

Made curious by the continual claims of politicians and industry executives that stronger copyright leads to more innovation, I went to the library early Freshman year to see if there was any corroborating research. I was unable to find any, so I went to a historical index of statistics. However, that only had data until 1970, so I extracted the more recent data from the annual *Statistical Abstract(s) of the United States*.

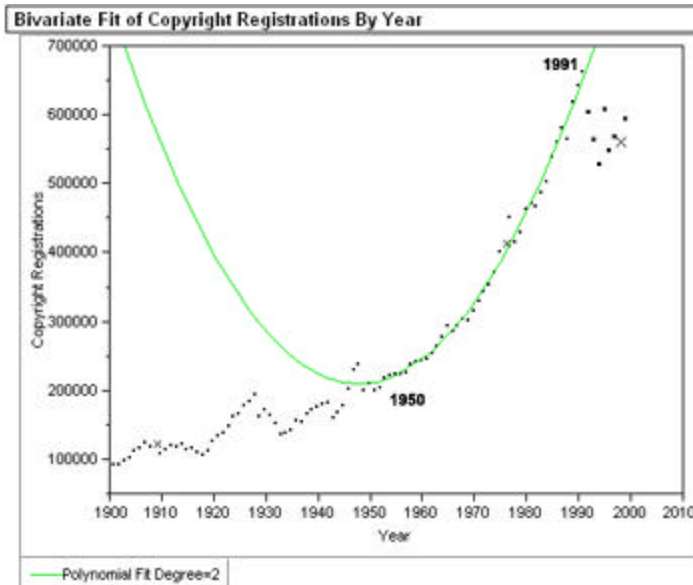
The trends are fascinating, especially in a field where a surprising amount of innumeracy and overinterpretation appears from people who should know better. For instance:

"We did a survey in April that asked people the reasons why they downloaded, and 65% said because it was free," a BPI spokeswoman [said](#).

They are, of course, absolutely correct. But they leave it up to the reader to infer that those respondents are displacing purchases with free music. In effect, however, what is happening is price discrimination. Those who are willing to tolerate lower-quality music are paying less (nothing) for it. Those who are not pay more. Society gains, the industry loses--and then only assuming [recent studies](#) showing that downloads serve as a form of music sampling, a free preview for users that later buy music, are incorrect.

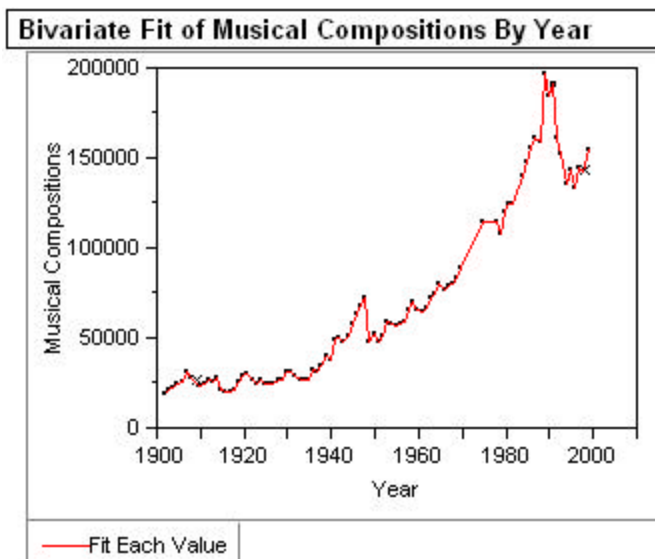
Now, on to the data. Some of this pertains directly to copyright, others directly to the RIAA.

Most interesting to me was one trend that my statistics professor, [Professor Wyner](#), pointed out. From the early 1950's until 1991, copyright registrations rise exponentially. In fact, a simple quadratic fit shows an Rsquare of over .99 .



