

## NBER WORKING PAPER SERIES

## PIRACY ON THE SILVER SCREEN

Rafael Rob  
Joel Waldfogel

Working Paper 12010  
<http://www.nber.org/papers/w12010>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
January 2006

Waldfogel thanks the Mack Center and the Wharton Electronic Business Initiative for financial support. Sarah Waldfogel provided excellent assistance with data input. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

©2006 by Rafael Rob and Joel Waldfogel. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Piracy on the Silver Screen  
Rafael Rob and Joel Waldfogel  
NBER Working Paper No. 12010  
January 2006  
JEL No. L8, K2

### **ABSTRACT**

New information technology has reduced marginal production and distribution costs of information goods to negligible levels and promises to revolutionize many industries. Unpaid copies of digital products can be as good as paid first-generation copies, and their availability can undermine the ability of sellers to cover first-copy costs. As a result, unpaid distribution has emerged as a major issue facing the music and movie industries in the past few years. Using survey data on movie consumption by about 500 University of Pennsylvania college students, we ask whether unpaid consumption of movies displaces paid consumption. Employing a variety of cross-sectional and longitudinal empirical approaches, we find large and statistically significant evidence of displacement. In what we view as the most appropriate empirical specifications, we find that unpaid first consumption reduces paid consumption by about 1 unit. Unpaid second consumption has a smaller effect, about 0.20 units. These estimates indicate that unpaid consumption, which makes up 5.2 percent of movie viewing in our sample, reduced paid consumption in our sample by 3.5 percent.

Rafael Rob  
University of Pennsylvania  
rrob@ssc.upenn.edu

Joel Waldfogel  
Public Policy and Management  
University of Pennsylvania, Wharton School  
3100 Steinberg Hall-Dietrich Hall  
Philadelphia, PA 19104-6372  
and NBER  
waldfogj@wharton.upenn.edu

New information technology, which has reduced marginal production and distribution costs of information goods to negligible levels, promises to revolutionize many industries. Zero marginal costs can allow higher profits and lower prices. Moreover, costless distribution allows sellers to use innovative marketing strategies such as free samples with partial functionality (Shapiro and Varian, 1999). But once the genie of negligible marginal costs is out of the bottle, the same technological changes that empower sellers can also enable unpaid distribution and redistribution. With digital goods, unpaid copies can be as good as paid first-generation goods, and the willingness and ability of consumers to share copies can undermine the ability of sellers to cover first-copy costs. Unless unpaid distribution stimulates demand for complementary goods which can continue to generate revenue, the information industries may ultimately suffer.

Over the past five years, unpaid distribution (variously known as “file-sharing” and “piracy”) has emerged as a major issue facing producers of information products such as music and movies. After revenue growth for many years, US music industry revenue has declined from 2000 to 2003 and was flat in 2004, and many in the industry blame file-sharing<sup>1</sup>. The motion picture industry is similarly concerned about piracy, although overall Hollywood revenue continues to rise even if theatrical box office receipts have been flat over the past few years.<sup>2</sup>

While it is perhaps surprising at first blush, theory gives no unambiguous guidance on the effect of unpaid consumption on sales. On one side, unpaid copies are

---

<sup>1</sup> See the RIAA’s “AntiPiracy” at <http://www.riaa.com/issues/piracy/default.asp> as well as the RIAA’s “2004 Yearend Statistics” at <http://www.riaa.com/news/newsletter/pdf/2004yearEndStats.pdf> (both accessed December 15, 2005).

<sup>2</sup> The Motion Picture Association of America estimates \$3.5 billion in losses to various forms of piracy in 2002. See the MPAA’s “2003 Piracy Fact Sheet Overview” at <http://www.mpaa.org/PiracyFactSheets/PiracyFactSheetOverview.pdf> (accessed December 15, 2005).

close substitutes for paid copies, particularly in music, suggesting that unpaid consumption would simply displace paid consumption. While MP3 files contain less information than, say, CDs, many listeners find the compressed files to contain nearly the same information. And possession of an unpaid music file allows its user all the activities available to the owner of a purchased CD.

Movies are different in a few respects that suggest less displacement by unpaid use. First, while an unpaid copy of a movie, on CD or DVD, contains the same information as a legal DVD and allows a very similar experience to home video ownership, home video is an imperfect substitute for theatrical viewing. Second, as economic theorists have emphasized, information sharing can stimulate paid consumption if the sharing remains small-scale<sup>3</sup>. The logic is simple: two isolated persons valuing a movie DVD at \$11 each will not purchase it for its \$20 price. If they plan to share, then the sum of their valuations exceeds the price, and they will purchase it together. The large size of video files has kept peer-to-peer sharing at small scale, largely through traditional “sneakernet” file transmission.<sup>4</sup> The possibility for file-sharing to stimulate paid demand therefore seems more real in movies than in music.

To date there have been numerous academic studies of the effect of file-sharing on revenue in the music industry.<sup>5</sup> Most, but not all, studies find that file sharing depresses sales, although much less than one for one. That is, much of the music consumption occurring through file sharing would not have occurred in the absence of

---

<sup>3</sup> See Bakos, Brynjolfsson, and Lichtman (1999) and Varian (2000).

<sup>4</sup> A user dons his sneakers and carries a DVD from his friend’s house to his DVD-burning computer. See <http://en.wikipedia.org/wiki/Sneakernet>, accessed December 15, 2005.

<sup>5</sup> See, for example Blackburn (2004), Oberholzer-Gee and Strumpf (2004), Zentner (2004), Rob and Waldfogel (2004), Bounie, Bourreau and Waelbrock (2005), Hong (2005), Hui and Png (2003), and Leibowitz (2003).

file sharing. As a result, file sharing tends to increase social welfare by more than the revenue loss to sellers (ignoring any effects on the supply of music). This study examines the displacement question in the context of the motion picture industry. Using survey data on movie consumption by about 500 University of Pennsylvania college students, we ask whether unpaid consumption (of burned or downloaded copies) stimulates or depresses demand for various forms of paid consumption, including theatrical, television, rental, and DVD ownership. The surveyed population is far from representative of the US population – respondents are richer and are more likely to have broadband Internet connections as well as access to DVD burning hardware. While results are therefore not representative of the entire US, the effects documented for this population are suggestive of effects that might operate more broadly as technologies continue to diffuse.

Using a variety of cross-sectional and longitudinal empirical approaches, we find large and statistically significant evidence of displacement. In what we view as the most appropriate empirical specifications, we find that unpaid first consumption reduces paid consumption by about 1 unit. Unpaid second consumption has a smaller effect, about 0.20 units. These estimates indicate that unpaid consumption, which makes up 5.2 percent of movie viewing in our sample, reduced paid consumption in our sample by 3.5 percent.

