

This final winnowing of publishers indicated that the prior regressions were more informative when they included publisher dummy variables since the copyright variable with the most refined data indicated a small and insignificant positive impact of copyright on the price of books, more in line with the regression results including publisher dummy variables than with those excluding the publisher dummy variables.

We also found fairly consistent evidence that the impact of copyright was positive for inexpensive books but negative for expensive books.

There are several possible conclusions that might be drawn from these empirical results:

1. Copyright has no consistent positive impact on the price of books.
2. Copyright raises the price of some categories of books (e.g., mass market paperbacks and other inexpensive books) but has a very small overall impact..
3. Copyright raises the price of all books, but the book categories are too poorly defined to reveal this, except for mass market paperbacks.

We shall discuss the implications of these possible conclusions one at a time.

A. *Copyright Has no Overall Impact on Price*

In light of basic economic theory, this result seems counterintuitive if not downright impossible. How can a monopoly, even only a sometimes economic monopoly, fail to raise the price? Even if old best-sellers are thought to no longer contain any economic monopoly power, when new best-sellers, which certainly must contain real economic monopoly power, are included in the analysis there still fails to be much support for copyright causing a higher price.

Since royalty payments are paid on a per book basis (beyond the advance) the marginal cost curve for copyrighted books would seem necessarily to be above that for a non-copyrighted book, all else equal. That said, the profit maximizing price for copyrighted books would be expected to be higher than the non-copyrighted price, contrary to the empirical results of a zero impact on price. It is possible that recent best-sellers have lower average costs because of their large quantity print runs, but the empirical finding holds equally well for only old best sellers.

Nevertheless, one can provide an economic explanation consistent with these findings. It might appear at first to be *ad hoc* in nature, but the conditions required for this explanation appear to be met in several related industries, so the explanation may appear ad hoc only because there have been few investigations of uniform pricing in such industries and because such pricing is not well understood.

The explanation is based upon the behavior of firms in some copyright industries with characteristics similar to those in the book market. These are real markets operating in the real world, and in these markets the prices of products, such as books, often do not deviate from one another based upon demand. For example, in North America, when you

go to a movie theater multiplex showing several different movies, the prices for the movies will generally be identical, even when it is generally well known in advance which movies are likely to have large demands and which will have small demands (Orbach and Einav 2007). The most successful movies may be shown on two or three screens at one time, to accommodate demand, but the price will not change. It is quantity that adjusts to demand differentials, not price. Similarly, music CDs from well known musical performers do not seem to command price premiums over CDs from less well known groups. The pricing on Apple's iTunes, where each and every song sells at the identical price to every other song, is the strongest example of this type of pricing. The same is apparently true for video rentals (Dana and Spier 2001).

Clerides (2002) takes note of the lack of price dispersion in books once adjustments are made for cost differences. Clerides suggests that this lack of dispersion in the price of books could be explained if the demands for different titles were linear and all had the same price intercept. In that case, every demand curve would have the same profit maximizing price (if the marginal cost were zero or constant). That explanation cannot work in our data set because copyrighted books have higher marginal costs (they pay royalties) than non-copyrighted books (although to be fair, Clerides' appears to have been discussing only copyrighted books). It is also unclear why demand might have this particular structure.

Orbach and Einav run through several possible explanations for uniform pricing in movies and provide an in depth analysis of that industry, but no single compelling explanation is found for such behavior in that antitrust-plagued industry.

McMillan (2005) has what is probably the most complete discussion of the 'uniform price puzzle' although his examples do not come from entertainment markets but from supermarkets. He points to menu costs as being sufficiently high relative to revenue gains as to make uniform pricing optimal.

Whether there is a single overriding explanation of uniform prices or different explanations for different types of markets and products is still unknown as this important topic is still in its infancy.

It may be that retailers of these products, whether they are movie theaters or record stores, believe that they are selling 'movie entertainment' or 'music entertainment' and the exact title that is at the moment fulfilling that purpose for the retailer is not particularly important to the retailer. In this case the seller may think that it is important to set a predictable price to consumers, implying the same price for all titles. Or perhaps it is easier to predict the size of the audience wishing to read a book or see a movie more than it is to predict the price that the audience is willing to pay, a result that leads us to a scenario like that presented by Clerides.

B. *Different Impact on Different Segments*

Alternatively, these results might be interpreted to imply that the high priced portion of the market is negatively impacted by copyright while the low price portion is positively impacted. This is certainly what the regressions for the bifurcated sample seem to be telling us, although when industry dummy variables are included even this result becomes small and not terribly convincing.

