Piracy and Box Office Movie Revenues
Evidence from Megaupload

Peukert, Christian; Claussen, Jörg; Kretschmer, Tobias

Document Version
Accepted author manuscript

Published in:
International Journal of Industrial Organization

DOI:
10.1016/j.ijindorg.2016.12.006

Publication date:
2017

Piracy and Box Office Movie Revenues: Evidence from Megaupload

Christian Peukert, Jörg Claussen, and Tobias Kretschmer

Journal article (Accepted manuscript*)

Please cite this article as:
Peukert, C., Claussen, J., & Kretschmer, T. (2017). Piracy and Box Office Movie Revenues: Evidence from Megaupload. International Journal of Industrial Organization, 52, 188-215. DOI: 10.1016/j.ijindorg.2016.12.006

DOI: [10.1016/j.ijindorg.2016.12.006](https://doi.org/10.1016/j.ijindorg.2016.12.006)

Piracy and Box Office Movie Revenues: Evidence from Megaupload*

Christian Peukert[1], Jörg Claussen[2], and Tobias Kretschmer[2,3]

[1]University of Zurich
[2]LMU Munich
[3]CEPR London

Accepted version, January 17, 2017

Abstract

In this paper we evaluate the heterogeneous effects of online copyright enforcement. We ask whether the unexpected shutdown of the popular file hosting platform Megaupload had a differential effect on box office revenues of wide-release vs. niche movies. Identification comes from a comparison of movies that were available on Megaupload to those that were not. We show that only movies that premiere in a relatively large number of theaters benefitted from the shutdown of Megaupload. The average effect, however, is negative. We provide suggestive evidence that this result is driven by information externalities. The idea is that online piracy acts as a mechanism to spread information about product characteristics across consumers with different valuations for the product. Our results question the effectiveness of blanket public anti-piracy policy, not only from a consumer perspective, but also from a producer perspective.

Keywords: Piracy, Movie Revenues, Megaupload, Natural Experiment
JEL No.: L82, M37, D83

*Support for this research by NBER's Economics of Digitization and Copyright Initiative is gratefully acknowledged. We thank Tobias Lauinger for sharing his data and Oliver Falck, Alex Shcherbakov, Pai-Ling Yin, Itai Ater and two anonymous referees for very useful comments. The paper has benefited from feedback of conference participants at EARIE, IIOC, MaCCI, Workshops on Media Economics (Tel Aviv), Economics of Digital Piracy (Rennes), Digital Media and the Modernisation of Copyright in the EU (Brussels), and seminar presentations at Copenhagen Business School, IPTS Seville, Paris School of Economics, LMU Munich, the University of Southern Denmark, the University of Texas at Arlington and the University of Zurich. We are grateful to Sandra Huber for research assistance. Peukert (corresponding author): christian.peukert@uzh.ch, Claussen: j.claussen@lmu.de, Kretschmer: t.kretschmer@lmu.de.

1 Introduction

Online piracy is generally considered a threat to profits of producers of information goods, such as software, music, books, and movies, which, in the long run, may reduce incentives to invest in the production of such products. Consequently, governments have repeatedly taken action to enforce copyright online.

Empirical research has reached some consensus that such efforts can increase producer surplus. Such results are typically based on data on relatively successful products (Danaher and Waldfogel, 2012; Danaher and Smith, 2014), or work that estimates average effects (Gopinath et al., 2013; Adermon and Liang, 2014; Danaher et al., 2014). That is, although theory has developed a good understanding of product differentiation and network effects in the piracy context (see reviews in Peitz and Waelbroeck, 2006 and Belleflamme and Peitz, 2012), empirical evidence studying heterogeneous effects is still scarce. A small number of studies does address product heterogeneity to some degree (Oberholzer-Gee and Strumpf, 2007, Bhattacharjee et al., 2007, Zhang, forthcoming). However, these studies typically do not observe variation in the strength of online piracy at the product level, which limits the extent of the conclusions that can be drawn.

In this paper, we study product heterogeneity in the context of the shutdown of *Megaupload* and its effect on movie box office revenues. We use weekly revenues of a set of movies in 14 countries in 2011 and 2012 from *Boxofficemojo*, a commercial provider of industry statistics, and obtained data on a movie's availability on *Megaupload* by accessing archived versions of the linking website *movie2k.to*. Our identification strategy builds on a standard difference-in-difference approach where the first difference comes from the shutdown and the second difference from the availability of specific movies on *Megaupload.* That latter piece of the identification strategy is what lets our data speak to the question of product heterogeneity.

Quite surprising in light of previous empirical work, we find that box office revenues of a majority of movies that have been available on *Megaupload* do not increase in response to the shutdown. Indeed, the average effect is negative. Looking closer we find that only movies that were on release in a relatively large number of theaters (wide release) benefited from the shutdown of *Megaupload*, while the effect on more narrowly released movies was neutral and even negative for niche movies. We subject these results to a number of robustness checks to rule out alternative explanations using different specifications and additional data.

One mechanism that can explain these prima facie counterintuitive findings is that piracy generates positive externalities through information about the quality (and existence) of an experience good spilling over from pirates to potential customers. Once it becomes significantly less easy to consume pirated content online, as was the case in our empirical setting, we would expect that at least some consumers revert to licensed consumption. At the same time, the positive externalities from pirates to non-pirates vanish, so that a number of prospective customers end up being less informed about specific titles, which reduces their likelihood of going to the theater. The net effect on a specific movie's revenues then depends on how important the information externality is for the performance of this movie. Large blockbusters (i.e., wide-release movies) may be able to compensate with large advertising budgets, while word-of-mouth is likely to matter more for movies targeted at smaller audiences.

The effect of the *Megaupload* shutdown has been studied previously in Danaher and Smith (2014), but with different emphasis and methods. They investigate how the number of licensed digital sales and rentals changes, while we focus on box office revenues; they find a positive effect, while the average effect in our sample is negative. Finding less or even zero displacement the theatrical setting could be explained by exogenous differences between the studied distribution channels, be it that theater-goers pirate less or home video is a closer substitute. The fact that we find the opposite average effect, however, may be related to different sampling strategies. Danaher and Smith (2014) work with aggregated data obtained from three major studios, while we analyze a disaggregated data set that includes movies produced by both major and independent studios. This is one reason why the movies in the respective samples are likely to be quite different, much more however, because movies typically become available on the home video market only after they ran in the theaters. Theatrical display is limited in time, making word-of-mouth an important factor for box office success, which in turn is highly correlated with home video sales (Prosser, 2002). The information effect of piracy therefore hits most directly when the movie is still running in the theaters, but carries over to the home video market indirectly with a lag that is most likely too long to show in the sample of Danaher and Smith (2014). In that sense, for licensed digital sales, word-of-mouth is unlikely to play a big direct and immediate role, which is why they do not discuss it explicitly. Finally, identification in Danaher and Smith (2014) comes from differences in the penetration of *Megaupload* in different countries, essentially exploiting within-firm, across-country variation in the severity of the shock over time. In our study, variation is at the movie level as we have information about the (global) availability of individual movies on *Megaupload*. This identification strategy is better suited to our needs as we are especially interested in how the effect of online copyright enforcement varies across different types of movies. Even given all these differences, we show that our most conservative estimates are quite similar to Danaher and Smith (2014) in terms of economic significance.

Our study speaks to the recent global debate on copyright in the digital society and suggest that a blanket policy on piracy may affect firms differently, depending on their product portfolio, business model and market presence. An implication of our results could be that private enforcement – i.e. increased and targeted efforts of content providers to make unlicensed consumption of their content more difficult (Reimers, 2016) – may be better able to alleviate the negative effects of online piracy for some products, while allowing firms to benefit from the positive effects with respect to other products.

2 Movie piracy and the Megaupload shutdown

2.1 Megaupload and online movie piracy

The increased availability of fast broadband connections in the last decade made online transfer of large files feasible, leading to an upsurge in video downloading and streaming over the Internet. This did not only open up a new potential distribution channel for the movie industry, but it also enabled users to access pirated movie content more easily and at a larger scale.

Peer-to-Peer (P2P) protocols such as *BitTorrent* originally played a leading role in the distribution of copyright infringing content. The decentralized hosting of files on private computers makes it difficult to shut down those protocols and no single operator has to incur costs for infrastructure and bandwidth. However, usage of P2P protocols requires installation of applications, configuration of network settings, and usually does not enable immediate streaming, limiting P2P movie piracy to a small group of expert, heavy users and rendering it inconvenient for less experienced computer users. The emergence of filehosters (or *cyberlockers*) made consumption of infringing movie content considerably easier: no installation of applications and network configuration is necessary and many filehosters even support direct video streaming, making use of these services as convenient as watching a video clip on *YouTube*.