## **I. Theoretical Background**

Prior to file sharing, each consumer has some willingness to pay for each movie. The distribution of these consumer valuations make up the demand curve for the movie.

The advent of file sharing has two kinds of effects. First, even holding constant the distribution of valuations, with file sharing, all kinds of consumers – with valuations above or below the market price – can now get access without paying. The sales displacement and welfare effects of this – worked out in section I.1 below – depend on whether the unpaid consumption would have manifested itself as paid consumption in the absence of file sharing.

### 1. Supply and Demand

Supply and demand provides a useful framework for thinking about the sales displacement and welfare consequences of unpaid consumption. For this discussion suppose that the marginal cost of production and distribution is zero.

First, consider the market prior to file sharing, in figure 1a. Sellers sell  $B_0$  units at a price of  $P$ . The area under the demand curve is divided into three regions, consumer surplus (the consumer valuations in excess of the price they pay), revenue ( $P * B_0$ ), and deadweight loss (foregone mutually beneficial transactions for units with consumer valuation above the marginal cost but below the price, so no transaction occurs). When unpaid consumption is feasible, its effect depends on which units are obtained without payment, whether they are those with valuation above or below  $P$ .

Start by considering an extreme case in which only units with valuation below  $P$  are obtained without payment. This results in the situation in figure 1b. Sellers continue to sell  $B_0$  units at a price of  $P$ . Another  $S$  units are obtained without payment. Hence, in this situation, unpaid consumption does not displace paid consumption, and  $\frac{\partial B}{\partial S} = 0$ .

Here, the deadweight loss prevailing prior to unpaid consumption becomes consumer

surplus, while sellers remain as well off as before. File sharing makes society unambiguously better off.

Next, consider a case where all of the unpaid consumption is drawn from situation where consumer valuations exceed  $P$ . This is depicted in figure 1c. Now  $(B_0 - S)$  units are transacted at  $P$ , while  $S$  units are consumed without payment. The sum of producer and consumer surplus is the same as before, as is deadweight loss. But  $S*P$  is transferred from producers to consumers. Here, paid consumption falls by the exact amount of unpaid consumption, or  $\frac{\partial B}{\partial S} = -1$ .

A general lesson from this analysis is that the lower the rate of displacement ( $\frac{\partial B}{\partial S}$ ), the more welfare gain that file sharing brings to society as a whole. As the rate of displacement grows (in absolute value, toward -1), the social gain from file sharing falls to zero. The effect of unpaid consumption becomes a pure transfer from sellers to buyers. The main goal of our paper is to estimate  $\frac{\partial B}{\partial S}$ .

An important caveat bears mention: it is possible that unpaid consumption displaces sales and changes the nature or volume of movies that get made. To the extent that our estimates have welfare implications, they are implications for welfare, given which movies are being made.

## 2. File Sharing and Changing Demand

In addition to allowing unpaid access regardless of one's valuation, file sharing can also stimulate demand via two channels. First, small-scale information sharing can stimulate demand if it makes a group of consumers willing, as a group, to purchase a

product that no group member would have purchased alone. Two features of music file sharing prevent it from remaining small in scale, in turn making the suggestion that file-sharing would stimulate music sales rather far-fetched. First, file compression makes it fast to download music, particularly over broadband connections. A typical song can be downloaded in a few seconds. Second, with peer-to-peer (P2P), if one person purchases an album, it becomes available to all Internet-connected persons willing to download. As a result, file-sharing is potentially unlimited in scale.

File sharing in movies remains fundamentally different. Even compressed, a typical Hollywood movie is about 4 gigabytes. Using the newest downloading technologies (e.g. BitTorrent), downloading a movie takes between 30 minutes and 8 hours, depending on how many persons have already made the file available online.<sup>6</sup> Similarly, DVD copying requires about 90 minutes using a typical computer setup and DVD burner. DVD copying is generally quicker and results in higher quality copies than downloading. As a result, the dominant form of file sharing in movies is DVD copying, rather than downloading. By its nature, DVD copying is smaller in scale than P2P copying, making it more plausible that file-sharing could stimulate sales. On the other hand, while consumers listen to some albums repeatedly, they may wish to watch movies only once. That is, movies and music may differ in the extent to which initial consumption is a substitute or a complement for subsequent use. Perhaps even more than in music, the effect of file sharing on movie industry revenue is theoretically ambiguous.

---

<sup>6</sup> “Cable and DSL modem users can expect an average of 25 megabytes per hour, sometimes slower if the swarm is small with less than 2 seeders. On a good day with a big swarm, however, you can download a 5MB song within 3 minutes, and a 900MB movie within 60 minutes.”  
[http://netforbeginners.about.com/od/peersharing/a/torrenthandbook\\_4.htm](http://netforbeginners.about.com/od/peersharing/a/torrenthandbook_4.htm) , accessed 2/24/2005



Second, file-sharing may increase consumers' willingness to purchase by allowing him to sample the product prior to purchase. Shapiro and Varian (1999) emphasize that free samples of information products can stimulate subsequent paid consumption. The samples they envision are incomplete versions of the product. Even if unpaid access to movies stimulates interest in movies, it is not clear that consumers would be willing to pay for something that they have already obtained without payment. First, unpaid copies are generally the entire movie rather than an appetite-enhancing free sample. And, second, since unpaid DVD copies are nearly identical, there is no need to pay for an upgrade to a higher-quality experience to see the movie again. Still, it is possible that consumers, once they know they like a movie, would be willing to pay for a legal copy for to continue to watch.

### 3. Industry Background

In 1980 box office revenue provided the majority of revenue to the US motion picture industry. The box office share has declined steadily since then. The US film industry generated \$45 billion in revenue in 2004. The box office share in 2004 was 17 percent, while home video accounted for nearly half.<sup>7</sup>

Between 2000 and 2004, overall industry revenue grew 39 percent in real terms.<sup>8</sup> In contrast to the music industry whose concerns about piracy are motivated in part by sharp recent revenue declines, the movie industry continues to experience robust growth in revenue, driven largely by the growth in DVD sales. Concerns about piracy, while

---

<sup>7</sup> See <http://www.edwardjavepstein.com/mpa2004.htm>, accessed December 13, 2005.

<sup>8</sup> Ibid.

substantiated in many instances (see <http://www.mpa.org/anti-piracy/>), are not additionally fueled by falling industry revenue.