It also seems difficult to reconcile this claim with any theoretical understanding of how copyright might work. Why would copyright raise the price only of inexpensive books? One might argue that the existence of mass market paperbacks suggests larger sales volumes for a particular title and that this might have something to do with higher prices. For example, many authors receive an upfront payment known as an advance. This advance payment is made against future royalties, so the advance converts some royalty payments from variable costs to fixed costs. As sales increase, however, the advance normally gets covered and the author begins to earn royalties again. So perhaps the royalty rate will only have an impact on price for books that sell relatively large amounts, and since one sign of large sales is a title having inexpensive paperbacks, such books will have higher prices.

The problem is that even if royalties were entirely lump sum in nature, they would still need to be covered by publishers in long run equilibrium. Why should the publishers wish to load these royalty costs into mass market paperback sales and not hardcover sales?

It is also the case that royalty rates tend to be higher for mass market paperbacks (and presumably for other inexpensive books) than for more expensive paperbacks although less than for hardcover. This might be able to explain why prices are higher for inexpensive paperbacks relative to expensive paperbacks but much of this impact should be picked up in the mass market dummy variable. Also seemingly incompatible with this hypothesis is the fact that hardcover books also have higher royalty rates than the typical paperback and these are the more expensive books that do not have higher copyright-induced prices. Further, the size of the royalty rate difference is only a few percentage points, so it could not explain a copyright induced price difference that is much larger.

But this does not seem very persuasive, and the results hold even when mass market paperbacks are eliminated from the analysis.[23]

C. *Only Mass Market provides correct Results*

It might be argued that mass market paperbacks are so narrowly and well defined that there is almost no variation in the physical quality of the various titles in that segment. If so, this could provide a more correctly measured price comparison between books, particularly for the impact of copyright on the price of books. The other books, although the type of cover binding provides some information, are presumably far less precisely segmented, thus allowing wide variations that might corrupt the measurement of the impact of copyright on price. Therefore, according to this view, the positive impact of copyright for low priced books (or mass market) should be given far greater weight than the negative or zero impact of copyright on price found for the more broadly based book market.

[23] Although the mass market paperbacks received the bulk of the attention in terms of my decision to examine copyright's impact in a bifurcated manner, the bifurcated results hold even when mass market paperbacks are excluded. Although I do not report these results in the text, they are available upon request.

The problem with this explanation, of course, is that the greater variation in quality for other categories of books might lower the precision of the estimated impact of copyright on price, but there is no reason to think it would bias the estimate. Without some explanation for why imprecisely segmented categories of books will artificially bias the measured copyright impact toward zero or negative values, this claim loses most of its force.

Although I have some sympathy with the first part of this claim, I am at a loss to provide a reasonable explanation for the second part.

VIII. These Results in Perspective

These results seem generally consistent with a belief that copyright has no impact on the price of books. Nevertheless, one can grasp at the slender reeds provided in these estimates to claim that copyright's impact on book prices might be in the 5%-7% range found in Table's 4 and 5 when publisher dummy variables are included.

How would such results fit into the policy debates currently swirling around copyright?

First the size of the copyright-induced price increase obviously impacts the size of any deadweight losses due to copyright. Second, the price difference might give us some inkling about the gains or losses likely to occur from alternatives to copyright. Third, the price difference might tell us something about how the authors and publishers share any monopoly rents.

The implications depend on the nature of the market for books. Is the book market like markets in standard textbook models, or is it more as presented above, with a uniform retail price for book titles independent of demand? We will discuss these issues under both scenarios.

A. *Uniform Pricing of Books*

The alternative to textbook monopoly pricing is the uniform pricing already mentioned above. In this instance the price of copyrighted books would be the same as the price for non-copyrighted books which is consistent with the empirical estimates.

How, the reader may ask, can publishers of copyrighted books pay royalties to authors when they sell their books at the same price as publishers of non-copyrighted books? The answer is, to use a quaint aphorism, that they make up for their low prices with their higher volumes.

