Pirate and Chill: The Effect of Netflix on Illegal Streaming [1]

February 2021

Abstract

Over 188 million people in the United States use a subscription video streaming service, yet digital piracy remains prevalent and costs the U.S. economy an estimated $29.2 billion annually. This paper investigates the relationship between a movie's availability on Netflix, the largest video subscription service, and intent to illegally stream the movie. We leverage a contract dispute that caused Epix (a cable network company) to move all its movies from Netflix to Hulu, representing a substantial decrease in the legal streaming availability of these movies. Using a difference-in-differences design, we find that the removal of Epix movies from Netflix results in a 20% increase in piracy intent relative to movies that remained on Netflix, as measured by Google search volume. This study contributes to the understanding of the substitution between legal streaming services and movie piracy and has implications for content owners deciding what platform to offer their movie on.

Keywords: Piracy, Online Streaming, Digital Goods, Netflix, Google Searches

[1] We are grateful for the helpful comments and suggestions from Bill Even, Jing Li, Greg Madonia, and Robert Leonard and grateful to Matthew Morely for providing the Netflix movie archive data.

1 Introduction

Roughly 188 million US citizens use a subscription streaming service and over half of those individuals stream with Netflix (Aleksander, 2019). Even though Netflix offers over 5,000 titles for on-demand streaming, there are even more titles it does not offer, and the content is constantly in flux. For example, in January 2020, Disney removed all its content from Netflix to further their own new streaming service, Disney+. Consumers could respond by purchasing a Disney + subscription in lieu of or in addition to their Netflix subscription, buying or renting Disney content, or searching to watch the content for free online.

This paper explores the last possibility: How does the removal of a movie from Netflix affect intent to pirate that movie? It is plausible that substitution between a legal subscription streaming service (i.e. Netflix, Apple TV) and an illegal streaming website (i.e. Putlocker, 123 Movies) would be widespread given their resemblance to each other. Indeed, a 2017 survey of 300 individuals ages 18-35 found that of the 98% of participants subscribed to Netflix, 47% also used illegal streaming platforms. Their first most common instinct was to "open Netflix and hope the movie they want to watch is there," and the second most common instinct was to "search for an alternative way to stream that movie," (Dellinger, 2017).

Using a natural experiment that exogenously affected movie availability on Netflix, this paper sheds light on the causal effect of removing movies from Netflix on movie piracy search rates. When Epix, a content producer, moved all their movies from Netflix to Hulu on October 1st, 2015, it represented an overall decrease in availability of streaming this content in the U.S. We measure the effect of reduced movie availability on piracy rates, using the proxy response of Google searches to watch a movie on an illegal streaming website. Our data contains piracy search information on over 1300 movies: 141 treatment movies and 1170 control movies. The panel nature of our data allows us to implement a difference-in-differences design with movie fixed effects. We also conduct a propensity score matching robustness test using individual movie characteristics such as release year and rating obtained from IMDb.

We find that the removal of Epix movies from Netflix results in a 20-22% increase in intent to pirate those movies, compared to movies that remained on Netflix. There are distinct heterogeneous effects by movie release year; older movies experience almost three times the increase in piracy following their removal from Netflix compared to newer movies. This finding is likely driven by the lower number of alternative legal viewing options online for older movies compared to newer movies (Follow, 2018).

To our best knowledge, this study is the first to investigate the effect of decreased availability of paid subscription video-on-demand on piracy streaming, and our findings contribute to literature on the interaction between piracy and paid movie viewing. Most prior studies focus primarily on the effects of digital piracy via illegal downloads or peer-to-peer file sharing on box office revenue or DVD sales (De Vany and Walls, 2007; Ma et al., 2016; Peukert et al., 2017). This study's two major departures from the previous literature are in the type of piracy and legal video consumption we look at: illegal streaming host sites and paid subscription video-on-demand services, respectively. These distinctions are important, as both subscription video-on-demand and piracy streaming continue to grow. Subscription streaming services like Netflix now have more U.S. subscribers than cable and over 80% of digital piracy now occurs through streaming (Blackburn et al., 2019).

Our results also contribute to the literature on factors that impact digital piracy behavior (Morris et al., 2009; Gunter et al., 2010; Nhan et al., 2020). The most similar paper, (de Matos et al., 2018), uses a field experiment to exogenously increase access to legal streaming. The experiment provides free legal streaming content to a sample of households that previously used BitTorrent for piracy. They find that the increase in *free* video-on-demand content only reduces piracy in households for whom the catalog of content aligns with viewing preferences. Our results build upon this intuition, demonstrating that reducing existing legal access to a catalog of films increases piracy intent in a real-world setting, even when the exact catalog of content is available for *paid* subscription video-on-demand elsewhere. These results are relevant to subscription video-on-demand providers purchasing streaming rights, and to content owners deciding how to distribute content. One striking result is that the increase in piracy intent persists for over one year following the shift of Epix films to Hulu. Therefore, it does not appear that the effect decays as one might expect if individuals slowly add new paid Hulu subscriptions as existing Netflix subscriptions expire. The persistent increase in piracy searches following movie removal from Netflix, combined with Herz and Kiljański (2018) finding that one unpaid viewing of a film displaces approximately 0.37 paid viewings, suggest that an effective crackdown on piracy could substantially increase the value of streaming rights to video-on-demand providers.[1]

2 Background

The last decade and a half has seen an immeasurable amount of offline information goods such as books, music, movies, and television become digitized. Digitization reduces the scarcity of these goods, decreasing the costs of production and making the goods easily obtainable (Khan et al., 2015). Consumers can now listen to their playlists on Spotify, download eBooks from Amazon Kindle, and stream movies and TV shows from Netflix.

While legal digitized content is easily accessible, it is not free, and digitization has also made illicit piracy of music, movies, and TV relatively easy and painless. Most college students do not even consider piracy to be morally wrong or hurting anyone like they would if stealing a physical DVD, because in the case of internet piracy, there is no tangible victim (Yu, 2012).

Movie piracy, specifically, does have potential downsides such as higher search costs, lower quality, and computer malware risks (Al-Rafee and Cronan, 2006). Another possible cost of piracy is legal repercussion, but this is not a strong deterrent because retribution against a single individual is rare. Given the drawbacks associated with piracy, most individuals do not use piracy as their main method of casually watching TV shows and movies. Instead, they subscribe to a

[1]This supposition is further bolstered by the Danaher et al. (2020) study which finds that when the UK blocked piracy sites, affected users decreased visits to piracy sites and increased their usage of legal subscription sites by 7-12%. Whereas Danaher et al. (2020) examines the effect of a reduction in piracy websites on legal streaming, we look at the reverse: the effect of a reduction in legal streaming content on piracy streaming intent.

streaming platform such as Netflix and only rely on pirated content when they want to watch something specific that is not available with their subscription.

As noted above, the majority of studies on movie piracy examine the interaction between piracy of digital content and DVD sales or movie box office revenue. This is a well-researched issue because the lack of effective copyright protection may reduce a film company's incentive for making new movies. The findings suggest that piracy affects box-office revenue through two opposing mechanisms: cannibalization and network effects. Ma et al. (2016) find that fully eliminating piracy would increase US box-office revenues by 16% and that any positive promotional aspects of piracy were minimal.

On the other hand, Peukert et al. (2017) find that a 2012 shutdown of Megaupload, an illegal file sharing site, had heterogeneous impacts on box-office revenue based on the type of movie. Movies distributed widely to a large number of theaters benefited from the shutdown to Megaupload, consistent with the narrative that illegal downloads were cannibalizing sales. However, niche movies opening in few theaters experienced a reduction in movie box-office revenue, plausibly due to the reduction in positive network effects of spreading information about movie characteristics to consumers with different valuations for watching the movie.

