Instrumental music measures

Eric Blair

6 April 2004

This is a review of this paper by Oberholzer and Strumpf on music downloads via peer-to-peer (P2P) networks and their effects on music sales. The claim by the RIAA is that music sales are down over the last several years because of file sharing, while this paper looks for such effects but finds none. Here is the response by the RIAA, cut and pasted from the New York Times:

> Amy Weiss, [RIAA Senior VP for Communications], expressed incredulity at what she deemed an "incomprehensible" study, and she ridiculed the notion that a relatively small sample of downloads could shed light on the universe of activity.
>
> The industry response, titled "Downloading Hurts Sales," concludes: "If file sharing has no negative impact on the purchasing patterns of the top selling records, how do you account for the fact that, according to SoundScan, the decrease of Top 10 selling albums in each of the last four years is: 2000, 60 million units; 2001, 40 million units; 2002, 34 million units; 2003, 33 million units?"
>
> [. . .] The industry response stresses that the new study has not gone through the process of peer review. But the response cites refuting statistics and analysis, much of it prepared by market research consultants, that also have not gone through peer review.

It should be noted that peer review is a painfully slow process in the social sciences. I have a paper I submitted to a reasonably well-known and together journal in November of 2002, and I still haven't heard back; if these authors were to wait that long for a full peer review, the downloading issue would be obsolete and legislation enacted by then. Meanwhile, I count as a peer of these authors, so here is my review of the paper.

Measuring the right causal story

The primary problem of measuring the causal effect of downloading on sales is that popular tracks see both many downloads and high sales at the same time. Therefore, if you just regress sales of a song on downloads of the same song, you'll get a positive correlation every time. You'd conclude that that downloads are free adverstising for an album, since more downloads mean more sales. But

this mucks up the causal story: the regression is driven by the fact that a track being popular leads to both downloads and sales, but our question was whether downloads lead to sales (or sales losses). We need more complexity to get at the causal story we're interested in. The authors are doing the RIAA a favor by providing a "incomprehensible" paper, since the naive method would have conclusively (and falsely) praised P2P.

The standard econometric response in this causally ambiguous situation is to use instrumental variables (IVs). An IV is a variable which is strongly correlated to one feature (downloads) but is not correlated to another (sales). Once you find a good set of IVs, you regress sales not on downloads but on the IVs.

So the story is:
A: the IV causes more downloading, and
B: more downloading causes fewer CD sales, therefore
C: the IV causes fewer CD sales.
We care about B, but can't directly measure it. So we find something for which A is true. If A is true and C is false, then we can reject B. OK—this isn't quite true: we can only measure correlations between these things, not causations. But with the right variables, correlation will plausibly imply causation.

This turns what was one issue into many substeps:
(1) work out whether the IVs cause downloads
(2) work out whether the IVs don't cause sales
(3) work out whether the IVs are correlated to downloads
(4) work out whether the IVs are correlated to sales

Causation can't be measured, so you're down to sort of eyeballing (1) and (2). The authors do a great job of coming up with creative IVs that clearly fit the first two criteria: they use measures of Internet traffic, which affects how frustrating it is to download a song, and a German school holiday, which led to a spike in the availability of songs. Intuitively, these variables will have a causal effect on the number of downloads, but there is no (non-convoluted) story as to why these would directly cause a rise or fall in sales at U.S. CD stores.

As for step (3), the paper's table 12, items IV and V, show that there is indeed a high correlation between these measures of the IVs (Net traffic and German file availablity) and downloads, so they are a decent measure. They overall correlation coefficient, $R^2 = 1.4\%$, is lousy, but not surprising for this sort of regression. I'll get back to this below. [Technical detail: one of the dozen coefficients goes the wrong way, saying that variability in download times increases downloads; I didn't see any explanation for this, but since this measure of internet frustration is correlated to other frustration measures in the same regression, this is not particularly crazy. Future studies should take this into account when designing their own instruments. As for this study, I wouldn't convict it over such a detail.]

Now for step (4): are these instruments correlated to album sales? O & S find that they are not correlated in any statistically significant way, which is the result reported by the media. So this effort to tell the causal story falls through, and the authors find no causal link between downloading and sales.

However, the IV method loses power, which needs to be borne in mind.

What this means is that the regression has more variability than it had before: you're looking at the variability in downloads plus variability in Net traffic—or maybe Net variability minus variability in downloads, if the two somehow cancel out. We don't know, and that means that the regression of sales vs the IVs doesn't tell us as much about sales vs downloads as we'd like.

Most academic papers aim to find some correlation, and if they do, then the fact that the IV method is stacked against finding a correlation means that we have still more confidence in the results. However, the authors here are reporting the null result that there is no correlation, and the IV method stacks the deck in favor of such a result. The low R^2 above means the deck-stacking is potentially relevant. It is possible that better IVs would find more of a relation between downloads and sales. Or maybe not.