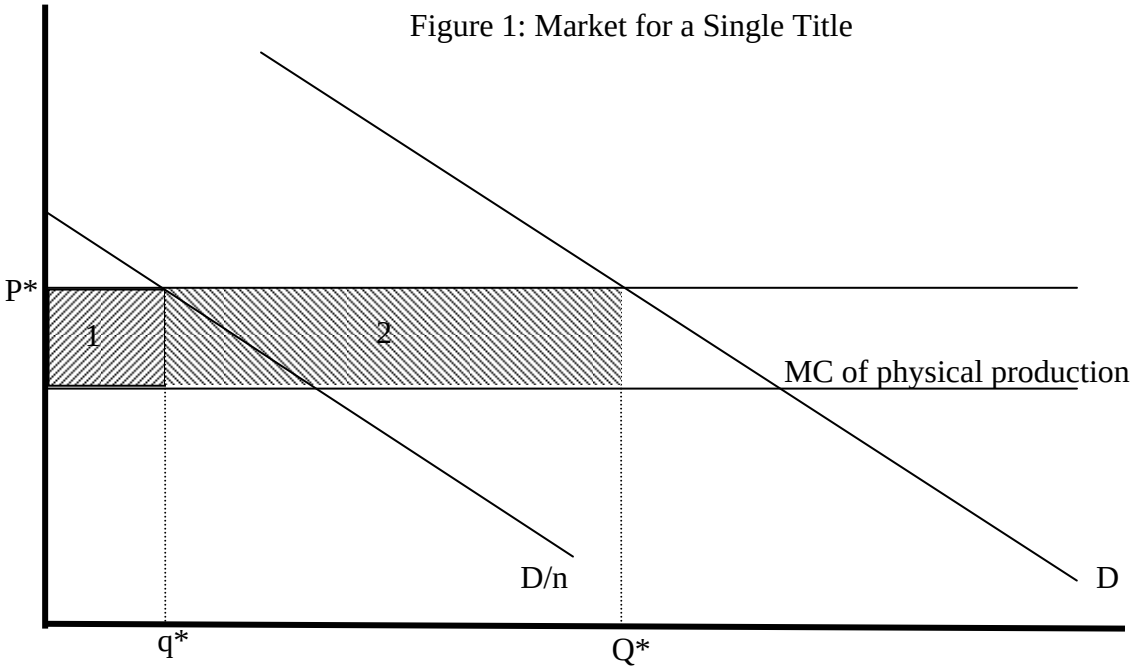
To see how this works, we need to consider a world where retail prices cannot change and derive how competition differs from monopoly in the selling of a single title.[24]

[24] If producers and retailers are separate entities, the producers of above average demand wholesale products can try to set a higher wholesale price, reducing the markup of the retailer. I effectively assume in the text that either publishers are so dependent on retailers to help sell their products that they would not want to lower the retailers margin.

In the competitive (non-copyright) case, competition drives above normal profits in the sale of a book title down to zero by having additional firms printing the book and reducing available sales to each incumbent producer of the title until the representative firm producing the title can generate only normal returns. In this world, instead of the price of the title falling as entry occurs, quantity per firm falls, as does profit per firm. For such a model to work there needs to be a fixed cost per publisher (title) so that additional publishers of the title increase average costs of producing this title even as prices remain constant. This assumption of fixed costs seem a realistic and important aspect of this industry. Clerides reports that even for pure production costs independent of manuscript preparation and book distribution, fixed costs, consisting of typesetting and plates, are important (the importance obviously depends on the size of the print run, but his estimates are for runs of several thousand units).

Figure 1 can be used to illustrate these ideas. Let the market demand for the title be represented by D. Since the price is exogenous, the industry quantity sold is determined only by the demand. As long as the marginal costs are constant and below the price the firm will sell as much as it can. Figure 1 shows the MC for producing copies of the book without any payment to the author.

Figure 1: Market for a Single Title



To simplify the diagram we treat the author's costs (and payments) as fixed since the question is whether there are enough total dollars generated to pay the author to produce the title and it doesn't matter whether such payments are made as a lump sum or as a percentage royalty since modeling this cost as fixed or variable will not impact the quantity sold by the firm.

The uniform price at which all copies of all book titles sell is P* which must be greater than the marginal (average) cost of producing and selling the physical copies of titles if

the book is produced. P* must be higher than MC if the producer(s) of copies of the title are to cover the fixed cost of producing copies of the title, which includes setting up the printing press, checking the document for errors, and so forth. If the book is copyrighted, the difference between P* and MC must also allow the (single) producer to cover the fixed costs of the author and the editing process.

Under copyright there is only one seller of the title. Q* units would be sold. Areas 1+2 represent the quasi rents that are used to pay off all the fixed costs. As long as areas 1+2 are greater than the fixed printing costs plus the fixed costs of the having the author create the title, the publisher-retailer-author nexus will make a positive profit and in a world where sales can be reliably predicted, the book will be written and published.

