

Movie Leaks, Box Office Success and Child's Play: An On-Line Game Is No Way to Measure Piracy

George S. Ford, PhD*

October 28, 2014

Introduction

Pirated copies of virtually every movie are available on torrent sites sometimes within days, if not hours, of their theatrical release. In a few instances, pirated copies of movies become available on-line prior to the official release date.¹ In 2007, a very good copy of the sequel film *Hostel Part II* was leaked a few weeks ahead of its release and the 2014 action film *The Expendables 3* appeared on torrent sites three weeks ahead of its premiere.² Both movies were box office flops, and piracy was assigned much of the blame.³ In 2009, an unfinished version of *X-Men Origins: Wolverine* was leaked thirty days before its theatrical release.⁴ While the movie did \$180 million in theaters,⁵ how much income was lost to piracy is unknown.

Certainly, the first-order expectation of piracy is reduced box office revenues,⁶ but pre-release piracy is a unique aspect that could exacerbate this revenue effect, especially since the first few weekends of cinematic release represent a huge share of a movie's box office returns.⁷ In a 2013 (and soon to be published paper) entitled *An Empirical Analysis of the Impact of Pre-Release Movie Piracy on Box-Office Revenue*, a group of professors from Carnegie Mellon University looked at box office revenues in an attempt to quantify the unique effects of pre-release piracy. Based on econometric analysis of those data, they concluded that pre-release piracy caused a 19.1% decrease in revenue. This effect is a

sizable reduction in revenues by any standard, but it is especially troubling when recognizing that this 19.1% reduction is only the add-on from pre-release piracy over-and-above the effect of post-release piracy.⁸ This study therefore implies that the box office effects of pre-release piracy are quite large.

I find the results and conclusions of the Strumpf Study to be without credibility. The on-line game data upon which Professor Strumpf centrally relies simply is not up to the task requested of it, and the statistical tests cannot quantify the effects of either pre- or post-release piracy independent of other effects. Also, my own analysis leads to some different conclusions than those reported in the study, which is worrisome.

If one wishes to quantify the effect of pre-release piracy on box office revenues, then it seems most practical to look at actual box office revenues. Academics, however, are not always so practical, which is part of their charm and value. In another recent study on the topic of

pre-release piracy, *Using Markets to Measure the Impact of File Sharing on Movie Revenues* (hereinafter the “Strumpf Study”), Professor Koleman Strumpf proposes to measure the effects of piracy using data from an on-line game where players guess at movie revenues prior to the cinematic release. It is a game without any financial consequences, so on its face it seems like a silly idea to use it to measure the effects of piracy. Is it possible that the responses of the players capture the effects of piracy? Professor Strumpf seems to think so, concluding “that file sharing has only a modest impact on box office revenues.”⁹

In this PERSPECTIVE, I provide an overview of the *Strumpf Study* and the data used therein to analyze pre-release piracy. I find the results and conclusions of the *Strumpf Study* to be without credibility. The on-line game data upon which Professor Strumpf centrally relies simply is not up to the task requested of it, and the statistical tests cannot quantify the effects of either pre- or post-release piracy independent of other effects. Also, my own analysis leads to some different conclusions than those reported in the study, which is worrisome.

The Strumpf Study

Using Markets to Measure the Impact of File Sharing on Movie Revenues—the *Strumpf Study*—is authored by Professor Koleman Strumpf of the University of Kansas. The study is labeled “Preliminary,” so it is a work in progress. Professor Strumpf has an active research agenda on the effects of piracy. Most notably, he was the co-author of one of the earliest publications on the issue in 2007—*The Effect of File Sharing on Record Sales: An Empirical Analysis*—which reached the highly-controversial and implausible conclusion that file sharing had no effect on record sales.¹⁰ Indeed, Professor Strumpf is no stranger to the counter-intuitive or controversy.¹¹

The goal of the *Strumpf Study* is to quantify the effects of piracy and not just pre-release piracy.

