Carnegie Mellon University
Research Showcase @ CMU

Heinz College Research | Heinz College

11-1-2008

Converting Pirates without Cannibalizing Purchasers: The Impact of Digital Distribution on Physical Sales and Internet Piracy

Brett Danaher
University of Pennsylvania

Samita Dhanasobhon
Carnegie Mellon University

Michael D. Smith
Carnegie Mellon University, mds@andrew.cmu.edu

Rahul Telang
Carnegie Mellon University, rtelang@andrew.cmu.edu

Follow this and additional works at: http://repository.cmu.edu/heinzworks

This Working Paper is brought to you for free and open access by the Heinz College at Research Showcase @ CMU. It has been accepted for inclusion in Heinz College Research by an authorized administrator of Research Showcase @ CMU. For more information, please contact research-showcase@andrew.cmu.edu.

Converting Pirates without Cannibalizing Purchasers:

The Impact of Digital Distribution on Physical Sales and Internet Piracy

Brett Danaher,[*] Samita Dhanasobhon,[†] Michael D. Smith,[†] and Rahul Telang[†]

bdanaher@wharton.upenn.edu, {sdhanaso, mds, rtelang}@andrew.cmu.edu

This Version: March 2010

Available From: http://ssrn.com/abstract=1381827

Acknowledgements: The authors thank Mel Stephens, Lowell Taylor, Joel Waldfogel and workshop participants at the 2009 International Conference on Industrial Organization, 2008 Mini-conference on Marketing and Innovation, the 2008 ZEW Web 2.0 Workshop, the 2008 Workshop on Information Systems and Economics, the University of Texas at Dallas, Carnegie Mellon University, the University of California at Irvine, and Arizona State University for valuable comments on this research. We thank Guillermo Jose Aguilar, Temi Awofisayo, Salahaldin Hussein, Suganth Pranesh, and Bryan Springer for excellent research assistance. The authors acknowledge the National Science Foundation for generous financial support provided through CAREER award IIS-0118767 (Smith) and CAREER award CNS-0546009 (Telang). Rahul Telang also acknowledges financial support of Alfred P. Sloan foundation industry studies fellowship.

[*] The Wharton School, University of Pennsylvania, Philadelphia, PA, 19104

[†] School of Information Systems and Management, Carnegie Mellon University, Pittsburgh, PA, 15213.

Converting Pirates without Cannibalizing Purchasers:

The Impact of Digital Distribution on Physical Sales and Internet Piracy

ABSTRACT

The availability of digital channels for media distribution has raised several important questions for marketers, notably whether digital distribution channels will cannibalize physical sales and whether legitimate digital distribution channels will dissuade consumers from using (illegitimate) digital piracy channels. We address these two questions using the removal of NBC content from Apple's iTunes store in December 2007, and its restoration in September 2008, as natural shocks to the supply of legitimate digital content, and analyzing its impact on demand through BitTorrent piracy channels and the Amazon.com DVD store.

To do this we collect two large datasets from Mininova and Amazon.com documenting levels of piracy and DVD sales for both NBC and other major networks' content around these events. We analyze this data in a difference-in-difference model and find that NBC's decision to remove its content from iTunes in December 2007 is causally associated with an 11.2% increase in the demand for pirated content. This is roughly equivalent to an increase of 49,000 downloads a day for NBC's content and is approximately twice as large as the total legal purchases on iTunes for the same content in the period preceding the removal. We also find evidence of a smaller, and statistically insignificant, decrease in piracy for the same content when it was restored to the iTunes store in September 2008. Finally, we see no change in demand for NBC's DVD content at Amazon.com associated with NBC's closing or reopening of their digital distribution channel on iTunes.

> "We can't compete with free. That's an economic paradigm that doesn't work."
> *James Gianopulos, Co-Chairman, Twentieth Century Fox, (quoted in Thompson 2003)*

> "You'll never stop [piracy]. What you have to do is compete with it."
> *Steve Jobs, CEO Apple Inc., (quoted in Goodell 2003)*

1. Introduction

The development of digital distribution channels has raised several important questions for marketers. For television and movie studios, two of the more important questions are (1) can paid digital distribution channels serve as an attractive alternative to consumption through ("free") digital piracy channels and (2) will digital distribution cannibalize DVD box set sales?

With respect to the first question, the quote above from James Gianopulos, Co-Chairman of Twentieth Century Fox, is representative of many in the industry who claim that it is difficult, if not impossible, to successfully use paid digital distribution channels to compete with a free (albeit illegal) piracy channel. On the other side of this argument, Steve Jobs, CEO of Apple Incorporated, claims that digital distribution channels, such as Apple's iTunes video store, offer studios the best opportunity to compete with piracy channels by mimicking the ease and convenience of pirated channels at a competitive price point.

With respect to the second question, there is ample evidence in the business press that DVD retailers feel that studios' distribution through digital channels will significantly cannibalize sales of DVDs. For example in late 2006 after Disney finalized a deal to distribute its movies through iTunes, press reports claim that a Wal-Mart executive visited Hollywood studios to tell them that "it will retaliate against them for selling movies on Apple's iTunes [store]" (Arango 2006). This report goes on to note that Wal-Mart, which makes up an estimated 40% of studios' DVD sales, made good on this threat by sending "'cases and cases' of DVDs back to Disney." Similarly,

Target, which makes up an estimated 15% of DVD sales, sent a letter to studios threatening them not to follow Disney into digital distribution (McBride and Marr 2006), and reportedly after sending this letter "ordered its stores to take down a multitude of internal signs steering customers to Disney products," and replaced Disney's end-cap promotional displays with displays for Disney's competitors (Menn 2006).

However, while these questions have received much discussion in the industry and in the press, we are aware of no studies in the literature that address the degree to which digital distribution of media content impacts demand for physical content and demand for Internet piracy. The goal of this paper is to analyze these questions through a quasi-experiment that occurred on December 1, 2007. During August of 2007, NBC expressed dissatisfaction with the Apple iTunes store's pricing policy. While NBC (and other media companies) wanted more flexibility in pricing, Apple was enforcing a one-price-fits-all policy across nearly all television for sale on iTunes. When negotiations broke down, NBC announced that they would remove all of their content from iTunes on December 1, 2007, a significant move since they reportedly supplied 40% of all video content on the iTunes store.[1] In response, Apple refused to offer NBC's fall 2007 season for sale starting in September 2007, and then on December 1, 2007 Apple removed all older NBC content from iTunes.

In our analysis, we use this event as an exogenous shock to legal digital supply of all older seasons of NBC television. This content was generally available for sale on DVD, on iTunes (prior to December 1), and through piracy, and thus we study a market with a physical sales channel, a digital sales channel, and a piracy channel. We account for the possibility of a time trend by observing changes in piracy and DVD sales two weeks before versus two weeks after

[1] Available from http://www.msnbc.msn.com/id/20531069 last accessed October 23, 2009.

December 1 for NBC's competitor networks: ABC, CBS, and Fox (all of which continue to offer their content and thus received no shock on December 1). We then contrast this time trend with the change in piracy and Amazon.com DVD sales for NBC, arguing that any differences for NBC content after December 1, over and above the difference for similar television networks, was caused by the removal of NBC content from iTunes. Finally, we ask if our findings are upheld in a second experiment on September 9, 2008, when NBC restored all of their content to the iTunes store.

We find that the removal of NBC content from iTunes caused a 11.2% increase in piracy for their content, which corresponds to 29 more pirated downloads per day per episode, or about 53,000 total additional pirated downloads per day. This number is nearly twice as high as the daily number of downloads of these episodes on iTunes in the two weeks prior to December 1, implying a fixed cost associated with the decision to pirate: once individuals start to pirate, they pirate more content than they would have originally purchased. This may also imply a spillover effect — that piracy of content on other networks could have increased as a result of NBC's decision to remove their own content from iTunes. While we cannot positively identify this externality due to lack of an appropriate counterfactual, our results are consistent with such an effect as non-NBC piracy increases by 5.8% over this time period (and thus the documented 11.2% increase in piracy may understate the true displacement of piracy by digital distribution). Finally, while studying NBC's return to iTunes in September 2008 is complicated by the start of a new season of television, our evidence suggests that the restoration of NBC content to iTunes caused a smaller, statistically insignificant drop in piracy, which is also consistent with a fixed cost to piracy.