We can repeat the above analysis except to remove copyright. Entry of new publishers lowers the demand available for each incumbent publisher of the title. For simplicity, assume that all publishers of the title share the output equally. In that case, the demand facing each (representative) firm is the firm's prorated share of the industry output, D/n. Quasi-rents for the representative firm equal area 1 when there are n firms. As long as area 1 is larger than the entrant's fixed costs of publishing, entry will occur. When area 1 equals the fixed costs of the representative firm, entry will stop. In equilibrium, the quantity of books sold per firm is q* which equals Q*/n where n is the number of firms such that the addition of a marginal firm would causes firm to earn negative returns. Industry output does not change since price does not change.

In such a world, the typical deadweight losses associated with copyright no longer apply. The title with copyright has an advantage of lower average (fixed) costs (and higher profit) that goes along with the larger volume. Copyright, in this case, prevents the duplication of unnecessary fixed costs and it is this freeing of fixed costs that allows the copyrighted book to earn quasi rents and so pay the author. In other words, because there are some fixed startup costs in printing a book it is inefficient to have more than one producer of a given title. In this world copyright enhances efficiency and allows authors to be paid from the efficiency savings.

Left out of the story so far is the competition between titles, which moves demand curve D for any given title. We would expect competition among titles to reduce the demand facing the representative title to a point where only normal returns can be earned. Book publishing, however, is thought to be one of several industries where most titles fail to earn any profits, analogous to a dry hole in drilling for oil, in part because the future sales of a book are not known with certainty and publishers are looking for the next gusher. Given these industry characteristics, the ex post demand for the typical copyrighted title is likely to merely allow negative economic profits to be earned from quasi-rents 1+2, after the author is paid. There will be a small number of authors whose talents are not easily duplicated, and the books from these authors will generate positive rents, in a sense subsidizing the poorly selling titles.[25]

[25] Because of the right of first refusal clauses in book contracts, authors cannot just get the highest payment that any publisher will make for their book. This means that even if book publishers compete their profits away, the successful authors do not get to keep all of the profits—the unsuccessful books will get some of those profits as well.

Because authors are paid as a function of output, however, and because their contracts contain escalator clauses in their royalty rates as output increases, authors of successful books are likely to earn a share of the quasi-rents. If publishers compete among themselves as perfectly competitive firms to sign an author, the author will, in fact, earn all the monopoly rents in the system.[26] We are not in a position to state that this is case, however.

The price of the retail product, therefore, may be largely exogenous for any individual title although it would obviously be endogenous to the entire set of titles and the setting of the price of titles clearly is a question worthy of examination, although it is not undertaken here. It seems on first blush that the uniform price might be higher or lower under copyright since the average cost is lower but the margin on average cost is higher.

Putting this issue aside for the moment, the great majority of titles purchased in a year are of copyrighted works. Decreases in copyright length, or putting more books into the public domain, which has been the goal of much of the recent criticism surrounding the Eldred case and the copyright extension under the Sonny Bono Act, are unlikely to change the uniform price and thus, given the empirical results of this paper, are unlikely to lower the price of copies of published versions of these no longer under copyrighted works.

B. *Books are Priced as in simple Textbook Models*

If the book market is best represented by a standard textbook model, ruling out such deviant discussions as that which takes place in Section VIIA and Section VIIIA, we would expect higher costs to lead to higher prices, all else equal. This would mean that publishers who must pay royalties to authors would have higher costs, for a given demand, and thus higher prices as well. This would be true regardless of whether the market for books was treated as competitive, monopolistically competitive, or as monopolistic.

In this world the interpretation of the empirical results that best fits, assuming that these empirical estimates bear any relationship to reality, would be the claim that copyright raises the price of books, perhaps by the 5%-7% already mentioned. It is also the case that this specific-sized increased price has some important implications for current controversies.

What would such an increase imply regarding the three questions posed above?

First, let's start with the copyright monopoly and its attendant deadweight loss. This deadweight loss occurs in the market for copies of a title and in spite of its name, only a portion of the standard deadweight loss is really a lost surplus. That portion of the deadweight loss, if any, which is unnecessary to have the work produced in the first place, is the real deadweight loss which Liebowitz and Margolis (2005) refer to as

[26] Unlike authors and individual titles, each of which can have a relatively high degree of product differentiation, mainstream publishers may be virtually undifferentiated from each other since many of them perform virtually identical services, although there will always be some publishers who might specialize in some aspect of the market and might differentiate themselves thusly.

unproductive deadweight loss. This is contrasted with the portion of the deadweight loss which is necessary, under copyright, to provide sufficient incentive to have a title produced in the first place, which they refer to as productive deadweight losses and which they argue should not be considered deadweight losses at all, at least not under a copyright-based system.