## II. Data

The data for this study are derived from a series of surveys administered to over 500 Penn undergraduates in Spring and late Fall 2005. In each survey, students see a list of the top 50 movies from each of the previous 3 years (2002-2004 in the Spring, 2003-2005 year to date for the Fall). For each movie they are asked whether they saw it, as well as the sequence by which they saw it. Respondents can have seen a movie by any (and all) of 4 paid methods: **theater**, **television**, **rental**, and **purchase**; as well as by either or both of two unpaid means: by viewing a **downloaded** copy, or by viewing a **burned** copy of a legally obtained copy (generally DVD). Thus, for each movie seen, a respondent enters a “1” under the mode they first saw it, a “2” under the mode by which they saw it next (if they saw it more than once), and so on.

In addition to their movie consumption, the survey also asks about family income, respondent race and age, speed of Internet access, and three variables related to their interest in movies: how often they go to movie theaters, how many movies they own, and their level of interest in movies (on a five-point scale). The sample includes nearly 21,000 instances of movie consumption. For analysis we include only respondents who made no mistakes in reporting their sequences.<sup>9</sup> This brings the usable sample down to 470 individuals.

---

<sup>9</sup> We discard data for individuals who indicate that they saw a movie by multiple means but do not indicate the sequence.

Table 1 describes the sample. The mean age is 18, and the age range is between 17 and 22. Nearly a quarter of the respondents are Asian, 6 percent are black, and 6 percent are Hispanic. Respondents are from families of above-average income: Nearly 30 percent of the sample reports family income in excess of \$250,000. Another 37 percent reports family income of \$100,000-\$250,000. Just 10 percent of respondents report family income below \$50,000.

Nearly half (46%) report a “typical” level of interest in movies. Nearly a third report above-average interest in movies. Only a fifth report below average interest in movies. Respondents own a mean (median) of 38 (20) movies, and they go to movie theaters 1.6 (1) time per month. Internet access is ubiquitous among the respondents: 98 percent of respondents report high-speed access in 2005, up steadily from 70 percent in 2002.

For the 150 movies in each survey, respondents report an average of 54.2 instances of paid consumption and 3.0 instances of unpaid consumption. Theatrical consumption is the most common form, accounting for nearly half, followed by rental (20 percent), television (16 percent) and purchase (15 percent). The two forms of unpaid consumption account for 5.2 percent of consumption, the majority (3.1 of 5.2) burned and the remainder (2.1 of 5.2 percent) downloaded.

A particular movie is frequently viewed through more than one medium by a respondent. The bottom part of Table 1 breaks consumption down by medium and by viewing order. While respondents engage in 57.2 episodes of consumption, they see on average only 44.5 separate movies. Virtually all theatrical consumption occurs on first viewing; and theatrical viewing is the way that nearly 60 percent of consumption

sequences begin. The other major medium for first viewing is rental (8.3 of 44.5, or 19 percent). The other media account for relatively few first viewings; on the other hand, first viewings via those media account for most of the viewings via those media. For example only 4 percent (1.8 of 44.5) first viewings use unpaid copies. But these viewings represent most of the unpaid consumption. To put this another way, unpaid consumption is rare; but when it occurs, it usually occurs on first viewing.

The vast majority of second viewings are on purchased copies, television, or rental copies. As with first viewings, unpaid consumption accounts for a small fraction of second viewings. At the same time, over a quarter of unpaid consumption occurs on the second viewing.

The first column of Table 2 shows consumption paths that begin with theatrical and rental consumption. The length of the sequences varies with the medium of initial consumption. Of the sequences beginning with rental, less than fifteen percent are viewed through another medium. Of the sequences that start in the theater, far more – over a third – continue to another medium, roughly 10 percent each to purchase, rental, and television with another few percent subsequently consumed via unpaid means.

### **III. Empirical Strategy**

Before beginning the detailed analysis, it is helpful to distinguish two questions of interest. First, to what extent does an instance of unpaid consumption displace paid consumption? Said another way, what is the *rate* of displacement? Second – if there is displacement – how much paid consumption is displaced by unpaid consumption? Or,

what is the *amount* of displacement? This is calculated as the volume of unpaid consumption times the rate of displacement.

We have two basic approaches to measuring the effect of unpaid consumption on paid consumption. First, we can make straight use of the cross section, asking whether people who engage in more unpaid consumption engage in less paid consumption. Implementing this approach requires a way of predicting how many movies respondents would have paid to consume in the absence of any unpaid consumption. The main challenge to this approach is that persons who like movies may like them through various media. That is, people who like movies may both like to pay for them and to consume them without paying. If it is not possible to control for all determinants of individuals' interest in movies, then unobserved heterogeneity will tend to induce a positive relationship between unpaid and paid consumption, even if the causal relationship is negative.

The structure of our data allows a second, panel approach. The data include individuals' movie consumption, by medium, yearly 2002(3)-2004(5). We can include individual fixed effects – which control for the individuals' vintage-constant interest in movies – then ask whether persons with higher unpaid consumption in some vintages have higher or lower paid consumption of movies from those vintages.

Our sequencing information allows us to use ask slightly more pointed questions. It makes no sense for a downloaded second viewing to affect first viewing. Indeed – to the extent that a second viewing by unpaid means mechanically follows paid theatrical consumption – naïve cross sectional regressions could perversely show positive relationships when analyzed without regard for sequence. We can more sensibly instead

ask whether unpaid first consumption increases or decreases paid first consumption. And whether it increases or decreases second consumption. These are the basic strategies we employ.

### 1. Cross Sectional Approaches

The cross sectional approach to the displacement question asks whether people with more unpaid viewings have fewer paid viewings. As a first pass at this approach we can simply plot first viewings by paid means against first viewings by unpaid means, in Figure 2. The line is a smoothed description of the relationship, and it is essentially flat. One interpretation of this flat relationship is that unpaid consumption does not displace paid consumption. Yet, persons who like movies more might simply engage in more paid *and* unpaid consumption, which would tend to mask any substitution that operates. Indeed, when we break our respondents down by self-reported interest in movies, the persons with high interest in movies engage in both more paid and more unpaid consumption. To meaningfully measure displacement using a cross sectional approach, we need to find variables that predict the counterfactual paid consumption absent any unpaid consumption. We have three main variables for this task, the respondent's interest in movies on a five-point scale, the size of his or her home video library, and how frequently he or she sees a movie in a theater.

To implement this approach in a regression, define:

$B_i^j$  = the number of paid movies that person  $i$  watches for the  $j^{\text{th}}$  time (i.e. the number of  $n^{\text{th}}$  viewings for this person),

$S_i^j$  = the number of unpaid movies that person  $i$  watches for the  $j^{\text{th}}$  time;

$X_i$  = characteristics of person  $i$  such as the level of interest in movies (fan dummy, size of movie collection, frequency of movie theater attendance);

$\varepsilon_i^j$  = unobserved determinants of tendency to make  $j^{\text{th}}$  viewings for person  $i$ .