In contrast to the strong correlation between legitimate digital distribution and piracy, we find no change in the Amazon.com sales rank of NBC television season box sets in the four weeks surrounding December 1 relative to the baseline change in non-NBC box sets, implying that while customers who cannot purchase digitally may turn to piracy, they do not consider DVD box sets — at least those sold on Amazon.com — as a substitute to digital downloads.

The remainder of this paper proceeds as follows. In Section 2, we review the relevant literature. In Section 3, we give a general description of the digital market for media and present a theoretical model of the consumer's choice between piracy, digital channel purchase, and physical channel purchase. In Section 4, we discuss our data. In Section 5, we present our empirical models and results. Finally, in Section 6, we discuss the implication of these findings, the limitations of the analysis, and areas for future research.

2. Literature Review

As this paper addresses the interaction between legitimate digital and physical distribution channels and the interaction between legitimate and illegitimate digital distribution channels, the paper fits into two main literatures: the marketing literature studying interactions between various distribution channels, and the economics and information systems literature on online piracy of digital goods.

With respect to the piracy literature, most existing studies examine the effect of online piracy on physical media sales. Most of these studies analyze the impact of piracy on music CD sales, with a few recent studies examining the impact of piracy on movie or television revenue. The challenge in this literature is typically identification, as the correlation between physical sales

and pirated downloads of each movie or song is predominantly driven by unobserved heterogeneity across goods.

Papers in this literature address the identification issue in several different ways: through cross-country variation, exogenous shocks to demand, or through survey results. With respect to cross-country variation, Zentner (2005), Hui and Png (2003), and Peitz and Waelbroeck (2004) use international panel data on music sales and pirated downloads with each study finding that piracy displaces CD sales to some extent. Similarly, Danaher and Waldfogel (2008) examine the impact of online piracy of Hollywood movies on international box office revenue and find evidence of displacement of ticket sales by online piracy.

In the context of exogenous shocks, one of the tests used by Oberholzer and Strumpf (2004) takes holidays in the German school system as exogenous demand shocks and find little or no displacement associated with music piracy, while Smith and Telang (2009) use the television broadcast of a movie as an exogenous demand shock for the DVD and find little or no displacement of DVD sales from piracy for movies broadcast on television.

Finally, in the context of survey data, Rob and Waldfogel (2006) use survey data from a population of college students, asking whether individuals who pirate music purchase less music, including controls such as stated valuations of the albums in question or personal taste for music. Rob and Waldfogel (2007) use a similar approach to study the effect of movie piracy on paid consumption of movies such as theater attendance, DVD rental, and DVD purchase. In both studies the authors find displacement of paid consumption by piracy.

Thus, the majority of existing studies in the literature find some degree of substitution of unpaid "pirated" consumption for paid consumption, which raises the question of how firms should

optimally combat the negative effects of piracy. Recent papers in the literature have examined this question in the context of litigation against pirates, protection of media content through Digital Rights Management (DRM) systems, and purposefully damaging the performance of file sharing networks.

Addressing the effectiveness of the first anti-piracy tool, Bhattacharjee *et al.* (2008) examine the impact of the RIAA's legal threats against individual file sharers during the summer of 2003 as a quasi-experiment and find that when the threat of litigation is higher, file sharing declines but availability of content is still substantial. In the context of DRM protection, Vernik (2008) uses an analytic model to argue that the presence of DRM may *increase* piracy by reducing the usability of the purchased files. In the context of degrading the performance of file-sharing networks, Christin *et al.* (2005) study the impact of several different "poisoning" strategies on four popular peer-to-peer file sharing networks and find that the injection of a few replicated decoys can strategically manipulate users' perception of content availability in the network.

However, another important tool that media companies may be able to use to reduce the impact of piracy is directly competing with piracy channels by adding legitimate digital distribution channels. To the best of our knowledge, there are no papers in the literature that examine the interaction between legitimate and illegitimate (piracy) digital distribution channels, and one contribution of this paper is to examine this question.

Our question is also related to the interaction among different distribution channels, which has been widely studied in the marketing literature. For example, Jeuland and Shugan (1983) show that coordination between distribution channels lead to higher profits. Extending this finding, Chiang *et al.* (2002), Chu *et al.* (2007), and Webb (2002) develop strategies for firms to manage

multiple distribution channels effectively. In the context of direct distribution channels, Balasubramanian (1998) uses analytic models to show that the presence of direct distribution channels, including Internet channels, yields higher returns when the product is well adapted to the channel. Other papers in the marketing literature also note that the more differentiated two channels are, the less likely they are to cannibalize one another (e.g., Friedman and Furey (1999) and Viswanathan (2005)).

However, there are relatively few papers in the literature that attempt to directly measure the effect of digital distribution on physical channel sales. One exception is Deleersnyder *et al.* (2002) which uses a sample of 85 British and Dutch newspapers who added digital distribution channels and finds that when newspapers make their content available online it only has a small impact on physical newspaper sales. Likewise, Biyalogorsky and Naik (2003) find that Tower Records' addition of an Internet distribution channel did not significantly cannibalize their retail sales. With respect to video distribution, Waldfogel (2007) uses survey data to show that authorized Youtube viewing of television content has only a small net displacement effect on over-the-air viewing and may achieve complementarities between the two channels.

In summary, the challenges of identification are significant in this domain. A key contribution of our paper is to use a natural quasi-experiment matched with data that we have collected to address the identification problem and show how legitimate digital distribution channels interact with physical and piracy channels. We believe that our paper is unique in both the data we have collected and the use of an event as a natural experiment to make causal inferences.

3. Theory

3.1. Digital Distribution and Piracy

Theory does not clearly predict the effect of a digital sales channel on consumption in a digital pirate channel. On one hand, iTunes purchases (by far the dominant legitimate digital channel for video purchases) and pirated downloads are similar in that both provide high quality, usually fast file downloads that can be viewed on a computer or, with some effort, a television or portable video device.[2]

Given these similarities, one might ask why anyone would purchase through a digital distribution channel if piracy is free. For this to occur, there must be some non-financial cost to piracy. There are several possible categories into which that cost could fit:

i) There may be a cost to learn to use BitTorrent, which would be akin to a fixed cost (especially since BitTorrent is generally considered to be easy to use once learned).

ii) Individuals may experience moral qualms about pirating, which could have the characteristic of a fixed or variable cost.

iii) Individuals may fear being caught and punished, a cost which is also variable with respect to downloads.

iv) It is possible that pirated downloads are viewed as less convenient (and lower quality) as compared to iTunes consumption (either due to the relative ease of use

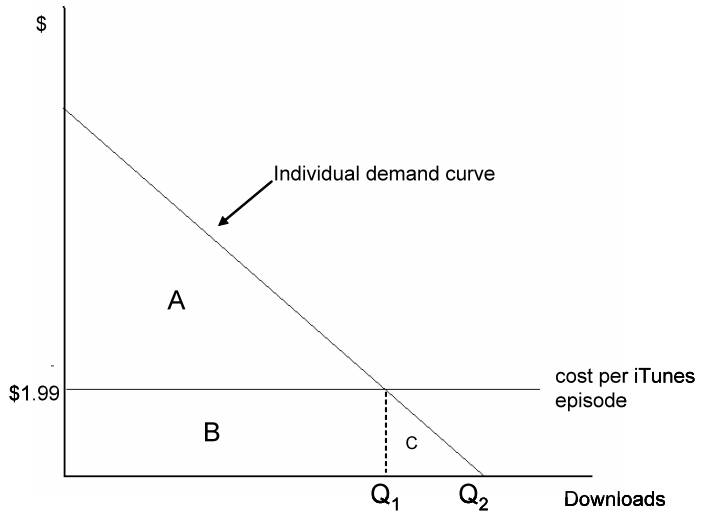
[2] Pirated files, of course, tend to be easier to share or use on a variety of devices (due in large part to the lack of Digital Rights Management restrictions on playback) whereas iTunes downloads tend to have more consistent quality.

of iTunes versus piracy sites or the variability in quality through piracy sites). This cost would also be variable with respect to the number of downloads.

The nature of these costs inform what we expect to happen to piracy when the digital distribution channel is removed (or introduced). Consider Figure 1 below.

Figure 1 represents an individual's demand curve for downloads of episodes of television. In this figure, we assume a typical downward sloping demand curve for episodes — the intercept would be the most highly valued show downloaded, and each successive download has diminished returns. In this figure, we assume that the non-financial cost of pirated downloads is completely fixed, equal to some number F.[3]

Figure 1



[3] For simplicity, we assume that there is no fixed cost to using iTunes, as it is a legal and less intimidating platform. Our analysis could easily be adapted to include a small fixed cost to using iTunes. We also assume a finite time horizon in which the pirated technology can be used. This assumption is motivated by the rapid rate of change in digital piracy technology.