If the goal, however, is to compare copyright with its alternatives, then this entire deadweight loss in the market for copies of a title can be considered a non-productive deadweight loss, at least in comparison to some idealized system of providing incentives to produce the work that has no deadweight loss of its own. More importantly, as long as we properly keep track of all the real-world deadweight losses in competing systems, we can deal with this issue at a final summing up stage.

Figure 2 shows the typical deadweight loss as represented in the market for a title. P_C is the marginal cost of making copies of a title, assumed to be constant, and also the price expected in a competitive market, in the standard model. A monopoly, on the other hand, looks for the profit maximizing price, P_M, and causes a deadweight loss equal to the shaded triangle. The benefit of eliminating copyright is thought to be the elimination of this deadweight loss, although the question of how the creators of titles are then to be paid is still left open.

The cost of creating the work is typically assumed to be fixed, which seems an appropriate classification of social cost although not necessarily the correct classification for the book publisher as they actually operate. Book publishers do not typically pay authors a fixed amount of money for their creations. Instead they pay royalties as a percentage of sales. Royalty rates appear fairly standard, and they appear to vary by book binding. Typically royalty rates escalate as sales increase. Hardcover fiction books tend to have initial royalties at 10% of the list price of the book, rising to 15% after some number of copies, usually five or ten thousand, are sold. Paperback books, largely thought to be non-fiction, command lower royalties, starting in the vicinity of 7.5%. Mass Market paperbacks are supposed to have a royalty rate starting in the vicinity of 10%.[27] Of course, these estimated royalty rates are based on current publisher behavior and the books in my sample are mostly very old. But these rates appear to have been approximately this size for a very long time. According to Kahn (2004), "In the 1840s, reputable authors received an average [royalty] of 10 percent, and later on, between 10 to 20 percent."

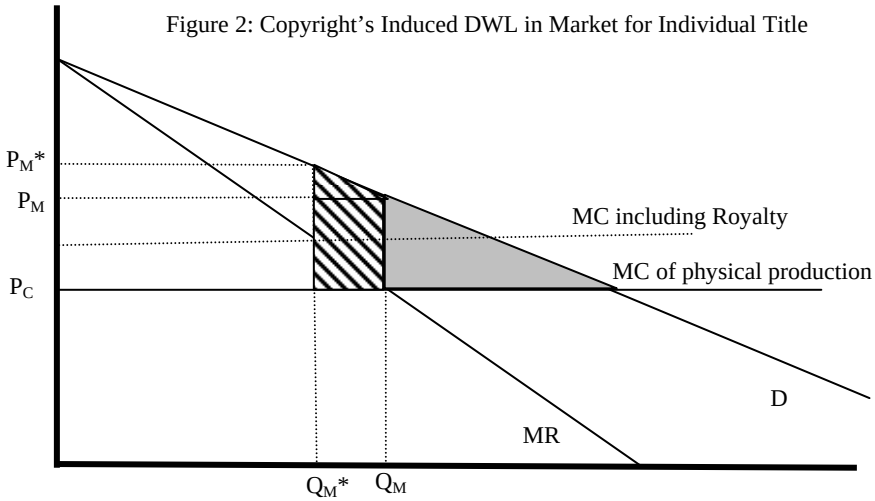
Although it is true that book publishers often pay an advance to authors, this is an advance against royalties and once sufficient sales to cover this advance have been met the author will then get the negotiated royalty rate for each book sold. Thus, if we assume that the publisher makes the pricing decision, it would be more appropriate to add the author's royalty to the marginal cost and to calculate the profit maximizing price on that

[27] For example see http://www.school-for-champions.com/writing/book_royalties.htm. This 15% number seems to be commonly used. It was used, for example, when the NY Times was estimating the royalty payments that Hillary Clinton was deriving from her book "Living History."

basis, which then rises to P_M*.[28] The deadweight loss triangle, in this case, should also include the diagonally striped area between Q_M* and Q_M.