Moving into the realm of online movie offerings, Danaher and Smith (2013) find that the same Megaupload shutdown increased licensed movie digital sales and rental charges. The option to buy a movie online still differs substantially from a movie streaming service in that purchasing a movie has a marginal cost associated with consumption, whereas a streaming service has a fixed cost for the subscription, but no marginal costs for movie consumption.

There are two closely related papers that study illegal TV downloads in response to (1) removing TV shows from a pay-per-view platform (iTunes) and (2) adding TV shows to a free, legal streaming service (Hulu[2]). Danaher et al. (2010) found that NBC's removal of its content from iTunes in 2007 caused a 11.4% increase in the demand for NBC's pirated television content compared to content from other networks which was not removed from iTunes. Danaher et al. (2015) showed that

[2]At the time of the study, Hulu did not have a subscription fee and gained revenue solely from advertisements.

ABC's decision to add its content to Hulu in 2009 decreased piracy of that content by 15-20% compared to content from other networks which was not added to Hulu.

The main difference between the above studies and this one is the type of piracy that is measured. One can pirate a movie online in three different ways: Peer-to-peer (P2P) file sharing, direct downloads, and streaming from an illegal website.[3] Both Danaher et al. papers measure piracy in terms of illegal downloads from BitTorrent, a P2P file sharing website, whereas we measure piracy in terms of piracy search rates which captures the most common method of piracy, streaming.

A likely reason streaming is the most common method of piracy is due to its technical legality. P2P file sharing and direct downloading are illegal because both methods involve either a permanent or public reproduction of copyrighted content, and thus could be prosecuted as a criminal or civil offense. In contrast, while streaming sites are engaging in illegal behavior by linking the public to copyrighted content, the individual utilizing the sites is only temporarily viewing the content and is not at risk of legal retribution (Supan, 2019).

Given the technical legality of streaming from an illegal site, it is not surprising that piracy is trending away from downloads and P2P file sharing, and towards streaming. According to a study from the piracy tracking company, Muso, illegal streaming websites made up 74% of 78.5 billion visits to access pirated film and TV content in 2015. In comparison, torrent-based sites (P2P file sharing) represented 17% of total visits to piracy sites, and direct-download sites made up the remaining 9%. Visits to torrent sites declined 19% in the second half of 2015 compared to the first six months of the year, and by 2019 the percentage of piracy attributed to streaming increased to 80% (Blackburn et al., 2019).

3 Data

The rising popularity of streaming as a form of piracy motivates using Google searches for "watch *movie title* free online" as the proxy variable to measure piracy. The top Google results for this search are links to well-known illicit streaming

[3]These well known sites avoid getting shut down by acting as a link to the pirated movie content which is stored on cyber-blocking websites (Ibosiola et al., 2018).

websites for the movie specified in the search. We measure piracy searches in the United States for “watch *movie title* free online” per month for each movie using Google Ads Keyword Search Planner. This Google Ads Keyword tool provides the absolute number of searches rounded to the nearest tens. Though this variable accounts for individuals intending to pirate a movie by streaming it from an illegal website, it is important to note that an individual may not actually pirate a movie after searching to do so.

There is a trade-off in the specificity of the keyword search to pirate a movie. Choosing a less specific keyword would result in a higher number of search results, but the keyword needs to be distinct enough to ensure the intent to pirate is captured only for the specific movie of interest. Thus, the keyword search is limited to the movie title and “watch free online”.[4] Movies that do not have a unique name, such as "Robin Hood" and "Phantom of the Opera" are removed from the sample, because over time newer versions of these movies have been released and the keyword does not discriminate between someone’s intention to pirate an older versus newer version of a movie.

To determine the list of movies removed on October 1st, 2015 we obtain data from *USA.Newonnetflix.info*, which archives the movies removed from Netflix each month. The list of movies removed in October 2015 are cross-referenced with movies owned by Epix, and only Epix movies are kept in the treatment group. Netflix movies that were also available on Hulu during this time are removed to ensure the streaming availability statuses for both the treatment and control group are the same before the event (i.e. both are available only on Netflix) but differ after the event (i.e. the control group remains available only on Netflix and the treatment group becomes available only on Hulu).

While the ideal situation would be to have piracy data for an extended period before and after the Epix switch occurred, Google ads only provides search data for the last 5 years. Consequently, we collect piracy search rates from July 2015 to December 2016 - three months prior to the switch and 15 months after the switch.

[4]Google Ads Keyword Search Planner does not allow special characters (e.g. ñ, !) or searches over ten words, so movie titles with these characteristics were either edited where obvious (such as removing an exclamation point) or dropped from the sample when no there was no obvious fix.

Table 1 displays the number of movies in observations throughout the study. There are 1,311 movies in the total sample: 141 in the treatment group and 1,170 in the control group.[5]

Table 1: Number of movies in sample for each month for the treatment and control group

	2015	2016											
	Jul-Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Treatment	141	141	141	141	141	141	141	141	141	141	140	140	140
Control	1170	1054	1006	949	901	859	714	691	666	623	578	546	501

Notes: In order for a movie to be in the control group it must have been accessible on Netflix from Jul-Dec 2015, after which content may fluctuate.

All control group movies remained on Netflix at least 3 months post October 1st, 2015, after which some control group movies are removed each month due to natural content changes. In the main results, we utilize control group movie piracy search data for months in which the movie remained on Netflix. Nonetheless, results are robust to using the 501 control movies that were not removed from Netflix during the entirety of the study time frame (see Appendix: Additional Robustness Check).

Additional movie data such as genre, number of reviews, telease Year, (PG-13, R, etc.), IMDb score, and box office revenue is collected from IMDb in order to conduct a propensity score matching analysis and to investigate heterogeneity by movie type. The summary statistics (Table 2) provide relative magnitudes of piracy searches and movie attributes. Of note, there is large variation in piracy searches between movies; less popular movies may vary between 0 and 10 piracy searches each month whereas popular movies may vary by thousands of piracy searches across months. Due to the large variation in movie searches and the frequency of "0" search observations, the main regression dependent variable is ln(Piracy Searches + 1).[6]

[5]The treatment group sample size decreased from 141 to 140 in September 2016 because during that month a new movie was released that had the same name as a movie in the treatment group, so

Table 2: Movie summary statistics by treatment

Control	Piracy Searches	Release Year	Score	Number Reviews	Gross USA ($Mil)
Mean	143	2006	6.36	47333	24.27
Median	30	2011	6.40	10839	2.85
SD	384	10.58	1.10	127194	43.85
Treatment	Piracy Searches	Release Year	Score	Number Reviews	Gross USA ($Mil)
Mean	315	2002	6.22	74528	42.44
Median	90	2005	6.40	27273	19.06
SD	854	14.59	1.02	128629	61.11

Notes: Movie characteristics are observed at the individual movie level and piracy searches are observed monthly for each movie. The "Piracy Searches" metrics are based on overall mean, median and standard deviation. Overall standard deviation is measuring $search_{it}$ variance around the overall mean.[7]

4 Empirical Strategy

A naïve approach to quantifying the causal relationship between Netflix movie offerings and piracy searches would be to compare piracy searches for a movie before and after it is removed from Netflix. This approach is problematic because an individual movie's availability on Netflix is usually influenced by endogenous factors such as movie popularity, contract costs, date of release, etc. For instance, one might find piracy searches are lower for movies that were removed from Netflix due to declining popularity compared to movies that were removed from Netflix randomly. This issue is addressed using quasi-experimental variation in Netflix movie availability which we will call the "Epix shock". On October 1st, 2015, there was a sudden cancellation of the movie contract between Epix and Netflix which

the keyword search no longer uniquely identified the movie.

[6]Estimates for our preferred specification (difference-in-differences using movie and month-year fixed effects) are similar when using negative binomial regression with searches as the dependent variable.

