Copyright : How Large are the Deadweight Losses?

Stan J. Liebowitz

School of Management

University of Texas at Dallas

Very Preliminary

May 2007

Abstract:

Copyright is thought to provide an incentive for authors to create new works. This benefit, however, is also thought to come at the expense of a monopoly deadweight loss since copyright appears to provide a monopoly to the copyright owner. Copyright, by definition, balances in some manner the value of the new works against monopoly deadweight losses although it is unclear whether the balance is anywhere near optimal. This paper empirically examines the book publishing industry to determine the impact of copyright on the prices of books to determine whether and to what extent the price under copyright contains a monopoly component. The empirical work so far indicates that the price differential does not appear to be that much higher than the royalties that are likely paid to the authors of these books. This implies that the majority of any monopoly profits from copyright tends to go to the authors, not the publishers. Thus the deadweight loss, if any, beyond that going to the monopoly talent of authors, seems very small. This information is crucial to the current debate over the efficient form of intellectual property protection.

When economists discuss intellectual property the discussion almost invariably involves the 'tradeoff' which lies at the heart of the debate about the optimal length of intellectual property laws. This balancing is sometimes known as the underproduction/underconsumption tradeoff, or in somewhat clearer nomenclature as the access/incentive tradeoff. The essence of this idea is that we want to give creators, whether inventors or authors, incentives to work on their creations while at the same time allowing access to these creations in a manner that maximizes social welfare.

Exactly what would be the best system for balancing these competing goals has been debated for centuries. Most developed countries have adopted and adapted and approach to balance these competing factors based on what we currently refer to as our institutions of intellectual property. The creators of original works are granted ownership over the rights to reproduce these works but only for a limited time. Ownership is supposed to provide the 'incentive' whereas the limited term of the ownership is supposed to enhance access.

Although the debate over the optimal balancing of these factors has simmered for decades, there has, in recent years, been an outpouring of legal and economic work suggesting that our current intellectual property laws need to be reexamined.[1] Some of these discussions revolve around the length of intellectual property protection whereas others question whether the current copyright and patent regimes, which are the two largest components of intellectual property law, are effective methods for maximizing social welfare.

Central to the economics of this debate is the concept that intellectual property laws create monopolies. If there were no monopolies then there would be no access or underconsumption issues to worry about. Without a monopoly there would be no copyright-induced inefficiency to balance against the gains from providing creators incentives to create.

The purpose of this paper is to examine this claim of monopoly power in the case of copyright. 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 limit the analysis to a single category of copyrighted works, books, which seems an appropriate place to start since that is the market granted the original copyright. The results are thus limited to this particular market, although it seems likely that they may apply to other copyright markets with similar characteristics. Whether there is any relevance for other forms of intellectual property, such as patents, is less clear.

I. *The Nature of the Copyright "Monopoly"*

Economists, along with almost everyone else, tend to equate intellectual property protection with monopoly. When economists discuss copyright (which is not very often)

[1] Lessig, Fisher, Cohen, Boldrin/Levine, Kremer,….

the term monopoly is commonly used just as it is used for patents. 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. Thus, although there is little doubt that copyright does grant the copyright owner a monopoly on making copies of the owners particular work, copyright provides the copyright owner no protection from competing works, independently created, no matter how close in concept they may be.

As already mentioned, copyright is normally thought to provide some sort of balancing between the two putatively competing goals of promoting new works and increasing consumption of those works. The critique of this balance that has achieved most prominence in the writings of copyright critics has tended to focus on whether a market incentive is needed to induce the optimal creation of creative works at all. Sometimes it is suggested that authors receive almost none of the revenues generated by copyrighted works.[2] Although this is largely true for the many works which achieve almost no market success, it is clearly false for those works which do have market success.

Sometimes it is suggested that authors do not require payment in order to be induced to create their works. Although it is true that some works are created for nonpecuniary or mainly nonpecuniary motives, there are a relatively small number of individuals who support themselves by writing full time and these individuals would not be able to devote the same energy or efforts if they were not compensated. These small numbers of individuals provide a very large share of the creative works that are consumed in most advanced countries. The difference in output that would occur if creators were not compensated is likely to be similar to the difference between professional athletes and amateur athletes. Although there is no shortage of amateur athletes, and many amateurs are very fine athletes, the quality of play would certainly deteriorate if athletes were not able to work full-time at their athletic endeavors by getting paid to do so. The same should be true of authors of creative works. There is little or no reason for economists to treat copyright markets as *sui generis* and to drop their normal assumptions that effort will only be undertaken when it is rewarded.

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 payments. Perhaps, the theory goes, some musicians could make sufficient revenues from concertizing that revenues from CDs would be irrelevant, but we have not seen successful musical groups providing free music in order to increase concert revenues. Nor do we see athletes willing to work for free so they can cash out in other markets, such as endorsements.