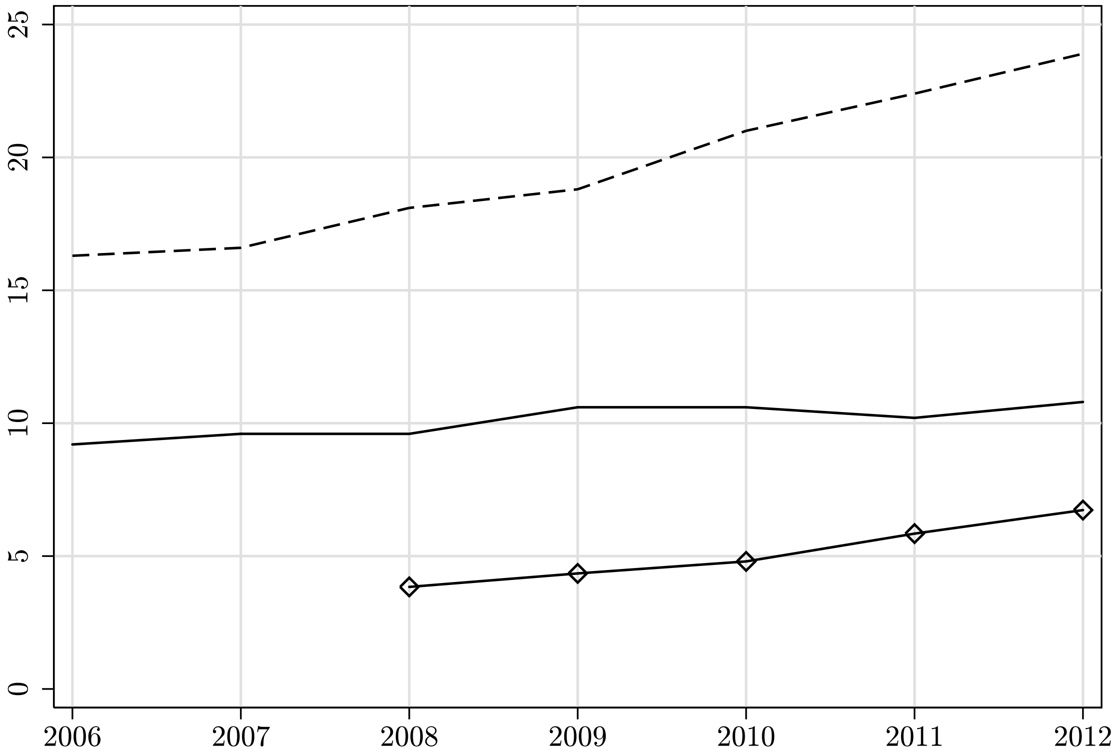
Founded in 2005, *Megaupload* grew to a dominant position in the filehoster market, by then the dominant channel to distribute pirated movie content. *Megaupload* made it easy for users to upload large files, which could then be made publicly available by distributing a link to the uploaded file. Movie files could then be either downloaded or directly streamed through *Megaupload*'s sister website *Megavideo*, launched in August 2007. Both services were financed through advertising revenues and premium subscriptions. In the free version of *Megaupload*, download speed was limited and video streaming was interrupted after 72 minutes for 30 minutes, which made it impossible for free consumers to watch a full-length movie in one go. *Megaupload* quickly became very popular and claimed to have more than 50 million daily visitors, more than 180 million registered users, and was capturing 4% of total Internet traffic at its peak.[1]

While direct visits to the *Megaupload* website usually did not bring up pirated content, movies could easily be located through search engines and to an even larger extent through link portals. Link portals act as link libraries and enable easy searching and browsing through links directed to filehosters.

The symbiotic relationship between *Megaupload* and the link portals created a grey area where legal responsibility was hard to attribute. The link portals claimed to be legal as they didn't host any content, while *Megaupload* claimed to be legal as they promised to take down unlicensed content when asked to do so. Still, *Megaupload* was widely

[1] See also Manhanti et al. (2012) who show that *Megaupload* and sister sites accounted for 68% of traffic to the top-10 cyberlocker sites in a 30,000 user campus network in 2009. This is equivalent to some 15 terabytes of data or 3% of overall HTTP traffic volume.

Figure 1: Global box office revenues



Revenues in billion US$
Box office: ---International, — US and Canada, Electronic Homevideo: -◇- US
Source: *MPAA Theatrical Market Statistics, 2010–2012,*
PwC Global Entertainment and Media Outlook (2013–2017).

considered to be at the very least a willing accomplice to extensive illegal filesharing. Chris Dodd, the chairman of the Motion Picture Association of America (MPAA) said: "By all estimates, Megaupload.com is the largest and most active criminally operated website targeting creative content in the world. [...] The site generated more than $175 million in criminal proceeds and cost U.S. copyright owners more than half a billion dollars."[2]

Interestingly, the development of box office and home video revenues in the US and Canada as well as in international markets in figure 1 shows no obvious downturn: revenues increase in international markets from 2006, while at the same time box office revenues in the American markets remain stable. Revenues from electronic home video purchases (IP-TV, over-the-top content, and streaming services such as *Netflix*) also increased in the US market since 2008.

[2]MPAA press release, available at `http://tinyurl.com/ktn3lhj`.

Of course, this descriptive analysis does not tell the full story because we lack a counterfactual that lets us compare revenues with piracy to revenues without piracy. In what follows, we propose a way to establish causality under some modest assumptions.

2.2 The shutdown of Megaupload as a natural experiment

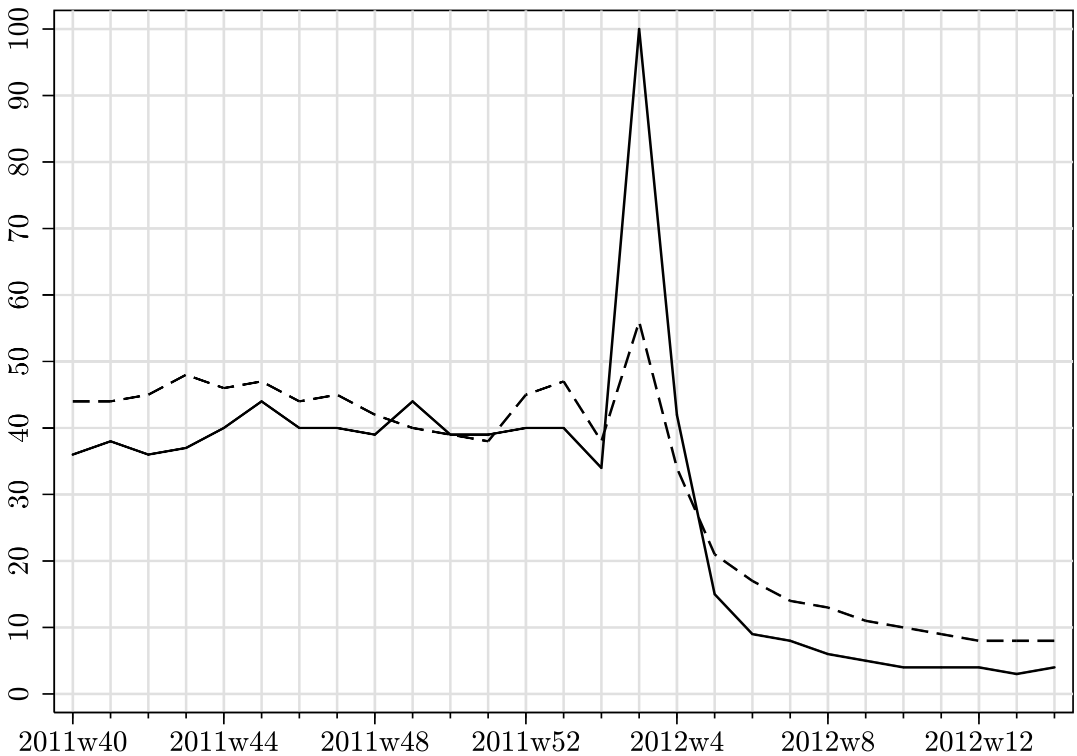
The shutdown of *Megaupload* is an exogenous shock well-suited to identify the effect of movie piracy on box office revenues. The *Megaupload* website was closed down on January 19th 2012 after an indictment by a federal grand jury. On the same day, raids were conducted in 8 countries, with search warrants being issued for 20 properties. Kim Dotcom, the founder of *Megaupload* and some of his managers were arrested in a spectacular dawn raid on his home in New Zealand and company assets were seized.[3]

On top of the largest filehoster being taken down, the events of January 19th, 2012 had additional consequences for the filehosting market as a whole. The shutdown was accompanied by massive press coverage, creating large public interest, visible in the spike in *Google* search queries in the third week of 2012 in figure 2. This had an effect on consumer awareness of what is illegal. In representative surveys among 10,000 Germans,[4] 86% stated that they know about legal consequences of up- and downloads of copyrighted material from news reports in 2012. Correspondingly, the percentage of consumers who believed it was legal to watch movies via link portals dropped from 24% in 2010 to 12% in 2012. Taking down the most successful filehoster suddenly and unexpectedly spilled over to the rest of the cyberlocker market. Even though *Megaupload* was not incorporated in the US, the lease of servers within the US was enough to make *Megaupload* liable to prosecution by US law. Many of *Megaupload*'s competitors anticipated similar legal action and reacted by shutting down or limiting their functionality. For example, in anticipation of possible prosecution, *Fileserve*, another popular filehoster, restricted downloads to the person who uploaded the file, rendering the platform useless for the distribution of pirated content.[5]

For the shutdown to let us reliably identify the causal effects of movie piracy, we have to be

[3]See for example New York Times, '7 Charged as F.B.I. Closes a Top File-Sharing Site', January 19, 2012, `http://tinyurl.com/87s6uzj`.
[4]See `http://tinyurl.com/kxpkqtc`.
[5]See `http://tinyurl.com/7fzykbc`.

Figure 2: Google search volume for Megaupload



Relative Weekly Worldwide Search Volume
— *Megaupload,* --- *Megavideo*
Source: *Google Trends*

confident that the event was indeed exogenous to the parties involved. As no reports about an expected shutdown leaked beforehand, it is plausible that the shutdown was exogenous to moviegoer demand. *Megaupload* itself did not implement any changes beforehand in possible anticipation of a legal intervention. Moreover, the management team did not try to relocate to a third country before their arrest, which they might have done had they expected the upcoming shutdown and the associated arrests. Finally, although industry organizations were seemingly involved in the investigations, it is hard to believe that they could have affected the exact timing of the shutdown. In particular, the more people are aware of the impending shutdown, the higher the risk of leakage, which would dramatically reduce the chances of success. On top of that, the long production cycles of movies makes strategic short-term supply-side reactions (e.g. delaying the release of a movie because of the imminent shutdown) unlikely.