[7]For control group movie searches, the "between" standard deviation ($\overline{search_i}$) is 353 and the "within" standard deviation ($search_{it} - \overline{search_i} + \overline{\overline{search}}$) is 170. Treatment movie search between standard deviation is 779 and within standard deviation is 355.

resulted in all Epix movies being removed from Netflix and added to Hulu.[8]

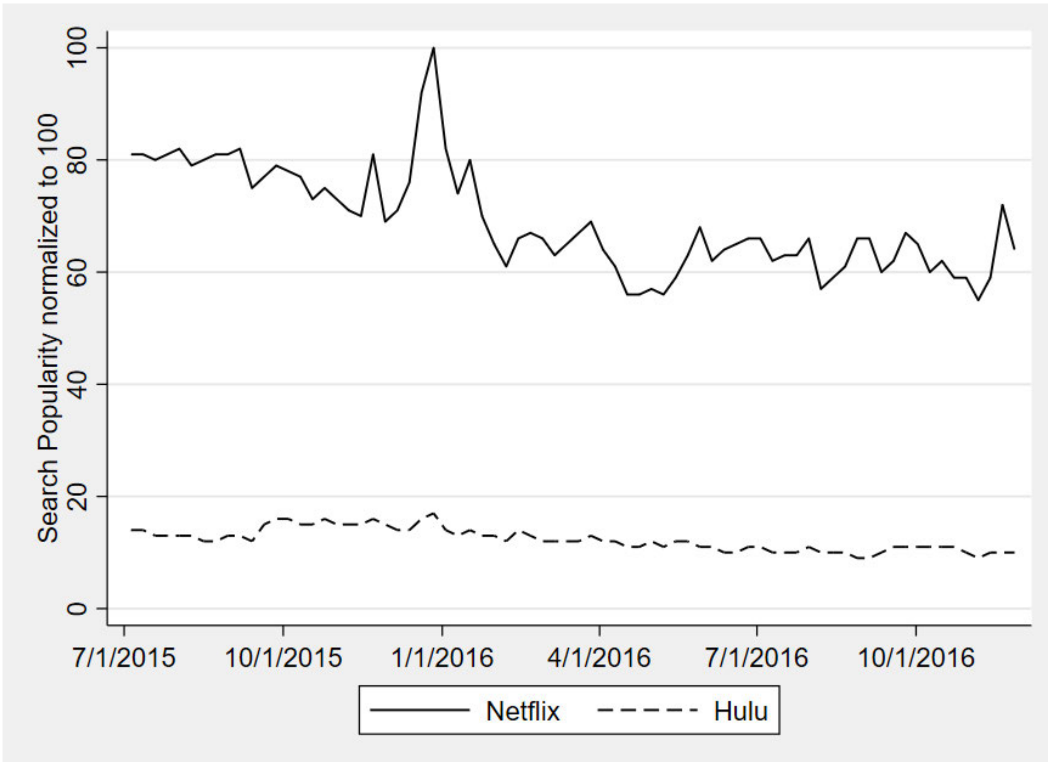
4.1 Event Description

Epix is an entertainment cable network that features movies and TV shows distributed by Paramount, Lionsgate, and Metro-Goldwyn-Mayer, and its movie content varies from large blockbusters such as *The Wolf of Wall Street* to smaller Indie films. Epix and Netflix upheld an exclusive licensing agreement from 2010 until the end of September 2015 when Netflix announced its decision not to renew the licensing contract with Epix, citing the company's plan to shift towards hosting its own original content (Netflix, 2015). In response to this, Epix entered into a multi-year agreement with Hulu (Hulu, 2015). Thus, all titles owned by Epix were removed from Netflix on October 1st, 2015 and appeared on Hulu for streaming that same day.

At the time of the switch, Netflix had roughly 4 times as many subscribers as Hulu, and Google trends for the number of searches for "Netflix" and "Hulu" between July 2015 and December 2016 (the time frame for this study) shows searches for "Netflix" were, on average, 5.6 times higher than searches for Hulu (Figure 1). Consequently, the Epix switch represents a substantial decrease in the availability of these movies for streaming in the United States.

[8] Yu et al. (2020) use this Epix shock to examine the relationship between streaming availability and physical DVD sales. They find the decreased streaming availability of Epix movies resulted in additional annual physical sales of $48,155 per movie title.

Figure 1: Trends in Google searches (normalized to 100) of "Netflix" and "Hulu" from July 2015 to December 2016



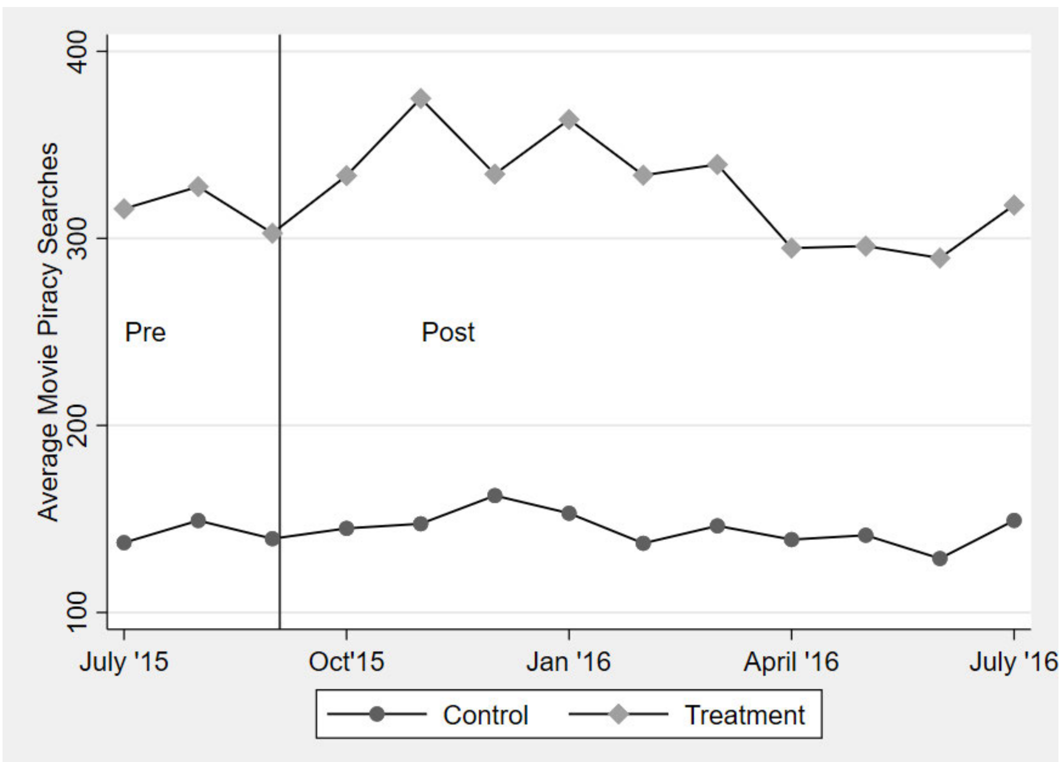
Notes: Numbers represent search interest relative to the highest point on the chart for the given region and time. (i.e. A value of 100 is the peak popularity for the term. A value of 50 means that the term is half as popular. The spike at 1/1/2016 is due to a change in Google's data collection system.)

4.2 Difference-in-Differences

The removal of Epix movies from Netflix provides an unusually attractive setting to measure the impact of reduced subscription streaming access on intent to illegally pirate a movie. First, there is a clear comparison group of content that remained on Netflix both prior to and after the removal of Epix films. Second, the removal of Epix content transpired primarily because the contract expiration coincided with Netflix's strategic shift to original content, rather than due to characteristics of individual movies. Third, the event occurred suddenly for a large share of content. On October 1st 2015, all Epix films that represented over 10% of the Netflix's movie portfolio were removed at the same time—providing adequate statistical power.