In this setup, if an individual chooses to download through iTunes, she will download Q_1 episodes and receive consumer surplus equal to area A. If the individual chooses to pirate, she will download Q_2 episodes (her satiation point for digital television) and receive consumer surplus equal to A+B+C–F. The consumer decides whether to pirate or purchase by comparing the surplus from iTunes (A) and piracy (A+B+C-F) and choosing the greater of the two. Thus, the individual will choose to download all episodes from iTunes if B + C < F — in other words, if the additional surplus gained from piracy is less than the fixed cost (or the perceived fixed cost) of pirating.

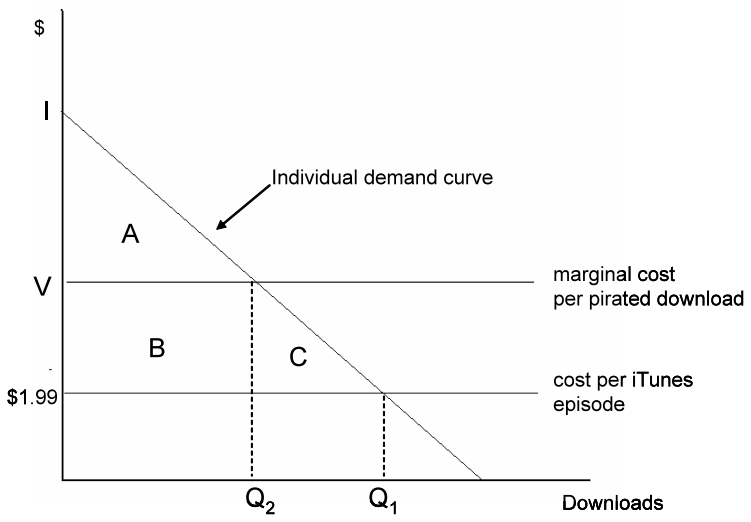
This leads to several predictions for our empirical analysis. For an individual who is using iTunes (B+C<F), if the content is removed from iTunes there are two possible outcomes. If A+B+C<F, then the individual will not turn to piracy and will no longer download the content (either no longer consuming it, or possibly purchasing the DVD box set if that is a substitute). However, if A + B + C > F, then if the iTunes channel is removed, the individual will download Q_2 episodes, and thus the increase in piracy will be greater than the original number of iTunes purchases.

As an alternative outcome, consider Figure 2, in which there is no fixed cost to piracy but instead a constant marginal cost equal to V. Note first that if V (the constant marginal cost of each pirated download) were greater than I (the demand intercept) the individual would never pirate regardless of the availability of a legal download source (a result similar to that from Figure 1 when there is a very high fixed cost). However, if \$1.99 < V < I as in the diagram,[4] then the individual will purchase Q_1 episodes from iTunes if they are available and receive surplus of A + B + C. If the legal purchase option is removed, then the individual will pirate Q_2 episodes and

[4] If V < \$1.99, then this is effectively equivalent to Figure 1 where we get the result that Q2 > Q1.

receive surplus equal to area A. Note that in this case, $Q_2 < Q_1$ so the increase in piracy is less than the original number of iTunes purchases before removal of the legal download alternative.

Figure 2



There are other possible structures for the non-financial cost of piracy, including a mix of fixed and variable costs as well as increasing or decreasing marginal cost.[5] However, in even the simple examples diagrammed here we have shown that when the digital purchase channel is removed, it is possible that an individual would either not turn to piracy, begin to pirate a number of episodes less than or equal to the number of purchases she had been making on iTunes, or even begin pirating more content than she had ever purchased.

Unfortunately, because we only observe aggregate changes in piracy levels after removal of the iTunes channel, we cannot use these models to analyze the perception of piracy costs at an individual user level. However, these models suggest that we can infer the general nature of

[5] In fact, it is possible to show that if the marginal cost is increasing then consumers might mix behaviors, pirating some episodes and purchasing others.

piracy costs perceived by consumers by observing changes in piracy after iTunes removal. Specifically, an increase in piracy after iTunes removal that is larger than the pre-removal iTunes sales would be consistent with a fixed cost to piracy among a substantial number of users, either through learning or moral costs.

3.2. Digital Distribution and Physical Sales

The theory behind the substitutability of iTunes purchases for DVD box sets is less clear than it was for piracy. DVD box sets may take longer to acquire and are often priced differently than the equivalent television content sold on iTunes. One might also speculate that DVD box sets appeal to a different market segment than the iTunes video store — possibly less technologically savvy viewers, viewers who prefer to watch content on their televisions, or perhaps gift givers. Finally, DVD's often contain "bonus" extra content that is lacking in iTunes files or (some) pirated copies.

Even if box sets and file downloads are substitutes for each other, it is possible that individuals who "go digital" will be disinclined to go back to the physical product (indeed, this would be consistent with a fixed cost associated with beginning to download/watch television online). However, it is also possible that some iTunes consumers would have otherwise purchased the DVD box set. Thus, when deciding whether and how to have a digital distribution channel, networks must balance the potential benefits (revenues from new consumers, regained revenues from pirates) against the potential costs (lost revenues from original channel sales).

In summary, iTunes customers may otherwise have been pirates, may otherwise have purchased the box set, or may otherwise not have consumed the content at all. We have shown that it is

even theoretically possible that the availability of content on iTunes could displace more pirated downloads than the number of episodes being purchased on iTunes.

Following this theoretical analysis, we ask the following two major empirical questions:[6]

i) What happens to the level of piracy of television content when that content is removed from iTunes (and when it is returned)?

ii) What happens to DVD sales of television seasons sold on the Internet when those seasons are removed from the iTunes store?

The answers to (i) and (ii) also provide evidence of the percentage of iTunes purchases that come from otherwise would-be pirates, the possibility that one network's decision to use (or not use) a digital distribution channel can influence piracy of content on other networks, and the shape of the non-financial cost curve associated with piracy.

4. Data

To address these questions we use panel data on consumption of pirated television content through major BitTorrent tracker sites, and panel data on sales of DVD season box sets at Amazon.com. We describe these data in more detail below.

4.1. Piracy Data

Following Smith and Telang (2009), we use the level of daily downloads of BitTorrent tracker files at Mininova.org as a proxy for piracy activity on the programs in our sample. The website Mininova is a search engine for torrent trackers — the files that allow you to link to other

[6] As we discuss in detail in the data section, it is important to note that of necessity we are asking these questions in relation to older, off-season content for most of our analysis.

computers and download a specific piece of content. BitTorrent serves as a useful proxy for video piracy as it was the most popular source of pirated video downloads during our study period (Smith and Telang 2009). Mininova is a useful proxy for download levels through the BitTorrent protocol because it was the most popular BitTorrent tracker site during our study period according to Alexa.com,[7] it posts a large number of television tracker files, and unlike some other sites it provides information on the cumulative number of downloads for all tracker files downloaded from its site.

Our piracy dataset contains the daily number of downloads for 5,200 unique episodes of television (corresponding to roughly 75 unique series) starting November 16, 2007. The data include the series name, season number, and episode number of each television program and the number of times that file was downloaded each day. We also added indicators for the network that owns the rights to the show, the genre of the show, and whether it is a series that is still producing new episodes (such as Heroes) or a "catalog" series (such as the original Star Trek).

This dataset was created from a larger dataset we collected monitoring all trackers posted to Mininova. We collected data at the torrent level starting in November 2007, obtaining roughly 210,000 records per day and yielding dataset of over 68 million observations for 180,000 torrents. We extracted the torrent file names from this dataset and interpreted the file names to code the series, season, and episode for our television data. When a file contains multiple episodes of a television show, we counted this as a download for each episode contained in the file. Because multiple files frequently map to the same episode of television (for example, there may be 6 different torrent files that contain, say, season 1 episode 4 of Grey's Anatomy), we then

[7] Available from http://www.alexa.com/browse?CategoryID=1316737, last accessed April 19, 2009.

collapsed the data to the episode level by adding the total daily downloads for an episode across all tracker files mapping to that episode.