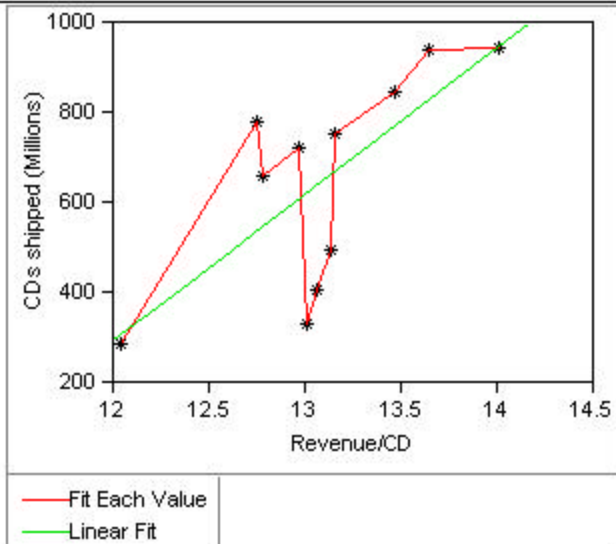
That a four-decade trend of such strength could reverse itself in a single year so dramatically--and without an apparent cause--is incredible. The fact that it happens across all categories of copyright suggests the effect is perhaps due to a change in the way the Copyright Office records entries. However, given that music registrations correlate well with overall registrations, it would have to have been a policy change for all copyright entries. The sheer precipitousness of the plummet belies many otherwise viable explanations. However, in 1992, Congress passed Public Law 102-307, making renewal automatic for works from 1964-1977. Depending on whether the Copyright Office was including renewals in its statistics, 1991 could be a break in analyzability for the data. Furthermore, if they did, indeed, include renewals, trends will be blurred and obfuscated by the lagging renewal registrations.

The single-category music registrations show the same plunge.



Also interesting is that, as the price of CDs increase, shipments increase. This trend is not nearly as strong as the former, and is only based on a decade of data [provided by](#) the RIAA. Possible explanations for this trend include that CDs are a luxury item--unlikely, I should think--or that the economy's rise during this period (1990-2000) lead to an increase in spending.

Bivariate Fit of CDs shipped (Millions) By Revenue/CD



Linear Fit

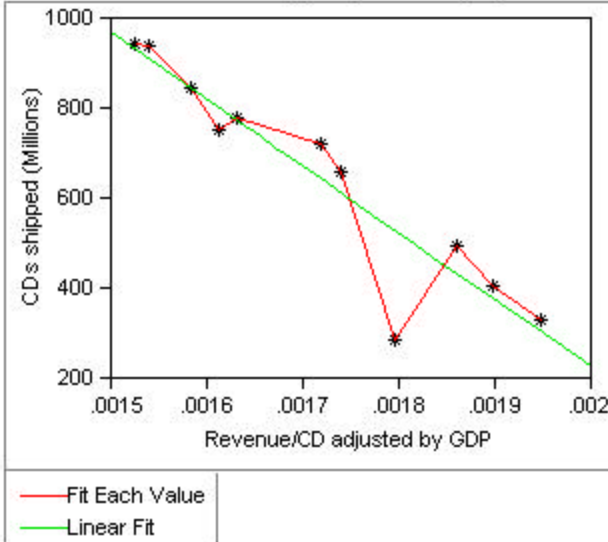
$$\text{CDs shipped (Millions)} = -3607.286 + 325.10319 \text{ Revenue/CD}$$

Summary of Fit

RSquare	0.504259
RSquare Adj	0.449177
Root Mean Square Error	174.946
Mean of Response	651.6455
Observations (or Sum Wgts)	11

And, in fact, it did. A classical Demand Curve. Not such a great mystery after all, as it turns out.

Bivariate Fit of CDs shipped (Millions) By Revenue/CD adjusted



Linear Fit

CDs shipped (Millions) = 3191.308 - 1481090.9 Revenue/CD adjusted by GDP

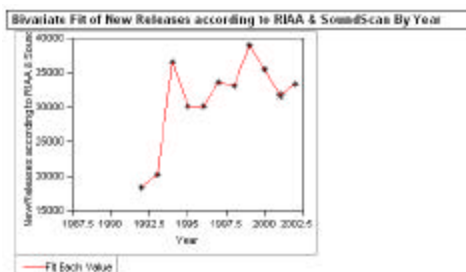
Summary of Fit

RSquare	0.861235
RSquare Adj	0.845816
Root Mean Square Error	92.5587
Mean of Response	651.6455
Observations (or Sum Wgts)	11

Since we are starting to analyze statistics provided by the RIAA at this point, I should mention that they have a nasty tendency to only release data which they can put a proper spin on. Consequently, analyzing becomes much more difficult and leads to kludges such as the 2002 CDs shipped data extrapolated from news of an 8.8% decline from previous years. If anyone would provide me with a complete set of Nielson SoundScan statistics this project would be much easier. If anyone disputes my figures *please* provide me with a better set. Many of these numbers took