To do so, he uses pre-release piracy as an experiment. Rather than look at actual movie revenues, however, Professor Strumpf tests whether pre-release piracy can be detected in the *predictions* of movie revenues. Such predictions are available from the on-line game Hollywood Stock Exchange (“HSX”), which is described as “an entertainment stock market where you can trade movies, stars, and more, just as you would stocks and bonds.”¹² Much like an event study on real corporate stock prices, Professor Strumpf posits that the effect of a significant event (say, pre-release piracy) may be detectable in these “stock” price movements. Specifically, Strumpf argues that if the players view pre-release piracy as important and detrimental to a movie’s revenue, then the stock price should fall after news of a leak. Strumpf’s proposal is clever enough for an academic exercise, but the credibility of the evidence depends critically on the beliefs about the quality and nature of the prediction data and the tests used to quantify relevant effects. I believe the study falters on all of these grounds.

Strumpf’s proposal is clever enough for an academic exercise, but the credibility of the evidence depends critically on the beliefs about the quality and nature of the prediction data and the tests used to quantify relevant effects. I believe the study falters on all of these grounds.

Empirical Setup

Strumpf’s approach to using the HSX data to capture the effect of pre-release piracy can be summarized as follows. Recall that the HSX stock prices are predictions of a movie’s revenues (through the first four weekends). Say that a movie earns revenue R_0 if it is not pirated and R_1 if it is pirated prior to its release. The

difference, $\Delta = R_1 - R_0$, measures the effect of the pre-release piracy (and in Strumpf's view, piracy generally). If the expectation is that pre-release piracy is exceedingly unlikely (and, thus far, it is very rare), then it may be sensible just to look at the price changes around the piracy release date (the "event") to see how the game's players incorporate such information into their predictions.

But, say that people have some beliefs about the likelihood of a movie leaking, and assign a probability p_t to it occurring at time t . If this probability was estimated to be very high (p_t is close 1.0) on (say) the day before the pirated copy is released, then the game's players would have already incorporated that information into the HSX prediction the day prior to the leak. As such, there would be no price effect observed on the day the pirated copy is made available. Alternately, if the assessed probability of a leak is low (p_t is near zero) prior to the actual leak, then the HSX price change will be close to Δ on the day the pirated copy is released because the leak is a surprise event.

Given this setup, it is not possible to get an accurate estimate of Δ without also having an accurate estimate of p_t . Like Δ , the probability p_t is not observed, so it must be estimated, adding to the complexity of the model and to the demands on the data. Strumpf estimates p_t based on the observed pirated-copy release dates relative to the theatrical release dates. He employs a statistical procedure to get a movie-specific values of p_t , though this specificity is provided only in relation to the releasing studio and the amount of non-movie file sharing on a particular date (which affects individual movies based only by its release date), both of which he takes to be exogenous to movie revenues.¹³ (However, the earlier study from Carnegie Mellon shows that movie revenues *are* related to the film's studio, weakening the case for exogeneity.¹⁴)

There are (at least) two important conceptual problems with Professor Strumpf's empirical approach and both are based on the fact that the HSX is a predictive market. First, as a prediction of future box office revenues, the observed pre-release HSX prices presumably account for all available information. Piracy is not a surprise event; it is effectively a guarantee after release (it is just a question of when). Knowing this, players bake the effects of post-release piracy into their predictions at the initiation of trading, since they are attempting to predict actual box office revenues. Thus, R_0 in Professor Strumpf's conceptual framework is not the "no piracy" price but the price including the influence of the near inevitable pirating of the movie after release. More formally, let R_X be the revenue accounting for post-release piracy (and R_0 the "no piracy" revenue). In practice, it is more legitimate to say that Strumpf's model is testing for $R_1 - R_X$ rather than $R_1 - R_0$. At best, then, the analysis in the *Strumpf Study* can only quantify the additional effect of a leak over and above the general effect of post-release piracy (like the study out of Carnegie Mellon properly acknowledges).¹⁵

Oddly enough, Strumpf's evidence on price responses to new information discredits his statistical tests of piracy. Plainly, whatever price response is observed after a leak cannot be said to measure the effects of piracy alone. Conflating effects is a serious indictment against an empirical test.