For our analysis, we focus on piracy among television programming for NBC and its subsidiaries (USA and the Sci-Fi Channel). We also analyze piracy for television programming from the other major television networks — ABC, CBS, and Fox — as a control. We removed all content from the 2007-2008 season from the data because, as noted above, the NBC content being sold on iTunes prior to December 1 only included episodes prior to the 2007-2008 season.[8] Thus, our analysis compares changes in piracy for older "out-of-season" content.

Finally, in our analysis we focus on the time period the two weeks before and two weeks after December 1, 2007 (as well as the two weeks before and after September 9, 2008) in order to best isolate the effect of the removal of NBC content from iTunes on piracy. Our main strategy will be to compare the change in piracy for NBC content after December 1 to the change in piracy for non-NBC content, arguing that any incremental NBC change over and above the non-NBC change is attributable to the removal of NBC content from iTunes.

Importantly, while December 1, 2007 was the official date of NBC's removal from iTunes and December 2 is the first entire day on which the iTunes store held no NBC content, Apple actually began the removal process on November 30 and continued through December 1. Thus we might expect to see some increase in piracy as early as November 30, but we conservatively remove December 1 from the data and code December 2 as the first day of the "post removal" period in the data. If, as our data shows, piracy began to increase as soon as the removal began, then our

[8] This is because, starting in August 2008, Apple did not add any new NBC content to the iTunes store in response to NBC's announcement that it would remove all of its content in December.

selection will lead to an underestimate of the change in piracy caused by the content's removal from iTunes.

It is also worth noting that past studies on Internet piracy have rarely made use of events or "quasi-experiments," because these events often occur with short notice and data collection on piracy cannot begin soon enough to match the event. Thus a contribution of our study is the method of data collection, which allows us to track a good index of piracy over time and analyze this data when shocks are observed.

4.2. DVD Sales Data

To analyze the effect of the December 1 experiment on DVD sales, we use a panel data on sales ranks of DVD season box sets on Amazon.com for the same date range: November 16 – December 15, 2008. We selected Amazon.com as our reference point for changes in DVD sales following the removal of a digital channel because Amazon.com has an estimated 90% share of online DVD sales in the United States (Netherby 2005) and it seems likely that users who are no longer able to purchase television content through iTunes would be disproportionately more inclined to purchase DVD box sets from an online retailer than from brick-and-mortar retailers. However, given that Amazon.com is only the fourth largest seller for DVDs in the United States, behind brick-and-mortar retailers Wal-Mart, Target, and Best Buy (DVD News 2006) a conservative interpretation of our results is that they apply only to consumption through Internet DVD channels.

To analyze changes in DVD sales at Amazon.com, we collect daily observations of the price and sales rank for each DVD Amazon sells. We then focus on DVD box sets of television content for the four major networks, capturing the television series, season, and network names.

We interpret the sales rank based on prior work,[9] which has shown that the relationship between sales rank and sales follows a Pareto distribution:

$$\times \tag{1}$$

Thus, following the experiment proposed by Chevalier and Goolsbee (2003), one can estimate the parameters of the relationship between Amazon.com sales rank and actual sales of the product. However, for the purpose of this study, we simply note that this implies that the relationship between price and sales rank is best modeled as an elasticity. Prior research[10] has dealt with the non-linear relationship between Amazon sales and Amazon sales rank by analyzing the effect of events or explanatory variables on the log of sales rank. If an effect is found, the experimental approach can allow us to translate the coefficient into the actual effect on sales, but this turns out to be unnecessary here as we find no statistical effect.

As in our analysis of piracy, we will compare the change in sales rank of NBC season box sets after December 1 to the change in sales rank of ABC, CBS, and Fox season box sets in order to determine the effect of the removal of NBC's digital distribution channel.

4.3. *Summary Statistics:*

The simplest possible analysis of this quasi-experiment is a comparison of means before and after the removal of NBC content from iTunes on December 1. So to start, we list some very high-level summary statistics and ask if the average number of daily downloads of an NBC episode is greater in the two weeks after December 1 than in the two weeks before. We also test the hypothesis of whether more unique NBC episodes become available through piracy when NBC content is removed from iTunes. Table 1 displays these summary statistics.

[9] Chevalier and Goolsbee (2003); Brynjolfsson, Hu, Smith (2003); Ghose, Smith and Telang (2006)
[10] Smith and Telang (2008)

Table 1: Piracy Data Summary Statistics
November 18 – December 15, 2007

	NBC Networks	*Non-NBC Networks*
1. Mean daily downloads before 12/1 (balanced panel)	259	405
2. Mean daily downloads after 12/1 (balanced panel)	265	396
3. Change	*6*	*-9*
4. Number of unique episodes available before 12/1	1,683	3,400
5. Number of unique episodes available after 12/1	1,812	3,383
6. Change	*129*	*-17*
7. Total number of all downloads prior to 12/1	6,366,662	19,612,479
8. Total number of all downloads after 12/1	6,576,952	18,903,649
9. Change	*210,290*	*-708,830*
	3%	*-4%*

In rows 1, 2, and 3, we report the mean download statistics for NBC and non-NBC shows. During our study period, average number of downloads for NBC increased while the average for non-NBC shows decreased. Thus, it provides basic trend for our results that NBC downloads have increased compared to the control panel. We use a balanced panel to construct these statistics: and in our subsequent regressions: i.e. we only use episodes that were available both before and after the event. If a new episode appears then it is a supply side shift and we ignore it to avoid the bias in the demand side shift. In rows 4, 5, and 6, we outline the supply sides of the effect of content removal from iTunes. In the two weeks after December 1 there were a net 129 more unique NBC episodes available for piracy than in the previous two weeks.[11] A more detailed analysis shows that 147 new episodes of NBC television became available through pirated channels in the two weeks after December 1, and 18 episodes ceased to be available. In contrast, for non-NBC content, we don't observe any new episodes becoming available, and we observe 17 episodes ceasing to be available. The number of new NBC episodes is striking, since we would expect older content to simply become less popular (and as a result less available) over time.

[11] In fact, just as we see some non-NBC torrents expiring, we also see some NBC torrents expire after December 1. So there are actually more than 128 new episodes of television being introduced on the piracy channel after December 1 – there are 147 new NBC episodes available after 12/1 that were not available before.

Examining the newly available episodes, we discover that entire seasons of some less popular NBC content — seasons that were not available on Mininova before December 1 but were available on iTunes — become available on Mininova after December 1. These series include, for example, a number of seasons of Saved by the Bell and Xena: Warrior Princess. None of these seasons had new or updated box sets released or new syndication deals during the time period of the study, supporting the inference of causality with NBC's removal from iTunes. In short, it is striking that these entire seasons of older NBC television became available for piracy immediately after the removal of the iTunes channel while no new non-NBC content becomes available. We conjecture that increased demand for these shows through piracy allowed a sufficiently large piracy swarm to exist after the content was no longer available on iTunes.

Another important metric is to simply compare the total number of pirated downloads for all episodes on each network before and after December 1. In rows 7, 8, and 9 of Table 1 we see that while the total number of pirated downloads on non-NBC networks decreased by 4% after December 1, total piracy for all NBC episodes increased by 3%. Together these summary statistics suggest that iTunes purchases may have been displacing some amount of piracy.

However, these simple statistics only tell a partial story as they do not control for important episode-level heterogeneity, nor do they reflect an appropriate model of the actual relationship between digital distribution and piracy. Thus, we turn to regression models for our main empirical analysis.