Given the empirical setting, we causally estimate the impact of reduced streaming availability on piracy using a difference-in-differences approach that compares the difference in piracy search rates between treatment and control groups after October 1, 2015 to the initial difference (prior to October 1, 2015). The key identification assumption is that Epix films would have continued on the same trajectory experienced by non-Epix Netflix films if they had not been removed from Netflix. We will provide both visual and statistical evidence to bolster the plausibility of this parallel trends assumption.

Figure 2: Average Piracy Searches Pre and Post Shock



Notes: Average piracy searches plotted from July 2015 to July 2016 for the treatment and control movies. Epix removal shock is indicated by vertical line.

To show sharpness of event effect and parallel trends, we graph average piracy searches before and after the Epix shock for the control and treatment movies. Figure 2 shows that treatment and control movies follow a similar trend in average piracy searches prior to the removal of treatment movies (as indicated by the vertical line). This lends plausibility to the parallel trends assumption that had Epix movies not been removed from Netflix, piracy search rates for Epix and non-Epix Netflix films would have continued on similar trajectories. After the shock, piracy searches for treatment movies increased sharply, whereas piracy searches for control group movies remains relatively unchanged.

Table 3 quantifies the difference in average search rates in the 3 months pre- and post-Epix shock for the treatment and control groups. In the post period, the average piracy searches for Epix movies increased by approximately 22 searches more than those observed in the control group.

Table 3: Difference in average piracy searches for three months prior and three months post Epix shock for the Treatment and Control group

	Average Piracy Searches		
	Pre	Post	Difference
Treatment	315.41	347.59	32.17
Control	141.97	151.63	9.66

Notes: The Pre and Post periods represent 3 months prior and 3 months post October 1st, 2015.

5 Results

To estimate the effect of removing a movie from Netflix on movie piracy search rates, we use the standard difference-in-differences model seen below:

$$ln(piracysearch_{it}+1) = \beta_0 + \beta_1 After_t + \beta_2 Treatment_i + \beta_3 After_t * Treatment_i + \alpha_i + \gamma_t + \varepsilon_{it}$$

with individual movie fixed effects, α_i, and month-by-year fixed effects, γ_t. Movie fixed effects are included because there is large variation between different movies' popularity and subsequent piracy searches, but an individual movie's popularity is relatively time invariant. Month-year fixed effects are included to account for changes in demand for movie streaming across time. Standard errors are clustered at the movie level to account for the fact that there are multiple observations for each film. Table 4 displays the value and inference of the main DID interaction term by regression specification with and without fixed effects.

Table 4: DID estimates of the impact of movie removal on piracy with fixed effects

	(1)	(2)	(3)
$Treatment_i$ x $After_t$	0.206*	0.160**	0.194***
	(0.10)	(0.05)	(0.05)
$After_t$	-0.144***	-0.098***	
	(0.03)	(0.02)	
$Treatment_i$	0.790***		
	(0.09)		
Movie fixed effects	No	Yes	Yes
Month-by-year fixed effects	No	No	Yes

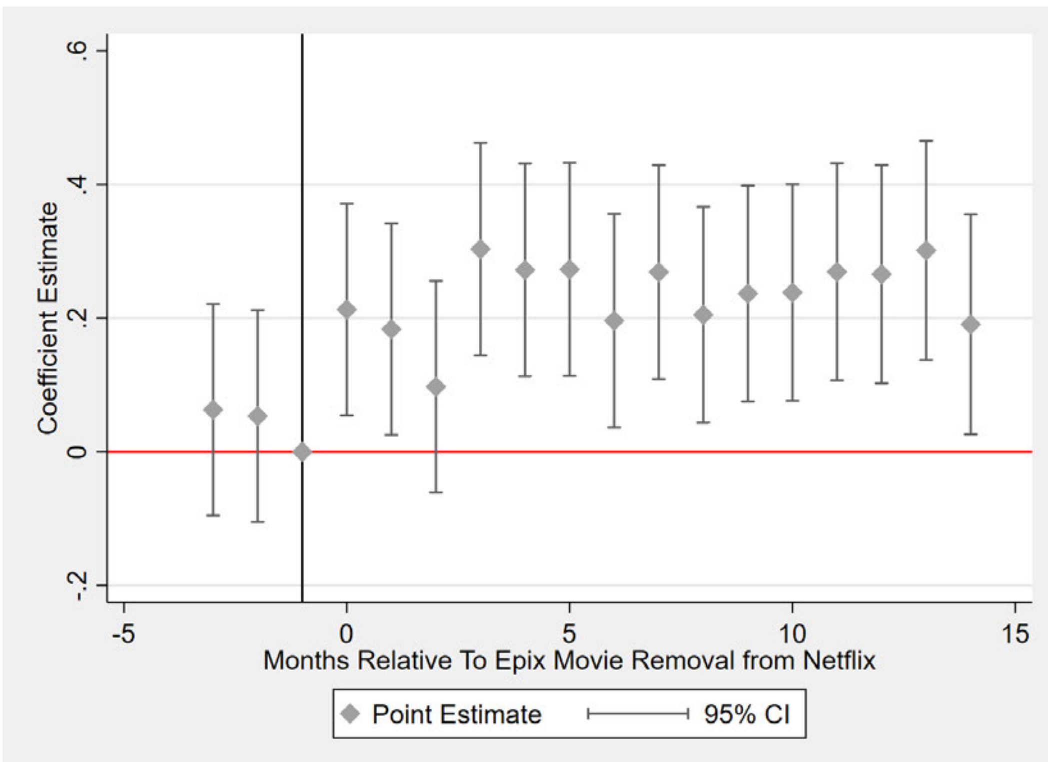
Notes: (1) The dependent variable is ln($piracysearch_{it}$+1). (2) This regression uses data from July 2015 - December 2016. (3) Standard errors clustered at the movie level are in parentheses. (4) *** p<0.001, ** p<0.05, * p<0.1

The interaction term in specification 1 (without movie or month-by-year fixed effects) indicates that removing a movie from Netflix increases average piracy searches for the movie by approximately 23% ($e^{0.206} = 1.228$). However, because this model does not account for time-invariant movie heterogeneity or movie watching seasonality, we prefer specification 3 which includes month-by-year and movie fixed effects. The interaction term in specification 3 indicates that removing a movie from Netflix increases average piracy searches for the movie by approximately 20%, and is statistically significant at the 0.01 alpha level.

To gain a deeper understanding of how treatment effects evolve over time, we present an event study in Figure 3. The reference period is set as the month directly prior to treatment – September 2015. Each dot represents the coefficient estimate on an interaction between the treatment group and a specific lag or lead. For instance, the coefficient of 0.21 for month 0 (October 2016) means that from September to October piracy searches increased by 21% more for Epix films relative to the change experienced by films remaining on Netflix.

The event study figure provides two key insights. First, it provides direct evidence supporting the assumption of parallel pre-trends. Prior to the removal in Epix content, there is no evidence of different trends in piracy searches between treatment and control films. Second, Figure 3 provides striking visual evidence that the increase in piracy searches is both immediate and persistent. Notably, the removal of Epix films from Netflix still causes an increase in piracy searches for Epix films of 19%, even 15 months after the removal.

