

Do Search Engines Influence Media Piracy? Evidence from a Randomized Field Study

Liron Sivan, Michael D. Smith, Rahul Telang
{lsivan@andrew.cmu.edu, mds@cmu.edu, rtelang@andrew.cmu.edu}

School of Information Systems and Management
Heinz College, Carnegie Mellon University

September 2014

Available From: <http://ssrn.com/abstract=2495591>

Acknowledgements: Author names are in alphabetical order. We thank Itai Ater, Felix Oberholzer-Gee, David Reiley, Hal Varian, and participants at the National Bureau of Economic Research Summer Institute, the University of California at San Diego, Carnegie Mellon University's Information Networking Institute, and University of Texas McCombs School of Business for valuable comments on this research. This research was conducted as part of Carnegie Mellon University's Initiative for Digital Entertainment Analytics (IDEA), which receives unrestricted (gift) funding from the Motion Picture Association of America. This research was conducted independently without any oversight or editorial control. All findings and errors are entirely our own.

Do Search Engines Influence Media Piracy? Evidence from a Randomized Field Study

Abstract

There is an ongoing public policy debate regarding the role search engines can play in the fight against intellectual property theft. However, for any sensible policy discussion it must first be the case that changes in search results can influence consumers' decisions to pirate or purchase legally. Surprisingly, there is little empirical evidence on this important policy issue.

To analyze this question, we design a customized search engine and conduct experiments on a general population of users and on college-aged users where we manipulate the positions of infringing and legal sites in users' search results.

Our data show that users are more likely to purchase legally when legal links are promoted in search results, and users are more likely to pirate when pirate links are promoted. Together our results suggest that reducing the prominence of pirated links can be a viable policy option in the fight against intellectual property theft.

1. Introduction

While the Internet has created many new opportunities for media companies to reach consumers through digital channels, it has also given rise to widespread media piracy. The potential impact of piracy on sales and overall social welfare has given rise to a variety of anti-piracy efforts initiated by governments and industry participants. However, these efforts are complicated by the fact that piracy is a multifaceted phenomenon with multiple stakeholders. Specifically, Internet intermediaries such as Internet Service Providers (ISPs), web-hosting services, and search engines all play a role in delivering information and content to the end user. While end users are ultimately responsible for their content consumption, many in the creative industries argue that Internet intermediaries should play a more active role in reducing piracy.

One active policy question is whether search engines can, and should, play a role in the fight against piracy by reducing the prominence of pirated links in search results displayed to users. Not surprisingly, this issue has strong views on both sides. Search engines argue that they are already doing a great deal to stop piracy. For example, in September 2013, Google released a report titled “How Google Fights Piracy”¹ which detailed Google’s efforts to reduce piracy by removing infringing links from their search results based on notices sent by copyright owners. This report also notes that in August 2012 Google started taking valid copyright removal notices into account when ranking sites in search results.² In fact, Google’s “Transparency Report” identifies more than 25 million copyright removal notices that they responded to in June 2012, affecting nearly 40,000 unique Internet domains.³

At the same time, many content owners argue that search engines’ efforts have been ineffective and that more is needed. For example, a recent study funded by the MPAA⁴ analyzed the role that search results play in the discovery of television and movie content, arguing that current efforts by search engines to reduce the availability of pirated content are insufficient and that most of the search queries that result in viewing infringing links do not contain keywords that indicate specific intent to view this content illegally.

¹ How Google Fights Piracy. 2013.

(<https://docs.google.com/file/d/0BwxyRPFduTN2dVFqYmISUENUeUE/edit>)

² See also Google’s press release on this program: <http://insidesearch.blogspot.com/2012/08/an-update-to-our-search-algorithms.html>

³ See <http://www.google.com/transparencyreport/removals/copyright/>

⁴ Understanding the Role of Search in Online Piracy. Prepared by Millward Brown Digital for the MPAA. (<http://www.mpaa.org/resources/38bc8dba-fe31-4a93-a867-97955ab8a357.pdf>)

The key question though is whether pirate and legal results are sufficiently close substitutes in the eyes of consumers that making one harder to find in search results will increase consumption of the other. In an editorial commenting on the RIAA's desire for Google to make piracy harder to find, the technology blog *Techdirt* summarizes the view of many in the technology industries on this question:

"The RIAA might not like it, but the simple fact is that when people are searching for [artist] [track] mp3 and [artist] [track] download, chances are they're not looking to *buy*, but to download for free. ... Even if Google magically did show them Apple, Amazon and Emusic as the top results for every [artist] [track] mp3 and [artist] [track] download, the people doing those searches *wouldn't go there*, because they're not looking to buy."⁵