In conclusion

There is no such thing as statistical proof that downloads do or do not cause sales losses, only persuasion one way or the other.

The question of how persuasive this paper is can be rephrased to: what is the likelihood that somebody will come along with a new method and reverse the results given here?

This paper is persuasive to me. The IV method they use is the best that Econometrics has to offer. It is standard, well-accepted fare among econometricians, despite its obscurity to Senior VPs of Communications. The quality of an instrumental variable paper depends mainly on the quality of the IVs chosen, and how well they meet the criteria above, and I personally can't think of any better measures than those used here. The R^2 on the IV-to-download regression is not great; there is some concern that the IV method loses power, favoring a null result, but this is a problem that all null-result papers using standard methods face, and I'm dubious that a better IV will come along and reverse the null result.

Competing papers

Amy Weiss has also said: "Countless well-respected groups and analysts, including Edison Research, Forrester, the University of Texas, among others, have all determined that illegal file sharing has adversely impacted the sales of CDs." So I tried to track down these papers.

Stan Liebowitz wrote the University of Texas paper, which uses a time-series method. It is a careful rendering of the RIAA argument above: draw a graph of music sales and downloads, and sales go down while downloads go up, therefore downloads cause losses in sales. This is wholly unpersuasive, and only demonstrates one of the primary lessons I've learned in economics: never trust a time series. Why not draw a graph of cell phone sales, unemployment rates, or how far Britney is past sixteen? With the appropriate specification, we could prove conclusively that all of these are the coffin nail in music sales.

Stan is much more reasonable than this, and gives me the impression that he's sincerely trying to find the truth. He's since changed his mind somewhat. But personally, I think this is indicative of the whole problem with time series analysis: it's too easy to redo it more carefully and find a new result. Sorry, but time series analysis has been so abused through the ages that I can't look at such a paper anymore without getting the creeps.

I tried registering at Forrester research, but had technical problems, and if it's the study I think it is, it costs $195 to view anyway.

I believe this is the study by Edison Research (but it could be another study on this page). It found the results the RIAA reports in its e-z time series: more people downloading, fewer people buying CDs. It also found a large number of people who have burned or downloaded a CD instead of buying it—but it also found "Many downloaders have gone on to buy a CD from an artist after downloading a track for free from the Internet." and "9. Some downloaders, especially in the older [demographics], report purchasing more music since they began downloading music files." Edison seems bent on reporting only those findings that it can honestly back up, meaning that there is no statement in the report that says that it has "determined that illegal file sharing has adversely impacted the sales of CDs" as Ms. Weiss claims that this study did above.

In the "among others" category, The RIAA often cites a Field Research Corporation survey of 2,500 college students finds a negative correlation between Napster usage and CD sales. Google couldn't find me the survey itself, and I have no information about whether the RIAA directly commissioned the study or not. Lacking the study itself, I can not say anything about whether the study is valid. My primary questions would be about how the sample was gathered and whether it is truly representative. [Amy Weiss, any complaints about small-sample studies?]

So I couldn't find two out of four of the studies I've seen cited by the RIAA; one of the two I could find refuses to make any causal statements, and therefore does not back up or refute either side of the debate above; while the other uses a method which I believe leads to easily mutable results—and the author of the study has confirmed this impression with his actions.

Meanwhile, the results by Oberholzer and Strumpf above are not perfect, but are carefully done, open to scrutiny, and are probably the best effort to find a causal link that Econometrics will be able to do. My experience with past papers is that it's easy to throw another variable into a time series like those supporting the RIAA position and throw the results entirely askew, while it is difficult to do the same with an IV study such as O & S's study.

My editorial

One thing I'd like to see, based on my intuition as a music buyer, is a distinction between major and minor acts—Britney vs the local bar band. As mentioned by some speakers at the recent South by Southwest conference, indie music get good publicity from free downloads, even if acts which appear on the Superbowl have no particular need for grassroots publicity. Thus, I would expect that downloads

help sales of small acts but hurt sales of big acts. Notice that the figures cited by the RIAA above are for the top 10 best-sellers, not for CD sales as a whole.

Oberholzer and Strumpf don't say much about this. They find that twice as many sales lead to more downloads, but less than twice as many. This could be due to selection bias in the sample (downloaders are too hip for U2), or the fact that you can just go next door and borrow the physical CD from a friend. Edison Research's result #9, that older demographics buy more when they download more, seems to point to a superstar v bar band distinction, since the superstar acts are typically aimed at kids.

O & S's null result of no effect overall may be due to substitution effects, where more downloads shift sales from the big acts to the smaller acts. If this were true (and the paper sheds no light on my claim either way) this means that the RIAA's argument is not that downloads hurt sales, but that downloads favor small acts over the acts the RIAA primarily represents. Legislation outlawing P2P networks wouldn't save the music industry as a whole, but it would save the big acts from losing sales to the small acts.