5. Results

5.1. Impact on NBC Piracy

First, we need to given evidence in support of our identifying assumption, that non-NBC piracy is a good counterfactual for NBC piracy. As evidence, we test whether non-NBC piracy and NBC piracy have similar time trends before NBC content is removed from iTunes. Specifically, we first estimate the following difference-in-difference model:[12]

$$\ln Downloads_{it} = \beta_0 + \beta_{1t}\Phi_t + \beta_{2t}NBC_i * \Phi_t + \mu_i + e \tag{2}$$

where NBC_i is an indicator variable equal to 1 if episode i is broadcast on NBC, Φ_t is a vector of date fixed effects, μ_i is a vector of episode level fixed effects, and $Downloads_{it}$ is the total number of pirated downloads of episode i on day t for the time period November 16 (the first date in our data) through December 24. We use a log transformation on downloads because the relationship between pirated downloads and the presence of a legal digital sales channel is not expected to be linear. Rather, torrents that are already highly downloaded will be more attractive to BitTorrent users because they will generally have faster download times owing to the nature of the BitTorrent protocol in creating "swarms" of users around specific files.[13]

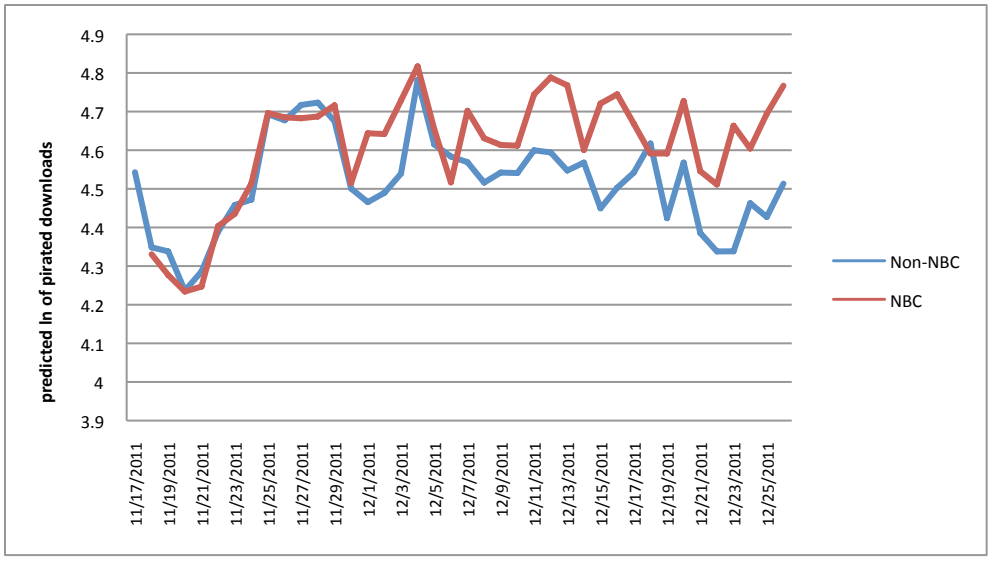
This model serves two purposes for our analysis. First, it allows us to examine whether non-NBC piracy is a good control for NBC piracy in the period prior to the removal of NBC content from iTunes. Second, assuming that non-NBC piracy is an adequate control, it allows us to visually analyze what happens to NBC piracy relative to non-NBC piracy in the period after removal. We

[12] Note that we do not include a main effect in this equation for the NBC dummy. This effect cannot be identified, as it would be subsumed entirely by the episode fixed effects.

[13] We also ran a Box-Cox test on the model to determine the best fitting transformation of downloads – the test produced a $\theta = .04$, indicating quite strongly that the log transformation produces the best fit to our data. However, if we run the linear version for our model, the coefficient of interest is sufficiently similar that none of our interpretations would change.

plot the resulting coefficients in Figure 1 where non-NBC piracy levels are given by $\beta_0 + \beta_{1t}$ and NBC piracy levels are given by $\beta_0 + \beta_{1t} + \beta_{2t}$ for all days in our sample.

Figure 1: NBC vs. Non-NBC Piracy Surrounding December 1, 2007



If non-NBC piracy is is a good control for NBC piracy, then we would expect β_{2t} to be equal to 0 for all dates before the treatment. One can see from the chart above that non-NBC piracy tracks NBC quite well until November 30th, which is the very first date that NBC began removing their episodes from iTunes. More formally, a Wald test of the null hypothesis that all β_{2t} are jointly equal to zero for all t<13 could not be rejected at the 5, 10, or even 20% levels. Conversely, it is quite clear from the graph that NBC piracy increases beyond non-NBC piracy immediately following the removal from iTunes and, for all but a few dates, remains at this increased level for the 25 days following the removal of NBC content from iTunes. The t-statistics for nearly all β_{2t} with t>13 indicate that the difference between NBC piracy and non-NBC piracy is significant at the 5% level suggesting that NBC piracy increased significantly relative to non-NBC piracy after iTunes removal.

While we consider the regression and chart above strong evidence that the removal of the digital distribution channel increased piracy — and that this increase was maintained for at least the time period of our data — it does not conveniently allow us to measure the overall average increase in piracy caused by the event. Given this, next we run the following regression as standard difference-in-difference model to estimate the aggregate effect of the removal of NBC content from iTunes:[14]

$$\boxed{\quad\quad\quad\quad} \tag{3}$$

These models are similar to model (2) above, except that here Φ_t is a single indicator variable equal to 1 if the observation occurs in the two weeks after December 1, 2007 and equal to 0 if it is in the two weeks before. Thus β_2 captures the average effect the event on NBC's piracy relative to control group's piracy. A positive value indicates that NBC piracy has increased by 100* β_2 percent in the period after removal, relative the piracy of non-NBC channels. As before, we include episode level fixed effects in the form of μ_i and daily level fixed effects in the form of phi.

One would worry that downloads of episodes within a season or series may be correlated. For example, an increase in piracy of episode 1 of Heroes season 1 may be correlated with episode 12 of Heroes season 2. We therefore cluster our standard errors at the series level. Therefore, all episodes within our 76 unique series are allowed to be correlated. We also estimate robust standard errors to allow for heterogeneity across series.

[14] Observations occurring on December 1 are removed from the data as NBC was in the process of removing episodes on this date.

The results of model 3 are displayed in columns (i) and (ii) of Table 2. Column (i) reports results for model 3, while column (ii) adds date fixed effects, producing very similar results. The variable of interest in the regression is β_2, as it indicates the percentage change in pirated downloads for NBC over and above any change for non-NBC programs.

The coefficients from the regressions in columns (i) and (ii) show that while non-NBC piracy increased by 5.8% during this time period, NBC piracy increased by an additional 11.2% over and above this level. Thus, the removal of NBC content from iTunes caused an 11.2% increase in piracy over and above the change in the non-NBC "control group." This shows a significant substitution between legitimate digital distribution and piracy channels.

Table 2: OLS Regressions of Log of Daily Downloads
November 16 – December 15, 2007 and 2008

	2007		*2008*	
	(i)	*(ii)*	*(iii)*	*(iv)*
After 12/1	0.058†		0.337**	
	(0.049)		(0.029)	
After 12/1 *	0.112**	0.113**	-0.055	-0.052
Removed	(0.043)	(0.043)	(0.058)	(0.059)
Constant	4.513*	4.314*	4.296*	3.806*
	(0.011)	(0.027)	(0.013)	(0.109)
Date Fixed Effects	No	Yes	No	Yes
Observations	161,784	161,784	170,556	170,556
R^2	0.28	0.16	0.15	0.37

The dependent variable is ln(total pirated downloads before/after September 9. T-statistics are listed in parenthesis; **, *, and † denote significance at 0.01, 0.05, and 0.10 respectively. Regressions include episode level fixed effects with standard errors clustered at the series level.

In columns (iii) and (iv) we run the same models for the same time period in 2008 (November 16 through December 15, 2008), when there was no event or treatment at iTunes. In these regressions, β_2 is negative and insignificant, indicating that the increase in NBC piracy over and above non-NBC piracy is not somehow typical during this time of year — further evidence of a causal relationship between the 2007 increase and the removal of NBC content from iTunes.

The increase in non-NBC piracy observed here could, however, derive from some other outside factor or be a general time trend for all piracy during this period, in which case the difference in difference model is an accurate estimation of the effect on NBC piracy. However, we have not been able to determine any outside factors during this timeframe that might result in an increase in television piracy demand unrelated to NBC's iTunes decision. An alternate explanation, and one that derives from the theory section, is that removing the digital distribution channel could have a spillover effect if the non-financial cost of piracy is largely fixed. Thus the 5.8% increase in non-NBC piracy found in column (i) could be a result of the December 1 NBC treatment and not a general time trend. If this were the case, then our results would understate the displacement of piracy by the iTunes channel, because the change in non-NBC piracy would no longer be an appropriate counterfactual to predict what should have happened to NBC piracy in the absence of the December 1 event. However, while the evidence is consistent with the possibility of a spillover effect, this study cannot identify this effect precisely due to lack of an appropriate counterfactual. In fact, the results above show a significant increase in non-NBC piracy during the same time period in 2008.

To explore the source of the increase in NBC piracy, in Table 3 we break down our results by type of program and run model (3) for four separate genre groups: Drama, Action, Comedy, and Science Fiction. Our results show that the piracy increase for drama programming is only slightly positive and insignificant. The increase for action is about 11% and significant at the 10% level, while comedy and sci-fi have large increases above 20%. We note that these results are consistent with the conventional wisdom that action, comedy, and sci-fi appeal to demographic groups that have disproportionately high rates if piracy (i.e., young males).