Figure 3: Event Study - Piracy Searches



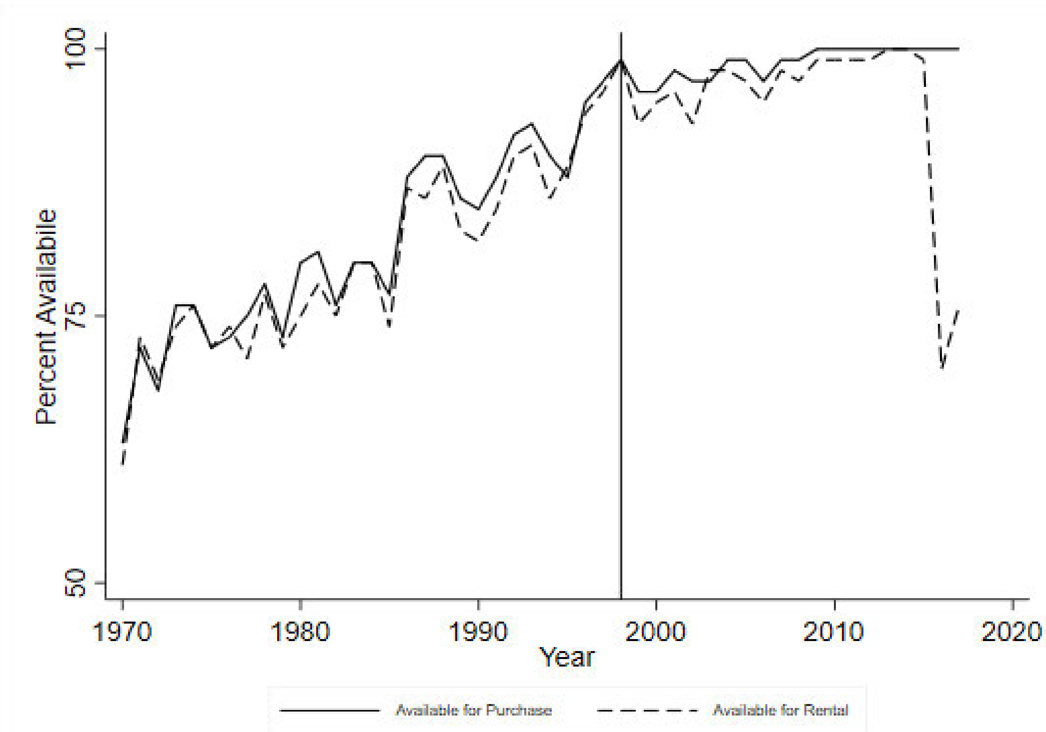
Notes: Each dot represents the coefficient estimate of an interaction between a treatment indicator and month relative to treatment. September 2016, the last pre-treatment month, is the omitted category. Regression includes movie and month-by-year fixed effects.

To further probe the robustness of the core results, we conduct tests to (1) show the statistically significant difference in piracy searches between the pre- and post-periods for the treatment and control groups and (2) confirm that there is no statistically significant difference in piracy searches during the months prior to the Epix shock (see Appendix Table 2A). As expected, there is no statistically significant difference in average piracy search when the pre-period is July 1st - 31st, 2015 and the post-period is August 1st – September 31st, 2015.

5.1 Heterogeneity by Movie Characteristics

Next, we look at the heterogeneous effects of movie removal from Netflix on movie piracy search rates by movie release year. This movie characteristic is relevant in the context of piracy because movie age is correlated with the availability of legal movie viewing alternatives—an important substitute which effects movie piracy demand.

Figure 4: Online Purchase or Rental Availability for Top US films by Year



Notes: Source is Stephen Follows: Film Data and Education. Vertical axis ranges from 50 to 100% and horizontal axis ranges from 1970 to 2017. Vertical line at 1998.

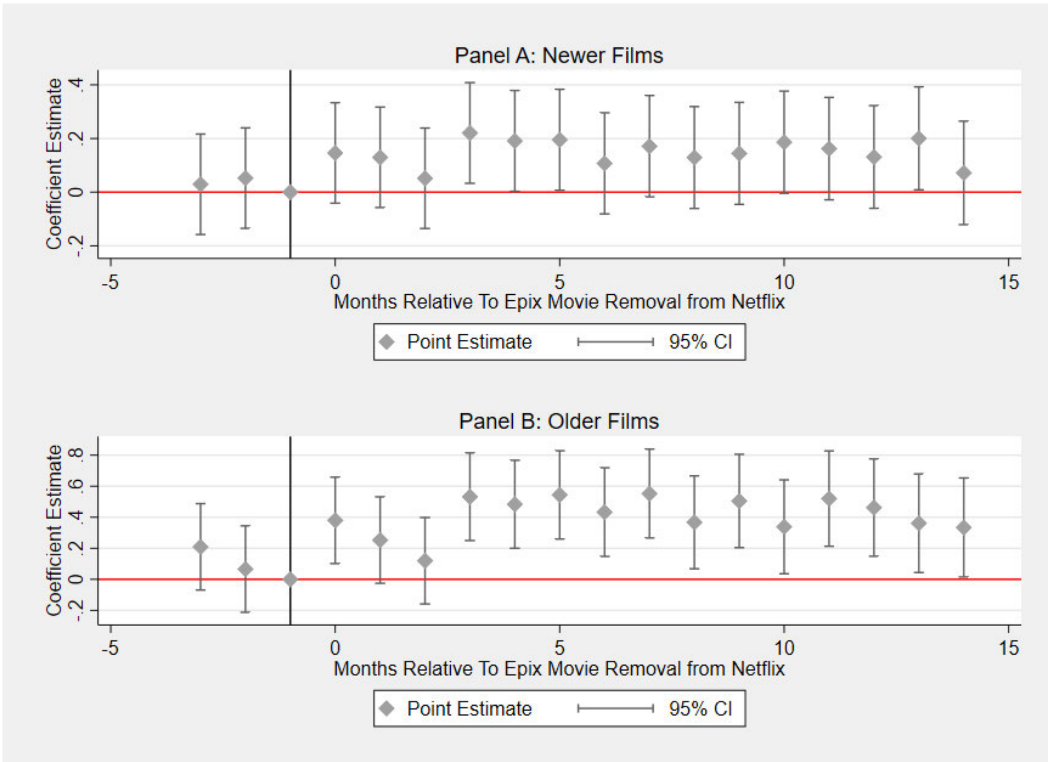
Stephen Follows' 2018 data of online legal streaming, rental, and purchase availability for the top 100 grossing movies in each year from 1970 to 2017 shows that there is a drop off in the availability to purchase or rent movies that were released prior to late 1990s[9] (Figure 4). Additionally, the number of top-grossing movies not available on any platform for streaming, rental or purchase increases as movie release year decreases for movies released prior to 1998.[10]

Given this finding, we split our data into "new" movies that were released after 1998 (panel A) and "old" movies that were released prior to or during 1998 (panel B). There are 37 treated movies in panel A and 104 treated movies in panel B (Figure 5). Theory would predict that if there are fewer legal viewing alternatives for older movies compared to newer movies, the effect of movie removal from Netflix on piracy searches would be higher for older movies. Indeed, older movies have an approximately 20 percentage point larger effect of movie removal on piracy search rates compared to the full sample, whereas recent movies experience a modest 12% increase in piracy searches due to movie removal.

[9]The mean percentage of online purchase availability for movies released before or during 1998 is 82% whereas, on average 99% of movies released after 1998 are available for online purchase.

[10]Only 117 of the movies in our data set are also in Follow's data set which, unfortunately, does not leave us with enough statistical power to use Follow's movie availability data in our study.

Figure 5: Event Study - Piracy Searches for Films Released Before Vs. After 1998



Notes: Each dot represents the coefficient estimate of an interaction between a treatment indicator and month relative to treatment. September 2016, the last pre-treatment month, is the omitted category. Regression includes movie and month-by-year fixed effects.

We also look at heterogeneous effects by (1) the two most common types of movie genre, comedy and drama and (2) movie popularity as determined by the median movie score. Drama movies and movies with a high score tend to have higher increases in piracy rates after being removed from Netflix than their counterparts, comedy and low-scored movies. These differences are not as striking and, in the case of the genre split, not statistically significant at the 0.01 alpha level due to the high variation in piracy searches and the smaller sample size. (See Appendix Table 3A and 4A for these results.)