Of course, there are few participants in the copyright debates actually suggesting that creators should have the fruits of their efforts withheld from them (although this is not a null set). Instead, particularly when economists get involved, the concern focuses on the monopoly inherent in the grant of copyright. 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

[2] Ku, U Chicago Law Review.

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.

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 so forth). Clearly, if enforcement is easy, this provides a literal monopoly on the work in question.

But this does not mean that there is necessarily any economic monopoly or any deadweight loss, as Kitch (2000) has reminded us. After all, every firm has a monopoly on their own name or brand. 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 ownership of his output doesn't provide any economic monopoly. Similarly, we each have a monopoly over our own labor. For most of us, however, this provides no advantage in the market. When workers are all basically the same as one another, the monopoly that each has over their own identity is not worth anything in the labor market.

What really counts is whether there are close substitutes or not for our nominal monopoly. Tiger Woods has a monopoly of sorts in the activity of playing golf. Other golfers are not identical, nor as good as is he. 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 golfer earn a zero economic return. It is another question entirely whether there is a monopoly restriction in the golf market and whether the government should break Tiger Woods fingers and eliminate his monopoly power to increase competition and promote economic efficiency.

The question we must ask is whether the book market is more like the wheat market, the golf market, or some other market. To do this is not a simple task. The book market consists of many individuals producing hundreds of thousands of works every year. Many of these individuals are trying to move from the amateur stage to the professional level where they can support themselves with their creative efforts. Even if we limit ourselves to those individuals who have secured contracts with professional publishers, distributors, agents, and so forth, the great majority are not likely to wield much if any economic monopoly. Nevertheless, there is a small minority of artists who wield clout by virtue of their difficult-to-imitate talent and this talent provides them monopoly power.

Of more direct interest in this paper is the question of whether the publishers wield any monopoly power. Do they wield any monopoly power (vis-à-vis consumers) for the majority of their publications created by 'ordinary' creators? Do the publishers wield any monopoly power when they publish works by authors who are the most successful?

There appears to be free entry into the market for the creation of artistic works, including the authorship of books. It is estimated that over 150,000 new titles are produced in the US alone, every year.[3] When there is free entry into a market we do not expect to find

[3] Greco.

monopoly power except for those cases where the entrants are unable to produce an equivalent quality product as efficiently as the incumbent.[4] In the case of book publishing, such monopoly power may survive free entry if the quality of the work is such that imitation is very difficult. 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. The cause of the monopoly in this case is not the copyright law but the inherent talent of the author although the copyright is required to bring this monopoly to fruition.

Nevertheless, those rare cases of exceptionally well executed books have market shares based on revenues generated that are far above their share of titles. It is possible, therefore, that the publishing industry does, in fact, exhibit economic monopoly in a non-trivial manner because the small number of very successful titles has a large impact on the overall market.

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

On the other hand, this is not how most copyright industries are normally portrayed. A more popular view of copyright industries is that there are powerful large corporations that usurp most of the industry rents, leaving the creative artists with little to show for their efforts.

Regardless of one's view of the industry, however, economists have been unanimous in their depiction of a monopoly deadweight loss on the consumption side of this market. It is this loss which provides the economic rationale for keeping copyright length less than infinite. The deadweight loss found in economic models involves consumers who would like to consume the work but are restricted from doing so by the supra competitive price in the market. This will be the focus of the current analysis. Some legal scholars, on the other hand, focus heavily on users who are not interested in consuming the work but instead are interested in reusing portions of the work in their own endeavors. This is a concern less with the price of the work than with the fact that there is any price for permission to use portions of the work. This is a debate about fair use and the impact of ownership on future creation which will not be the focus of this paper.

With a focus on the 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,

[4] 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. See Yu, …Economists spent decades trying to unravel the welfare implications of monopolistic competition and we will leave these aside since in our empirical work we will have one group of works that are identical and competing with one another versus another where each title is unique.

i.e. the increase in price that causes consumption to decline below the optimal level. We could, if we then wanted to go further, make some assumptions about the elasticity of demand and determine an estimate of the actual deadweight loss.

The basis of the analysis, therefore, will be to measure by how much the price is higher, if at all, for copyrighted works.

II. *The Structure of the Publishing Industry*

Book publishing is the oldest of the industries relying on copyright. [a little history].

Book publishing is normally classified into several different components. Trade books (typical fiction and nonfiction for the mass market) 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 or pricing schemes.

There appear to be relatively low barriers to entry in the publishing industry and the barriers appear to be falling. Printing on demand allows small start ups to produce books with relatively small initial investments. It is for these reasons that Greco finds 53,000 publishing houses in the United States. With the advent of the Internet and eventually ebooks, the entry barriers will fall even further. Sales of ebooks are still just a glimmer in the eyes of the publishing industry, waiting for the device that will provide portability and convenience equivalent to that of paper books.