Second, in addition to the effect of pre-release piracy (not piracy generally), price responses to a leak also represent an improvement in information. A leak is very much like the

release of a trailer in this regard. Say, for example, players expect a movie to be a stinker so the HSX prices are low. However, a viewing of the pre-release leak suggests that the movie is a good one, so the HSX prices rise in response. Critically, this positive response is not a “promotional” effect nor does it suggest that pre-release piracy is good for box office revenues. Rather, the price increase reflects an improvement in information to the HSX players who are making *predictions* about box office revenues. It is an information-driven correction of a prediction, not an increase in box office revenues. In the same way, if the expectation of a good movie is dashed by the leak, then there will be a HSX price reduction for the movie.

Put bluntly, HSX is just a game and the predictions it produces are of questionable and unproven quality. Moreover, my own analysis of the data conflicts with some results reported in the Strumpf Study.

What’s important about these two effects for Professor’s Strumpf’s empirical tests is that they are inseparable—his tests cannot distinguish between the two. If HSX prices fall upon the leak, then it is not possible to attribute that reduction solely to the leak, since the new price also embeds increased information about the movie that is unrelated to piracy. Or, as in the example above, if HSX price rises after the leak, then it is not legitimate to say that piracy increases box office revenues, since the rising price may simply reflect more information about the movie’s quality that leads to a correction in a *prediction* of box office revenues.

In fact, Professor Strumpf’s own empirical analysis provides evidence of the importance of increased information by showing that HSX prices respond to the release of trailers. A pre-

release leak is akin to the release of a very thorough trailer, and Strumpf’s analysis indicates that prices may rise or fall upon this new information depending on the signal it provides. Oddly enough, Strumpf’s evidence on price responses to new information discredits his statistical tests of piracy. Plainly, whatever price response is observed after a leak cannot be said to measure the effects of piracy alone. Conflating effects is a serious indictment against an empirical test.

The HSX Data

To be clear, *HSX is a game, not a real financial market*. Participants accumulate or lose points by betting for or against stars and movies; there is no real financial consequence of guessing well or poorly. Thus, the HSX data is perhaps more akin to an opinion survey than financial results, but a survey where the participants are self-selecting based on their interest in the game.

On the HSX exchange, a movie’s “stock” is traded daily, often years before its wide theatrical release.¹⁶ The “stock price” is a prediction of a movie’s gross revenues through (in most cases) the fourth weekend of its release. However, if a film is released in fewer than 650 theaters and remains in limited release, then the prediction is for the gross revenues of the film through its 12th weekend. While most of the top 150 films (which is the focus of the *Strumpf Study*) are released in more than 650 theaters (or eventually hit that number), it is worth noting that the number of theaters in which the film appears is not known for certain prior to its release, so it is not always clear what is being predicted.

Also, the number of weeks of trading following release is treated somewhat loosely in the HSX data. For many movies, the four-week period is not measured from the official release date of the movie but rather from a later date when the movie is more widely released. It is not uncommon for a movie to open to limited release but eventually, given its popularity or

the release plan, be released more widely at a later time. Some movies, therefore, are in theaters for weeks before the (four- or twelve-week) final trading period begins. Such occurrences implies the “pre-release” trading on some movies reflects significantly more information than for other movies. In light of these particulars, it is clear that the game’s players do not always have a clear idea as to what they are predicting (that is, is it four or twelve weeks of revenue being predicted) and the quality of information available across movies is often substantial.¹⁷

*HSX is a type of predictive market, the use of which is a burgeoning field in economics. *** For predictive markets to be useful, however, they must be a reliable proxy for real markets, responding to information like real markets do. There is absolutely no evidence that the HSX is a reliable predictor of the effect of piracy or, for that matter, any other “event” in a movie’s pre-release experience.*

HSX is a type of predictive market, the use of which is a burgeoning field in economics. Many economists are critical of the use of predictive markets, but others researchers are hopeful for their use.¹⁸ The power of predictive markets remains an open question. For predictive markets to be useful, they must be a reliable proxy for real markets, responding to information like real markets do. There is absolutely no evidence that the HSX is a reliable predictor of the effect of piracy or, for that matter, any other “event” in a movie’s pre-release experience.¹⁹ While the HSX predictions may be correlated with actual box office figures, correlation does not imply the estimates are

unbiased or precise, and, more importantly for Strumpf’s analysis, accurate predictors of specific information-improving events.