Copyright, then, engenders two differences compared to idealized competition. First, there is a lack of competition in the sales of copies of the title since copyright provides the legal monopoly. Second, there is a payment to the owner of the copyright, the author. The actual price difference between the hypothesized book under copyright or not then would be measured by the price difference between P_M* and P_C, which is what the empirical estimates would presumably be measuring if this is the proper representation of how the market for books works. The difference between P_M* and P_C, therefore, would appear to be between 5% and 10%.

Figure 2: Copyright's Induced DWL in Market for Individual Title



Since the books in my sample were all bestsellers, and frequently by authors already famous, we would expect the royalty payment to be at the high end or in excess of the standard form rate. Although an increase in marginal cost will lead to a smaller percentage increase in price, due to demand being flatter than marginal revenue, when the price increase due to monopoly is added in, it seems likely that copyright would increase price by more than the increase in marginal costs. Yet almost none of the estimates of the increase in the price of the book due to copyright were at or above this 10%-15% royalty rate. Thus the empirical results seem inconsistent with this model.

Nevertheless, sticking with this model, we can ask what the size of the deadweight loss from copyright might be. And how would the deadweight loss it compare to alternative mechanisms for rewarding authors?

There are a few simplifying assumptions that make calculating the deadweight loss easier. First, I continue to assume a horizontal marginal cost curve, which I do not think will elicit too many howls of protest since almost everyone seems to believe that

[28] This might more properly be shown as a step function, equal to the MC of physical production up to the output where royalty payments cover the advance.

information industries have relatively low and relative constant marginal costs. I also going to assume that the demand curve, at least in the vicinity of the market equilibrium, is essentially linear.

In this case the deadweight loss, compared to some sort of idealized alternative, is merely a right triangle with the height equal to the percentage increase in price and the base equal to the percentage decrease in quantity. The area, which is half the product of the base and the height, then provides the deadweight loss measured as a share of initial total revenue.

Table 6 below indicates the size of the deadweight loss for various assumptions about elasticity and the copyright price impact. Elasticities are chosen in a range from .5 to 3. Bittlingmayer (1992) estimates the elasticity of demand for books in Germany to range from 1.5-3 although Clerides find elasticities as high as 4.

Table 6: Deadweight Loss as % of Revenue Compared to Perfection

Copyright Price Enhancement / Arc Elasticity of Book Demand	2.5%	5%	7.5%	10%
0.5	0.016%	0.063%	0.141%	0.250%
1	0.031%	0.125%	0.281%	0.500%
2	0.063%	0.250%	0.563%	1.000%
3	0.094%	0.375%	0.844%	1.500%
4	0.125%	0.500%	1.125%	2.000%

I do not provide any calculations for elasticities greater than 4 because elasticities beyond this range seem unlikely to occur. It would mean that revenues could be easily increased by lowering price and given that the marginal costs are thought to be quite low, such increases in revenues would likely increase profit. Since we are assuming a textbook-like price setting mechanism, it should be easy for individual publishers to lower the price of a title if that would increase profit.[29] Further, in the world of digital distribution, ebooks or iTunes music for example, the marginal cost of distribution actually approaches the royalty rate and profit maximization would, by definition, occur near an elasticity of 1.

The question that naturally arises is whether these are large losses or not. The deadweight losses (compared to revenues) are generally considerably less than 1%, so they seem small, but when talking about size it is useful to know what the deadweight loss is being compared to.

Fortunately there is a fairly natural metric to use as a comparator to these measures, one that arises in the context of current discussions and debates about the efficacy of copyright and its future direction.

Critics of copyright (e.g., Fisher) suggest replacing it with a form of licensing system whereby a government organized agency would collect tax revenues to pay for

[29] Bittlingmayer (1992) argues that almost all production costs, including printing are fixed since most books only have a single printing which occurs prior to any sales.

copyrighted works and use the proceeds from such taxes to pay copyright owners based on measures of usage. This means that the tax receipts would go to the record companies or the book publishers in lieu of market revenues. These proposals suggest emulating the workings of music performing rights societies which currently distribute billions of dollars to copyright owners that they collect from licenses sold to radio and television broadcasters, and other major users of copyrighted works.[30]

Under such a system publishers of copyrighted works would not have to pay anything directly to the copyright owner and thus, in principle, the price of copyrighted works would be driven down to the costs of delivering the work to the consumer. The underconsumption, or access, problem associated with copyright would then be eliminated. This still leaves the underproduction, or incentive problem, to be dealt with. The hope in these proposals is that the government can in some way, perhaps through the use of Ouija boards, or equivalently, by requesting economists to write reports on the subject, divine how much money should be collected through taxation and paid to creators (Liebowitz, 2005). Two additional problems are how to distribute the tax proceeds among the thousands, nay millions, of copyright owners and the size of the deadweight loss caused by the taxes used to generate the revenues use to pay the copyright owners.