In this view if a user is looking for infringing content, they have already made up their mind about whether to consume legally, and minor changes in search results will not change that behavior. Thus, while it is well known that when users are choosing between close substitutes changing the order of search results can change what sites users select, in our setting pirate and legal links may be sufficiently differentiated that users looking for pirate content will not substitute toward legal links and vice versa. In the words of the technology blog, ExtremeTech:

"In general, if someone wants to pirate something, they're going to pirate it, even if they have to click through a couple of pages of search results."⁶

Thus, while many copyright holders have called for Google and other major search engines to reduce the prominence of pirate links in search results, a fundamental question remains: Are users' media consumption decisions for infringing versus legitimate sources influenced by search results? In particular, are users more or less likely to consume content from an infringing source if search results increase or decrease their exposure to infringing content? If anything, these

⁵ "RIAA: Google Isn't Trying Hard Enough To Make Piracy Disappear From the Internet," Mike Masnick, February 21, 2013. Available from <https://www.techdirt.com/articles/20130221/07560622055>, last accessed November 24, 2014. Italics in original text.

⁶ "Google finally decides to demote 'notorious' piracy sites in search results," Sebastian Anthony, October 20, 2014. Available from <http://www.extremetech.com/extreme/192471-g>, last accessed November 24, 2014.

questions have become more important recently, given Google's recently announced changes to its search algorithm to more aggressively demote pirated links.⁷

However, despite the importance of these questions for both managers and policymakers, we are aware of no studies that empirically analyze the role of search results on the consumption of legal versus infringing content. One reason for the few studies in this area may be that assigning causation is tricky in this context. This is because, by design, the top results are also likely to be the most "relevant" results to the user. If a user is searching for infringing content and sees infringing results listed first, it is impossible to disentangle whether she clicks on that content because of her interest or because of its placement. In short, a user's desire to look for particular content shapes their search terms and subsequent behavior, making it impossible to use observational or archival data alone to show whether search results influence user choices.

In this paper, we address this methodological challenge through randomized field experiments, using the experimental design to avoid contamination due to user intent. We then implement this design in two experiments. The first experiment uses participants drawn from a representative panel of the U.S. population, and the second uses a panel of college-aged participants. In both cases, the participants are recruited through an independent company maintaining large survey panels. We expose these users to a potential task of finding a movie through online channels, and encourage users to use our custom-built search engine in place of the search engine they would normally use (e.g., Bing, Google).

In the first experiment (representative panel), our search engine displays results to users in one of three randomly assigned conditions: in the control condition users see the same results that would be displayed from a major search engine, in the first treatment condition infringing links are artificially promoted in the search results, and in the second treatment condition legal links are artificially promoted in the search results. The second experiment (younger users) adds two additional treatment conditions to test milder legal and milder infringing search manipulations.

We avoid the user intent fallacy by randomly assigning users to one of these treatment conditions. We also record and examine their search choices through our interface, providing additional detail on their behavior within the experiment. Finally, we ask them to complete a questionnaire at the

⁷ See for example <http://googlepublicpolicy.blogspot.com/2014/10/continued-progress-on-fighting-piracy.html>, last accessed December 9, 2014.

end of the experiment to measure their attitudes regarding piracy. Our results suggest that the prominence of search results can play an important role in users' subsequent choices about whether to pirate content or consume through legal channels. In the control condition about 80% of users choose to buy the product legally versus 94% who choose to purchase legally in the legal treatment condition and 56% in the infringing treatment condition.⁸

We see similar results in a second experiment with a college-aged population (18-24 year old users): 62% of users in the control condition choose to purchase the product legally, versus 92% in the legal treatment condition and 39% in the infringing treatment condition. Our second experiment also shows that stronger treatments (in terms of number of pirated or legal links) lead to stronger outcomes: In the "mild" legal treatment the number of purchases drops to 76% (versus 92% in the regular legal treatment), and in the "mild" infringing treatment the number of purchases rises to 48% (from 39%).

Our data also allow us to infer user intentions from their initial search terms. When we classify user intentions based on the content of their initial search terms we find that users who initially express an intention to consume legally are less likely to do so in the infringing treatment condition than in other conditions, and conversely that users who initially express an intention to consume pirated content are less likely to do so in the legal treatment condition than in the other conditions.

Together, our results suggest that user behavior is elastic to search results when it comes to media consumption decisions, and that reducing the prominence of pirated links in search results can be a viable strategy for fighting intellectual property theft for both a general population of users and for younger (college-aged) users and for both "undecided" users and users with a pre-existing preference for legal/pirated content.