3 Methods and data

3.1 Identification strategy

The fact that the shutdown of *Megaupload* was exogenous to all market participants is a necessary, but not sufficient condition for being able to identify the causal effect of movie piracy on box office performance. We further want to compare the actual outcome after the shutdown to some plausible counterfactual that would have occurred in the absence of the shutdown. Of course, it is impossible to observe box office performance of the same movie in both states of the world. However, it is possible to define a control group of movies that were not affected by the intervention. Under the assumption that the group of affected movies would have followed a similar sales trend in the absence of the intervention, we can compare the performance of affected and unaffected movies before and after the shutdown of *Megaupload*. We provide some support in favor of this assumption by looking at differences in pre-intervention trends of affected and unaffected movies in section 4.1.
The empirical implication of this identification strategy requires us to find a way to observe whether a movie was available for downloading or streaming on *Megaupload* before the shutdown. We can then compare the box office performance of movies that were available (treatment group) to movies that were not (control group), before and after the shutdown by estimating δ in a standard difference-in-differences model defined as

$$\ln(R_{ijt}) = \alpha + \beta After_t + \delta(After_t * Mega_i) + C'_{ijt}\gamma + \mu_i + w_t + y_t + \nu_j + \varepsilon_{ijt}. \qquad (1)$$

In this model, R_{ijt} denotes box office revenues of movie i in country j at time t, $Mega_i$ indicates availability on *Megaupload* before the shutdown, and $After_t$ indicates the time period after the shutdown. Country suffixes are not necessary for the shutdown dummy ($After_t$) and *Megaupload* availability ($Mega_i$) as the website was accessible (and subsequently non-accessible) in all countries.
We make use of the panel structure of our data and include a number of fixed effects to account for unobserved heterogeneity. First, movies obviously differ in their inherent appeal to audiences. Some characteristics that determine a movie's appeal may be observable, e.g. previous commercial success of actors and directors, while others remain

unobservable. We control for such time-invariant heterogeneity by including movie fixed-effects μ_i. As a result however, we cannot separately identify a coefficient for a specific movie's pre-shutdown availability on *Megaupload*, as this variable is also time-invariant. Second, seasonality matters for the box office performance of movies, so we include week fixed-effects w_t and year fixed-effects y_t. Finally, we account for unobserved time-invariant differences between countries in ν_j, which for example allows us to control for heterogeneity coming from differences in the size of population and therefore the number of movie theaters.

Because the intervention hits them at different stages of their life-cycle, different movies experience different revenue growth patterns before and after the shutdown by virtue of their age. We account for this by controlling for the time since the movie has premiered in a given country. This variable is included in the vector C_{ijt}.

Assumptions about the error term ε_{ijt} are standard, but we allow for heteroscedasticity within observations of the same movie by clustering. Serial correlation of error terms is an obvious concern in our setting, therefore we report results from a number of different specifications following Bertrand et al. (2004) in section 4.2.

3.2 Data

We construct a rich dataset from a variety of public sources, starting with weekly country-level box office revenue data from *Boxofficemojo*, a commercial provider of movie industry statistics. We match this with information obtained from *IMDb*, the leading Internet platform for movie meta information, which lets us observe international titles and whether the movie is a remake or part of sequel. It is not possible to directly observe a list of movies available on *Megaupload*, mainly because the website never had a user interface displaying the uploaded content to a public audience. The way consumers accessed files stored on *Megaupload* was via link portals that provided a catalogue of available movies. Much like licensed services such as *Netflix* or *Amazon Prime*, a link portal provides meta-information and links to downloads or streams hosted by a cyberlocker, often organized by genres, popularity or release years. Information about the contents of linking sites therefore provides an indirect way to measure (at least a lower bound of) availability on a cyberlocker.

We collect information on whether a movie was available for streaming on *Megavideo* from the linking site *movie2k.to*. With about 144,000 daily pageviews, *movie2k.to* ranked 58th on *Alexa*'s list of the most popular websites in Germany in the end of 2011.[6] We obtain 16,773 snapshots of *movie2k.to* content pages from December 13th, 2010 to January 18th, 2012 from the *Internet Archive*.[7] This lets us observe 21,943 links to 16,212 movies, out of which 8,234 (37.5%) point to *Megavideo*.[8] For a majority of movies (60%), *movie2k.to* provides links to the Internet Movie Database (*IMDb*), where users can access information about cast, critics and awards, along with trailers. *IMDb* has a unique ID for each movie, which we can use to unambiguously match the *movie2k.to* data to revenue data obtained from *Boxofficemojo*. In our final sample, 24% of the movies from *Boxofficemojo* can also be found on *Megavideo* as according to *movie2k.to*. Because we do not consistently observe an *Internet Archive* snapshot, say every week, for each movie-specific page on *movie2k.to*, we cannot use the time dimension of the archival data. Hence, we define a movie as available on *Megavideo* if we observe at least one link on a page archived before January 19, 2012. We discuss potential measurement errors with this data in section 4.5.

Our final sample includes weekly observations of 308 movies in 14 countries (see table 1) from 2011 to 2012. Descriptive statistics of all variables used in the analysis are reported in table 2, along with a mean difference comparison between treated and untreated observations. Across almost all variables, we do not see a significant difference between affected and unaffected observations. We do observe relatively more treated movies after the intervention.[9]

Our main variable of interest is weekend box office revenues, measured in US dollars. The definition of "weekend" differs across weeks and countries. *Boxofficemojo* sometimes

[6] See `http://tinyurl.com/kysykqm`.

[7] The Internet Archive (`http://www.archive.org/web`) provides a database of more than 450 billion web pages saved over time.

[8] In the majority of cases, a movie can be downloaded or streamed from a variety of sources. *Megavideo* has by far the largest market share, followed by *Stream2k* (18.3%) and a number of much smaller linking sites and cyberlockers, namely *Dixstage*, *Movshare*, *Novamov*, *Ovfile*, *Putlocker*, *Royalvids*, *Sharefiles*, *Sockshare*, *Ufliq*, *Upafile*, *VideoWeed*, *Vidxden*, *Xvidstage*, *Xvidstream*, and *Zalaa*.

[9] We investigate whether this is driving our results by carrying out a series of regressions where we remove the two latest weeks of a random selection of 10% of the treated movies. The baseline results reported in section 4.2 are robust to this. Across 1,000 runs we find the negative coefficient of *After*Mega* to be significant at the 10% level in 93.5% of the cases.

Table 1: Countries in the sample

	Frequency	%
Australia	362	4.68
Austria	905	11.71
Belgium	756	9.78
Denmark	114	1.47
Finland	364	4.71
France	229	2.96
Germany	826	10.68
Israel	271	3.51
Italy	660	8.54
Japan	220	2.85
Netherlands	337	4.36
New Zealand	781	10.10
Norway	404	5.23
United States	1,502	19.43
Total	7,731	

reports revenue figures based on two, three, four or five days. We therefore control for the number of days in a given weekend and country. Because the variable is heavily skewed (mean: $623,132; median: $42,080), we use the log in the regressions. We measure a movie's country-specific age by counting the number of weeks since the launch in a given country. The average lifetime is some 7 weeks, but there are also some movies that run for more than 30 weeks (most of which are narrowly released ("small") movies in the genres animation, documentary, and short film; the maximum is 213 weeks). We therefore use the log of weeks active in our estimations. The shutdown of *Megaupload* occurred on Thursday, January 19th, 2012, i.e. in the third calendar week. Revenue data for the third calendar week in 2012 refer to January 20th to 22nd. We therefore define the post shutdown period as after 2012w2 and construct a corresponding dummy variable.

4 Effects of the shutdown on box office revenues

4.1 Testing the identifying assumption

Before we present the results of our various specifications, we investigate if treated and untreated movies follow similar pre-intervention trends. This is the testable part of our identifying assumption, i.e. that revenues of both types of movies would have developed

Table 2: Descriptive statistics

	Total		Non-Mega		Mega		Diff	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Coeff.	S.E.
ln(Weekend Gross)	10.376	2.337	10.294	2.342	10.526	2.322	-0.231	0.199
After	0.055	0.228	0.029	0.167	0.104	0.306	-0.076**	0.032
Weekend Days	3.560	0.747	3.556	0.747	3.567	0.748	-0.011	0.031
ln(Weeks Active)	1.488	0.892	1.511	0.897	1.446	0.882	0.065	0.073
Release Intensity	0.289	0.243	0.288	0.247	0.289	0.236	0.000	0.045
ln(Reviews)	4.564	1.491	4.517	1.568	4.651	1.333	-0.134	0.251
Remake	0.152	0.359	0.161	0.367	0.136	0.342	0.025	0.074
Sequel	0.262	0.440	0.279	0.449	0.229	0.420	0.050	0.099
PirateBay	0.920	0.272	0.916	0.277	0.926	0.262	-0.010	0.033
Indie	0.299	0.458	0.334	0.472	0.233	0.423	0.101	0.076
Observations	7,731		5,018		2,713			

Note: Standard errors (clustered at the movie-level) in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

in the same way had the shutdown not taken place.