While it has been shown that the HSX prices do adjust to various events likes trailer releases and actor changes (indeed, Strumpf provides some examples), we have no idea whether or not these responses are unbiased estimates of the true effect. Certainly, there is no evidence that the game’s players have any special skill at predicting the revenue impact of piracy (see Table 1 below). It is not sufficient merely to assume the HSX is a good proxy for real markets; it must be demonstrated, and it has not been so. (In fact, one might argue that if the HSX predicted a 19% revenue effect, then its reliability would be demonstrated.) For now, there are simply too many unanswered questions about the accuracy of the HSX predictions to do so; it’s a leap of faith.²⁰ Being dependent on the HSX data, the *Strumpf Study* thus lacks credibility, being of academic interest alone with no direct policy relevance.

Another significant concern with the *Strumpf Study* is the sample size of pre-release pirated movies. Nowhere in the paper does Strumpf provide a specific count of leaked movies, which is an important omission. Pre-release piracy is very rare event. These few observations are likely to make up a trivial share of his total sample of 1,057 movies (less than 5%), which could prove problematic for econometric estimation.²¹

Also, the extreme diversity of the leaked films makes estimating average effects from such a small sample even more challenging. The sample of leaked films probably includes, for example, *Iron Man* (leaked a day or so before release and eventually grossing \$318 million), *88 Minutes* (leaked 15 months prior to release and grossing \$17 million), to *Skinwalkers* (leaked 10 days before release and grossing \$1 million). It appears from the descriptive statistics reported in the *Strumpf Study* that at least one movie was

pirated about two-years prior to its theatrical release (637 days), an odd circumstance suggesting that some of the already scarce observations may not be terribly useful for measuring the effects of piracy. Many of the pre-release pirated copies are made widely available only a few days prior to the theatrical release, another special situation requiring attention. The useful sample of pre-release observations is likely to be very, very small, and asking these few observations to support the econometric model employed by Strumpf is risky.

While Strumpf concludes there “was little change in the stock price when the movie [The Wolverine] became available for illicit download,” my analysis of the returns data does not support this conclusion. The price changes on the day of and after the April 1st leak were the two largest negative returns over the sample’s four-month period ...

The Results (or Lack Thereof)

Turning to the primary results on predicted revenue effects, there isn’t much to talk about. Strumpf’s estimation model finds very little, if any, effect from pre-release piracy in the HSX data, concluding: “The estimates indicate that the displacement effect is quite small, both on a movie level and in aggregate. The effect is precisely estimated. This is perhaps not surprising given the low quality of early file sharing releases and the lack of amenities such as theater sound and video systems.”²² For a number of reasons I’m not terribly surprised by the findings, including not only the reasons Strumpf provides but also his use of predictive data of unknown quality and the application of

a complex estimation approach to a paucity of data on pre-release events.

As discussed above, what Strumpf labels as a “displacement effect” also includes an information effect, though Strumpf does not recognize this fact. So, even though the average effect is estimated to be “quite small,” the effect of piracy may be quite large though it is masked, on average, by the information effect. While I cannot conclude that this is empirically true, it is certainly a conceptual concern with Strumpf’s analysis.

Another Look at the HSX Data

Another concern is that my analysis of the HSX data presents a different picture than that offered by the *Strumpf Study*. Consider Strumpf’s “motivating example” – the leak of *X-Men Origins: Wolverine*. An unfinished version of the movie was posted and available on a torrent site April 1, 2009, a full month prior to its theatrical release. In analyzing the HSX data around the leak, Strumpf concludes that there “was little change in the stock price when the movie became available for illicit download.”²³ But I find the evidence suggests otherwise.