5.2 Propensity Score Matching

The DID model does not require Epix movie and control group movie attributes to be the same because it identifies the difference in changes in trend over time. However, one might be concerned that an attribute of Epix movies that differs from control movies might also affect the change in trend in Epix movie piracy searches in a way that does not affect the control group trend. We address this concern by replicating the main analysis using the subset of potential control movies that most closely resembles Epix movie attributes based on propensity score matching.

Using IMDb movie information, a propensity score (hereafter p-score) is calculated for each movie by estimating a probit model with treatment as the outcome and release year, review number, IMDb score, rating, and genre as explanatory variables (Dehejia and Wahba, 2002). The control group sub-sample is identified by matching the p-scores of the Epix movies with the closest p-scores of the control group movies. Matching is done without replacement and with a caliper of 0.01 (i.e. matched movies must have a p-score difference less than 0.01). After matching there are 139 movies in the treatment group and 139 movies in the control group.

The probit regression results in Table 5 displays the statistical significance of movie attributes on the probability a movie is in the treatment group before and after matching. Prior to matching, on average, Epix movies had older release years and more mature ratings. The fact that Epix movies proportionally have more R-rated movies compared to the control group might be problematic if it causes piracy search rates to increase more in October and November compared to the control group's piracy searches. A possible cause is that Halloween movies that people tend to watch in late October contain more gore or violence which would give it an R rating. Such an increase in piracy might be erroneously attributed to the removal of the movies from Netflix because both occurred in the same time period. After matching, all the movie attribute coefficient estimates are statistically insignificant at the 1% level. This shows that Epix movies and control group movies are similar in all the attributes included in the p-score regression, and if there is a spike in demand for certain type of movie, it will affect Epix and control group movies similarly.

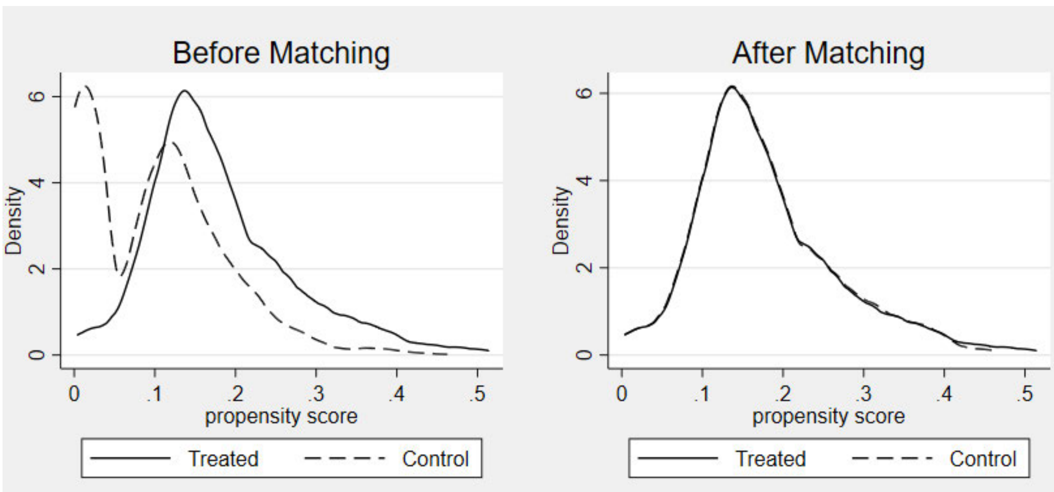
Table 5: Results of probit regression before and after matching

	Treatment Dummy			
VARIABLES	Before Matching		After Matching	
	Coefficient	Std. Error	Coefficient	Std. Error
Number Reviews	0.000	0.00	0.000	0.00
IMDb Score	-0.081	0.06	-0.019	0.09
Release Year	-0.018***	0.00	0.000	0.01
Action	-0.446	0.52	0.205	0.61
Adventure	-0.273	0.55	0.307	0.68
Animation	-0.594	0.61	0.456	0.84
Biography	-0.368	0.56	0.008	0.67
Comedy	-0.4	0.52	0.23	0.6
Crime	-0.461	0.55	0.538	0.7
Documentary	-0.777	0.57	0.377	0.79
Horror	0.193	0.53	0.374	0.63
Drama	-0.448	0.52	0.398	0.62
Family	0.277	0.87	0.061	1.1
Musical	.	.	.	.
Romance	.	.	.	.
Fantasy	.	.	.	.
R	0.977***	0.19	0.172	0.38
PG-13	1.185***	0.19	0.097	0.39
PG	0.937***	0.23	0.375	0.42
G	-0.371	0.47	.	.
Not Rated	.	.	.	.
Constant	34.984***	9.07	-0.495	13.26

Notes: (1) Before Matching includes the full sample and After Matching includes the reduced sample of 139 treatment movies and 139 control movies, (2) The omitted groups are represented with a dash, and in the case of "G" rating differs after matching due to the diminished number of observations, (3) *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Common support graphs (Figure 5) show that even before matching there is sufficient overlap in the distribution of the treatment and control group p-scores to support the full sample analysis in the main results sections and lend credence to the shock exogeneity assumption. In other words, Epix movies were not removed based off of individual attributes because even before matching, the p-scores for both groups are similar. After matching, the p-scores of the treatment and control group follow a very similar distribution. This implies that control movies in the matched subsample are equally likely to be an Epix movie as the Epix movies themselves based off of movie characteristics.

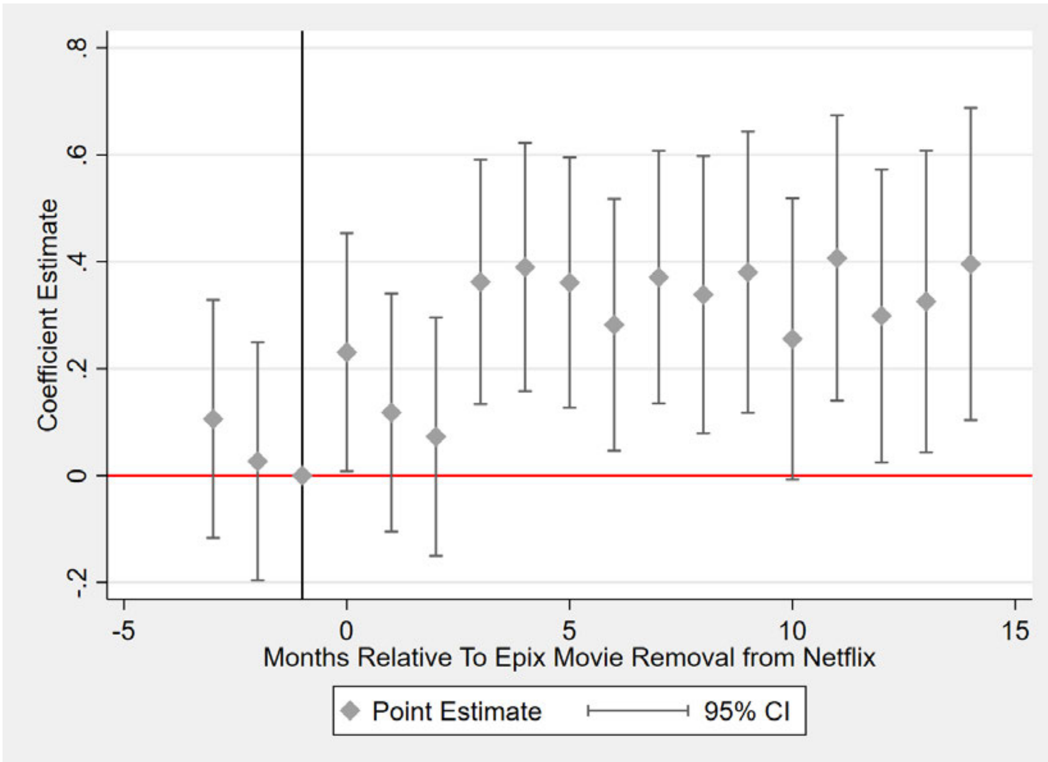
Figure 5: P-score density before and after matching



Notes: There are 139 movies each in the matched treatment and control groups.