It is not clear that the concentration could be described 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 largest 4 firms accounted for approximately 38% of the revenues. More recent work by Greco indicates that this level of concentration is overestimated due to an undercounting of small publishers.[5] Nevertheless, even the higher levels of concentration is not terribly high relative to many other industries and is non inconsistent with expectations of competitive behavior.

The Internet has had only a moderate impact on the publishing market. According to USA Today in 2004 sales of hardcopy books over the Internet accounted for 8% of industry sales. Ebooks have yet to have had a serious market impact and the digitization of books is not yet a threat to the book industry the way that digitized music on file-sharing networks has been a threat to the recording industry.

5

We 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

III. *The Data*

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 so simple.

How does one find sets of similar books where some are under copyright and others are not and then compare them? I decided to use current books in order to be able to take advantage of the Internet's ability to check data quickly and because the current market is the one we are most familiar with.

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.

Examinations by Landes and Posner and Rappaport 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 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.[6] Since these books generally disappeared altogether from the market one cannot examine their current prices for new copies.

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 we looked. Liebowitz and Margolis also mention that bestsellers made up a surprisingly large share of the total market (on the order of xx% for fiction).

The process in this paper began by having research assistants examinee best seller lists to find the titles which had the largest sales in a given year. This was the initial starting point in generating the data set. Once these titles were 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.

[6] As reported in Landes and Posner, Indefinitely Renewable Copyright,

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 later, 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 books with the data and checking each title by hand.

The data set, therefore, was constructed of bestsellers from 1895 to 1940, essentially several decades prior to 1922 and several decades after 1922. [I will probably bring the years forward to balance the sample].

Besides the copyright variable, we collected list price, the number of pages, the type of binding, the publisher, whether the book was available to purchase, the selling price on Amazon, category of book content, ISBN number, and sales rank on Amazon. Not all books had complete data for every variable.

The data set originally consisted of 2445 separate observations although many of these titles had multiple observations when there were multiple publishers of a book no longer under copyright or when there were different bindings (hardcover, paperback) for books still under copyright. In many cases the same book was listed with two or three different publication years although these are being weeded out of the data set. There were 603 unique titles among these observations. There were also 280 separate publishers. Of these publishers, 124 had only one book in the sample and 64 publishers had only two books in the sample.

IV. *Empirical Approach*

Since all books in the sample were bestsellers, they had market acceptance as a common trait. There was some concern that publishers might specialize in both a style of book and also specialize in titles under copyright or not under copyright (e.g., a publisher specializing in deluxe leather bound editions of famous works in the public domain). Although the publisher dummy variable might pick this impact up, some of it might also go to the copyright category in which the publisher specialized. To avoid this problem we limited our sample to publishers who published both copyrighted and non-copyrighted works. This restriction should reduce the possibility that publishers who specialized in one or the other had unusual pricing where the copyright variable might pick up some of this pricing that more appropriately belonged with the publisher. This opened up an interesting issue about whether publishers of copyrighted books actually honor the copyright, a topic that we will return to later in the paper.

In order to include publishers likely to have a variety of both copyright and non-copyrighted works we limited our sample to publishers with more than 10 books (who account for 75% of all titles in the sample), of which there were 23. The sample was at

that point reduced to 1837 duplicate titles. Then we eliminated publishers who either specialized or almost specialized (in our sample) either in books in copyright or out of copyright, which left a total of 12 publishers. These publishers were responsible for 872 titles in the sample, but a few hundred of these observations appeared to be duplicates upon a closer inspection or else proved difficult for us to retrieve data on [data cleansing is still going on, particularly for the larger data set]. In the end the data set consisted of slightly over 500 usable observations but even these did not always have sufficient information to be included in the regressions. The title with the largest number of observations was Arthur Conan Doyle's "Hound of the Baskervilles" with 21 observations. It was published in 1902 and thus is no longer under copyright.

The dependent variable was the price of the book. We obtained data on both the list price and the selling price on Amazon, but had far more observations using list price. Since the correlation between the list prices and the Amazon transaction prices was very high (.95) it probably doesn't make much difference which is used and we will use list price in the main as our dependent variable. There are several reasons to prefer list price. First, we had information on transaction prices for a smaller number of books, and that information came from only one seller, Amazon. Second, Amazon appears to base discounts on popularity of the book and large price changes can result when a book loses popularity. Such pricing could provide misleading results. Third, one can easily go back to 'Books in Print' to get the list prices for previous years. One cannot easily retrieve old Amazon data in a useful fashion to keep the prices all in a similar time period.

The explanatory variables that were available included the number of pages and the type of binding. Bindings consist of trade paper, cloth (hardcover), library binding (hardcover), mass market (inexpensive paperback) and a few other less important categories.