Using the performing rights organizations as a proxy would imply that the costs of running a parallel system to standard copyright would tend to be in the vicinity of about 15%-20% of revenues, as is the case for performing rights organizations.[31] Although these organizations spend some resources policing users to make sure they purchase required licenses (e.g., restaurants and bars that use copyrighted music) the major costs of operation consists of monitoring usage of music, determining who gets paid and how much, and lobbying for higher payments. These major costs would all exist in any new organization meant to replace copyright.[32]

Even ignoring the welfare costs from having too much or too little creation of copyright works and the deadweight loss from the taxes, this approximate 20% cost of collecting and distributing the revenues (and the inevitable rent seeking involved in trying to increase the size of government controlled monies) appears to be considerably larger than the likely deadweight loss from copyright found in Table 6.[33]

---

[30] ASCAP and BMI, the two best known music performing rights organizations in the U.S., are often held out as examples.

[31] For example, these numbers can be found in Besen and Kirby who examine several international performing rights societies and are similar to estimates made for SOCAN, the Canadian equivalent of ASCAP and BMI. These numbers appear to be independent of the size of the organization.

[32] The suggestion put forward by Terry Fisher and other legal scholars is that large surveys of usage of copyrighted works be conducted. The surveys would need to be large to gather useful information on the large number of not terribly successful works. Although Fisher is largely talking about digital distribution he admits that without surveys copyright owners can try to game the system to increase their measured share of downloaded or consumed works.

[33] The astute reader will note that I am ignoring the costs of keeping track of sales and paying copyright owners in the current system. Although true, this is probably not a gross oversight. The costs of keeping track of sales figures and royalty payments and then sending out the checks would seem to be a very small

Since the likely deadweight losses from a copyright-alternative system, including those due to taxes and incorrect choice of size, is likely to be considerably greater less than just those involved in running such a system, we can view this conclusion is being conservative. If the estimates of copyright price increases for music or DVDs turns out to be similar to that for books, then the presumed benefit from removing the impediment to optimal consumption would appear to be far smaller than the deadweight losses that would be engendered by a switch to this different system.

IX. Conclusions

To my knowledge this is the first time that an examination of the impact of copyright on the price of copyrighted works has been undertaken. The prices of current editions of best-sellers from the late 19th and early 20th century, some of which are still under copyright and some of which have had copyright expire, were compared The results indicate that there is no clear evidence that copyright increases the price of books, a finding that I believe is likely to be surprising to most students of these topics. To the extent that copyright might increase overall price, the price increase would be very small, less than 10%. An interesting but unexplained finding is that copyright seems to raise the price of low priced books but lower the price of high priced books.

I proposed two possible explanations consistent with these results. First, to explain how copyright might not have any impact on price I appealed to the uniform pricing which is often found in products in copyright-based industries, such as movies, music, and books. If retailers sell different titles at the same price then it is likely that the publisher of any individual title takes that uniform market price as a given. Copyright increases the profit of the publisher by increasing the quantity sold relative to a similarly situated non-copyrighted title, but the increased profit does not depend on price changes. By amortizing fixed costs over a larger quantity the copyright owner has lower average costs and thus, like any natural monopoly, it is more efficient to have only a single producer of the title. Therefore the deadweight loss associated with higher prices normally caused by monopolies does not apply.

The second explanation uses more traditional textbook view of copyright as a monopoly, and concludes that the price increase causes by copyright is no more than 5%- 10% which are the at the higher end of the regression coefficients found for various formulations of the empirical problem. In this case it is possible to estimate the range of likely deadweight losses that are possible. The size of the deadweight loss was unlikely to be more than 1% of the sales revenues. Not only does this seem like a small number, but compared to proposed alternatives to copyright, it appears to be very small in comparison to the deadweight losses likely under the major proposed alternative.

I should note that this is the first empirical analysis of its kind and as such we should be careful not to put too much stock in the results until other researchers have been able to confirm them. Nevertheless, the implication of these results, if they turn out to be robust,

---

component of the total cost which includes looking for new titles, editing the book, publicizing the book, printing and shipping the book and so forth. None of these costs go away in the copyright-alternative world.

is quite clear—the copyright monopoly may be far less serious a problem than has been normally stated in the vast literature on the topic. The results in this paper imply that the underconsumption impact and its attendant deadweight loss is much smaller than the importance attached to it in the literature might have suggested.