Returning to our main results in model (3), we note that there are a few ways to interpret the overall percentage change in piracy resulting from the removal of iTunes content. The first is to calculate the implied average unit increase in piracy per episode as the average number of pirated NBC downloads per episode prior to December 1 (259) multiplied by the estimated increase in piracy in our model (11.2%) to obtain an average increase of 29 pirated downloads per episode attributable to NBC's decision to remove the iTunes distribution channel. Since there were 1,683 NBC episodes available for piracy prior to December 1 and the average episode experienced an increase of 29 pirated downloads, using this method we would conclude that the removal of NBC content from iTunes caused a total increase of about 49,000 pirated downloads per day of NBC content.

Table 3: OLS Regressions of Log of Daily Downloads by Genre
November 18 – December 15, 2007

	Drama	*Action*	*Comedy*	*Sci-Fi*
After 12/1 * NBC	0.011 (0.062)	0.112† (0.060)	0.223* (0.092)	0.203* (0.070)
Constant	5.050** (0.049)	3.704** (0.055)	4.920** -0.029	4.210** (0.076)
Observations	27,378	51,830	40,734	11,738
# Groups (episodes)	1014	1925	1510	435
R^2	0.21	0.24	0.20	0.25

The dependent variable is ln(total pirated daily downloads). T-statistics are listed in parenthesis; **, *, and † denote significance at 0.01, 0.05, and 0.10 respectively. Includes both episode and date fixed effects, with standard errors clustered by series.

Another way to look at this increase is to compare it to iTunes purchases of NBC episodes before the removal of the content. To do this, we were able to obtain data showing that NBC sold slightly over 320,000 episodes of its content through iTunes in the two weeks prior to December

1, 2007, or about 23,000 episodes per day. Thus, our results suggest that the unit increase in piracy was about twice as large as the pre-removal sales on iTunes. This result, while surprising at first, was predicted by the model if the fixed cost of piracy were significant, and we discuss this in further detail in the final section of the paper.

Importantly, our model is measuring the increase in piracy across episodes in percentage terms. One practical question might be whether or not the percentage increase that we've witnessed is being driven disproportionately by the most popular or least popular episodes, as networks are more likely to care about television that has a larger audience.[15] To investigate this, we first categorized all episodes into two groups: top 20% most downloaded episodes prior to the event were classified as being in the "head" while the remaining 80% episodes were classified as being in the "tail". Our classification follows the widely used "80/20" Pareto principle; though a 90/10 or 70/30 split yields similar results. Using this split, we ran a triple difference model, to ask if the increase in NBC piracy, over and above non-NBC piracy, was different for the head (the most popular 20% of episodes) than for the tail (the remaining 80%).

(5)

In model 5, Φ_t is again the dummy variable equal to 1 if the date is after December 1, and Head_i is now a dummy variable indicating if the episode is in top 20% in terms of number of downloads. Thus, β_2 represents the increase in piracy caused by NBC's removal from iTunes for the less popular tail episodes, while $\beta_2 + \beta_4$ indicates the causal increase in piracy for the head — the most popular episodes.[16] The results are reported in Table 4.

[15] We thank a reviewer for pointing this out.

[16] As before, several main effects and pairwise combinations of these effects have been left out of the model as they would be subsumed by the episode and date fixed effects.

These results suggest that NBC's less popular episodes experienced a 9.2% increase in piracy over and above changes in non-NBC piracy. The coefficient on the three-way interaction term is directionally negative, but statistically insignificant, suggesting that on a percentage basis the change in piracy for the most popular episodes was statistically no different than that for the tail.

Table 4: Analysis of Piracy Change for Head and Tail Titles

	DVD Box Sets
After 12/1	0.100* (0.036)
After 12/1 * NBC	0.092** (0.044)
After 12/1 * Head	-0.124* (-0.044)
After 12/1 * Head * NBC	-0.043 (0.065)
Constant	4.513* (-0.010)
Observations	161,784
# Groups (Series)	6,005
R^2	0.35

The dependent variable is ln(iTunes Downloads). T-statistics are listed in parenthesis; ** and * denote significance at 0.01 and 0.05 respectively. Regressions include episode-level fixed effects with standard errors clustered at the series level.

However, ignoring significance for a moment, the point estimates in Table 4 suggests that the causal increase in piracy for the tail was 9.2%, while for the head it was just under 5% (9.2% - 4.3%). The pre-removal average daily pirated downloads of NBC episodes in the tail was 42, so these episodes received 3.9 more downloads per day (42 * 9.2%) as a result of the removal from iTunes. In comparison, these episodes received 5 downloads per day on iTunes on average before NBC's removal. Thus, each iTunes purchase for these tail titles converted into about 0.8 more pirated downloads.

On the other hand, the top 20% of NBC episodes were pirated 1,765 times per day on average before December 1, and thus downloads of these episodes increased by 88 per day (1765 * 5%) as a result of NBC's removal from iTunes. These same episodes were legally downloaded an

average of 52 times per day from iTunes before their removal. Thus, each iTunes purchase for the head titles converted into about 1.7 more pirated downloads.

This finding is interesting for a number of reasons. Our primary argument for a fixed cost to piracy is the fact that the increase in piracy caused by the removal of the iTunes channel was larger than the size of the iTunes market, pre-removal. A possible objection to this is a story of variable cost — in particular, when the iTunes channel is removed, some percentage of iTunes customers turn to piracy. This additional demand becomes supply and makes download speeds faster, reducing the variable cost of downloading and attracting additional new pirates (who may not have even been iTunes customers).

However, we can test whether the increase in piracy is due to a decrease in download speed and thus due to a reduction in variable cost. Less popular episodes are indeed likely to become easier to pirate as a result of additional downloaders from the removed iTunes channel. Because of the nature of the number of connections maintained in BitTorrent "swarms," episodes that were already popular with many downloaders will receive very little, if any, increases in download speed from the addition of more pirates. For example - there is a large difference in download speed between an episode with 20 current downloaders and an episode with 40 current downloders, but there is little or no difference in download speed between an episode with 400 current downloaders and an episode with 420 current downloaders because peers typically only connect to 50 other peers in the BitTorrent protocol. This "diminishing returns from increased swarm size" property is an easily observed fact and is widely known among most pirates.

Thus, if decreasing variable cost is causing the increase in piracy to be larger than the size of the iTunes market, we should only observe this phenomenon for less popular episodes (where

additional downloaders do indeed decrease variable cost) and not for episodes that were already very popular (where additional downloads have little to no effect on download speeds). However, in our previous analysis of head vs. tail piracy, we saw just the opposite. The increase in piracy caused by NBC's removal from iTunes for the most popular "head" shows was larger than the size of the iTunes market for these shows, despite the fact that there is little reason to believe that the variable cost of pirating these shows was changing. And in the tail shows, where we'd most likely expect to see the effects of a decreased variable cost, the increase in piracy was not larger than the size of the pre-removal iTunes market. Because of this, we conclude that a change in variable cost stemming from increased download speeds is unlikely to be responsible for our observations, and we thus maintain that our evidence indicates the presence of a fixed cost to piracy.

Finally, we note that the main results reported throughout this section are robust to a variety of alternative specifications. A linear model also shows an increase in NBC piracy over and above non-NBC piracy. Changes to the time frame, the removal of the Sci-Fi and USA networks from the analysis, and the inclusion of controls for the age of each episode do not materially impact the results. Coding November 30 as the first day of the “after removal from iTunes” period produces even stronger results, likely because NBC started removing content on November 30. Additionally, one might worry that the early announcement of NBC’s removal from iTunes (announced several months before it happened) might have caused people to switch to piracy before the actual removal. If this were the case, piracy would have spiked before December 1 causing our results to underestimate the true effect. Thus, the model appears to produce robust, and if anything somewhat conservative, results.

5.2. Impact on Amazon.com DVD Box Set Sales

To determine the degree to which the iTunes digital distribution channel displaces purchases of DVD box sets sold on the Internet, we employ similar tests to those above. The dependent variable in this case is the Amazon.com sales rank, and thus a decrease in a DVD's rank indicates an increase in sales of that DVD. Table 5 compares means for sales ranks of NBC and non-NBC box sets before and after December 1, 2007.