We can then estimate models of the form:

$$B_i^1 = X_i\beta^1 + \alpha^1 S_i^1 + \varepsilon_i^1, \text{ or} \quad (1)$$

$$B_i^2 = X_i\beta^2 + \alpha^{22} S_i^2 + \alpha^{21} S_i^1 + \varepsilon_i^2. \quad (2)$$

In effect a model like (1) asks whether people with more unpaid first viewings have fewer paid first viewings, after controlling for determinants of paid first consumption.

The second model is similar but examines second paid consumption, with possible relationships between both unpaid first and second consumption on paid second consumption. That is, not only can decisions to engage in second unpaid viewings affect paid second viewings. Unpaid first viewings, if they reduce the tendency for a second viewing, can also affect the tendency for a paid second viewing.

Unless the controls are exhaustive, unobserved heterogeneity is a looming concern with models of this sort: persons who watch movies without paying may simply like movies a lot, even beyond the capacity of our control variables to distinguish among persons in the sample. If so, then our estimate of the extent of displacement (the  $\alpha$  coefficients) will be biased upward, that is, against a finding of sales displacement.

Table 3 reports results, overall and disaggregated by each paid medium. The first five columns are regressions of the number of first paid viewings in this medium on the number of first unpaid viewings (either downloaded or burned). Controls include the fan

dummy, the movie library, and the individual's monthly theater visits. People with larger libraries, more frequent theater patronage, and higher levels of movie fandom have more paid first viewings. The coefficient on unpaid first viewings is negative but insignificantly different from zero. The second column examines theatrical first viewing. Here the coefficient of interest is -0.12, below its standard error of 0.16 and therefore not statistically significant.

Columns (6)-(10) are regressions of second viewings on controls and the numbers of unpaid first and second viewings. Persons with one more unpaid first viewing have 0.34 fewer second paid viewings. The coefficients for television, rental, and purchase are all negative, and the rental coefficient is large and significant (-0.26).

Column (6) also indicates that persons with more unpaid second viewings have more paid second viewings as well (0.42 and significant). Most of this reflects additional rental: persons who watch more unpaid second viewings engage in more rented second viewings.<sup>10</sup>

If one takes these models literally – and causally – then the effect of each additional first unpaid viewing is the sum of 0.01 and -0.34, or a reduction of 0.33 in paid viewings. Each second-time unpaid viewing, on the other hand, raises paid viewing by 0.42. There are many reasons to doubt a causal interpretation, however. First, as already emphasized, there is unobserved heterogeneity. People who like movies will tend to engage in more paid and unpaid consumption. This can look like a stimulating effect of unpaid consumption on paid consumption even if it is not.

---

<sup>10</sup> We also estimated models that disaggregate unpaid consumption into distinct downloaded and burned components. The distinct categories have statistically indistinguishable effects, so here and below we report only models with overall unpaid consumption as an explanatory variable.



More convincing estimates require ways of circumventing the unobserved heterogeneity. A natural strategy is instrumental variables. Our survey includes information on broadband and DVD-burner access which seem like promising instruments for unpaid consumption (see Rob and Waldfogel, 2004; Zentner, 2004). In our data these technological access variables bear no significant relationships with unpaid consumption. These variables fail as instruments for three reasons. First, broadband access is virtually ubiquitous, so there is little variation. Second, even if there were more variation in broadband access, most unpaid consumption copies are obtained via copying rather than downloading, which reduces the usefulness of access speed in predicting unpaid consumption. And while access to DVD-burning hardware would seem a promising predictor, whether one has access to this technology may be less important than whether one's friends have such access. We are left without a promising IV approach.

The sequential nature of consumption allows a different strategy. We can restrict attention to second viewings for the subset of movies with a first paid consumption, then ask how much paid second consumption is displaced by an unpaid second consumption. On average a third of movies first viewed at the theater are subsequently viewed by another means (see Table 2). By restricting attention to the movies first viewed, say, at the theater, we grow closer to holding constant individuals' interest in the movies in question. We can then ask how the number of paid second viewings, per movie first seen in the theater, relates to the number of unpaid second viewings, again relative to the number first seen in the theater.

Table 2 offers a simple implementation of this approach. The table divides survey respondents into two groups, those who engage in no unpaid consumption during the sample, those who engage in some. Table 2 also reports these sequences separately for persons who report no unpaid consumption as well as those engaging in unpaid consumption. A comparison of these groups' consumption sequences provides some suggestive evidence of displacement following a first (paid) viewing.

For persons with no unpaid consumption, 34.9 percent of movies first viewing in a theater are viewed a second time by another paid means. For the persons who engage in some unpaid consumption, only 30.1 percent of movies first viewed in a theater have a second, paid viewing, while 4.5 percent have an unpaid second viewing. That is, both groups follow up a first theatrical viewing with the same tendency to view the movie through a second medium. But the paid viewing by those who sometimes steal is lower by almost exactly the amount of their unpaid consumption. This suggests one-for-one displacement.<sup>11</sup> A similar exercise to second consumption following first viewing of a rental copy yields similar implied displacement.<sup>12</sup> While suggestive, these calculations control for nothing.

Table 4 implements a regression analogue to the comparison in Table 3. Column 1 is a regression of the ratio of paid second viewings to theatrical first viewings on the ratio of unpaid second viewings to theatrical first viewings. The coefficient on the unpaid ratio is  $-0.27$  and nearly statistically significant, indicating that an unpaid second viewing displaces  $0.27$  paid second viewings. The second column regresses the number

---


$$^{11} \frac{\partial \text{paid}}{\partial \text{unpaid}} = \left( \frac{.301 - .349}{.045 - 0} \right) \approx -1.$$

$$^{12} \frac{\partial \text{paid}}{\partial \text{unpaid}} = \left( \frac{-0.043}{0.125 - 0.097} \right) \approx -1.5.$$

of paid second viewings following a theatrical viewing on the number of unpaid second viewings following a first theatrical viewing, controlling for the number of theatrical first viewings. The displacement coefficient is again  $-0.27$  and significant. Columns (3) and (4) repeat the exercise for movies with a first paid rental consumption. Given that subsequent consumption follows initial rental in only 15 percent of consumption sequences, we do not have much scope for finding relationships among the movies first viewed through rental. And, indeed, second unpaid consumption of movies first seen through rental has no statistically discernible displacement effect on second paid consumption.