Figure 3 plots week-treatment group coefficients θ_k estimated in a model defined as

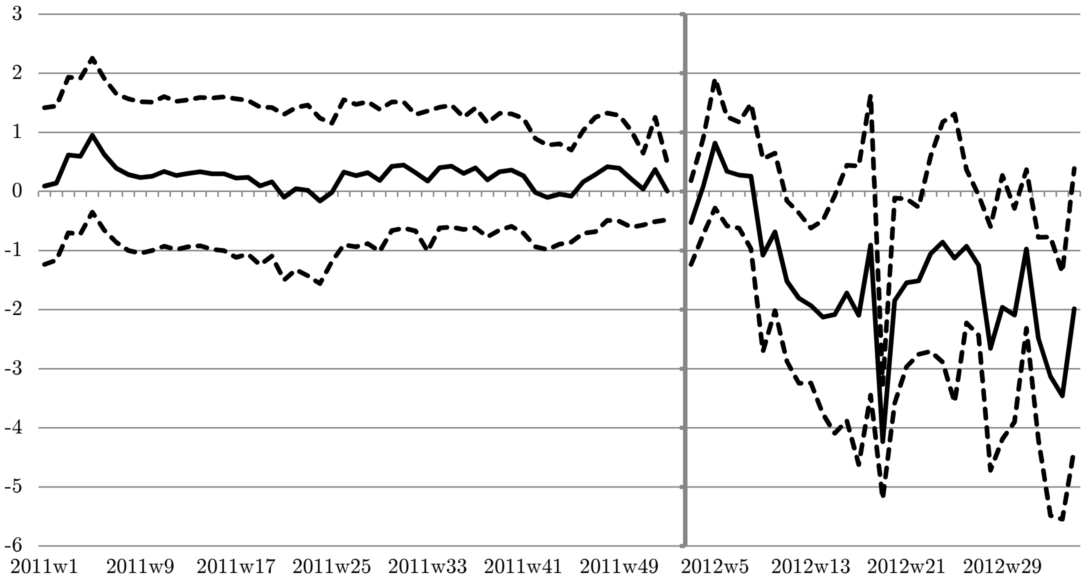
$$\ln(R_{ijt}) = \alpha + \sum_{k=1}^{t}(\beta_k w_k + \theta_k(w_k * Mega_i)) + C'_{ijt}\gamma + \mu_i + y_t + \nu_j + \varepsilon_{ijt} \qquad (2)$$

as a test for the identifying assumption of similar pre-intervention trends of treatment and control group. We look at weekly differences in revenues across the two groups of observations. For the whole pre-intervention period, both groups follow trends that are indistinguishable from zero at the 90% significance level. Further, the results show a clear change after the shutdown of *Megaupload* on January 19th, 2012. Hence, after the intervention, treatment and control group start to have diverging trends. This gives a strong indication that our identification strategy is sound for estimating the causal effect we are after.

4.2 Baseline results

Results of the main regressions are presented in table 3. Across all columns we report results of models that include year, calendar week (or month, respectively), country, and

Figure 3: Treatment and control Group, pre-and post intervention



OLS Coefficients, *— Week*Mega, ---90% Confidence Bands, Vertical line indicates shutdown. Reference category is the second week of 2012.*

movie fixed effects. Because serial correlation could result in incorrect inference, we report different specifications as proposed by Bertrand et al. (2004). Standard errors are clustered at the movie-level in the first column, at the country-level in the second column, and at the movie- and country-level in the third column. We report estimates using block bootstrapping with 5,000 replications (with country-movie blocks) in the fourth column. Throughout all model specifications, the coefficient of *Weekend Days* is negative, mainly because *Boxofficemojo* tends to report longer weekends in smaller countries.[10] The life-cycle follows the expected decreasing non-linear trend. In the fifth column we estimate a model that neglects most of the time dimension by using averages at the month-level (Bertrand et al., 2004).

Across all specifications, the estimated coefficient for the difference-in-differences is negative. Although not estimated very precisely (significant at the 10% level), the effect is of sizable magnitude, suggesting that weekend revenues of the average movie decreased by about two thirds.[11]

[10] For example, five day weekends are only observed in France and Belgium. Most weekends reported for the US are three days long.

[11] (Exp(-1.131)-1)*100=-.677.

Table 3: Results: Baseline specification

	Standard error cluster				
	Movie	Country	Movie-Country	Movie-Country Block Bootstrap	Monthly Average
After	1.045*	1.045***	1.045*	1.045*	0.529
	(0.587)	(0.342)	(0.568)	(0.589)	(0.600)
After*Mega	-1.131*	-1.131**	-1.131*	-1.131*	-1.261*
	(0.655)	(0.378)	(0.600)	(0.635)	(0.711)
Weekend Days	-1.366***	-1.366***	-1.366***	-1.366***	
	(0.159)	(0.425)	(0.444)	(0.494)	
ln Weeks Active	-1.402***	-1.402***	-1.402***	-1.402***	
	(0.051)	(0.074)	(0.037)	(0.039)	
Year Effects	Yes	Yes	Yes	Yes	Yes
Week Effects	Yes	Yes	Yes	Yes	No
Month Effects	No	No	No	No	Yes
Country Effects	Yes	Yes	Yes	Yes	Yes
Movie Effects	Yes	Yes	Yes	Yes	Yes
Observations	7,731	7,731	7,731	7,731	2,856
$\overline{R}^2$	0.782	0.782	0.782	0.782	0.635

Dependent variable: ln Gross Weekend Revenues.
Note: Constant not reported. Standard errors in parentheses. Block bootstrap results in column 4 are based on 5000 replications, and column 5 reports estimates on a sample that is averaged on the movie-country-month level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.3 Effect heterogeneity

4.3.1 Different types of movies

To investigate our somewhat surprising findings in more detail, we explore heterogeneity in the effect we identify. We are particularly interested in testing whether the intervention affects differentiated products (as measured by movies targeted at different audience sizes) differently.[12]

We measure a movie's targeted audience size using information about the exhibition intensity of a movie in its first week in a given country. The number of opening screens is an important strategic variable and is closely related to the targeted audience and expected demand for a movie (Roos and Shachar, 2014). We do not directly use absolute numbers or market shares per country and week because such measures are endogenous

[12] A number of alternative (policy and strategy relevant) sources of heterogeneity – like genre, MPAA rating, presence of movie stars, production budget – are covered by movie-fixed effects.

Table 4: Results: Heterogeneity

	Release Intensity			
	Linear		Non-linear	
After	1.397**	(0.653)	0.942**	(0.374)
After*Mega	-1.216	(0.747)	-0.716	(0.443)
Release Intensity	2.694***	(0.353)		
After*Release Intensity	-4.530***	(1.670)		
Mega*Release Intensity	0.027	(0.425)		
After*Mega*Release Intensity	3.143	(1.969)		
Wide Release			0.456***	(0.125)
After*Wide Release			-2.717***	(0.358)
Mega*Wide Release			-0.107	(0.243)
After*Mega*Wide Release			1.315**	(0.563)
Narrow Release			-0.990***	(0.270)
After*Narrow Release			0.909	(1.046)
Mega*Narrow Release			0.649	(0.401)
After*Mega*Narrow Release			-1.895*	(1.121)
Weekend Days	-0.543***	(0.209)	-1.139***	(0.301)
ln Weeks Active	-1.429***	(0.049)	-1.388***	(0.051)
Year Effects	Yes		Yes	
Week Effects	Yes		Yes	
Country Effects	Yes		Yes	
Movie Effects	Yes		Yes	
Observations	7,731		7,731	
$\overline{R}^2$	0.799		0.791	

Dependent variable: ln Gross Weekend Revenues.
Note: Constant not reported. Standard errors (clustered on the movie-level) in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

when theater owners can for example quickly adjust the number of screens as a response to (unexpected) changes in demand. Using the exhibition intensity – the "width" of a release – in the first week as a measure of expected overall demand (De Vany and Walls, 1996) can mitigate this issue.

For most countries *Boxofficemojo* reports the total number of screens per movie and weekend, while for some countries we observe the number of theaters.[13] This is not the same since one theater location may play a movie on several screens.

To ensure that we are not picking up this artifact in the estimations, we relate the first-week

[13]These countries are Australia, France, Germany, and Italy. Excluding these countries gives the same qualitative results, albeit at reduced significance.

screens (theaters) to the maximum number of screens (theaters) in a given country. The resulting measure is a percentage where 1 indicates that the movie has the biggest starting week in a given country between August 2007 and December 2012.[14] Not surprisingly, the distribution of this variable is skewed, with a median of 0.24 and a mean of 0.29. The 95th percentile is 0.76 and the 99th percentile is 0.93.

The first column of table 4 shows results of a specification where we interact our measure of release width with the difference-in-differences indicator. The main effect retains a similar magnitude but is estimated less precisely. Most importantly, however, we do not find evidence for a linear effect of release width. While the number of screens in the first week is informative about the overall box office performance of a movie, it is also an indicator of one of two generic exhibition strategies. High levels of initial exhibition intensity are generally employed by major studies for intensely promoted movies featuring star actors, while smaller independent distributors usually choose a smaller number of screens in the first week (Sawhney and Eliashberg, 1996: 119). The results reported in the first column could therefore be a reflection of two contrary effects that cancel out in the averages.

To investigate this, we define "wide release" movies as within the 90th percentile, "narrow release" movies as those below the 10th percentile, and the remaining movies are those between the 10th and 90th percentile. We use these "medium release" movies as the omitted category. For the US market, examples of wide-release blockbusters are *Cars 2* (Animation/Adventure/Comedy), *Harry Potter and the Deathly Hallows: Part 2* (Adventure/Fantasy/Mystery), and *X-Men: First Class* (Action/Adventure/Sci-Fi); medium release movies include *The Help* (Drama, Oscar for best performance by an actress in supporting role), *The Adjustment Bureau* (Romance/Sci-Fi/Thriller) and *Winnie the Pooh* (Animation/Adventure/Comedy). Examples of narrow release movies in the US market are *Midnight in Paris* (Comedy/Romance, Oscar for best writing), *Senna* (Documentary about race-car driver Ayrton Senna) and *The Guard* (Comedy/Crime/Thriller).

Our results are reported in column (2) of table 4. The difference-in-differences coefficient *After*Mega* remains negative, but the estimate is not significantly different from zero. We find a positive and significant coefficient for *After*Mega*Wide Release* and a negative and

[14] We choose this time period as it marks the start of Megavideo. However, as 94% of all movies in our data are not released before the first week of 2011, choosing a shorter time period does not affect the results.

significant coefficient for *After*Mega*Narrow Release*. Hence, after the shutdown, revenues of movies that open on a relatively small number of screen decrease, while revenues of movies that are on wide release increase.