In Table 1 you will find summary statistics for the data. First note that the period of time covered by our books runs from 1895 to 1940. The average book is 367 pages long, sells for $25.58 and has a list price of $30.51. One third of the books are still under copyright. Most of the books are classified as "trade paper" although "library binding" is also well represented. There are 11 publishers in the sample. Kessinger has a disproportionately large share of books and we will want to make sure our results are not overly impacted by a single publisher. Most of the books in the sample are fiction, although this is a variable we are in the process of checking.

Table 1					
Variable	Obs	Mean	Std. Dev.	Min	Max
year	517	1915.723	12.32454	1895	1940
noofpages	517	366.9923	133.8133	48	896
listprice	517	30.51201	19.59471	1.5	199
sellprice	448	25.58248	18.69845	1.5	159
nocopyright	415	0.66988	0.470824	0	1
libr	442	0.187783	0.390981	0	1
massmkt	442	0.045249	0.208085	0	1
paper	442	0.024887	0.155957	0	1
prebound	442	0.004525	0.067191	0	1
regular	442	0.004525	0.067191	0	1
trdcloth	442	0.067873	0.251813	0	1
trdpap	442	0.665158	0.47247	0	1
amereon	517	0.059961	0.237645	0	1
buc	517	0.05029	0.218755	0	1
classictxts	517	0.025145	0.156717	0	1
dover	517	0.032882	0.178501	0	1
kessinger	517	0.555126	0.497433	0	1
penguin	517	0.085106	0.27931	0	1
reprintser	517	0.05029	0.218755	0	1
sagebredu	517	0.042553	0.202043	0	1
simonshuster	517	0.065764	0.248109	0	1
thorndike	517	0.029014	0.168007	0	1
turtleback	517	0.003869	0.062137	0	1
fiction	516	0.804264	0.397152	0	1

More information about the various publishers is shown in Table 2. There we show the percentage of books for each publisher that are copyrighted and the total number of books in the sample for each publisher. This sample, containing books for which information on the number of pages could be found, is smaller than that in Table 1.

Table 2		
Publisher	Copyrighted Share	Frequency
Amereon Limited	50.92%	27
Buccaneer Books	54.17%	24
Classic Textbooks	60.00%	10
Dover Publications	30.00%	10
Kessinger Publishing Co	25.81%	217
Penguin Books	36.11%	36
Reprint Services Co.	31.82%	22
Sagebrush Education Re	14.29%	21
Simon & Schuster	53.13%	32
Thorndike Press	42.86%	14

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 with the variables that are supposed to explain the price as the covariates. The form of the price variable

that seems to make the most sense for our purposes is to take natural logs of the prices. There are two reasons for this. First, the payment to the author is generally calculated as a percentage of revenues and one key difference between books under copyright and books no longer under copyright is that the latter group is generally under contract to pay the author whereas the latter group is not. The price differential is thus not a fixed amount but a (relatively) fixed percentage.[7] Second, any monopoly power is likely to be best expressed as a percentage of price.[8]

In the first instance we run a simple OLS regression and then we run the regressions with the publisher dummies included. In each of these instances we also run a robust version of each of these regressions using Stata's built-in routine to weaken the weight of potentially influential observations. The results are found in Table 3.

Table 3: Explaining log of list price				
	OLS	RREg	OLS	RREG
copyright==yes	0.0498	-0.0132	0.1267	0.0111
	(0.429)	(0.253)	**0.000**	**(0.058)**
No. of Pages	0.0019	0.0015	0.0012	0.0015
	0.000	**0.000**	**0.000**	**0.000**
binding==trade cloth	0.1327	-0.2090	1.0494	0.9275
	(0.029)	**0.000**	**0.000**	**0.000**
binding==mass marke	-1.2628	-1.5946	-0.3239	-0.3565
	0.000	**0.000**	**0.000**	**0.000**
binding==library bind	0.3592	-0.0966	0.9031	0.9339
	0.000	**0.000**	**0.000**	**0.000**
publisher dummies	no	no	yes	yes
Constant	2.4445	2.9330	2.4393	1.9156
	0.000	**0.000**	**0.000**	**0.000**
Observations	359	359	359	359
R-squared	0.285	0.942	0.741	0.996
Robust p values in parentheses				

The first two columns provide results for simple OLS regressions. The next two columns include the publisher dummies (the coefficients for the publishers are not shown). The columns with the term RREG in the heading provide results for a special Stata routine which lowers the weight of observations that appear to be outliers.[9] We will quickly pass over the OLS regressions and merely note that the copyright variable is essentially zero when publisher dummies are not included.

[7] Royalty rates differ by author and even for a single book where the rate increases as the quantity of sales increases, although for bestsellers the rate is presumably at the top.