Figure 6 presents the event study results for the propensity matched sample of 278 movies. These results are strikingly similar to those of the full sample, providing additional evidence that changes in piracy rates were not driven by differences in film characteristics that were correlated with the timing of the Epix movie removal. While estimates are less precise due to the diminished sample size, removal from Netflix leads to an immediate and persistent increase in piracy searches relative to the narrower matched comparison group.

Figure 6: Event Study - Piracy Searches for Propensity Matched Sub-sample



Notes: There are 139 movies each in the matched treatment and control groups. Each dot represents the coefficient estimate of an interaction between a treatment indicator and month relative to treatment. September 2016, the last pre-treatment month, is the omitted category. Regression includes movie and month-by-year fixed effects.

6 Discussion

Some studies suggest that people rarely pirate when there is a legal and cheap alternative (Jacobs, 2012) and that piracy would not be as prevalent if consumers were provided with a high-quality, paid streaming service (Greenberg, 2015). However, this study shows that some consumers who can afford Netflix will still pirate a movie if what they want to watch is not available on Netflix. Past research has also indicated that the main motivation for piracy is to "see rare and new movies" (Jacobs, 2012). Our results contradict this and indicate there is a large increase in piracy searches after a movie is removed from Netflix if the movie is older, but no significant effect if the movie is newer, likely driven by the lack of alternative online viewing options for older movies.

Knowledge of the degree of substitution between legal and illegal streaming platforms is essential for a content owner who is comparing the profitability of making a movie available for online purchase in lieu of streaming on Netflix. If a content owner ended her contract with a subscription streaming service and switched to making the movie available for online purchase or rental, average earning per movie view would increase from $0.41 per movie view to $5.54 per movie view (Blackburn et al., 2019). Of course, the marginal cost for consumers would no longer be zero, and thus the number of people viewing the movie would also decrease. If the content owner knew the price sensitivity of movie watchers, she could do a cost benefit analysis of keeping the movie on Netflix. The results would be incorrect, however, if she did not take into account the significant number of consumers with a low willingness to pay who would switch from watching the movie on Netflix to pirating the movie if the movie is taken off of Netflix.

Future research may benefit from studying how the effect of removing content from a streaming platform on piracy differs based on piracy method and content type. Previous studies (Danaher et al. 2010, 2015) have found a weaker effect of the removal of a TV show from an online platform on BitTorrent downloads compared to this study. This piracy method, unlike streaming, comes with an additional risk of legal punishment, and thus the difference in our results could be indicative that concern of punitive action from copyright enforcement is a somewhat effective deterrent for some individuals intending to illegally download content. We cannot, however, definitively say that the difference in piracy type is driving the difference in our results because the type of content (movies versus TV shows) also differed.

Our results indicate that the removal of a movie from Netflix causes the US piracy searches for the movie to increase by roughly 20% compared to if the movie remained on Netflix, but how would the magnitude of substitution differ if instead the availability of content on illegal streaming sites was reduced? Danaher et al. (2020) exploits the blocking of 53 piracy sites in the UK in 2014 to measure the reduction in availability of illegal streaming sites on legal viewing, and while the US has not had any similar court mandated blocks of piracy streaming sites, the legal situation surrounding piracy in the US remains in flux. The Covid-19 relief bill passed on December 21st, 2020 included new legislation that increased punitive measures for individuals hosting (and profiting from) illegal streaming sites. The crime class was increased from a misdemeanor to a felony and persons convicted could now be subject to 10 years imprisonment. Along with the Covid-19 lockdown, there as been a surge in new streaming service subscriptions (Leatherby and Gelles, 2020) and an increase in the amount of time people spend watching content at home (Dixit et al., 2020). This new legislation may significantly reduce piracy streaming options in the US, and presents an excellent opportunity for future research studying the substitution between legal and illegal streaming.

We are in an era of battling of streaming services. With so many existing players: Netflix, Disney+, Hulu, Amazon Prime Video, and ESPN+ the market is becoming over-saturated, and most consumers are not willing to pay for subscriptions to all these platforms. As platforms work toward providing original content not accessible on other platforms, it is essential to take into account the ever-present alternative method of viewing content—piracy.

References

S. Al-Rafee and T. P. Cronan. Digital piracy: Factors that influence attitude toward behavior. *Journal of Business Ethics*, 63:237–259, 2006.

S. Aleksander. The most comprehensive netflix statistics of 2019. *Techjury*, 2019.

D. Blackburn, J. A. Eisenach, and D. Harrison. Digital piracy: Factors that influence attitude toward behavior. 2019.

B. Danaher and M. Smith. Gone in 60 seconds: The impact of the megaupload shutdown on movie sales. *International Journal of Industrial Organization*, 33, 03 2013.

B. Danaher, S. Dhanasobhon, M. D. Smith, and R. Telan. Converting pirates without cannibalizing purchasers: The impact of digital distribution on physical sales and internet piracy. *Marketing Science*, 29:1138–1151, 2010.

B. Danaher, S. Dhanasobhon, M. D. Smith, and R. Telan. Understanding media markets in the digital age: Economics and methodology. In *Economic Analysis of the Digital Economy*, pages 385 – 406, 2015.

B. Danaher, J. Hersh, M. D. Smith, and R. Telang. The effect of piracy website blocking on consumer behavior. *MIS Quarterly*, 44(2):631–659, 2020.

M. G. de Matos, P. Ferreira, M. D. Smith, et al. The effect of subscription video-on-demand on piracy: evidence from a household-level randomized experiment. *Management Science*, 64(12):5610–5630, 2018.

A. S. De Vany and W. D. Walls. Estimating the effects of movie piracy on box-office revenue. *Review of Industrial Organization*, 30(4):291–301, 2007.

R. Dehejia and S. Wahba. Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics*, 84: 151–161, 2002.

A. Dellinger. Illegal streaming: More than half of millennials are still watching content illegally. 2017.

A. Dixit, M. Marthoenis, Y. Arafat, P. Sharma, and S. K. Kare. Binge watching behavior during covid 19 pandemic: A cross-sectional, cross-national online survey. *Psychiatry Research*, 289, 2020.

S. Follow. What percentage of movies are available to stream, rent or buy online?, 2018. URL `https://stephenfollows.com/how-many-movies-are-available-to-stream-rent-or-buy-online/`.

M. Greenberg. The economics of video piracy. *PIT*, 2015.

W. D. Gunter, G. E. Higgins, and R. E. Gealt. Pirating youth: Examining the correlates of digital music piracy among adolescents. *International Journal of Cyber Criminology*, 4, 2010.

B. Herz and K. Kiljański. Movie piracy and displaced sales in europe: Evidence from six countries. *Information Economics and Policy*, 43: 12–22, 2018.

Hulu. Hulu and epix announce extensive multi-year agreement. *Businesswire*, 2015.

D. Ibosiola, B. Steer, Álvaro García-Recuero, G. Stringhini, S. Uhlig, and G. Tyson. Movie pirates of the caribbean: Exploring illegal streaming cyberlockers. *ICWSM 2018*, 2018.

R. S. Jacobs. Digital movie piracy: A perspective on downloading behavior through social cognitive theory. *Computers in Human Behavior*, 28: 958–967, 2012.

S. Khan, S. Khan, and M. Aftab. Digitization and its impact on economy. *International Journal of Digital Library Services*, 5(2):138–149, 2015.