2. Related Literature

Our research primarily relates to studies in the economics and information systems literatures analyzing the impact of piracy on sales and analyzing the effectiveness of anti-piracy measures. Within these literatures, the vast majority of papers find that piracy harms sales. For example, Rob and Waldfogel (2006) show that overall piracy reduced per capita expenditures on music by

⁸ As we note in more detail below, our experiment is designed to identify relative differences between the two treatment conditions and the control condition (as opposed to absolute levels of piracy).

about 20% and Zentner (2006) shows that piracy reduced music sales by about 7.8% in 2002 (for reviews of the literature, see Liebowitz (2008), Oberholzer-Gee and Strumpf (2010), and Danaher, Smith, and Telang (2014)).

With piracy's impact on sales well established in the literature, several recent papers have analyzed the effectiveness of regulatory and industry-led efforts to reverse the impact of piracy. These papers generally find that making content available in legal digital channels reduces the incidence of piracy (see, for example Danaher et al. (2010) and Danaher et al. (2014)) and that anti-piracy policies targeting the demand-side of piracy (e.g., the HADOPI anti-piracy law in France as studied by Danaher et al. (2014)) and the supply-side of piracy (e.g., the shutdown of Megaupload studied by Danaher and Smith (2014)). Beyond these legislative interventions, Reimers (2014) shows that industry-led notice and takedown strategies for eBooks can be effective at increasing legal sales, and Bhattacharjee et al. (2006) find that RIAA lawsuits lowered levels of file sharing, with a much greater impact on high-level sharers.

Another important question is whether piracy impacts social welfare. Here the results are less conclusive. On one hand, Abhishek, Telang, and Zhang (2013)) use analytical models to show that firms' incentives to invest decrease in the presence of piracy, especially when users do not readily migrate back to legal channels. Further, using the Indian motion Picture Industry, Telang and Waldfogel (2014) find that media piracy had a strong negative impact on movie production. However, Waldfogel (2012) finds that the overall impact of technological change (including piracy) from 1980 to 2010 did not reduce the quantity or the overall quality of music released.

Our paper extends these results by studying anti-piracy efforts in the context of search results, and specifically whether reducing the prominence of pirated links in search results can be effective at changing user behavior and user satisfaction with search results.

In addition to the piracy literature, our research relates to the academic literature on consumer behavior at search engines and overall consumer search costs online. Previous studies of consumer behavior at search engines examined how the prominence of search results influence user behavior, showing that the position of search results can have a significant impact on click-through rates and conversion rates for both sponsored search results (e.g., Yang and Ghose 2010, Agarwal et al. 2011) and organic search results (e.g., Baye et al. 2012).

Specifically in the context of organic search results, Brooks (2004) showed that searchers are more likely to click links that are placed higher in the results listing. This is because the top results are perceived to be more relevant and because the search costs of exploring results with lower rank is relatively high. Similarly, an eye tracking experiment performed by Pan et al. (2007) revealed that college students have substantial trust in Google's ability to rank results by their true relevance, such that users' decisions are strongly biased towards links higher in position even if the abstracts themselves are less relevant. Further, using click-through observational data, Chesnes, Dai, and Jin (2014) study how the ban of non-NABP-certified pharmacies from Google's sponsored search listings affects consumer search on the Internet, finding that non-NABP-certified pharmacies receive fewer clicks after the ban than they did before the ban.

An important difference between these studies and our study is that our setting analyzes how search costs affect users' substitution pattern between legal content with infringing content. Thus our study relates to an active policy question, while also addressing a setting where the tradeoffs between different (legal and pirate) links could be quite high. Unlike most other studies which focus on organic search results or sponsored links, the perceived difference between legal content and infringing content could be large, particularly for consumers who may have a stated preference for one type of content or the other.

3. Experiment 1 — General Population, 2 Treatment Conditions

Experimental Design

In this experiment we utilize a custom-built search engine to test the impact of search results on media consumption choices among the general population of Internet users. The participants for the experiment were recruited from the general population by an independent company maintaining a large survey panel. The study was performed online and consisted of three parts: (1) a screening phase in which the participants were asked to choose a movie they wish to acquire; (2) a search phase in which the participants were asked to use our custom search engine to search for a source to acquire the movie; and (3) a post-experiment questionnaire.

In order to determine if you qualify to participate in this experiment, you are asked to answer the following screening question.

Out of the following list of movies, is there a movie you wish to watch and have not watched yet (and do not currently own, physically or digitally)? If your answer is 'Yes', choose the movie you wish to watch the most. If your answer is 'No' Choose the last option ('None of the above').