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

[9] This routine first eliminates observations with levels of Cook's D that are above a particular threshold, then it iteratively lowers the weight for observations with large absolute residuals until a convergence threshold is reached.

When publisher dummies are introduced into the regression we get results indicating that copyrighted books have higher prices than non-copyrighted books. In the ordinary publisher dummy regression we find (1) that copyrighted works have higher prices than non-copyrighted works; (2) that the difference in price appears to be in the vicinity of 12.7%; and (3) that the 95% confidence interval around this point estimate (not shown) ranges from 5.6% to 19.7%. The outlier reduced (RREG) version of this regression, however, indicates that copyright has a far smaller impact on price, only 1.1% so although it is still on the border of statistical significance it is no longer economically significant.

These regressions also show that the number of pages impacts the price of a book, with each hundred pages increasing the price of a book by 12%-15% (although this variable is probably better estimated in a regression where the dependent variable is not logged). Binding also plays an important role, with cloth and library bindings have prices about double that of trade paper (the left out category) and mass market paperbacks having prices about 30% lower than trade paper.

V. *What might these results mean?*

If we were to base our conclusions on these results we would be hard pressed to conclude that copyright leads to higher prices at all. Although the pure model with publisher dummies indicates a positive and significant impact of copyright, it is only 12.7% which is barely above the standard form contract of royalty payments. Since these books were all bestsellers, and frequently by authors already famous, we would expect the royalty payment to be handily in excess of the standard form rate. When the impact of outliers is reduced, the impact of copyright appears to be essentially zero (1.1%), which is far below expected royalty payments being made to the author.

The former results would imply that the publishing market is, for all intents and purposes, perfectly competitive even though the market for titles would not be. If these numbers were correct it would mean that copyright allowed authors to generate payments and perhaps some monopoly rents, but that publishers merely recovered their costs of production. All the rents, if any, would go to the author.

The latter results, zero price difference between copyrighted and non-copyrighted books, on the other hand, certainly would appear difficult to reconcile with normal textbook models of supply and demand. Since royalty payments are paid on a per book basis 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 copyrighted equilibrium price would be expected to be higher than the non-copyrighted price, contrary to the empirical results of a zero impact on price.

Before we can dismiss this zero result as being inconsistent with basic economic understanding, however, we need to look a little more carefully at this market. In the typical economic model of competition, profits vanish when the entry of firms increases industry output, causing the price of the product to fall until the demand facing the firm becomes tangent to the typical firm's average cost curve. Note as well, that in the typical model of markets, higher demand lead to higher prices.

Many copyright/entertainment markets do not seem to function in this manner. Prices do not seem to differ between different titles, even when demand is known to be greater for product A than for product B. It is not clear, for example, that there is a generally accepted explanation for the common price found at movie theaters for each movie being shown even though some are clearly bigger hits than others. Instead, the theater runs showing of the more popular movies on multiple screens, adjusting only the quantity supplied to demand variations instead of adjusting price as well.

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. In this case the seller may think that it is important to set a predictable price to consumers and that might explain the pricing in this market. The goal is to keep the theater filled. Weekly changes in price and variations across theaters may scare off consumers who prefer predictability. The price of the retail product, therefore, may be largely exogenous for any individual title although it is endogenous to the entire set of titles. Similarly, CDs from well known groups do not seem to command a price premium over CDs from less well known groups. The pricing on Apple's iTunes, where each and every song sells at the identical price, is the strongest example of this type of pricing this.

How does competition operate in a world where retail prices cannot change? In such a world the sellers of above average demand wholesale products can try to set a higher wholesale price, reducing the markup of the retailer, or they can accept that volume (quantity) will be the sole variable that determines their profit, with price a variable that is beyond their control.

In such a world, above normal profits are driven down to zero in competition by additional firms or titles entering the industry and reducing available sales to each remaining title until the average title can only generate normal returns. Instead of price falling for individual firms as entry occurs, quantity falls. For such a model to work there needs to be a fixed cost per publisher (title) so that additional publishers (titles) increase industry average costs even as prices remain constant, but fixed costs seem a realistic and important aspect of this industry.

In such a world, the typical deadweight losses associated with monopoly are no longer valid. The only advantage of a title with monopoly power relative to a title that has almost perfect substitutes is that the former has lower costs per unit (and higher profit) that goes along with the larger volume. In this case, there is no price advantage for consumers when a book is no longer under copyright and thus there is no restriction in consumption and no deadweight loss. Copyrighted books, where the author is paid require larger quantities to be sold, everything else equal, relative to non-copyrighted books.

Is this the explanation for the results found above, particularly the results indicating no difference in the price of books whether under copyright or not? Perhaps, but we do not have enough information to know for sure, as the next section makes clear.