To assess the effect of the leak, I test for large price changes around the leak, setting a two-day window on each side of the reported April 1, 2009, leak (a five-day event window).²⁴ As a counterfactual, I use the daily returns computed from a sample of HSX price data for the film over the 120 days prior to the event window (a total of 125 days of trading). While Strumpf concludes there was “was little change in the stock price when the movie became available for illicit download,” my analysis of the returns data does not support this conclusion. The price changes on the day of and after the April 1st leak were the *two largest negative returns* over the sample’s four-month period, reducing the price a total of 6.5% over the two days. The tenth largest negative return was observed on the day before the leak. Over a three-day event window the negative return was 8%, and over the five-

day window it was 5%.²⁵ In light of these relatively large price reductions around the leak, the evidence seems to suggest a fairly robust response to the leak, a result contrary to the claims of the *Strumpf Paper*.²⁶

The same analysis can be applied to another leaked movie. Consider the recent, high-profile leak of *The Expendables 3*, which was leaked three weeks (on July 24, 2014) prior to its official release (on August 15, 2014). The price changes for the two days following the leak were the two largest negative price changes over the entire sample (125 trading days). The movie's HSX price also fell by very large amounts two days prior to the leak (the third and fourth largest negative price changes), suggesting that the information available on the leak perhaps predated the reported leak date. Over the five-day event window, the HSX stock price fell 14%.²⁷ Again, the data seem to suggest a response to the leak. Also, given the large negative returns a few days prior to the claimed leak date, this example indicates that the exact leak date may be difficult to pinpoint (and the statistics should account for that possibility).

My analysis of the HSX data for 1,400-plus movies over the period 2003 through 2013 indicates an average (absolute) prediction error of about 36%. At best, the HSX data is a very noisy signal of box office revenues.

Two other cases reveal the potential problems with using the HSX data to quantify the effects of piracy. Consider first the pre-release leak of Michael Moore's *Sicko*. News reports about the movie's leak appeared on June 14 and 15, 2007, a few weeks prior to its July 3, 2007, release. On June 15, the HSX stock price for the movie fell by 6.9% followed by a 3.6% decline on the

following day. These were large negative changes (one being the eighth largest in the sample), but not the largest changes as in the two prior examples. A closer look at the data revealed that *Sicko's* stock price was highly variable, having an absolute average daily price change of 5.2%. Given such high variability, it would be difficult to statistically detect even a large price response to a pre-release leak (with 5% being typical, the leak response must be very large to be "abnormal").²⁸ A price change of over 15% would be required for a statistical test to signal a "problem." Also, over the relatively short 120 day window prior to its leak, the predicted box office revenue for *Sicko* ranged from \$1.16 million to \$30.02 million. This huge range of predictions over a four-month window makes me question the relevance of the HSX data for piracy research.²⁹

The second case involved Clint Eastwood's 2009 sleeper film *Gran Torino*, which was leaked nearly a month prior to its box office release. At the time of the movie's leak, the HSX prediction of gross revenue was about \$28 million, having risen by nearly 20% over the preceding week. Even after the estimated leak date, the predictions continued to rise for the next few days, so unlike the prior examples there was not a sizable reduction in the stock prices after the leak. In fact, there was no obvious effect on HSX prices around the leak date (unlike the other examples mentioned), but there is another serious issue with the HSX data for this movie. Specifically, at initial release, HSX players predicted the movie would gross \$26 million, but at the end of its fourth week the movie had grossed \$110 million.³⁰ The prediction error was massive (123%, using the arc formula).³¹

In fact, large prediction errors are quite common in the HSX data. Table 1 summarizes the prediction errors for a sample of leaked 2009 movies.³² Most of the prediction errors are very large (averaging 41%), casting doubt on the accuracy of the HSX predictions and the usefulness of the data for research purposes.