These estimates that come closer to controlling for unobservable tendencies to buy movies provide some evidence of displacement. Unfortunately, this approach does not provide an estimate of the displacement effect of first unpaid viewing.

## 2. Longitudinal Approach

The problem with the cross sectional estimates is that the error term, containing the individual's unobserved tendency to pay for movies, is correlated with his tendency to engage in unpaid consumption. The solution is to purge the error of this component. Longitudinal data allow this, provided that the unobservable determinant of individuals' paid consumption is constant across vintages. Because these estimates avoid this important form of unobserved heterogeneity, this is our preferred estimation approach, particularly for estimating the effect of first unpaid consumption on first paid consumption.

The movies in the survey represent the 50 most popular (highest box office) films over each of the three years prior to the survey. We can calculate the number of paid and unpaid consumption episodes among the films of each vintage. We can then use the data as a panel, including a fixed effect to purge the error term of individuals' vintage-invariant interest in buying movies. In effect, we ask whether persons who watch more unpaid 2005 movies, relative to 2003 movies, watch fewer paid 2005 movies, relative to 2003 movies.

Figure 3 shows the relationship between the differences in paid and unpaid consumption of movies between the 2003 and 2005 vintages (for the Fall sample). The smoothed line summarizes the relationship, and it is clearly negative. Persons who consume more unpaid movies from 2005 relative to 2003 consume fewer paid movies from 2005 relative to 2003. This looks like sales displacement.

To test this more systematically in a regression framework, define:

$B_{it}^j$  = number of paid  $j^{th}$  (first or second) consumption episodes by person  $i$  for vintage  $t$  movies,

$S_{it}^j$  = number of unpaid  $j^{th}$  (first or second) consumption episodes by person  $i$  for vintage  $t$  movies,

$\mu_i$  = unobserved determinants of an individual's tendency to buy movies,

$\phi_t$  = vintage effect, and

$\varepsilon_{it}$  = idiosyncratic error.

We can then estimate:

$$B_{it}^1 = \alpha^1 + \beta^1 S_{it}^1 + \mu_i + \phi_t + \varepsilon_{it}^1 \quad \text{and} \quad (1)$$

$$B_{it}^2 = \alpha^2 + \beta^{22} S_{it}^2 + \beta^{21} S_{it}^1 + \mu_i + \phi_t + \varepsilon_{it}^2. \quad (2)$$

Table 5 reports results.<sup>13</sup> As before, the first five columns refer to first consumption episodes. The coefficient on unpaid consumption in the first column, -0.76 (s.e.=0.10) indicates that persons with one additional unpaid first consumption engage in 0.76 fewer paid first consumption episodes. Each of the by-medium coefficients is also negative and significant, the largest being -0.27 (theater) and -0.29 (rental).

The last five columns of the table examine the relationship between unpaid first and second consumption and paid second consumption. Here, the coefficient on unpaid second consumption is -0.19 (s.e.=0.09), while the coefficient on unpaid first consumption is -0.24 (s.e.=0.06).

Interpreting the coefficients causally, a first unpaid consumption reduces paid consumption by 1 unit (-0.76 + -0.24). Recall that most first consumption episodes are theatrical and that a third of these are followed by a second paid consumption episode. Hence the maximum displacement exceeds 1.

The longitudinal displacement estimate of unpaid second consumption of -0.19 is similar to the Table 4 estimates of roughly -0.27 following theatrical first consumption.

What do these estimates mean? On average in our data, consumption sequences include 1.3 viewings, so the maximum amount of conceivable displacement exceeds one.<sup>14</sup> An unpaid first consumption reduces paid consumption by 1 unit. This is not 100 percent displacement, but it is nearly 80 percent of the maximum possible displacement.

---

<sup>13</sup> The models are estimated combining data from the Spring (“early”) and Fall (“late”) surveys. Because early respondents report on movies 2002-2004 while late respondents report on movies 2003-2005, each group gets its own vintage dummies in these models.

<sup>14</sup> To see this imagine a first consumption in the theater followed, sometimes, by rental and/or purchase. Alternatively, one might begin by watching an unpaid copy and then engage on no additional paid consumption. A decision to begin with an unpaid copy can therefore displace more than one paid consumption.

An unpaid second consumption reduces second paid consumption by only about 0.2, or 20 percent of the maximum possible displacement, by contrast.

In our sample individuals have an average of 1.8 unpaid first consumption episodes. According to our longitudinal estimates, these unpaid first consumption episodes displace 1.4 paid first consumption episodes ( $=1.8*0.76$ ) per person and 0.4 second consumption episodes ( $=1.8*0.24$ ). Our respondents have an average of 0.9 unpaid second consumption episodes, which displace 0.2 ( $=0.9*0.19$ ) paid second consumption episodes. Overall, then, unpaid consumption displaces an average of 2.0 paid consumption episodes per person in the sample. To put this another way, unpaid consumption reduces paid consumption episodes in the sample from a counterfactual 56.2 to the observed 54.2, or by 3.5 percent.

## **Conclusion**

Unpaid consumption of movies is a small share of consumption, even in a sample of highly Internet connected and technically sophisticated college students from high-income families. Moreover, with the rapid growth of DVDs over the past few years, film industry revenue has grown sharply, even as theatrical attendance has been flat. Unlike the music industry, which has actually experienced a pronounced downturn which is arguably based on file sharing, the movie industry faces smaller outward signs of a problem.

Yet, the evidence in this paper indicates that the file sharing that occurs has sharp displacing effects on paid consumption. Our best estimate indicates that an unpaid first consumption reduces paid consumption by one unit, and an unpaid second consumption

reduces paid second consumption by about a fifth of a unit. The movie sales displacement rates that we estimate are much larger than displacement estimates that we estimate for music (Rob and Waldfogel, 2004). The large rate of displacement suggests that the main effect of unpaid consumption in movies is simply to transfer revenue from sellers to buyers. Unlike in music, where much unpaid consumption would not have occurred without file sharing, in movies unpaid consumption simply displaces paid consumption. This suggests that unpaid movie consumption does not reduce deadweight loss.

Why would displacement rates be so high in movies? File sharing in movies remains clumsy. Persons can choose between slow downloads and relatively easy to defeat DVD file protection. If the means available for copying movies become easier to use, file sharing may become a very serious threat to the film industry. The difficulty of obtaining unpaid copies may explain the sharp displacement. Perhaps only highly motivated consumers – those placing high valuations on the movies they are copying – engage in file sharing today. If file sharing were easier (and therefore more nearly costless), it is possible that even persons placing low valuations on movies – and who would therefore not otherwise consume the movie if paying – would share. Then the displacement would be smaller.