VI. *Problems with the publishers*

When trying to find information about some of the smaller publishers, we often found claims on the web site of these publishers that they honored copyright obligations. One potentially troubling aspect of this analysis, however, is that we came across stories on the web claiming that some of these publishers do not pay for rights to works still under copyright.[10] We reran the regressions removing the small number of publishers specifically mentioned in stories as not respecting copyright and found little change in the results results.

It is apparently very easy for a 'publisher' to set up business these days. Most of the publishers on our list are minor publishers. Some of them may fall into the 'fly by night' category. If there are publishers in our data set violating copyright then we would expect to find that the expected increase in price for copyrighted works may not show up for the simple reason that many copyrighted works are being treated by these publishers as if they were not copyrighted, allowing them to charge the same price as for non-copyrighted works.

Since every publisher claims to be legitimate, it is difficult to solve the potential problem described above. One method of assessing such a problem would be to limit our sample to only those publishers that are well established and who therefore are unlikely to knowingly violate copyright. There are only three of our publishers that seem to fit this description of having been in existence for many years and having a substantial set of books in print: Penguin, Simon & Shuster, and Dover. The downside of using just these publishers, however, is that we have a much smaller data set.

Table 4

dependent variable	log of price		price	
	OLS	RREG	OLS	RREG
copyright==yes	0.2298	0.2656	1.3973	2.8897
	(0.0230)	**(0.0130)**	(0.2253)	**(0.0013)**
No. of Pages	0.0014	0.0013	0.0208	0.0095
	0.0000	**0.0000**	**(0.0017)**	**(0.0059)**
Constant	1.6938	1.7165	2.4342	6.0561
	0.000	**0.000**	(0.277)	**0.000**
publisher dummies	yes	yes	yes	yes
binding dummies	yes	yes	yes	yes
Observations	72	72	72	71
R-squared	0.721	0.702	0.7728	0.8078

[10] These accusations were leveled against Kessinger, Amereon and Buccaneer. 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. Although this may merely be boilerplate attempting to deal with issues that arise beyond the contract reached with a putative copyright owner, it almost sounds like Kessinger doesn't know which of its publications may be violating copyright, which seems strange. This story about Kessinger by an unhappy copyright owner makes the point very well: http://new.businesscommonsense.com/enews/fullStory.bsp?sid=32029&var=story .

The results in the first two columns represent the regressions with the natural log of the price as the dependent variable. For this smaller sample (72 observations) copyright has a statistically and economically significant coefficient whether we try to minimize the impacts of influential observations or not. We find that copyright appears to increase the price of a book by 23%-27%, which is considerably larger than before. Nevertheless, given that these authors have written best sellers and are generally well known, we would expect the royalty rate to be greater than 10% and although these price differences are greater than the standard rate royalty, and probably also above the higher actual royalty rate would have been for such authors, the price differential for the publisher would seem to be a minority of this price differential. Except for the payment to the author, it seems unlikely that the increased price paid by the consumer is more than ten percent above a competitive rate.

The second two columns present the results when the actual price is used as the dependent variable. In the standard OLS results the price is only $1.40 higher for these books and the result is not statistically significant. Weakening the impact of influential observations restores the copyright coefficient to statistical significance and raises the coefficient to $2.89. Since the average price of a book for this sample is approximately $10, it is simple to calculate the percentage increase caused by copyright for these regressions. The regular OLS regression has a smaller difference when using price instead of log price as the dependent variable, only about 14%. When influential observations are removed, however, the coefficient is about the same but actually slightly larger than with the logged dependent variable.

VII. *Analysis*

Although these results are still preliminary, we might ask what it would mean if these results hold up as we improve the data. At this point we have two results somewhat at odds with each other. The first is that copyright has no impact on the price of books. The second is that copyright appears to increase the price of books by about 25%. If the first result were correct we would need to rethink how we model the copyright tradeoff and whether there is any sort of deadweight loss associated with it. This would require a major rethinking of our understanding of the economic consequences of copyright. We also would want to determine whether and to what extent these results hold for other copyright industries. Most likely, if this result holds, copyright becomes an ideal method for paying the author.

The second result indicates that copyright does raise the price of books and by somewhat more than we might expect to be the payment to the copyright owner, although more work on this specific question, i.e., what royalty rate did authors of bestsellers normally get during this period of time, is clearly called for. Since the standard royalty rate appears to be about 10%, and there are usually escalator clauses raising that rate as sales increase we can conservatively assume that at least 12.5% goes to the authors of these best selling works leaving 12.5% for the publisher. Less conservatively, we could assume that all of it goes to the author.

We will assume, from this point, that there is a price increase for copyrighted books, i.e., that the second set of results is more informative than the first. What, then, might the size of the deadweight loss from copyright be? And how does it compare to alternative mechanisms to reward authors?