**Table 1. HSX Prediction Errors
(2009 Leaked Movies)**

Title	Prediction	Actual	Error
Taken	\$44.38	\$95.03	72.7%
The Reader	15.94	23.01	36.3%
Milk	27.35	28.12	2.8%
The Wrestler	23.44	21.58	8.3%
Revol. Road	28.97	20.52	34.1%
Frost/Nixon	19.93	16.56	18.5%
Defiance	16.18	25.20	43.6%
Gran Torino	26.16	110.18	123%
Wolverine	224.36	165.16	30.4%

My analysis of the HSX data for 1,400-plus movies over the period 2003 through 2013 indicates an average (absolute) prediction error of about 36%.³³ At best, the HSX data is a very noisy signal of box office revenues. While noisy data may be useful in some cases, in many cases it is not. For example, data with a low signal-to-noise ratio is unlikely to be useful in small samples, and the sample of leaked movies is quite small. Certainly, more work on the data must be done, including tests for bias, assessments of Type II error rates, and other empirical considerations before it can be relied upon to quantify the effects of piracy (or anything else).

To be clear, the few examples discussed above are a select sample of all the leaked movies and I am not arguing the change in prices prove anything in particular. Anecdotally, however, there appears to be a response in prices around leak dates for some films, but not others. The prediction error calculations indicate that the HSX predictions can be very poor, and with the small sample of leaked movies, these large errors could be a problem for statistical analysis.³⁴

There is also the question of whether or not the game players have adequate information to be relied upon to make adjustments to their predictions in instances of leaked movies. Many of the leaked movies are not blockbusters and

information on the leaks is difficult if not impossible to find on the Internet, suggesting perhaps the information was not widely disseminated.

And, while attention to piracy has heightened in recent years, it's not clear that the gamers were really cognizant of the effects of piracy in years past or had any evidence by which to assess the potential box-office impact of such piracy. Over much of Strumpf's sample period (2003-2009) there was little, if any, published or otherwise available evidence on the revenue effects of piracy (i.e., most of the evidence is recent). Without such evidence, it is not clear that the gamers would know how to respond to piracy or leaks, and thus there is no reason to expect the responses to be accurate reflections of the true effects of piracy. (Even with such evidence, the gamers' response would likely be equal to the published predicted effects and thus add no new information on the effects of piracy.) Though most researchers now acknowledge that piracy is harmful to sales, there remains an active research effort to discover the finer details of the mechanisms through which piracy influences revenues (such as the timing of the piracy). Certainly, we should not expect that the HSX traders have already solved this problem.

While the Strumpf Study's use of the on-line game data is clever as an academic project, its credibility for public policy is limited by the quality and nature of the data and the inability of the statistical test to cleanly assess the effect of piracy.

Conclusion

While the piracy of copyrighted material is rampant, quantifying its economic impact has proven challenging. Researchers evaluating

actual box office revenues report that pre-release piracy is associated with a 19% reduction in box office revenues. In contrast, the *Strumpf Study*, using predictions from an on-line game where players guess about movie revenues, finds only a very small effect (if any). While the *Strumpf Study's* use of the on-line game data is clever as an academic project, its credibility for public policy is limited by the quality and nature of the data and the inability of the statistical test to cleanly assess the effect of piracy. Put bluntly, HSX is just a game and the predictions it produces are of questionable and unproven quality. Moreover, my own analysis of the data conflicts with some results reported in the *Strumpf Study*. At this time, I can't help but assign low credibility to the *Strumpf Study*; I don't think it has any policy relevance.