Table 5: DVD Sales Rank at Amazon.com (Comparison of Means)
November 11 – December 15, 2007

	NBC Networks	*Non-NBC Networks*
Amazon.com sales rank prior to Dec. 1	24,553	35,384
Amazon.com sales rank after Dec. 1	26,056	38.785
Change	*1,503*	*3,401*
	6%	*10%*

We see from this table that the mean rank for non-NBC box sets increased by 10%, meaning that fewer non-NBC television series box sets were sold after December 1 than before. The increase in rank for NBC box sets was only 6%, which could indicate that the removal of NBC content from iTunes caused some additional purchases of DVD box sets. However, as with our analysis for piracy, this comparison of means does not account for changes in price that may occur during this time period (especially with the approaching holidays), nor is a linear model appropriate when predicting sales rank. Thus, we run a similar difference-in-difference model to the one we ran for piracy, specified as follows:

(6)

where $Rank_{it}$ is defined as the Amazon.com sales rank of season box set i on day t. Φ_t is an indicator variable equal to one in the 2-week period after December 1, π_{it} is the log of the price of box set i on day t and μ_i is a vector of fixed effects for each season box set. We log transform

the Amazon sales rank as well as the daily Amazon price consistent with prior literature and based on the explanation provided in the data section.[17] We present results for this regression in Table 6.

Table 6: OLS Regressions of Ln Sales Rank of Season DVD Box Sets
November 18 – December 15, 2007

	DVD Box Sets
Ln(Amazon Price)	**1.727 (0.103)
After 12/1	-0.023 (0.020)
After 12/1 * NBC	0.000 (0.048)
Constant	2.925 (0.338)
Observations	9555
# Groups (Series)	397
R^2	0.29

The dependent variable is ln(Amazon sales rank). T-statistics are listed in parenthesis; ** and * denote significance at 0.01 and 0.05 respectively. Regressions include DVD-level fixed effects. Adding date fixed effects does not materially change the results.

We note that raising price has the predicted effect of decreasing sales, as observed by increased rank. We also observe no significant change in sales rank for NBC box sets relative to non-NBC box sets after December 1. Thus, the removal of the digital sales channel did not seem to increase sales of DVD box sets sold on the Internet.[18] Given the prior marketing literature on channel differentiation (e.g. Viswanathan 2005), this finding could suggest that consumers consider illegal digital downloads a much stronger substitute for legal digital downloads than legal physical purchases are for legal digital downloads. We discuss this result in more detail below.

[17] See Chevalier and Goolsbee (2003), Ghose et al. (2006), Smith and Telang (2008, and 2009).

[18] We analyzed DVD box sets at the level of box set sales per day, because price is a significant predictor of sales and changes by box set and by the day. However, if we were to cluster our standard errors at the series level, it would merely inflate them and thus our results would remain close to zero and insignificant.

5.3. NBC's Return to iTunes

The results presented above represent the best experiment we can find to determine the relationship between piracy, digital distribution, and physical sales. However, another experiment occurred on September 9, 2008 when, after reaching an agreement with Apple, NBC restored all of their content to the iTunes store. Unfortunately, this date also coincides with the new fall season of television in 2008, which presents complications for our empirical analysis. The premiere of a new season of television undoubtedly increases demand for older seasons of the same show, creating two problems in our data. First, there are large day-to-day swings in piracy of older content that correspond to the premieres of new shows. Second, this naturally implies that non-NBC piracy may not be a good control for NBC piracy as the schedule of fall premieres is not constant across networks.

However, it still may be valuable to examine the changes in piracy around this time period to see if our earlier findings are supported. In Section 5.1 we noted that when NBC removed their content from iTunes, NBC pirated downloads increased by more than the number of iTunes purchases previously made, and we interpreted this as evidence of a fixed cost to piracy. If the non-financial cost of piracy is largely fixed, then we would not expect to see as many customers return to iTunes as left when the content becomes available again there. We list some summary statistics surrounding this event in Table 7 below.

A simple comparison of means shows that non-NBC shows averaged 286 downloads per day for the two weeks prior to September 9, and this increased to 307 downloads for the two weeks after. As noted above, this change likely results from increased demand for older seasons generated by the premieres of new seasons. NBC piracy also increased, with a daily mean of 141 downloads per episode before September 9 and 168 daily downloads after. The fact that NBC piracy did not

decrease is consistent with our prior finding of a fixed cost to piracy, but admittedly the increases in piracy due to the new season may be masking other effects.

Table 7: Piracy Data Summary Statistics
August 26 – September 22, 2008

	NBC Networks	*Non-NBC Networks*
1. Mean daily downloads of each episode before 9/9	141	286
2. Mean daily downloads of each episode after 9/9	168	307
3. Change	*27*	*21*

To explore these changes in more detail, we run model 3 again for the 4 weeks surrounding September 9, 2008, reporting these results in Table 8.

Table 8: OLS Regressions of Log of Daily Downloads
August 26 – September 22, 2008

	(i)	*(ii)*
After 12/1	0.063	
	(0.049)	
After 12/1 * NBC	-0.077	-0.077
	(0.096)	(0.096)
Constant	4.306	4.275
	(0.021)	(0.053)
Date Fixed Effects	No	Yes
Observations	129,720	129,720
Number of Groups	4,709	4,709
R^2	0.015	0.074

The dependent variable is ln(total pirated downloads before/after September 9. T-statistics are listed in parenthesis; **, *, and † denote significance at 0.01, 0.05, and 0.10 respectively. Standard errors are clustered at the series level.

The glaring problem with this chart is the size of the standard errors and the lack of explanatory power as indicated by a low R^2. This suggests that our concern about excess noise in piracy around the beginning of a new season of television is justified. However, in spite of this, the data suggest that when NBC content was returned to iTunes, piracy of NBC content was reduced by 7.7% over and above any change for non-NBC. This coefficient is not significant. As well, piracy levels were lower at this time than in December, and if we multiply the 7.7% decrease by

the average number of daily downloads of an NBC episode before September 9, our results imply an average decrease of 11 downloads per day for NBC episodes.

Thus, with the caveat that measuring piracy at the beginning of a new television season does not provide a clean experiment, we note that returning content to iTunes may decrease piracy of that content to some extent, but our results suggest that the reduction in piracy when the content is restored to iTunes is lower than the increase when the content was removed. This result is consistent with our finding of an elasticity/substitution between piracy and digital sales and it is also consistent with the theoretical predictions arising from a fixed cost to piracy.

6. Discussion

Our results represent the first test we are aware of that quantifies the effect of a legal digital distribution channel on both online piracy and online sales of physical products. As such, they offer decision-makers at media firms some much-needed evidence regarding the ability of legal digital distribution channels to compete with illegal piracy channels and physical distribution channels. In this final section, we discuss the results and their implications, as well as the limitations of the work, and how these limitations might be overcome by future research.

In this study, we used NBC's decision to remove its content from the iTunes music store on December 1, 2007 as a quasi experiment and found that the removal of NBC's primary digital sales channel caused an 11.2% increase in piracy of that content over and above any change experienced by competitor networks ABC, CBS, and Fox over the same period.

An 11.2% increase in piracy corresponds to about 29 more downloads per day per episode, or 49,000 additional pirated downloads of all NBC content per day. To put this number in

perspective, it is about twice as large as the number of daily iTunes sales NBC received in the two weeks before December 1. This estimate is conservative due to the fact that Mininova piracy, while arguably a good proxy for overall BitTorrent piracy, represents only a portion of overall BitTorrent television piracy.

Further, this large jump in piracy (larger than the size of the iTunes market) is predicted by theory when there is a significant fixed cost to piracy but only a small (or no) variable cost. In other words, iTunes purchasers may avoid piracy because the fixed cost to learn to use BitTorrent (or the fixed moral/stigma cost of illegal behavior) makes piracy less attractive than iTunes. However, when the digital sales channel is not available, these individuals turn to piracy, pay the fixed cost and, owing to the seemingly low marginal costs of additional downloads, begin to consume much more content through piracy than they had previously purchased. Moreover, this phenomenon seems most prevalent for the most popular episodes of television, which suggests that it is not driven by a decrease in the variable cost of piracy driven by more pirates participating in the BitTorrent swarm.

We note that a fixed cost effect could even have a spillover effect for other networks that have a digital sales channel, since once the fixed cost is paid for NBC it is likely paid for, say, ABC as well. The data is consistent with this possibility — the model shows a 5.8% increase in non-NBC piracy when we might have expected a decrease due to decreasing interest — but this result should be treated cautiously owing to a lack of strong counter-factual evidence for non-NBC piracy. Either way, these results should sound an alarm to content providers, because once the fixed cost of piracy is sunk it may be difficult to get pirates to return to legal options. Indeed, while the return of NBC content to iTunes presented some analytical challenges due to coinciding with the start of the fall television season, we observed a smaller and statistically

insignificant decrease in piracy for NBC content (as compared to non-NBC content) when it was returned to iTunes.