However, even if the cost of obtaining movies without paying falls, movies remain different from music in their cost of consumption. Watching a movie requires a few hours of undivided attention, which is costly to people. Music, on the other hand, can occupy divided attention. Hence, a person can get some use out of music he values very little (he can give it an experimental listen while doing something else). To the

extent that movies are costly to consume even when free, the extent of displacement per unpaid movie viewing may remain quite high.

Because the amount of unpaid consumption in the sample is small (about 5.2 percent), the amount of displacement (3.5 percent) is small as well. While the amount of displacement in the sample is small, our large estimates of the rate of displacement may give the movie industry a cause for concern as the technologies allowing unpaid consumption diffuse widely. Sharp sales displacement effects raise difficult questions about the marketing of information goods that future research will need to address.



## References

Bakos, Yannis, Erik Brynjolfsson, and Douglas Lichtman. "Shared Information Goods." *Journal of Law & Economics* XLII (April 1999): 117-155.

David Blackburn (2004), "On-line Piracy and Recorded Music Sales" available at [http://www.economics.harvard.edu/~dblackbu/papers/blackburn\\_fs.pdf](http://www.economics.harvard.edu/~dblackbu/papers/blackburn_fs.pdf)

David Bounie, Marc Bourreau, Patrick Waelbroeck "Pirates or Explorers? Analysis of Music Consumption in French Graduate Schools." (2005)  
<http://www.ecare.ulb.ac.be/ecare/ws/honorvictor/papers/waelbroeck.pdf>

Hong, Seung-Hyun. "The Effect of Digital Technology on the Sales of Copyrighted Goods: Evidence from Napster." August, 2005.  
<https://netfiles.uiuc.edu/hyunhong/www/napster.pdf>

Hui, Kai-Leung and Ivan Png. "Piracy and the Legitimate Demand for Recorded Music." *Contributions to Economic Analysis*. Vol. 2, no. 1, article 11 (2003).

Liebowitz, Stanley J. "Copying and Indirect Appropriability." *Journal of Political Economy*, vol. 93, No. 5 (Oct 1985): 945-957

Liebowitz, Stan "Will MP3 Downloads Annihilate the Record Industry? The Evidence So Far." Unpublished Manuscript. University of Texas at Dallas. June 2003.

Oberholzer-Gee, Felix and Koleman Strumpf. "The Effect of File Sharing on Record Sales: An Empirical Analysis." Mimeo. Harvard Business School. January 2004.

Shapiro, Carl and Hal Varian. *Information Rules*. Harvard Business School Press. Cambridge, MA. 1999.

Varian, Hal. "Buying, Sharing, and Renting Information Goods." *Journal of Industrial Economics*. Vol. XLVIII, no. 4, Dec 2000, pp. 473-488

Zentner, Alejandro. "Measuring the Effect of Online Piracy on Music Sales." Unpublished Manuscript. University of Chicago. December 2003.

**Table 1: Sample Characteristics**

movie interest	a lot less	less	same	more	a lot more
N=469	6.0%	15.4%	46.3%	23.9%	8.5%
family income (000)	under \$25	\$25-\$50	\$50-\$100	\$100- \$250	over \$250
N=441	3.4%	7.0%	22.7%	37.0%	29.9%
	mean	median	N		
home video library	37.6	20	463		
age	18.7	18	470		
monthly theater visits	1.6	1	461		
DVD burner	no access	access w/o ownership	own one		
N=541	21.1%	24.5%	54.4%		
race	white	Asian	black	Hispanic	
	59.6%	24.5%	5.5%	5.5%	
Viewing Modes on N <sup>th</sup> Viewing					
viewing mode	all	first	second	third	
theater	25.5	25.4	0.1	0.0	
television	8.9	4.9	3.3	0.6	
rental	11.2	8.3	2.7	0.2	
purchase	8.5	4.0	3.8	0.6	
total paid	54.2	42.7	9.8	1.4	
downloaded	1.2	0.8	0.3	0.1	
burned	1.8	1.0	0.6	0.2	
total unpaid	3.0	1.8	0.9	0.3	
total	57.2	44.5	10.7	1.7	

**Table 2: Consumption Sequences, Overall and by whether Engage in Unpaid Use**

	Overall	No Unpaid Use	Some Unpaid Use
<b>total</b>	20687	9315	11372
<i>Sequences Starting With</i>			
theater	56.9%	58.5%	55.6%
television	11.0%	11.5%	10.6%
rental	18.7%	19.2%	18.3%
purchase	9.1%	10.0%	8.3%
downloading	1.7%	0.0%	3.0%
burning	2.3%	0.0%	4.1%
<i>Following First Viewing in Theater</i>			
no subsequent	65.3%	65.1%	65.4%
television	9.8%	10.7%	9.0%
rental	9.7%	9.8%	9.7%
purchase	12.8%	14.4%	11.4%
downloading	0.8%	0.0%	1.5%
burning	1.6%	0.0%	3.0%
Total paid	32.3%	34.9%	30.1%
Total unpaid	2.4%	0.0%	4.5%
<i>Following First Viewing through Rental</i>			
no subsequent	86.7%	87.5%	85.9%
theater	0.1%	0.1%	0.1%
television	6.4%	5.9%	6.9%
purchase	4.5%	6.5%	2.8%
downloading	0.6%	0.0%	1.1%
burning	1.8%	0.0%	3.3%
Total paid	11.0%	12.5%	9.7%
Total unpaid	2.3%	0.0%	4.3%