There are a few simplifying assumption that we are going to make when calculating the deadweight loss. First, we 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 information industries have relatively low and relative constant marginal costs. This is particularly true for digital goods, where the marginal cost is thought to be constant at a value of essentially zero. We are 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 total revenue.[11]

Table 5 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. The deadweight loss rises as the elasticity rises for the simple reason that the quantity reduction is greater for any increase in price when the elasticity is higher.

Table 5: Deadweight Loss as % of Revenue Compared to Perfection

Copyright Price Enhancement	12.50%	20%	30%	40%
Arc Elasticity				
0.5	0.39%	1.00%	2.25%	4.00%
1	0.78%	2.00%	4.50%	8.00%
2	1.56%	4.00%	9.00%	16.00%
3	2.34%	6.00%	13.50%	24.00%

I do not provide any calculations for elasticities greater than 3 because it seems unlikely that the market would be that elastic. 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. Although individuals firms cannot raise price to achieve an increased monopoly profit it is easy for individual firms to lower price if that would increase their profit. Any elasticities larger than 3 seem unlikely, for that reason.[12] Further, in the world of digital distribution, ebooks or iTunes music for example, the marginal cost of distribution actually approaches zero and profit maximization for the industry occurs at an elasticity of 1 and competition would force the elasticity below 1.

[11] A cite to Harberger, maybe?

[12] Some estimates of price relative to marginal cost would be nice here.

It is possible to separately calculate the loss for the publisher component of the monopoly assuming that the payment to the authors is a necessary deadweight loss.[13] By this we mean a deadweight loss (relative to an ideal) that must occur in the real world to generate any surplus in the market at all. Even when the higher price is split fairly evenly between the publisher and the author the publisher will be seen to have a disproportionately large share of the deadweight loss in these calculations because I assume that the payment to the authors occurs first and the deadweight loss from publishing builds on the already distorted prices. In other words, the marginal deadweight loss from the first lost unit due to publisher monopoly is a finite number, not one that approaches zero.

That said, Table 6 provides these deadweight loss calculations. Although lower, the differences between Table 5 and Table 6 are not very large so this distinction doesn't seem to have important consequences.

Table 6: Publisher Portion of Deadweight Loss as % of Revenue				
Deadweight Loss from Publisher Monopoly Power compared to Ideal				
Copyright Price Enhancement	12.50%	20%	30%	40%
Arc Elasticity				
0.5	0.00%	0.61%	1.86%	3.00%
1	0.00%	1.22%	3.72%	6.00%
2	0.00%	2.44%	7.44%	12.00%
3	0.00%	3.66%	11.16%	18.00%

The question that naturally arises is whether these are large losses or not. To what are they to be compared? Perfection is not possible and these deadweight losses are calculated relative to perfection. Is a deadweight loss of 9%, which seems to be about the maximum that we might expect based on our analysis, large enough to worry about? How about a more likely value of 4%-5%? Fortunately there is a fairly natural metric to use as a comparator to these measures.

Critics of copyright (e.g., Fisher) prefer to put in its place a form of licensing system whereby the government would collect tax revenues to pay for copyrighted works and the proceeds from such taxes would be distributed to copyright owners based on measures of usage. Users of copyrighted works do not have to pay anything directly to the copyright owner and thus 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 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.

[13] For a discussion of the concept of necessary deadweight loss see Liebowitz and Margolis, 2005 Harvard Journal of Law and Technology.

Ignoring the issue of how large the payments should be, there is obviously an issue about how to distribute the tax proceeds among the thousands, nay millions, of copyright owners. There is also the question of the size of the deadweight loss caused by the taxes used to generate the revenues with which to pay the copyright owners.[14] Most proposals suggest emulating the workings of performing rights societies which currently distribute hundreds of millions of dollars to copyright owners that they collect from licenses sold to radio and television broadcasters, and other major users of copyrighted works.[15]

Although these organizations spend some resources policing users to make sure they purchase required licenses (e.g., restaurants and bars that use copyrighted music) and lobbying for higher payments, the major cost of operations consists in monitoring usage of music and determining who gets paid and how much. Those costs would remain in any new society meant to replace copyright.[16]

The costs of running performing rights organizations tend to be in the vicinity of about 15%-20% of proceeds. This number is fairly constant, even for very different sized organizations.[17]

Even ignoring the social costs from having too much or too little creation of copyright works, the costs merely of distributing the revenues (and the inevitable rent seeking involved with government mandated monies) appears to be larger than the likely deadweight loss from copyright, as indicated in Table 5 since the maximum potential price increase is thought to be less than 30%, leading to a maximum deadweight loss of 13.5%, which is less than the costs of running an alternative *distribution* system.[18] Since the likely deadweight loss is considerably less than this maximum estimate we can view this conclusion is being conservative.