Digital distribution's impact on sales of DVD box sets sold through Internet channels presents a different story, however. When NBC removed their archived seasons of television from iTunes, we found no significant change in the Amazon.com sales rank for NBC's DVD sales relative to the trend that we saw for non-NBC box sets. One possible interpretation of this finding is that digital downloads and DVD's are simply not substitutes in the short term, and thus adding a digital distribution channel would not lead to a short-term displacement in DVD box set sales. A similar interpretation is that there is a fixed cost to digital viewing of television, and once a consumer has "gone digital" she is unlikely to come back.

We note that there are several limitations to our study. First, and most notably, our findings represent a test of short-term elasticity between legitimate digital distribution, pirated digital distribution, and physical distribution channels. The long-term presence of a digital distribution channel likely has a much stronger effect on physical channel sales than those observed here in the short term. However, it is important to note that there is little that media firms can do to forestall the penetration of digital channels given the increased ease, speed, and flexibility associated with obtaining media in digital environments. Rather, our results suggest that media companies would be best served by competing with piracy through digital distribution rather than hoping that the lack of a legitimate digital distribution channel will drive consumers away from the digital channel and back to physical purchases. In addition to this limitation, we also note again that the spillover results mentioned above should be interpreted cautiously owing to the lack of an appropriate counterfactual. Further, we note that our results represent a snapshot in

time for a single media type. It would be useful for future studies to analyze competition between legitimate and illegitimate digital channels in other settings to confirm our results.

Another interesting potential direction for future research relates to the "moral" cost of piracy. In the theory section, we argued that the decision to purchase or pirate rests largely on the shape of the non-financial cost curve associated with piracy. While part of that cost may be related to learning or to the (sometimes) diminished quality of the pirated copies, these costs may approach zero in the future as pirates become more sophisticated in their methods, consumers become more technologically savvy, and client software for piracy becomes even more user-friendly. We suspect that a large part of antipiracy efforts in the future may need to rely on the consumer's "moral" cost associated with piracy. In order to price competitively, digital distribution efforts would benefit from knowing more about the shape of this moral cost. As such, we believe that user studies — whether in the laboratory or in the field — aimed at revealing the nature of the moral cost of piracy for different types of consumers is a fertile area for future research.

References

Arango, T. 2006. Retail-iation. New York Post, September 22, page 31.

Balasubramanian, S. 1998. Mail versus Mall: A Strategic Analysis of Competition between Direct Marketers and Conventional Retailers. *Marketing Science*, 17(3) 181-195.

Bhattacharjee, S., R. Gopal, J. Marsden, R. Telang. 2008. The Effect of Digital Sharing Technologies on Music Markets: A Survival Analysis of Albums on Ranking Charts. *Management Science*, 53(10) 1359-1374.

Biyalogorsky, E., P. Naik. 2003. Clicks and Mortar:The Effect of On-line Activities on Off-line Sales. *Marketing Letters*, 14(1) 21-32.

Brynjolfsson, E., Y. Hu, M. Smith. 2003. Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety. *Management Science*, 49(11) 1580-1596.

Chevalier, J., A. Goolsbee. 2003. Measuring prices and price competition online: Amazon.com and BarnesandNoble.com. *Quantitative Marking and Economics*, 1(2) 203-222.

Chiang, W. Y., D. Chajed, J.D. Hess. 2002. Direct marketing, indirect profits: A strategic analysis of dual-channel supply-chain design. *Management Science*, 49(1) 1–20.

Christin, N., A. Weigend, J. Chuang. 2005. Content Availability, Pollution, and Poisoning in File Sharing Peer-to-Peer Networks. *Proceedings of the 6th ACM conference on Electronic commerce*, pp. 68-77.

Chu, J., P.K. Chintagunta, N. Vilcassim. 2007. Assessing the Economic Value of Distribution Channels – An Application to the PC Industry. *Journal of Marketing Research*, 44(1) 29-41.

Danaher, B., J. Waldfogel. 2008. Reel Piracy: The Effect of Online Movie Piracy on Film Box Office Sales. Working paper. University of Pennsylvania, Philadelphia, Pennsylvania.

Deleersnyder, B., I. Geyskens, K. Gielens, M.G. Dekimpe. 2002. How Cannibalistic is the Internet Channel? A study of the newspaper industry in the United Kingdom and The Netherlands. *International Journal of Research in Marketing* 19(4) 337-348.

DVD News. 2006. Top DVD Retailers: Market Shares. *DVD News Magazine* 10(41), November 30.

Friedman, L. G., Furey, T. R. 1999. The Channel Advantage: Going to Market with Multiple Sales Channels to Reach More Customers, Sell More Products, Make More Profit. Oxford ; Boston Butterworth-Heinemann.

Ghose A., M D Smith, R Telang R. 2006. Internet Exchanges for Used Books: An Empirical Analysis of Product Cannibalization and Welfare Impact. *Information Systems Research*, 17(1), 3-19.

Goodell, J. 2003. Steve Jobs: The Rolling Stone Interview. *Rolling Stone Magazine*. Available from: http://www.rollingstone.com/news/story/5939600/steve_jobs_the_rolling_stone_interview. Last accessed March 15, 2009.

Hui, K., I. Png. 2003. Piracy and the Legitimate Demand for Recorded Music. *Contributions to Economic Analysis and Policy*, 2(1).

Jeuland, A.P., S.M. Shugan. 1983. Managing Channel Profits. *Marketing Science* 2(3) 239–272.

McBride, S., M. Marr. 2006. Target, a Big DVD Seller, Warns Studios Over Download Pricing. *The Wall Street Journal*, Page A1, October 9.

Menn, J. 2006. Target, Disney in DVD Truce. *The Los Angeles Times*, November 18.

Nagin, D.S., R. Paternoster. 1991. On the Relationship of Past To Future Participation in Delinquency. *Criminology*, 29(2) 163-189.

Netherby, J. 2005. Amazon.com Dominates in Online DVD Sales. *Reed Business Information Inc*. Gale Group, Farmington Hills, Michigan, December 19.

Oberholzer, F., K. Strumpf. 2007. The Effect of File Sharing on Record Sales. An Empirical Analysis. *Journal of Political Economy,* 115(1) 1-42.

Peitz, M, P. Waelbroeck. 2004. The Effect of Internet Piracy on Music Sales: Cross-Section Evidence. *Review of Economic Research on Copyright Issues*, 1(2) 71-79.

Rob, R., J. Waldfogel. 2006. Piracy on the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students. *Journal of Law and Economics*, 49(1) 29-62.

Rob, R., J. Waldfogel. 2007. Piracy on the Silver Screen. *Journal of Industrial Economics*, 55(3) 379-393.

Smith, Michael D., Rahul Telang. 2008. Internet Exchanges for Used Digital Goods. Working Paper. Carnegie Mellon University, Pittsburgh, Pennsylvania.

Smith, M., R. Telang. 2009. Competing with Free: The Impact of Movie Broadcasts on DVD Sales and Internet Piracy. *Management Information Systems Quarterly*, **33**(2) 312-338.

Stetler, B. 2008. Serving Up Television Without the TV Set. *The New York Times*, March 10.

Thompson, A. 2003. Tinseltown Follies. The New York Magazine. May 5.

Vernik, D. 2008. Digital Music Set Free: The Flip Side of DRM. Working Paper. Duke University, Durham, North Carolina.

Viswanathan, S. 2005. Competing Across Technology-Differentiated Channels: The Impact of Network Externalities and Switching Costs. *Management Science*, 51(3) 483-496.

Waldfogel, J. 2007. Lost on the web: Does Web Distribution Stimulate or Depress Television Viewing? Working Paper. University of Pennsylvania, Philadelphia, Pennsylvania.

Webb, K. L. 2002. Managing channels of distribution in the age of electronic commerce. *Industrial Marketing Management*, 31(2) 95-102.

Zentner, A. 2005. File Sharing and International Sales of Copyrighted Music: An Empirical Analysis with a Panel of Countries. Topics in Economic Analysis & Policy 5(1), Article 21. Available at: http://www.bepress.com/bejeap/topics/vol5/iss1/art21.