**Table 3: Cross Sectional Displacement Estimates for First and Second Viewings**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Viewing:	first buy	first theater	first tv	first rental	first purchase	second buy	second theater	second tv	second rental	second purchase
Unpaid First Viewings	0.0062	-0.1200	0.0926	-0.0030	0.0367	-0.3351	0.0120	0.0222	-0.2636	-0.1057
	(0.1965)	(0.1622)	(0.0766)	(0.1150)	(0.0677)	(0.1100)**	(0.0064)	(0.0695)	(0.0499)**	(0.0624)
Unpaid Second Viewings						0.4186	-0.0070	0.0136	0.5347	-0.1227
						(0.1587)**	(0.0092)	(0.1002)	(0.0720)**	(0.0900)
fan==2	9.6677	4.1441	1.8201	4.0491	-0.3456	1.4561	-0.0532	0.8772	0.4692	0.1630
	(3.4942)**	(2.8841)	(1.3620)	(2.0455)*	(1.2045)	(1.8299)	(0.1062)	(1.1556)	(0.8300)	(1.0380)
fan==3	16.9425	9.8594	1.6728	5.1814	0.2289	3.5706	-0.0086	1.0674	1.4291	1.0827
	(3.1999)**	(2.6412)**	(1.2473)	(1.8732)**	(1.1030)	(1.6757)*	(0.0973)	(1.0582)	(0.7601)	(0.9506)
fan==4	24.2538	15.5082	3.8026	5.5389	-0.5959	6.9474	-0.0038	3.4337	1.2330	2.2845
	(3.4618)**	(2.8574)**	(1.3494)**	(2.0266)**	(1.1933)	(1.8137)**	(0.1053)	(1.1454)**	(0.8227)	(1.0288)*
fan==5	28.1352	17.7141	3.9585	7.8630	-1.4004	6.5046	-0.0696	3.6289	0.5197	2.4256
	(4.0252)**	(3.3224)**	(1.5690)*	(2.3564)**	(1.3875)	(2.1080)**	(0.1224)	(1.3312)**	(0.9562)	(1.1958)*
own	0.0288	0.0045	-0.0029	-0.0061	0.0333	0.0320	0.0002	0.0008	-0.0001	0.0311
	(0.0120)*	(0.0099)	(0.0047)	(0.0071)	(0.0042)**	(0.0064)**	(0.0004)	(0.0040)	(0.0029)	(0.0036)**
often	3.2080	3.2493	-0.5044	0.6606	-0.1975	1.1235	0.0114	0.1493	0.6326	0.3302
	(0.6959)**	(0.5744)**	(0.2713)	(0.4074)	(0.2399)	(0.3655)**	(0.0212)	(0.2308)	(0.1658)**	(0.2074)
Constant	19.2457	10.1857	3.3753	2.4537	3.2310	3.0394	0.0525	1.2220	0.6058	1.1591
	(3.0131)**	(2.4870)**	(1.1745)**	(1.7639)	(1.0386)**	(1.5780)	(0.0916)	(0.9965)	(0.7158)	(0.8951)
Observations	454	454	454	454	454	454	454	454	454	454
R-squared	0.28	0.25	0.04	0.04	0.13	0.22	0.01	0.06	0.18	0.21

Standard errors in parentheses. \* significant at 5%; \*\* significant at 1%.

**Table 4: Displacement on Second Viewing, Given First Paid Viewing**

	(1)	(2)	(3)	(4)
	Paid 2nd / First at Theater	# Paid 2nd after Theater	Paid 2nd / First thru Rental	# Paid 2nd after Rental
Unpaid 2nd / Theater	-0.2743 (0.1454)			
# Unpaid 2nd after Theater		-0.2726 (0.1608)		
# First in Theater		0.3592 (0.0213)**		
Unpaid 2nd / Rental			-0.1641 (0.1305)	
# Unpaid 2nd after Rental				0.0182 (0.0506)
# First thru Rental				0.1195 (0.0111)**
fan==2	0.0172 (0.0476)	-0.0757 (1.2739)	-0.0848 (0.0499)	-0.3767 (0.4733)
fan==3	0.0532 (0.0435)	0.0661 (1.1817)	-0.0644 (0.0460)	-0.4225 (0.4383)
fan==4	0.1123 (0.0471)*	0.9524 (1.3001)	-0.0511 (0.0496)	-0.3488 (0.4701)
fan==5	0.0575 (0.0548)	-0.1557 (1.5111)	-0.1185 (0.0568)*	-0.1298 (0.5443)
own	0.0009 (0.0002)**	0.0260 (0.0044)**	0.0003 (0.0002)**	0.0021 (0.0016)
often	0.0015 (0.0095)	0.0606 (0.2638)	-0.0012 (0.0098)	0.0079 (0.0940)
Constant	0.2170 (0.0410)**	-2.1084 (1.1135)	0.1606 (0.0437)**	0.1148 (0.4147)
Observations	452	452	397	397
R-squared	0.11	0.52	0.03	0.28

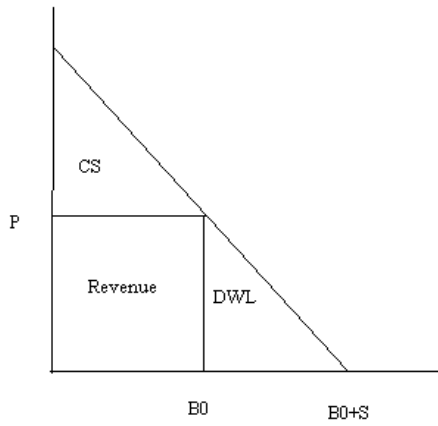
Standard errors in parentheses. \* significant at 5%; \*\* significant at 1%

**Table 5: Longitudinal Displacement Estimates for First and Second Viewings**

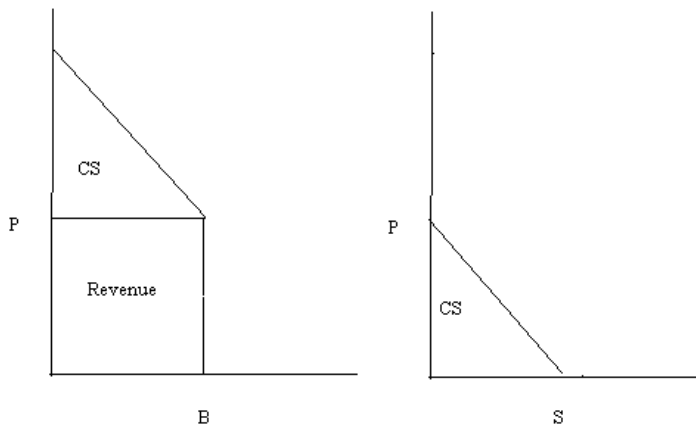
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Viewing:	first buy	first theater	first tv	first rental	first purchase	second buy	second theater	second tv	second rental	second purchase
Unpaid First Viewings	-0.7579	-0.2652	-0.1186	-0.2947	-0.0793	-0.2443	0.0013	-0.1188	-0.0351	-0.0917
	(0.0958)**	(0.0694)**	(0.0507)*	(0.0594)**	(0.0373)*	(0.0585)**	(0.0037)	(0.0415)**	(0.0283)	(0.0360)*
Unpaid Second Viewings						-0.1944	-0.0034	-0.0423	0.0176	-0.1663
						(0.0862)*	(0.0055)	(0.0611)	(0.0417)	(0.0530)**
early_d2003	0.9531	0.3567	0.2336	0.1297	0.2331	0.5010	-0.0136	0.1813	0.1098	0.2236
	(0.2818)**	(0.2041)	(0.1490)	(0.1748)	(0.1097)*	(0.1722)**	(0.0109)	(0.1220)	(0.0833)	(0.1058)*
early_d2004	-1.9830	-0.3221	-0.4335	-0.9954	-0.2320	-1.2785	-0.0138	-0.5000	-0.2274	-0.5373
	(0.2822)**	(0.2044)	(0.1493)**	(0.1750)**	(0.1099)*	(0.1718)**	(0.0109)	(0.1217)**	(0.0831)**	(0.1056)**
late_d2004	-0.2637	0.0741	-0.5439	0.1750	0.0311	-0.2853	0.0147	-0.3420	0.0669	-0.0248
	(0.2577)	(0.1867)	(0.1363)**	(0.1598)	(0.1003)	(0.1567)	(0.0099)	(0.1110)**	(0.0758)	(0.0963)
late_d2005	-6.3243	-0.9934	-1.7435	-2.3790	-1.2084	-3.2709	0.0149	-1.1645	-0.8565	-1.2648
	(0.2246)**	(0.1627)**	(0.1188)**	(0.1393)**	(0.0875)**	(0.1365)**	(0.0086)	(0.0967)**	(0.0660)**	(0.0839)**
Constant	13.9864	7.5657	1.9289	3.0669	1.4249	3.7833	0.0189	1.3420	0.9495	1.4729
	(0.1334)**	(0.0967)**	(0.0706)**	(0.0828)**	(0.0520)**	(0.0824)**	(0.0052)**	(0.0583)**	(0.0398)**	(0.0506)**
Observations	2164	2164	2164	2164	2164	2164	2164	2164	2164	2164
Number of individuals	541	541	541	541	541	541	541	541	541	541
R-squared	0.43	0.04	0.14	0.24	0.15	0.35	0.00	0.11	0.15	0.19