NOTES:

* **Dr. George Ford is Chief Economist of the Phoenix Center for Advanced Legal and Economic Public Policy Studies. The views expressed in this PERSPECTIVE do not represent the views of the Phoenix Center or its staff.**

- 1 See, e.g., *6 Major Films that Leaked Online*, TOTAL FILM (April 3, 2009) (available at: <http://www.totalfilm.com/features/6-major-films-that-leaked-online>).
- 2 *6 Major Films, id*; T. Spangler, "Expendables 3" Leaks Online, Pirated Copy Downloaded 189,000 Times in 24 Hours, VARIETY (July 25, 2014) (available at: <http://variety.com/2014/digital/news/expendables-3-leaks-online-pirated-copy-downloaded-189000-times-in-24-hours-1201268947>).
- 3 *Id.*; B. Lang, "Expendables 3" Flop Shows Danger of Pre-Release Piracy, Expert Says, VARIETY (August 18, 2014) (available at: <http://variety.com/2014/film/news/expendables-3-piracy-box-office-1201285413>).
- 4 *Wolverine Leak Means Corporate Humiliation and Potential Bad Buzz*, VARIETY (April 1, 2009) (available at: <http://variety.com/2009/digital/news/wolverine-leak-means-corporate-humiliation-and-potential-bad-buzz-33434>).
- 5 <http://www.the-numbers.com/movie/X-Men-Origins-Wolverine#tab=summary>.
- 6 *Cf.*, G.S. Ford, *Piracy and Movie Revenues: A Critical Review of "A Tale of the Long Tail"*, PHOENIX CENTER POLICY PERSPECTIVE NO. 13-05 (September 16, 2013) (available at: <http://www.phoenix-center.org/perspectives/Perspective13-05Final.pdf>).
- 7 For example, the movie *Ironman 3* grossed \$409 million with 85% of that coming in the movie's first three weeks in theaters (<http://boxofficemojo.com/movies/?page=weekly&id=ironman3.htm>).
- 8 L. Ma, A. Montgomery, P. Singh, and M. Smith, *An Empirical Analysis of the Impact of Pre-Release Movie Piracy on Box-Office Revenue* (July 8, 2014) (hereinafter "*Carnegie Mellon Study*") (available at: <http://ssrn.com/abstract=1782924>) and forthcoming in INFORMATION SYSTEMS RESEARCH.
- 9 K. Strumpf, *Using Markets to Measure the Impact of File Sharing on Movie Revenues*, Working Paper (July 11, 2014) (available at: http://www.unc.edu/~cigar/papers/P2PandHSX_HSX+NBER.pdf).
- 10 F. Oberholzer-Gee and K. Strumpf, *The Effect of File Sharing on Record Sales: An Empirical Analysis*, 115 JOURNAL OF POLITICAL ECONOMY 1-42 (2007). Compelling critiques of this paper include: S.J. Liebowitz, *How Reliable is the Oberholzer-Gee and Strumpf Paper on File-Sharing?*, Working Paper (September 1, 2007) (available at: <http://ssrn.com/abstract=1014399> or <http://dx.doi.org/10.2139/ssrn.1014399>) and S.J. Liebowitz, *The Oberholzer-Gee/Strumpf File-Sharing Instrument Fails the Laugh Test*, Working Paper (May 1, 2010) (available at: <http://ssrn.com/abstract=1598037> or <http://dx.doi.org/10.2139/ssrn.1598037>).
- 11 See, e.g., N. Haring, *No Comment, Please*, HANDELSBLATT (July 23, 2008) (available at: <http://www.handelsblatt.com/politik/oekonomie/nachrichten/steven-levitt-blocks-an-undesired-statement-no-comment-please/2976444.html>) (discussing how an independent referee for the JOURNAL OF POLITICAL ECONOMY sided with Liebowitz, concluding that "[Oberholzer-Gee and Strumpf's] conclusions are not warranted given the analysis and evidence that they provide"); D. Glenn, *Dispute Over File Sharing's Harm to Music Sales Plays Again*, CHRONICLE OF HIGHER EDUCATION (June 17, 2010) (available at: <http://chronicle.com/blogs/wiredcampus/dispute-over-file-sharings-harm-to-music-sales-plays-again/24881>). Significantly, despite requests, Oberholzer-Gee and Strumpf have refused to make their data available for replication purposes.
- 12 <http://www.hsx.com/help/?page=basics#1>.
- 13 If a factor affected both the probability of a leak and box-office revenues, then using that factor to predict the probability, which is then used to explain revenues, would introduce bias.
- 14 *Carnegie Mellon Study, supra* n. 8 at Tables 6 and 7.
- 15 The reasons why the estimated effect is an add-on to the average effect of piracy are a bit different. In the *Carnegie Mellon Study*, the control group is infected with the effects of post-release piracy. In the *Strumpf Study*, there's a question about whether the predictions of movie revenues already incorporate the effects of piracy at the time the movie is leaked, in which case both the treated and control groups are infected.