Bittlingmayer, G., "The Elasticity of Demand for Books, Resale Price Maintenance and the Lerner Index" Journal of Institutional and Theoretical Economics 148 (1992) 588-606.

?Boldrin M, Levine DK "The case against intellectual monopoly" International Economic Review 45 (2): 327-350 May 2004

Brief of George A. Akerlof et al. as Amici Curiae in Support of Petitioners at 12, Eldred v. Ashcroft, 537 U.S. 186 (2003) (No. 01-618).

Clerides, Sofronis K. "Book value: intertemporal pricing and quality discrimination in the US market for books" International Journal of Industrial Organization 20 (2002) 1385–1408.

Connolly, Marie & Krueger, Alan B., 2006. "Rockonomics: The Economics of Popular Music," *Handbook of the Economics of Art and Culture*, Elsevier.

Dana, James D. and Kathryn E Spier (2001) Revenue Sharing and Vertical Control in the Video Rental Industry *Journal of Industrial Economics* 49 (3) 223–245.

Fisher W.W. (2004), Promises to Keep: Technology, Law, and the Future of Entertainment, Palo Alto, California: Stanford University Press.

Greco Albert N., "Market Concentration Levels in the U.S. Consumer Book Industry: 1995–1996," Journal of Cultural Economics 24: 321–336, 2000.

Greco Albert N., The Book Publishing Industry, 2nd edition, Lawrence Erlbaum Associates, New Jersey 2005

Khan, B. Zorina (2004) "Does Copyright Piracy Pay? The Effects of U.S. International Copyright Laws on the Market for Books, 1790-1920" Working Paper 10271 http://www.nber.org/papers/w10271

Kitch, Edmund W. "Elementary and Persistent Errors in the Economic Analysis of Intellectual Property" 53 *Vanderbilt Law Review*, November, 2000, p. 1727.

Kremer, M. "Patent Buyouts: A Mechanism for Encouraging Innovation", Quarterly Journal of Economics, November 1998, Vol. 113, No. 4, Pages 1137-1167.

?Klein, Benjamin, Lerner, Andres V., Murphy Kevin M. "Intellectual Property: Do We Need It? The Economics of Copyright 'Fair Use' in a Networked World," *American Economic Review*, 206-208, May 2002

Ku, R. S. R. (2002), "The Creative Destruction of Copyright: Napster and the New Economics of Digital Technology", University of Chicago Law Review, 69; 263-324.

Landes, William M. and Posner, Richard A. The Economic Structure of Intellectual Property Law Belknap Press, Cambridge, MA 2003.

Lessig, Lawrence *Free Culture*, The Penguin Press, New York 2004.

Liebowitz, Stan J. "Copying and Indirect Appropriability: Photocopying of Journals," *Journal of Political Economy*, 93-5 October 1985, Pp. 945-957.

Liebowitz, Stan J., and Stephen E. Margolis "Seventeen Famous Economists Weigh in on Copyright: The Role of Theory, Empirics, and Network Effects" *Harvard Journal of Law and Technology* 2 Vol. 18, (Spring 2005) pp. 435-457.

Liebowitz, Stan J. "MP3s and Copyright Collectives: A Cure Worse than the Disease?" Developments in the Economics of Copyright: Research and Analysis, edited by Wendy Gordon, Lisa Takeyama and Ruth Towse, Edward Elgar, 2005.

McMillan, Robert S. (2005), "Estimating Demand for Differentiated Products with Continuous Choice and Variety-Seeking: An Application to the Puzzle of Uniform Pricing," Ph.D. Dissertation, Stanford University.

Orbach, Barak Y. and Liran Einav Uniform prices for differentiated goods: The case of the movie-theater industry International Review of Law and Economics, Volume 27, Issue 2, June 2007, Pages 129-153

Rappaport, Edward Cong. Research Serv., Copyright Term Extension: Estimating The Economic Values (1998), available at http://countingcalifornia.cdlib.org/crs/pdf/98-144.pdf).

Sorensen, Alan T. "Bestseller Lists and Product Variety" Journal of Industrial Economics 2007 55:4 715-738

Yoo Christopher S. "Copyright and Product Differentiation" New York University Law Review 2004 Vol. 79: 212-280.