Standard errors in parentheses. \* significant at 5%; \*\* significant at 1%

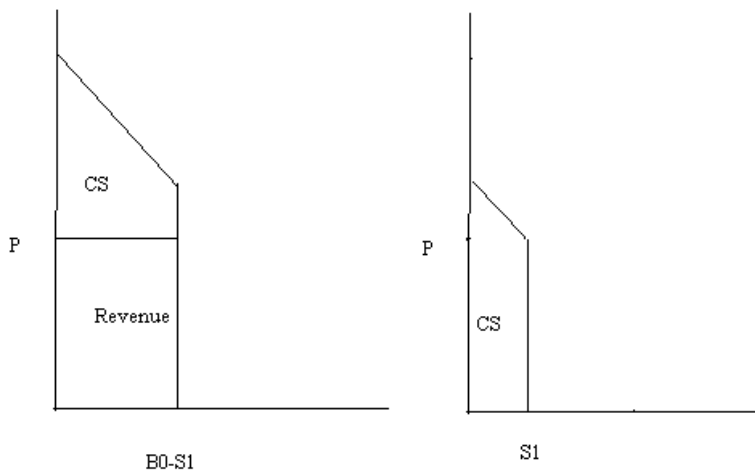
**Figure 1**  
a: No Unpaid Consumption



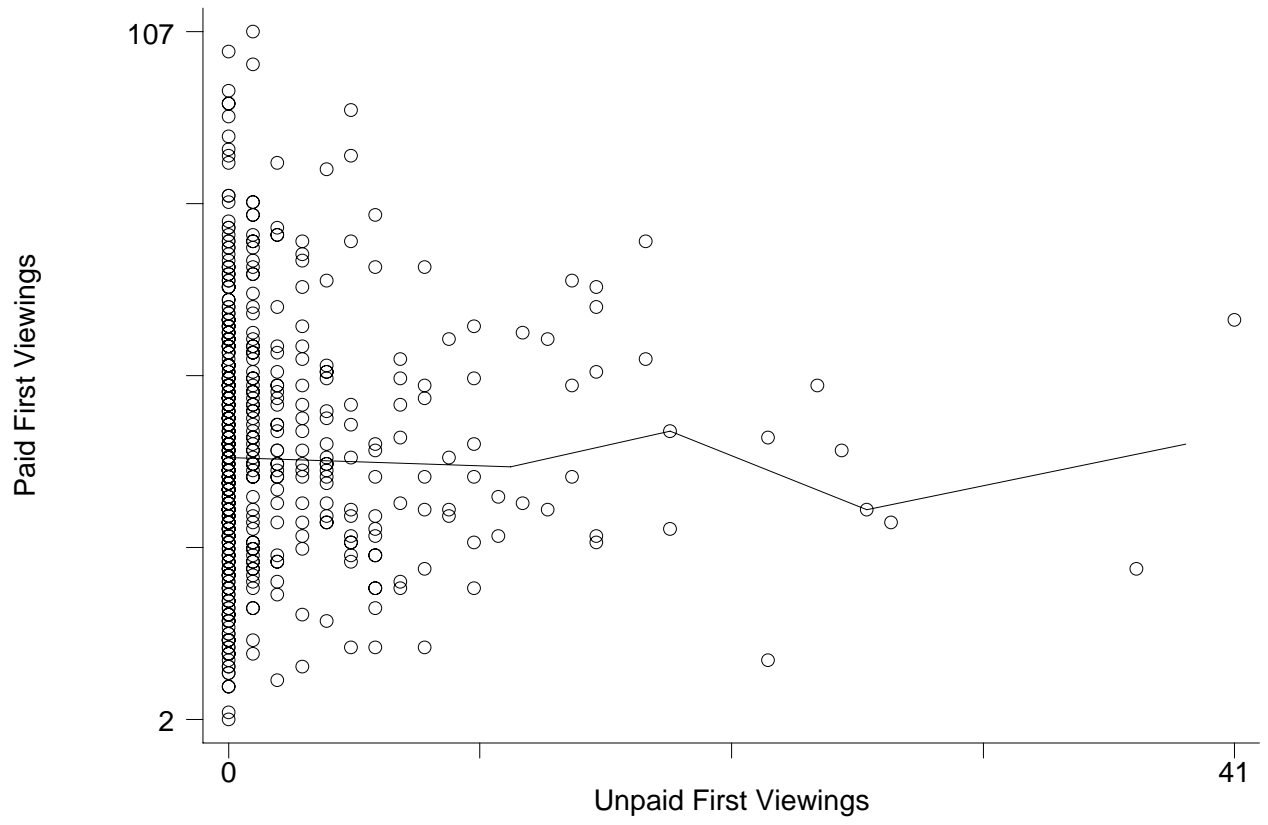
1.b Paid and Unpaid: Low Valuation File Sharing



1c. High Valuation File Sharing



**Figure 2: Paid and Unpaid First Viewings**





**Figure 3: Changes in Paid and Unpaid First Viewings, 2003 to 2005 (Fall Sample)**