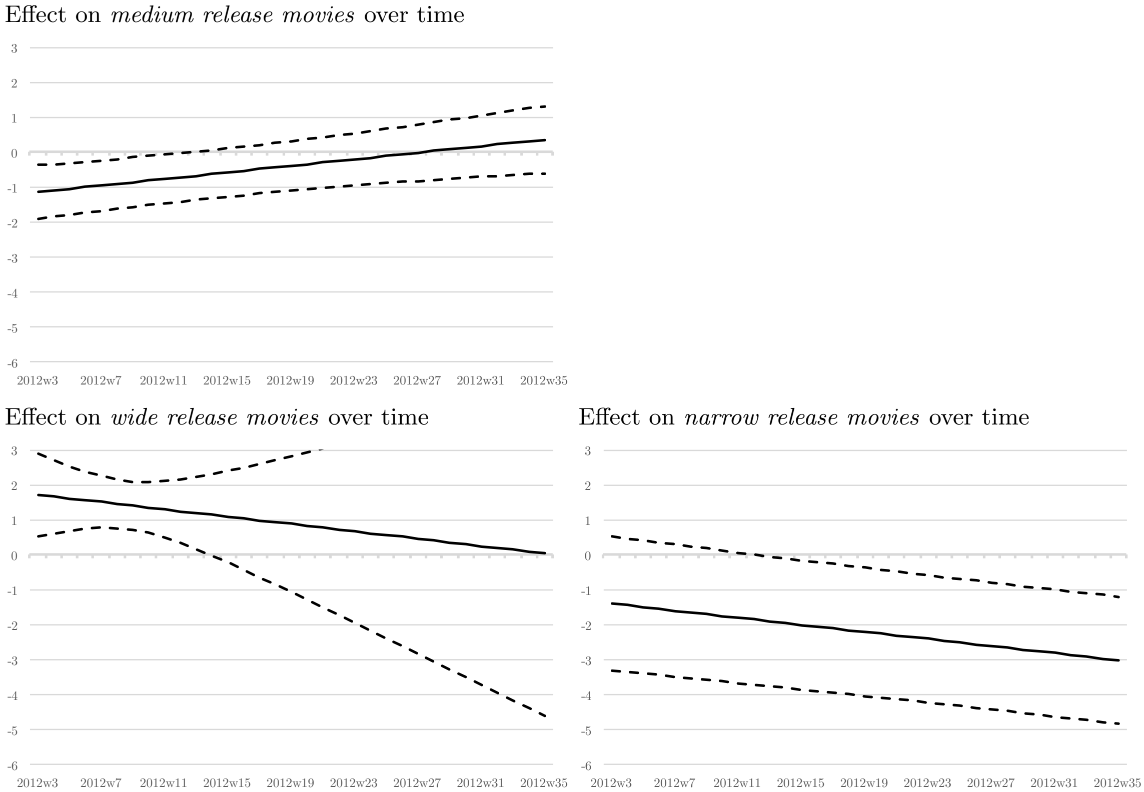
The coefficients are perhaps too imprecisely estimated to draw overly broad quantitative conclusions. The effect sizes are strikingly large but also come with large confidence bands. Revenues of narrow release movies decrease by about 85% ($\approx e^{1.315} - 1$) after the intervention compared to medium-sized movies. The 90% confidence band spreads from -98% to -4%. Revenues of widely released movies increased by some 273% ($\approx e^{-1.895} - 1$), with a 90% confidence band from 47% to 843%. Expressed in dollar units this translates a 85% decline into $640,800 and a 273% increase into $4,108,000 respectively.[15] Taking into consideration that our coefficient estimates are not extremely precise, we prefer to use our most conservative estimates of -4% and 47%, which lets us arrive at a decline worth $30,000 and an increase worth $707,000.

4.3.2 Heterogeneity over time

Figure 3 suggests that the effect becomes significant only after a couple of weeks. To understand this, it is useful to remind ourselves that identification comes from a comparison of the same movie before and after the shutdown. That is, the effect for weeks relatively long after the shutdown is identified by movies that run relatively long in the theaters. It turns out that the average length of the theater lifecycle is largely determined by the release strategy (Sawhney and Eliashberg, 1996; Chen et al., 2013). Wide release means that the movie runs simultaneously in a large number of theaters across the country, while a narrow release strategy implies that the movie opens in a limited number of theaters sequentially in selected cities. As a consequence, revenues of wide release movies typically follow a L-shaped form, with a peak within the first few weeks and a fast decay thereafter. In contrast, sales of narrow release movies are much less concentrated in the first weeks and more evenly distributed over time, decaying at a much lower rate. This naturally implies that the lifecycle of wide release movies is typically shorter than the lifecycle of

[15] Subtracting country-specific means, the average pre-shutdown weekend revenue is $753,892 for narrow release movies, i.e. 753,892*0.85 ≈640,800. For wide release movies it is $1,504,787, i.e. 1,504,787*2.73≈4,108,000.

Figure 4: Shutdown effect and linear time trend



Note: *Same specification as in table 4 plus linear time trend.*
— *After*Mega*[Wide, Narrow]+After*Mega*[Wide, Narrow]*Time*
--- *90% Confidence Bands.*

narrow release movies, which is also evident in our sample. The average number of weeks we observe a movie that we define as a wide release is 7.37, which is significantly smaller than the average number of weeks we observe a movie we define as narrow release (12.85). Hence, the identification of the shutdown effect in the first weeks after the shutdown comes from all types of movies, while it increasingly comes from more narrow release movies in later weeks.

To see more closely how this operates, we augment the model in table 4 with a linear time trend that we interact with $After * Mega$, $After * Mega * WideRelease$ and $After * Mega * NarrowRelease$. In figure 4 we plot the corresponding effect estimates as function of time, e.g. $After * Mega + After * Mega * Trend$. In the first couple of weeks we observe a significant negative effect on medium release movies, significant positive effect on wide release movies and an non-significant negative effect on narrow release movies. In later weeks, the effect on both medium and wide release movies becomes insignificant (with a very large standard error in the latter case), while the negative effect on narrow release movies becomes significant. Hence, when the positive and negative effect cancel out in the first couple of weeks, and narrow release movies dominate the sample in the longer run, this is consistent with the insignificant short-run, but significantly negative long-run effect that figure 3 suggests.

4.3.3 Effects on revenue distribution

Having established that the policy intervention affects different types of movies in different ways, we now focus on the effects on the distribution of revenues across firms and movies. The *Megaupload* shutdown can potentially affect market structure significantly if especially movies that are released widely benefit while revenues of more narrowly released productions are hurt.

First, we begin by looking at effect differences between major and independent distributors.[16] Our results in the first column of table 5 suggest that the shutdown of *Megaupload* did not disproportionately benefit larger movie distributors. The interpretation is that the effect is happening at the movie-, not the distributor-level. Put differently, larger companies sometimes release movies narrowly, while smaller companies sometimes put their movies on wide release. In columns two and three of table 5 we look more specifically at the distribution of revenues in a given week. First we investigate whether the top of the distribution has changed as a result of the intervention by looking at the difference between revenues of the most and second most grossing movie in a given week and country. The coefficient of *After*Mega* is not significant, suggesting that the intervention did not accelerate a superstar effect. However, looking at the difference between the biggest and smallest movie in a given week and country (measured in box office revenues) carries an important insight. The estimated coefficient of *After*Mega* is positive and significant, which implies that the shutdown of *Megaupload* increased the gap between the top and least grossing movie. As illustrated in figure 5, this indicates that the shutdown of *Megaupload* made the market more concentrated, shifting mass towards the head of the distribution.

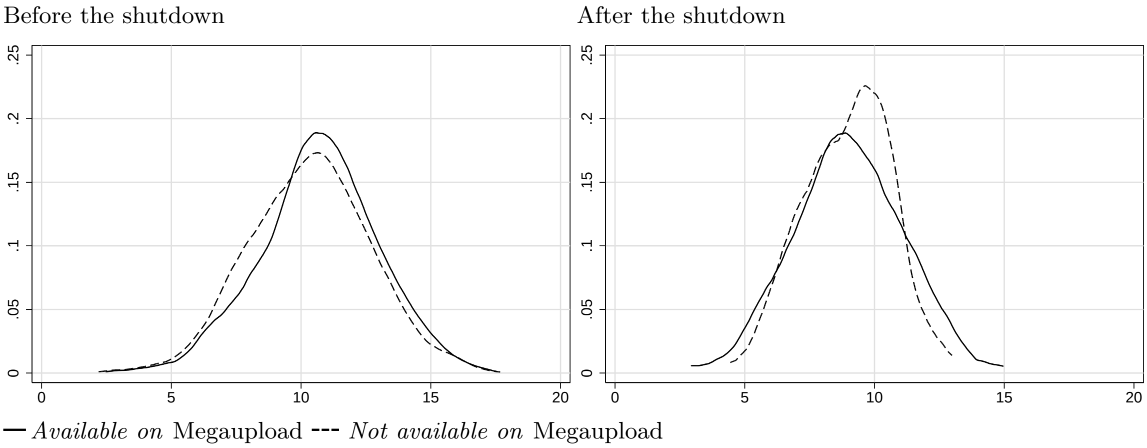
[16] We classify movies as independently distributed if *Boxofficemojo* does not indicate the distributor as *Buena Vista*, *Dreamworks*, *Fox*, *Lionsgate*, *New Line*, *Paramount*, *Samuel Goldwyn*, *Sony*, *Summit*, *UTV*, *Universal*, *Warner* or *Weinstein*.