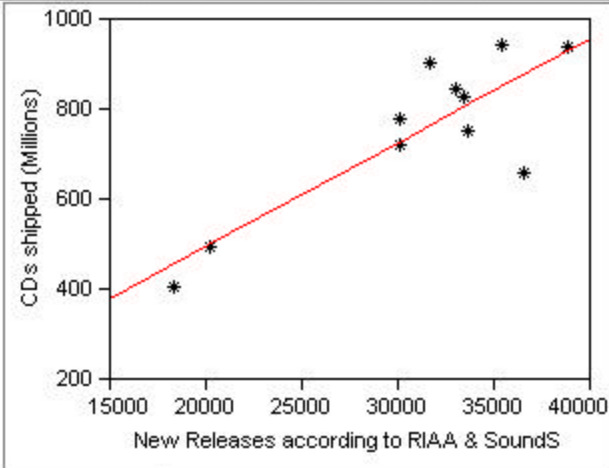
hours to find, here from one source, there from another. Fortunately, most of the time there was some overlap in data provided, so I was able to see that the numbers were directly comparable.

That said, the numbers are interesting. The RIAA has been shipping fewer CDs in the last few years, by all accounts. The most recent (and most contested) numbers come from SoundScan, and so should be pretty reliable. The rest come from the RIAA itself, which does not provide data for those years.



As one might expect, offering more music produces more sales. Here, CDs shipped is used as a proxy for CDs sold because the data is more available. Again, if anyone has data available for sales, please send it to me.

Bivariate Fit of CDs shipped (Millions) By New Releases according to RIAA & SoundS



— Linear Fit

Linear Fit

CDs shipped (Millions) = 32.784997 + 0.0231547 New Releases according to RIAA & SoundS

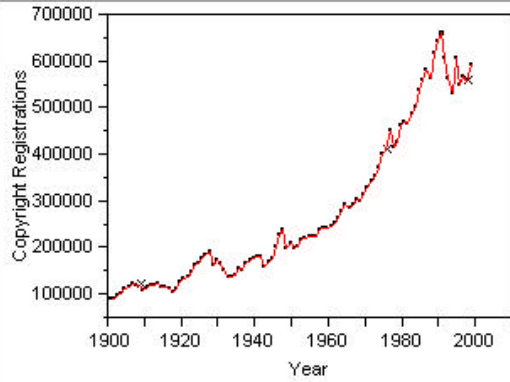
Summary of Fit

RSquare	0.717636
RSquare Adj	0.686262
Root Mean Square Error	97.64667
Mean of Response	752.8818
Observations (or Sum Wgts)	11

Based on this data, one might well expect to see a decline in sales given the approximately 10% decline per year in releases from 1999-2001. From 1999 to 2000, CDs released fell 8.70%. In 1999, the linear model predicts 933 million CDs shipped. In 2000 it predicts 855 million shipped, a decline of 8.36%. The actual value is 942.5 million. However, sales that year may have been particularly bad relative to units shipped because the industry overshipped based on past sales, not taking into account the decline in new releases. Again, if anyone has the CD sales data by year I would love to have it.

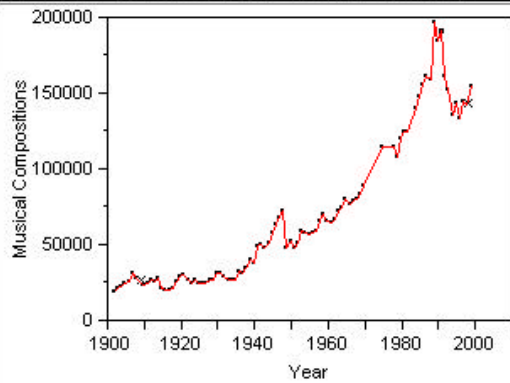
Now for the pretty data:

Bivariate Fit of Copyright Registrations By Year



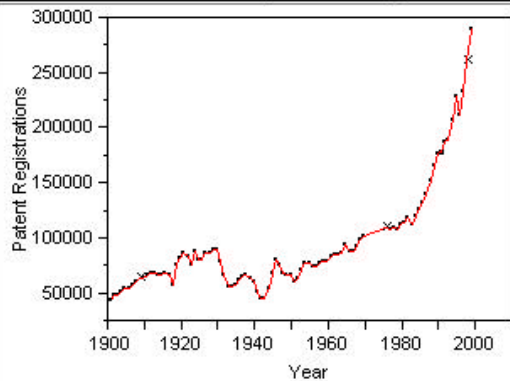
— Fit Each Value

Bivariate Fit of Musical Compositions By Year



— Fit Each Value

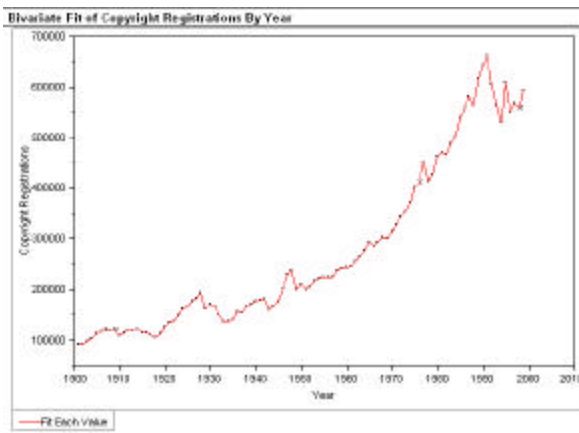
Bivariate Fit of Patent Registrations By Year



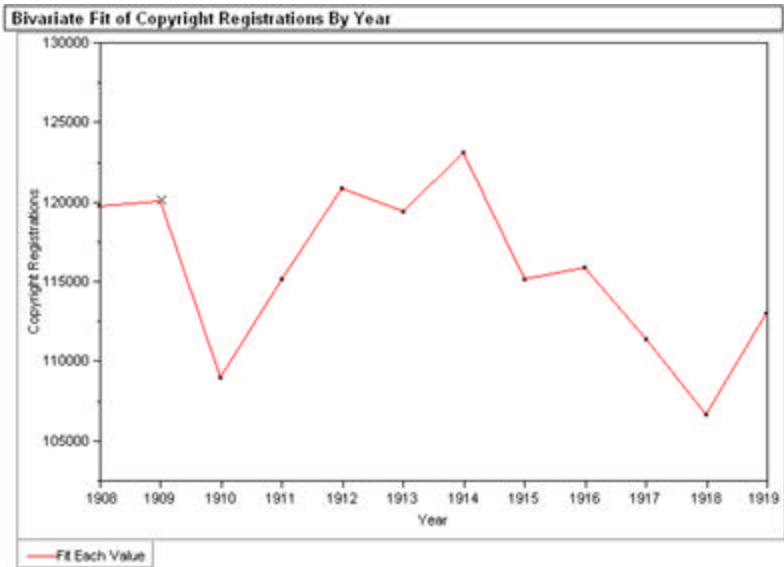
— Fit Each Value

The patent data I gathered to try to answer Professor Wyner's question: why do copyrights plunge after 1991? However, all the additional data did was increase my puzzlement. Not even a hint of a plunge. Curiouser and curiouser!

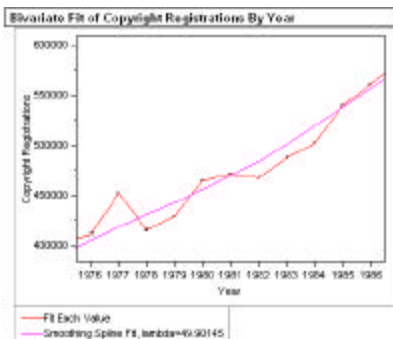
Perhaps the most amazing thing about the copyright data itself is how much it is affected by major historical events. The Great Depression is clearly visible, along with the post-war speculative boom leading up to the market crash. Continuing onwards, one can see World War II and post-war expansion, followed by the Korean War and an economic adjustment. After this, the aforementioned 40-year trend begins. The 1909, 1976, and 1998 "X" markers indicate expansions of copyright law.



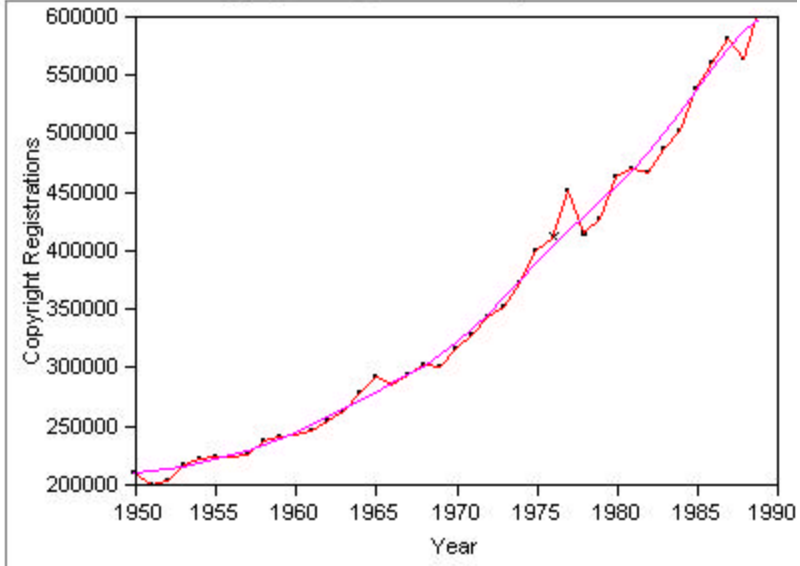
The expansion of copyright law in 1909 did little to encourage innovation as the Constitution proscribes it must. In fact, for a decade afterwards the best that could be said of the law was that it maintained the status quo. This is particularly interesting because it happened in the middle of a boom following the Panic of 1907.