NOTES CONTINUED:

- ¹⁶ It is not clear how useful this early data is for empirical work, and Strumpf limits his analysis to no more than a year prior to release.
- ¹⁷ The data does not include indicators of key dates, such as release dates and the adjustment date (when the price is reset to a scale of actual revenues following the first weekend of wide release).
- ¹⁸ See, e.g., *Prediction Markets: Don't Bet On It*, ECONOMIST MAGAZINE (December 1, 2012) (available at: <http://www.economist.com/news/finance-and-economics/21567382-intrade-retreats-american-regulators-dont-bet-it>); C. Manski, *Interpreting the Predictions of Prediction Markets*, NBER WORKING PAPER 10359 (March 2004) (available at: <http://www.nber.org/papers/w10359>); E. Rosenbloom and W. Notz, *Statistical Tests of Real-Money versus Play-Money Prediction Markets*, 16 ELECTRONIC MARKETS 63-69 (2006); E. Servan-Schreiber, J. Wolfers, D. Pennock, and B. Galebach, *Prediction Markets: Does Money Matter?* 14 ELECTRONIC MARKETS 243-251 (2004).
- ¹⁹ Strumpf cites two papers for support, but neither study evaluates the accuracy or bias of the responses to events. See D. Pennock, C. Lawrence, and F. Nielsen, *The Real Power of Artificial Markets*, 291 SCIENCE 987-988 (2001); A. Elberse, *The Power of Stars: Do Star Actors Drive the Success of Movies?* 71 JOURNAL OF MARKETING 102-120 (2007).
- ²⁰ *Id.*
- ²¹ *Carnegie Mellon Study*, *supra* n. 8, finding a little over fifty pre-release movies.
- ²² *Strumpf Study*, *supra* n. 9 at p. 16.
- ²³ *Id.* at p. 13.
- ²⁴ A similar approach is used in A. Elberse, *The Power of Stars: Do Star Actors Drive the Success of Movies?*, 71 JOURNAL OF MARKETING 102-20 (2007).
- ²⁵ Three of the five event dummies are statistically significant at the 10% level or better, and the joint tests of the coefficients and the cumulative effects are also statistically significant at better than the 5% level. These tests are based on a simple regression of a constant and event dummies on the stock returns. Note that the price change two days after the leak was the largest positive return over the period, which explains the difference between the 3-day and 5-day return.
- ²⁶ There were no other major events on these dates (see <http://www.traileraddict.com/xmen-origins-wolverine#2>).
- ²⁷ Four of the five event dummies are statistically significant at the 10% level or better, and the joint tests of the coefficients and the cumulative effect are also statistically significant at better than the 5% level. These tests are based on a simple regression of a constant and event dummies on the stock returns.
- ²⁸ The data has a high Type II error rate for even large changes in prices in response to piracy (or any other event).
- ²⁹ In some models, Strumpf uses trades for a year prior to release and in others fifty days prior to release.
- ³⁰ *Gran Torino* had a limited release on December 12, 2008, and a wide release on January 8, 2009. I use the former date as the release date (<http://boxofficemojo.com/movies/?page=weekly&id=grantorino.htm>).
- ³¹ The arc formula uses the average of the prediction and realization in the denominator of the percentage calculation.
- ³² Predictions are equal to the HSX movie stock price the day before release. Actual revenue is the stock price at expiry.
- ³³ The prediction is computed as the average prediction over the five days prior to the movie's release.
- ³⁴ Of course, before making broad claims about the accuracy of the HSX predictions, a more thorough analysis is desirable.