Table 5: Results: Market structure

	Market Structure					
	Indie vs. Major		Top1-Top2		Top1-Last	
After	0.197	(0.565)	-0.043	(0.586)	-2.912***	(0.768)
After*Mega	-0.637	(0.617)	-0.045	(0.423)	2.433***	(0.269)
After*Indie	0.991	(0.822)				
After*Mega*Indie	0.334	(0.922)				
Weekend Days	-1.313***	(0.159)	0.028	(0.581)	-2.719***	(0.503)
ln Weeks Active	-1.390***	(0.050)				
Year Effects	Yes		Yes		Yes	
Week Effects	Yes		Yes		Yes	
Country Effects	Yes		Yes		Yes	
Movie Effects	Yes		No		No	
Observations	7,731		1,327		1,327	
$\overline{R}^2$	0.783		0.220		0.568	

Dependent variables: ln Gross Weekend Revenues, differences in the distribution of ln Gross Weekend Revenues.
Note: Constant not reported. Standard errors (clustered on the movie-, country-, and movie-country-level) in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 5: Revenue distribution before and after the shutdown

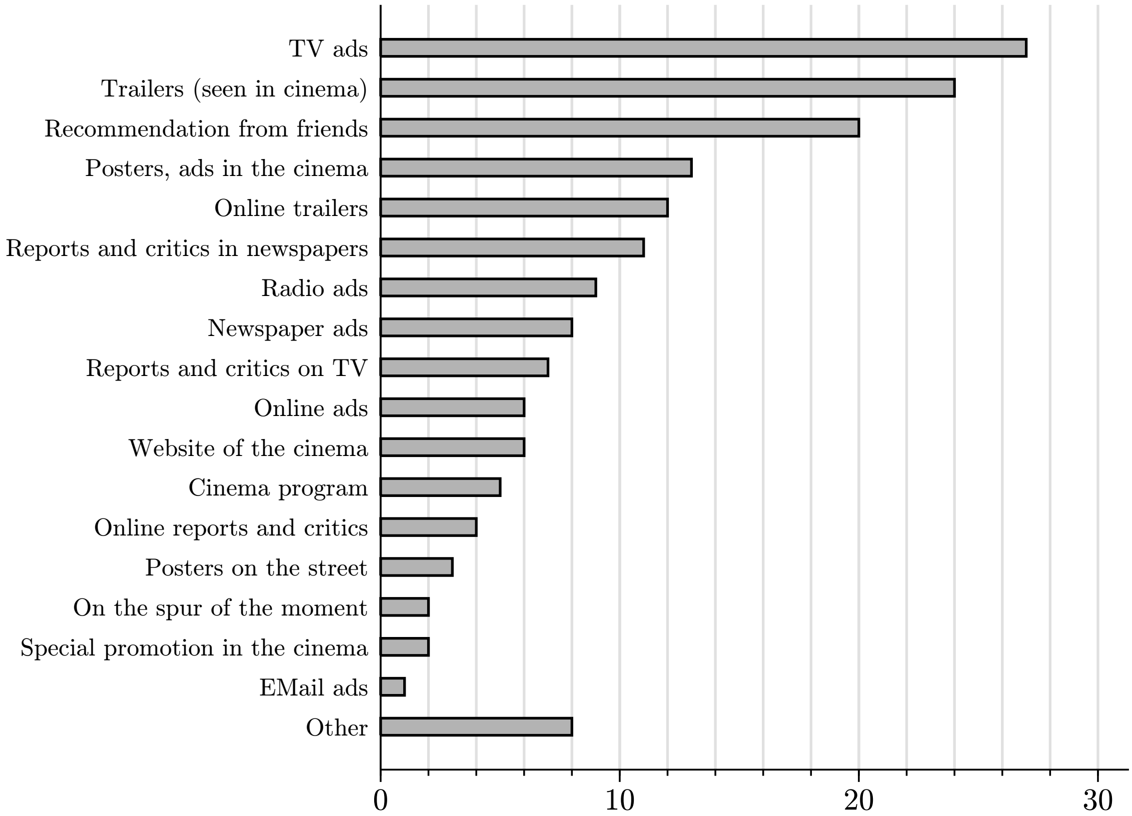


4.4 Product discovery and the role of word-of-mouth

4.4.1 Consumer awareness, word-of-mouth and advertising

The main result of our analyses is that the shutdown of a large supplier of unlicensed downloads did not have a positive effect across the board. We find that smaller and larger movies, as measured by the width of their initial release, were differentially affected by

Figure 6: Sources of awareness about movies



"How do you decide to go to the movies?"
Percentage of people answering with "yes".
Data from representative sample of 25,000 German individuals older than 10 years (GfK Panel, 2011)
***Source**: German Federal Film Board (FFA), "Der Kinobesucher 2011", p. 70*

the shutdown of *Megaupload*. While widely released blockbuster movies benefitted from the shutdown on average, the effect on revenues of most other movies is either statistically not distinguishable from zero or even negative. This result is surprising for two reasons. First, intuitively one would not expect that removing the option of free consumption can have a negative effect on firm revenues. And second, it is not immediately clear why the effect is negative for smaller movies but turns positive for larger ones.

One way of thinking about our first main finding that some movies may see a drop in revenues if piracy is reduced is that piracy may play an important role in how consumers become aware of products and collect information about product quality. Figure 6 showing survey data from a representative sample of 25,000 Germans suggests that the primary source of information is advertising, while at the same time purchase decisions are often influenced by recommendations of friends. Similar data from the US suggests that some 30% of persons aged 12–74 that attended at least one movie in a theater during 2011 have used social media to discuss movies.[17] This notion is confirmed by several studies that show that recommendations by direct contacts play an important role in the demand for a particular product (Oestreicher-Singer and Sundararajan, 2012), its evaluation (Lee et al., forthcoming), or even aggregate demand for a media category (Hosanagar et al., 2014). Thus, conventional advertising and online channels can interact in intricate ways (Joo et al., 2014).

To see this more explicitly, consider the following simple thought experiment. If recommendations can come from pirates and consumers who have watched the movie through a licensed channel before, eliminating piracy reduces the likelihood that a consumer will receive a recommendation for a movie, which results in a lower likelihood to watch a particular movie in the theater. Given that word-of-mouth communication is an important promotion channel, the net effect of this could be negative at least for some movies. For which type of movies would we expect the word-of-mouth effect to be especially strong and the revenue effect to turn negative? Note that advertising budgets for movies are correlated with production budget (Vogel, 2004), making up for up to 40% of the production budget (Prag and Casavant, 1994). Production budget is positively correlated with wide release movies, and negatively correlated with narrow release movies. If narrow release movies have smaller marketing budgets, word-of-mouth is a relatively more important factor for their overall success. If this word-of-mouth effect is reduced with the shutdown of unlicensed content, the performance of narrow release movies is affected negatively.

We can make use of variation in the timing of advertising expenditures in search for more direct evidence for this type of information externality. The literature suggests that advertisement for motion pictures is typically concentrated around the opening week (Eliashberg et al., 2000). The information effect of piracy should therefore not only vary by movie type, but also along a movie's lifecycle. We test this idea by augmenting the model in table 4. To see how the shutdown effect changes over the lifecycle of different types of movies, we add interactions with a linear and time-invariant measure of the movie's lifecyle

[17] See `http://tinyurl.com/mect3jy`.

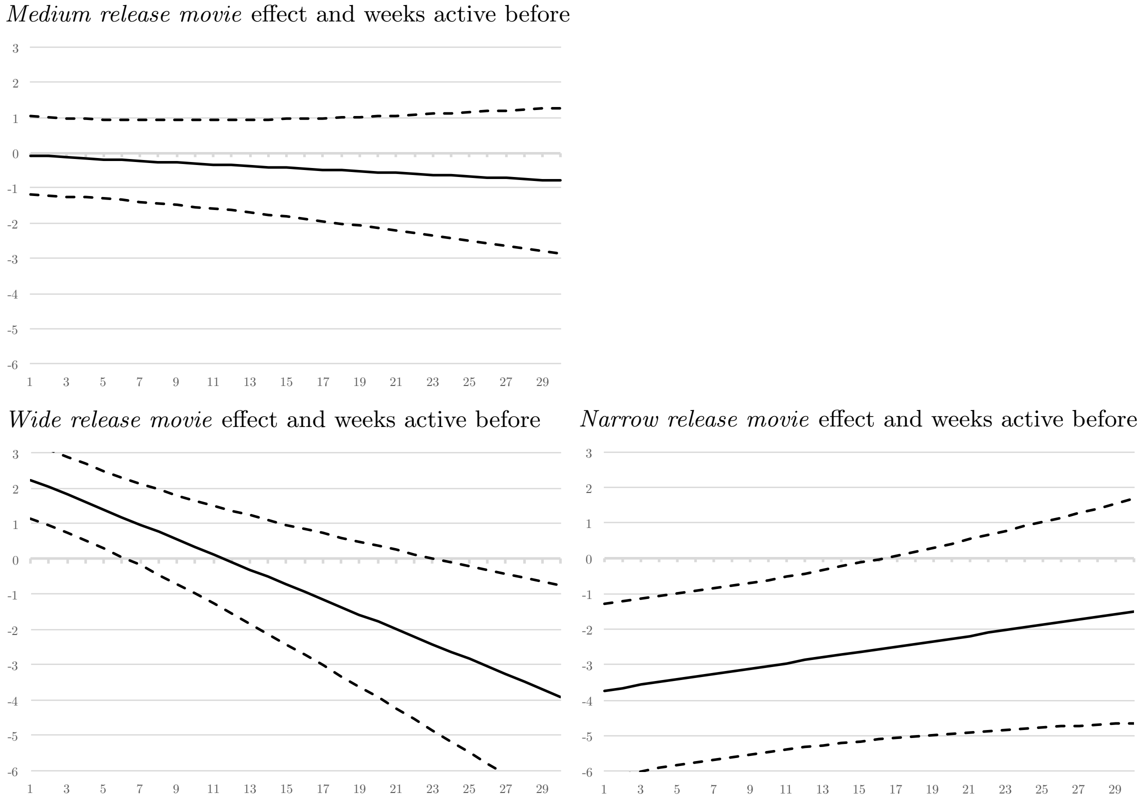
before the shutdown. In figure 7 we plot the corresponding effect estimates as function of this variable, e.g. $After * Mega + After * Mega * WeeksActiveBeforeShutdown$.[18] The upper left panel suggests that medium release movies at different stages of their lifecycle are not affected in a significantly different way. The story is different for wide release and narrow release movies (lower panels of figure 7). Box office revenues increase for wide release movies that had just premiered when the shutdown happens. This effect becomes insignificant for wide release movies that had already been running for a longer time when *Megaupload* disappears, and turns negative. Narrow release movies that haven't been exhibited in theaters for too long when the shutdown happens, experience a decrease in revenues. Again, this effect becomes insignificant for movies that are "older" when the shutdown happens. In summary, these results imply indirect evidence for the idea that advertising and piracy can have complementary effects. The types of movies that are likely to be advertised more than other types of movies, at the stage of their lifecycle where they are more likely to be advertised than at other stages of their lifecycle, are the only ones for which we can find evidence for a positive effect of the shutdown of *Megaupload* on box office revenues.