The next time Congress expanded copyright was in the middle of the dramatic increases in copyright registrations from the mid century almost to the turn of the millenium. The effect this time is much more qualified. The overall trend is upwards, but that certainly cannot be attributed to the law, as the decade shown deviates little from the overall 50-year trend.



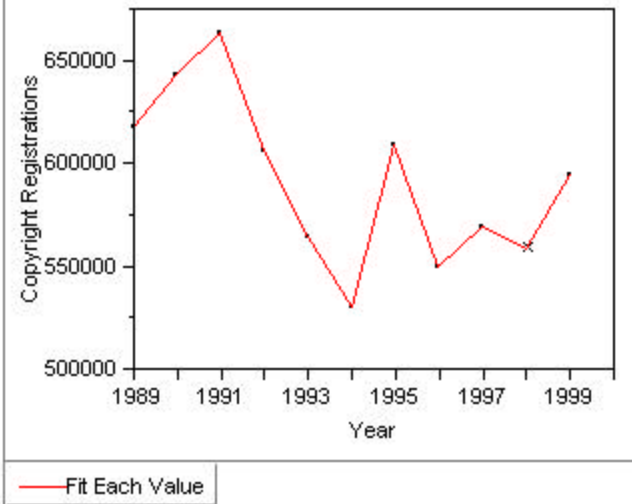
Bivariate Fit of Copyright Registrations By Year



The enormous hump in 1977 seems to be an artifact of looking at all copyright registrations. Since categories were added, a rush to copyright previously-unprotected works would have ensued. An analysis of a single category would provide further clarity. Unfortunately, I am missing the data for those years.

The most recent decade provides the least information of all, unfortunately. Many analyses have been done of the miniscule incentives to produce the latest round of copyright expansion gives in return for gutting the public domain. However, the actual effect of the law remains unmeasured to my knowledge. Consequently, it would be useful if enough data were available to draw a conclusion. However, that is simply not the case as of my research a few months ago, especially given the eccentricity of the data since 1991.

Bivariate Fit of Copyright Registrations By Year



What about the early years, you ask? I spend months searching, and found nothing. The Copyright Office is all-but-useless in gathering the statistics for the early years. In fact, the various people I talked with in my visit to the Library of Congress had very little of substance to say at all regarding serious research. The entire Copyright Office seems to be set up with the sole intent of helping would-be copyright holders do copyright searches. In addition, 1870 was the first year records were centralized in Washinton. Prior to that they were kept in the District Courts. The National Archive was somewhat helpful. "<Laughter> Good luck! You'll never find that," was the response I received before the librarian helpfully tried to find as much information as he could--painfully little, just as he predicted.

And now for the main question of this document: does increasing the length or protective powers of copyright has any effect on innovation as measured through the number

of registrations? The choice of metric is unfortunate because it says nothing about the quality of those works produced, as well as being affected by changes in the way such things are registered. However, given that hundreds of thousands of works are produced each year, one must assume that the sheer numbers involved evens out the effects of differing quality. So the premise remains valid. The conclusion is pretty clear as well, as seen from the decades following the passage of the 1909 and 1976 laws: the drastic expansions of copyright had little to do with increasing innovation in this country. As such, in future years they may well be ruled [unconstitutional](#), [Eldred v. Ashcroft](#) notwithstanding.

All graphics on this page were generated by [JMP-IN](#) from data I've collected over the past year. If you are a student, you can purchase a copy of JMP for about \$70. If not, I believe it's a few hundred, but well worth it.

[Data file 1.](#) [Data file 2.](#)

If you have any data to add to these files, please [e-mail me](#) to that effect. If you disagree with my numbers, please [e-mail me](#) to that effect. If you disagree with my analysis, please [e-mail me](#) to that effect.

Look for my symphony, which should be completed by the end of this summer.

Ari Friedman

Student

[The University of Pennsylvania](#), [1740](#)