L. Leatherby and D. Gelles. How the virus transformed the way americans spend their money, 2020. URL `https://www.nytimes.com/interactive/2020/04/11/business/economy/coronavirus-us-economy-spending.html`.

L. Ma, A. Montgomery, and M. D. Smith. The dual impact of movie piracy on box-office revenue: Cannibalization and promotion. *Information Systems Economics eJournal*, 2016.

R. G. Morris, M. C. Johnson, and G. E. Higgins. The role of gender in predicting the willingness to engage in digital piracy among college students. *Criminal justice studies*, 22(4):393–404, 2009.

Netflix. Why you'll see some high profile movies leave netflix us next month. 2015.

J. Nhan, K. Bowen, and A. Bartula. A comparison of a public and private university of the effects of low-cost streaming services and income on movie piracy. *Technology in Society*, 60:101213, 2020.

C. Peukert, J. Claussen, and T. Kretschmer. Piracy and box office movie revenues: Evidence from megaupload. *International Journal of Industrial Organization*, 52:188–215, 2017.

J. Supan. When is streaming illegal? here's what you need to know about pirated content, 2019. URL `https://www.allconnect.com/blog/is-streaming-illegal`.

C. Yu, P. Hailiang, C. Chih-Hung, and Y. K. Patrick. The causal effect of subscription video streaming on dvd sales: Evidence from a natural experiment. *Unpublished manuscript*, 2020.

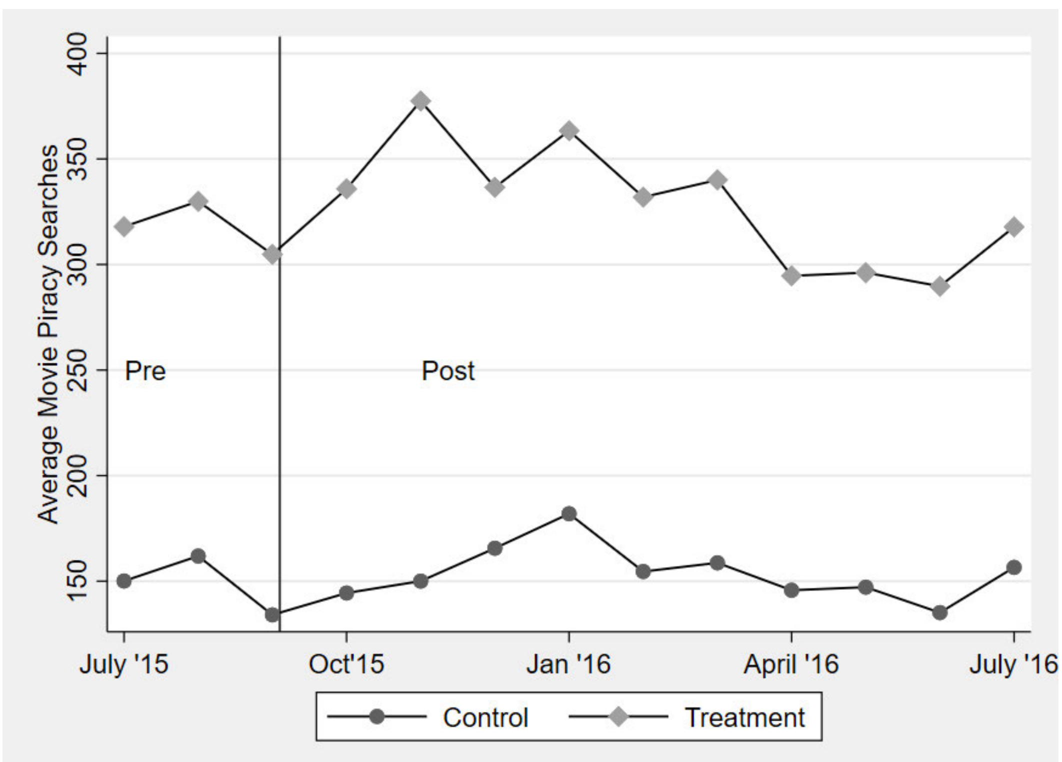
S. Yu. College students' justification for digital piracy: A mixed methods study. *Journal of Mixed Methods Research*, 6(4):364–378, 2012.

A Appendix

A.1 Additional Robustness Check

Netflix movies are frequently removed and replaced with new content and thus about half of the control group movies that were on Netflix in July 2015 were removed by December 2016. One might question if there is an important difference in movies that were removed during the observation period versus movies that were not removed. To address this, we replicate our results using a control group limited to the 501 control group movies that remained on Netflix from July 2015 to December 2016. Table 1A shows that the main results are robust to the limited control group.

Figure 1A: Average Piracy Searches Pre- and Post-Shock for movies that remained on Netflix for the full observation period



Notes: Average piracy searches plotted from July 2015 to July 2016 for the 140 treatment and 501 control movies. Epix removal shock is indicated by vertical line.

Table 1A: DID estimates of the impact of movie removal on piracy movies that remained on Netflix for the full observation period

	Oct '15	Jan '16	Apr '16	Jul '16	Oct '16
Treatment$_i$ x After$_t$	0.211***	0.183***	0.189***	0.189***	0.197***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)

Notes: (1) The dependent variable is ln(*piracysearch*$_{it}$+1). (2) This regression uses data from July 2015 - October 2016; a full year post shock for the 501 control movies that remained on Netflix for the full observation period (through December 2016). (3) robust standard errors in parentheses; (3) *** p<0.001, ** p<0.05, * p<0.1

A.2 Additional Tables

Table 2A: T-test for difference in average ln(*piracysearch*$_{it}$+1) for full sample and reduced sample

Time period	Mean ln(Piracy)	Std. Err.	|T-Stat|	P-Value
T-test with full sample				
Months 1-3	-0.087	0.0164		
Months 4-18	0.063	0.0469		
Difference	-0.1501	0.0501	2.9957	0.0028
T-test sensitivity analysis				
Month 1	0.0109	0.0193		
Months 2-3	-0.0254	0.04		
Difference	0.0363	0.0573	0.6334	0.5266

Table 3A: DID estimates of the impact of movie removal on piracy over time by movie popularity

	Oct '15	Jan '16	Apr '16	Jul '16	Oct '16
	Panel A: Popular Movies				
Treatment $_i$ x After $_t$	0.171*	0.156*	0.211**	0.228**	0.240**
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
	Panel B: Unpopular Movies				
Treatment $_i$ x After $_t$	0.169**	0.162**	0.145**	0.143**	0.145*
	(0.06)	(0.05)	(0.05)	(0.06)	(0.06)

Notes: (1) The dependent variable is ln($piracysearch_{it}$+1). (2) This regression uses data from July 2015 - October 2016; a full year post shock. (3) There are 703 high score movies of which 73 are treated (4) There are 607 low score movies of which 69 are treated (5) robust standard errors in parentheses; (6) *** p<0.001, ** p<0.05, * p<0.1

Table 4A: DID estimates of the impact of movie removal on piracy over time by movie genre

	Oct '15	Jan '16	Apr '16	Jul '16	Oct '16
	Panel A: Comedy Movies				
Treatment $_i$ x After $_t$	0.182*	0.074	0.102	0.123	0.104
	(0.08)	(0.07)	(0.08)	(0.08)	(0.09)
	Panel B: Drama Movies				
Treatment $_i$ x After $_t$	0.079	0.101	0.177*	0.211**	0.226**
	(0.08)	(0.06)	(0.07)	(0.08)	(0.08)

Notes: (1) The dependent variable is ln($piracysearch_{it}$+1). (2) This regression uses data from July 2015 - October 2016; a full year post shock. (3) There are 522 drama movies of which 51 are treated (4) There are 278 comedy movies of which 32 are treated (5) robust standard errors in parentheses; (6) *** p<0.001, ** p<0.05, * p<0.1