4.4.2 Information and heterogeneity across movies

A second piece of indirect evidence for the word-of-mouth logic emerges from the idea to exploit variation in the amount of information consumers have about movies, on top of advertising. We do this by distinguishing between different types of movies. Broadly, we can think of three types of movies: Original productions (e.g. *Inception* or *The American*), remakes of older versions (e.g. *The Three Musketeers* or *Gulliver's Travels*), and movies that are part of sequels (e.g. *Cars 2* or *The Hangover Part II*). The idea is that these types of movies can be ranked according to the amount of information the average consumer has about the characteristics of a movie (and therefore the likely match to her preferences). The most extensive information is available on the plot of remakes, because a consumer may have watched the older version of the movie or the plot is commonly known. Sequels

[18]The maximum number of weeks active before the shutdown we observe in our data is 107 for medium release movies, 27 for wide release movies, and 213 for narrow release movies. For illustration purposes we report results using a linear version of the variable, results are qualitatively unchanged when using the logarithm.

Figure 7: Shutdown effect and weeks active before shutdown



Note: *Same specification as in table 4 plus interaction term.*
— *After*Mega*[Wide, Narrow]+After*Mega*[Wide, Narrow]*WeeksActiveBeforeShutdown*
--- *90% Confidence Bands.*

share common characters and often continue the plot of a preceding movie, while the least is known about original movies.

We make use of this difference in available information across movie types by estimating three-way-interactions with our difference-in-differences indicator, leaving sequels as the omitted category. The results in table 6 confirm our basic intuition. While the baseline effect *After*Mega* remains negative and significant, the coefficient of *After*Mega*Original* is not different from zero and the estimate of *After*Mega*Remake* is significantly positive. Hence, the interpretation is that with less word-of-mouth, consumers will tend to stick to products about which they already have relatively more information.

Table 6: Results: Word-of-mouth

	Types of Movies	
Weekend Days	-1.360***	(0.158)
ln Weeks Active	-1.404***	(0.049)
After	-0.017	(0.310)
After*Mega	-1.385***	(0.270)
After*Remake	-1.988***	(0.345)
After*Mega* Remake	5.805***	(0.377)
After*Original	1.647**	(0.682)
After*Mega*Original	-0.159	(0.732)
Constant	20.456***	(0.637)
Year Effects	Yes	
Week Effects	Yes	
Country Effects	Yes	
Movie Effects	Yes	
Observations	7731	
$\overline{R}^2$	0.785	

Dependent variable: ln Gross Weekend Revenues.
Note: Constant not reported. Standard errors (clustered on the movie-level) in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.5 Robustness checks

4.5.1 Data quality: Availability on Megaupload

A potential issue could be that our source of *Megaupload* availability, *movie2k.to*, is not representative of the contents of *Megaupload*. As a result, our definition of the control group would be "contaminated" by false negatives, leading to estimation bias. The direction and severity of the bias would depend on the sign and significance of a correlation between movie type and wrong or missing *Megaupload* information.

We can think of three major reasons why our measure may not be entirely optimal. First, we don't observe information about a particular movie, because it is not possible to access the entire historical content of *movie2k.to* through the *Internet Archive*. Second, the movie was not listed on *movie2k.to* at all. Third, *movie2k.to* listed the movie, but didn't link to *Megaupload* although the movie was actually available.

To check the robustness of our results in this regard, we obtain similar historical information from a different source. Data used in Lauinger et al. (2013), which the authors shared with us, lets us observe all links to movies that were posted on the website *scnsrc.me* from

March 2009 to April 2012. In total, we observe 5,772 links corresponding to 3,828 movies, and whether these links point to *Megaupload*. We match this information to the box office revenue data using *IMDb* identifiers. It is noteworthy that the distribution of *Megaupload* links in the resulting sample (88%) is quite different to the *movie2k.to* sample (35%). Hence, it looks like the number of observations in the *movie2k.to* control group is indeed somehow inflated. Still, the parallel-trend assumption of the difference-in-differences model also holds with this data. Results in columns 3 and 4 of table A.1 are very similar to those obtained using data from *movie2k.to* in columns 1 and 2. In fact, although the point estimates are different, these differences are not statistically significant in the sense that 90% confidence bands overlap for all coefficients of interest. Hence, we conclude that data quality with regards to availability on *Megaupload* is most likely only a very minor issue.

4.5.2 Counterfactual: The Pirate Bay

An alternative explanation for our results could be that there is an unobserved event or process that triggers a shift in demand at the same time as the shutdown of *Megaupload* (e.g. seasonal fluctuation), and for some unobserved reason, this affects pirated movies, i.e. movies that were available on *Megaupload* in a different way than those that were not (e.g. because availability is correlated to box office success). To investigate this further, we run a counterfactual analysis by considering availability on a different pirating platform, and treating January 19th, 2012, the *Megaupload* shutdown date, as a placebo shutdown date for that website. If it is true that the shutdown coincided with some demand shock that is strongly correlated to the supply of pirated movies, we would expect to see the same results from this exercise. To do so, we collect information on availability on *BitTorrent* (a large peer-to-peer filesharing network) from the popular linking site *The Pirate Bay*.[19] For every movie in our sample (including country-specific titles) we obtain all links listed on *The Pirate Bay* along with the upload date. From this information we can construct an indicator of whether a particular movie has been available on the *BitTorrent* network from a given week onwards. We interact this variable with the post-shutdown dummy

[19] See `http://tinyurl.com/a68jwmz`.

to test whether the correlation between *BitTorrent* availability and movie revenues has changed after the shutdown of *Megaupload*.

Results reported in the first column of table A.2 show that this is not the case, i.e. that box office revenues of movies that were available on *The Pirate Bay* are not affected differently from movies that were not available, comparing pre- and post-intervention.

We further investigate whether consumers switched from *Megaupload* to *BitTorrent* after the intervention, which would result in a different shutdown effect of movies that were available on *Megaupload* as well as on *BitTorrent* compared to those that were only available on one of the two platforms. If this was true, we would expect a significant estimate of the three-way-interaction *After*Mega*PirateBay* in the second column of table A.2. The results show that this is not the case. Strikingly however, the estimated coefficient of *After*Mega* remains negative and significant, further supporting our baseline results.

Both tests combined clearly suggest that the observed effect is likely driven by the shutdown of *Megaupload* and not by other dynamics that affected the overall online piracy market at the same time. The underlying reason may be that users of services such as *Megaupload* are substantially different from users of services such as *The Pirate Bay* and *BitTorrent*. According to data from a representative sample of German individuals (GfK Media Scope, February 2011),[20] 80% of the consumers that use unlicensed services for movies and TV series consumption do so mainly via cyberlockers and streaming sites. Only 2% use mainly *BitTorrent*. The analysis of survey data from a representative sample of French Internet users in Arnold et al., 2014, Table 4 suggests that users of peer-to-peer networks (such as *BitTorrent*) are younger, poorer, and have less strong tastes for music and video compared to users of direct download services (such as *Megaupload*). As we have argued above, using *BitTorrent* is much less convenient to use, hence, it seems that users of *Megaupload* are more "casual" pirates, that encounter non-zero switching costs regarding a different (less convenient) service for their unlicensed consumption after the shutdown.

[20]See Studie zur digitalen Content-Nutzung 2011, `https://drive.google.com/file/d/0Bxe11iVXrXgsSjBGRFpqR2txVFk`.

5 Conclusion

We find that box office revenues reacted to the sudden shutdown of one of the main supply channels of unlicensed content, the cyberlocker *Megaupload*, in intricate ways. Specifically, the average movie reported *less* box office revenues after the shutdown. In further analysis, we find that the effect of the shutdown depends crucially on the breadth of a movie's release, i.e. the number of screens a movie on a movie's opening weekend. Movies on wide release, i.e. in the top decile of the number of opening screens, experience an increase in box office revenues, while movies in the bottom decile see a decline in box office revenues. We rule out a number of alternative explanations, which lets us pinpoint this effect on the nature of the movie (not type of movie publisher) and its availability on *Megaupload* and the associated drop in pirated supply following the shutdown.

A plausible mechanism for this differential effect is the relative importance of word-of-mouth as a marketing channel. As the breadth of a release is likely to be correlated to other movie characteristics such as the targeted audience (e.g. mainstream versus niche) and production and advertising budgets, publishers may accompany the wide release of a movie with a large budget for "conventional" advertising.

We provide two pieces of suggestive evidence. First, we show that the negative effect of less piracy is weaker for movies that are hit by the shutdown relatively early in their theatrical lifecycle – a period where firms invest most heavily in advertising. Second, we find that the effect of the *Megaupload* shutdown is less negative for remakes, compared to originals and sequels, i.e. for movies about which consumers are more likely to have more information. We cannot test the word-of-mouth mechanism more directly, as this would require micro-level data that lets us determine individual pre- and post-shutdown behavior. However, future research could model a word-of-mouth-effect in a structural empirical model in the spirit of Givon et al. (1995) and De Vany and Walls (1996, 2007).

It should be noted that we are studying the substitution patterns between one specific piracy channel and one licensed channel. This is in line with most of the recent literature which also capture only part of a movie's entire monetization chain (see, e.g. Danaher and Smith, 2014, and Zentner et al., 2013). Theatrical box office accounted for 36% of overall industry revenues in 2012, our study period, second to physical home video with 42%, while electronic home video contributed 21% to overall movie revenues (PwC Global Entertainment and Media Outlook, 2013–2017). Moreover, especially given that the movies we studied were still running in movie theaters, studying the substitution patterns between two concurrent channels (of which only one was licensed) generates important insights for the overall success of a movie, given that a movie's box office revenues are highly correlated to later monetization opportunities. Over time and for other distribution channels, we would expect the effect of word-of-mouth to subside.