Of course, this also ignores the problems with possibly (and likely) choosing the incorrect size of these markets as far as the money that the government will make available to creators. That imperfection is likely to be very large compared to the imperfections found here, given the imprecision of the Ouija Boards mentioned above. Thus, if the estimates of copyright price increases for music or DVDs turns out to be similar to that for books,

---

[14] Since this tax would be on top of other taxes already in place the deadweight loss might be quite large relative to the revenue raised.

[15] ASCAP, BMI, and SESAC are the major performing rights organizations in the U.S.

[16] 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.

[17] For example, these numbers are similar to those for SOCAN, the Canadian equivalent of ASCAP and BMI.

[18] The astute user will note that I am ignoring the costs of keeping track of sales and paying copyright owners based on those sales for the current system. Although true, this is probably not a gross oversight. The costs of keeping track of the sales figures and payments would seem to be a very small component of the total of 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.

then the presumed benefit from removing the impediment to optimal consumption would appear to be far smaller than the costs of removing copyright.

VIII. *Further Work*

The key work to still be performed consists of cleaning up the data and checking to determine whether the results reported here are in fact correct. When this is done I plan to look at the history of prices for books after they are published. One can go to books in print and check out the prices for best sellers in the year they are published and then follow those prices over time and see if they keep up with the prices of the then new bestsellers. Similarly, we can determine whether current best sellers have prices different than these very old best sellers.

My RA is German and he claims we can replicate this examination for German books. If so, it would be interesting to see how the results differed across the two countries. Eventually one might be able to perform this type of analysis on music and movies to see how they compare. I may leave it to others to perform those tasks.

IX. *Preliminary Conclusions*

I am reluctant to put much down in the way of conclusions given the preliminary nature of the analysis. What we have done is to perform a simple examination of the impact of copyright on the purchase price of the product. To my knowledge this is the first time that such an analysis has been performed. The results that we have found indicate that increases in price, to the extent that they exist at all, tend to be less than 25%.

Since a large chunk of this goes to authors, copyright appears to be a fairly effective mechanism for paying authors. The extent of deadweight losses, even including the author payment as a deadweight loss, seems likely to be considerably less than 9%. Thought of as a transaction cost of running a system this seems rather small and indicates that copyright may be efficient compared to its putative replacements.

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

Khan, B. Zorina "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

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

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

Greco 1999, J of Media Economics 12 3 pp165-181.

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

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

Liebowitz, Stan J. "Will MP3 downloads Annihilate the Record Industry? The Evidence so Far" *Advances in the Study of Entrepreneurship, Innovation, and Economic Growth*, V. 15, 2004, pp. 229-260. http://ssrn.com/abstract=414162

Liebowitz, Stan J. "The Elusive Symbiosis: The Impact of Radio on the Record Industry" *Review of Economic Research on Copyright Issues* Vol. 1, 2004a, pp.20-45. http://ssrn.com/abstract=520022

Liebowitz, Stan J. *Rethinking the Network Economy*, Amacom, New York, 2002.

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

Liebowitz, Stan J., "The Impact of Reprography on the Copyright System". Copyright Revision Studies, Consumer and Corporate Affairs, ISBN 0-622-11396-9, Canada 1981, http://ssrn.com/abstract=250082

Lyman, P. and H.R. Varian (2003), "How much Information 2003", www.sims.berkeley.edu/research/projects/how-much-info-2003.

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

Siwek, Stephen E., "Copyright Industries in the U.S. Economy: The 2002 Report" 2002, Economists Inc.

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

Varian, Hal "Copying and Copyright" *Journal of Economic Perspectives*, forthcoming.

Watt, Richard Copyright and Economic Theory: Friends or Foes? Edward Elgar: Cheltanham, 2000.

Watt, Richard "The Past and the Future of the Economics of Copyright" *Review of Economic Research on Copyright Issues*, 2004, vol. 1 (1), pp. 1-11.

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

****

Or just get prices for current best sellers and compare them to the data on very old bestsellers to see if there is a difference.***

****

Could it be that publishers of copyrighted books need to amortize losers over all winners, whether they are copyrighted or not? Could there be publishers who cream skim, have no losers, but specialize in famous old books? Would they have lower prices because they do not have to amortize losers? A dummy for publishers specializing in non-copyrighted books would capture this effect.

****

It might be argued that prices for old books still under copyright are no higher than those that have recently lost it because they are close substitutes for one another and the competitive price of the latter keeps the price of the former down. It is not clear whether this would be the case in terms of substitutability or not, although for styles of music, such as big band, one can see how this might make sense.

Can we do anything about this?

Kahn states, regarding average royalties that: “In the 1840s, reputable authors received an average of 10 percent, and between 10 to 20 percent. However, there was wide variation in contracts for unknown authors.”

You might want to check whether selling and list prices tend to have a fairly stable relationship.

Alibris glossary of book terms: http://www.alibris.com/glossary/glossary.cfm#cloth