It is interesting to compare our results to the findings by Danaher and Smith, 2014, who study the same question using data on (licensed) digital sales and rentals from three major motion picture studios. Their results show that revenues increased as a consequence of the *Megaupload* shutdown, while our paper finds a decrease in theatrical revenues for narrow-release movies. For a large majority of movies we do not find that theatrical revenues change in a statistically significant way. Even given these differences, the economic effect size found in Danaher and Smith (2014) is quite similar to what we find – at least when compared to our most conservative estimate. Danaher and Smith (2014) estimate the weekly increase in revenue to be 6.5 – 8.5%, which in dollar units is $222,000 – $285,000 for digital sales and $174,150 – $213,300 for digital rentals. Our most conservative estimate is an average weekly decrease of 4% or $30,000 for narrow release movies, and an average increase of 47% or $707,000 for wide release movies. Taking into account that the effect is not statistically different from zero for 80% of the movies in our sample, this results in an average effect for the entire sample of 4.3%.

Altogether, our results carry a number of implications. Most importantly, we have shown that heterogeneous products are affected differently by the availability of pirated substitutes. This suggests that firms may choose to protect their intellectual property more or less vigorously, depending on the balance of positive and negative effects they anticipate from piracy. As we have shown, these effects do not only depend on the type of product, but also on the product's stage of the life cycle. Hence, complementary effects of advertising and piracy need to be taken into account.

Moreover, we find that firm-level characteristics seem less important in determining the effect of piracy than product-level ones. For managers, this implies that positioning efforts and the associated IP strategy should take place at the product level and that, say, "Halo"-effects of an established brand are unlikely to substitute for the media-specific dynamics that arise from word-of-mouth promotional channels.

We contribute an alternative perspective to the emerging empirical literature on the effects of piracy. When online piracy has very different (even opposing) effects on heterogeneous products, that also depend on the distribution channel through which the products are being offered, blanket interventions aimed at reducing the negative welfare effects are difficult to implement because externalities may affect product variety and ultimately market structures.

Similar to Luo (2014), who finds that IP protection may affect different types of content producers differently in the market for ideas, we find that even post-release IP strategies can lead to different outcomes across heterogeneous products, and to different outcomes across distribution channels (Danaher and Smith, 2014). Hence, more research is needed to better understand how policies need to be designed across different dimensions and desired goals, which may include media diversity, sufficient niche content, but also overall welfare.

References

Adermon, A., and Liang, C.-Y. (2014). "Piracy and Music Sales: The Effects of an Anti-Piracy Law." *Journal of Economic Behavior & Organization*, *105*, 90–106.

Arnold, M., Darmon, E., Dejean, S., and Penard, T. (2014). "Graduated Response Policy and the Behavior of Digital Pirates: Evidence from the French Three-Strike (Hadopi) Law." *Working Paper*.

Belleflamme, P., and Peitz, M. (2012). "Digital Piracy: Theory." In M. Peitz, and J. Waldfogel (Eds.), *The Oxford Handbook of the Digital Economy*, 489–530, Oxford University Press, New York.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics*, *119*(1), 249–275.

Bhattacharjee, S., Gopal, R., Lertwachara, K., Marsden, J., and Telang, R. (2007). "The Effect of Digital Sharing Technologies on Music Markets: A Survival Analysis of Albums on Ranking Charts." *Management Science*, *53*(9), 1359–1374.

Chen, X., Chen, Y., and Weinberg, C. B. (2013). "Learning About Movies: The Impact of Movie Release Types on the Nationwide Box Office." *Journal of Cultural Economics*, *37*(3), 359–386.

Danaher, B., and Smith, M. (2014). "Gone in 60 Seconds: The Impact of the Megaupload Shutdown on Movie Sales." *International Journal of Industrial Organization*, *33*, 1–8.

Danaher, B., Smith, M. D., Telang, R., and Chen, S. (2014). "The Effect of Graduated Response Anti-Piracy Laws on Music Sales: Evidence from an Event Study in France." *Journal of Industrial Economics*, *Forthcoming*.

Danaher, B., and Waldfogel, J. (2012). "Reel Piracy: The Effect of Online Film Piracy on International Box Office Sales." *Working Paper*, SSRN–ID 1986299.

De Vany, A., and Walls, W. D. (1996). "Bose-Einstein Dynamics and Adaptive Contracting in the Motion Picture Industry." *The Economic Journal*, *106*(11), 1493–1514.

De Vany, A., and Walls, W. D. (2007). "Estimating the Effects of Movie Piracy on Box-office Revenue." *Review of Industrial Organization*, *30*(4), 291–301.

Eliashberg, J., Jonker, J.-J., Sawhney, M. S., and Wierenga, B. (2000). "MOVIEMOD: An Implementable Decision-Suppor System for Prerelease Market Evaluation of Motion Pictures." *Marketing Science*, *19*(3), 226–243.

Givon, M., Mahajan, V., and Muller, E. (1995). "Software Piracy: Estimation of Lost Sales and the Impact on Software Diffusion." *Journal of Marketing*, *59*(1), 29–37.

Gopinath, S., Chintagunta, P. K., and Venkataraman, S. (2013). "Blogs, Advertising, and Local-Market Movie Box Office Performance." *Management Science*, *59*(12), 2635–2654.

Hosanagar, K., Fleder, D., Lee, D., and Buja, A. (2014). "Will the Global Village Fracture Into Tribes? Recommender Systems and Their Effects on Consumer Fragmentation." *Management Science*, *60*(4), 805–823.

Joo, M., Wilbur, K. C., Cowgill, B., and Zhu, Y. (2014). "Television Advertising and Online Search." *Management Science*, *60*(1), 56–73.

Lauinger, T., Szydlowski, M., Onarlioglu, K., Wondracek, G., Kirda, E., and Kruegel, C. (2013). "Clickonomics: Determining the effect of anti-piracy measures for one-click hosting." *Proceedings of NDSS Symposium 2013, San Diego, CA*.

Lee, Y.-J., Hosanagar, K., and Tan, Y. (forthcoming). "Do I Follow My Friends or the Crows? Information Cascades in Online Movie Ratings." *Management Science*.

Luo, H. (2014). "When to Sell Your Idea: Theory and Evidence from the Movie Industry." *Management Science*, *60*(12), 3067–3086.

Manhanti, A., Carlsson, N., Arlitt, M., and Williamson, C. (2012). "Characterizing Cyberlocker Traffic Flows." *IEEE LCN Conference Proceedings*.

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

Oestreicher-Singer, G., and Sundararajan, A. (2012). "The Visible Hand? Demand Effects of Recommendation Networks in Electronic Markets." *Management Science*, *58*(11), 1963–1981.

Peitz, M., and Waelbroeck, P. (2006). "Piracy of Digital Products: A Critical Review of the Theoretical Literature." *Information Economics and Policy*, *18*(4), 449–476.

Prag, J., and Casavant, J. (1994). "An Empirical Study of the Determinants of Revenues and Marketing Expenditures in the Motion Picture Industry." *Journal of Cultural Economics*, *18*(3), 217–235.

Prosser, E. K. (2002). "How early can video revenue be accurately predicted?." *Journal of Advertising Research*, *42*(2), 47 – 55.

PwC Global Entertainment and Media Outlook (2013–2017). "Filmed Entertainment." http://tinyurl.com/ph6spgf.

Reimers, I. (2016). "Can Private Copyright Protection be Effective? Evidence from Book Publishing." *Journal of Law and Economics*, *59*(2), 411–440.

Roos, J. M. T., and Shachar, R. (2014). "When Kerry Met Sally: Politics and Perceptions in the Demand for Movies." *Management Science*, *60*(7), 1617–1631.

Sawhney, M. S., and Eliashberg, J. (1996). "A Parsimonius Model for Forecasting Gross Box-Office Revenues of Motion Pictures." *Marketing Science*, *15*(2), 113–131.

Vogel, H. L. (2004). *Entertainment Industry Economics – A Guide for Financial Analysis*. Cambridge University Press, Cambridge, 6th edn.

Zentner, A., Smith, M., and Kaya, C. (2013). "How Video Rental Patterns Change as Consumers Move Online." *Management Science*, *59*(11), 2622–2634.

Zhang, L. (forthcoming). "Intellectual Property Strategy and the Long Tail: Evidence from the Recorded Music Industry." *Management Science*.

Table A.1: Results: Different linking site

	movie2k.to		scnsrc.me	
After*Mega	-1.131*	-0.716	-1.438	-0.461*
	(0.655)	(0.443)	(0.892)	(0.238)
After*Mega*Wide Release		1.315**		0.888**
		(0.563)		(0.400)
After*Mega*Narrow Release		-1.895*		-2.716***
		(1.121)		(0.746)
Observations	7,731	7,731	12,039	12,039
$\overline{R^2}$	0.782	0.791	0.787	0.795

Dependent variable: ln Gross Weekend Revenues.
Note: Same specification as in tables 3 and 4, fixed effects and some coefficients not reported. Standard errors (clustered on the movie-level) in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Results: Counterfactual

	The Pirate Bay			
	Placebo Shutdown		Substitution	
After	0.662**	(0.296)	0.975***	(0.230)
PirateBay	0.202	(0.390)	0.208	(0.410)
After*PirateBay	-0.443	(0.376)	0.076	(0.597)
After*Mega			-0.571*	(0.291)
Mega*PirateBay			-0.046	(0.726)
After*Mega*PirateBay			-0.579	(0.716)
Weekend Days	-1.373***	(0.169)	-1.349***	(0.164)
ln Weeks Active	-1.410***	(0.051)	-1.402***	(0.051)
Year Effects	Yes		Yes	
Week Effects	Yes		Yes	
Country Effects	Yes		Yes	
Movie Effects	Yes		Yes	
Observations	7,731		7,731	
$\overline{R^2}$	0.780		0.782	

Dependent variable: ln Gross Weekend Revenues.
Note: Constant not reported. Standard errors (clustered on the movie-level) in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$