* The data in the following table are taken from IMDB.com

Movie	Genres	Storyline	Average Rating	Release Date
<input type="radio"/> 1. The heat 	Action Comedy Crime	Sarah Ashburn, an FBI agent, is extremely ambitious and has her eye on a promotion, but she doesn't get along with her co-workers. She is sent to Boston to uncover the identity of an elusive drug lord, Mr. Larkin, by tracking down his proxy, Rojas, and is told that she'll have a good shot at the promotion if she finds Larkin. When she arrives in Boston, she learns that Larkin has been eliminating his competition and taking over their operations. She learns that Rojas is in Boston PD custody and goes to see him to ask him what he knows about Larkin, but is warned that the cop who arrested Rojas, Shannon Mullins, is very territorial, and she is not exactly sociable. When the two meet they don't get along. When Mullins learns why Ashburn is in Boston, she decides to find Larkin herself. Ashburn is told by her boss to work with Mullins, but it won't be easy because Ashburn does things by the book while Mullins does things her way.	6.6/10	28 June 2013
<input type="radio"/> 2. The internship 	Comedy	Billy (Vince Vaughn) and Nick (Owen Wilson) are salesmen whose careers have been torpedoed by the digital world. Trying to prove they are not obsolete, they defy the odds by talking their way into a coveted internship at Google, along with a battalion of brilliant college students. But, gaining entrance to this utopia is only half the battle. Now they must compete with a	6.3/10	7 June 2013

Figure 1: Screening Question. The list of movies also included data on the Genre of the movie, its storyline, average reviewers rating, and its release date to the movie cinemas. These additional data were taken from IMDB.com, and the movies were chosen from IMDB's weekly top DVDs lists during the months prior to the experiment.

During the screening phase the participants were asked to choose a movie they wish to watch from a list of 50 alternatives (see Figure 1). We asked users to select a movie from a specific list of movies in advance to make sure that our manipulation conditions could be implemented. We asked users to select a specific movie they were interested in (as opposed to assigning them a movie) to ensure that the users were motivated to find a movie they wished to view. Consistent with our goal of observing user behavior in search for movies, we excluded from the experiment any participants who stated that they are not interested in watching any of the movies in our list.

Participants who chose a movie were given a \$20 prepaid virtual Visa card as compensation for their time in completing the experiment. The payment details were given to the participants before they were presented with the task in order for them to think of the money as if it was their own. The money could be spent however the participant chose and could be retained after the experiment (i.e., the money was not limited for use in purchasing the movie they selected, and could be used for purchases in any store that accepts Visa cards).

The participants were then given a task to search for a source from which they would like to acquire (download, stream, purchase, or rent) the movie. The participants were not told that we were examining their piracy behavior. Instead, the experimental instructions stated that the user was participating in an experiment to test the effectiveness of our search engine. We highlighted that our search engine was designed to efficiently search for movie related content (see Figure 2). Therefore, they were instructed to use our search engine in place of Google, Bing or other search engines when they searched for content. The instructions also stated that if the participant already had a specific website in mind to obtain the movie, they could go directly to that site after initially trying to search for it using our search engine. The participants were able to keep the movie they acquired, and keep any money left over after the task was completed. We also stated that their identity is unknown to us and their behavior cannot be traced back to them.

We have designed a search engine which is especially efficient at searching content related to movies (URL: search.co.be). The goal of this experiment is to understand if (and how) people use search engines to search for movie content.

As part of the experiment you are asked to find a website (source) where you would like to get this movie. Please use [our search engine](http://search.co.be) (do not use Google, Bing, etc.) to search on the Internet. You can also use your own installed programs (if you have any) to reach the desired source or directly go to the specific source to acquire the movie (by entering appropriate URL on a browser for example), but please make sure to search with [our search engine](http://search.co.be) first. Once you reach the desired source, you should download/stream/purchase/rent the movie. If you need to purchase the movie you can use the provided virtual credit card to do so.

Figure 2: Experiment Task.

Before the experiment started, each participant was randomly assigned to one of three search engine conditions: no manipulation, non-infringing (legal) content manipulation, or infringing (piracy) content manipulation.

- Condition 1 – No manipulation: In the control condition, the first 100 search results were retrieved from a major search engine and were displayed to the searcher without any manipulations. The search results consist of 10 pages with 10 search results on each page.
- Condition 2 – Legal content manipulation: In the legal treatment condition, the first 100 search results were retrieved from a major search engine and were displayed to the searcher such that the first 3 results on each of the 10 pages were replaced (if necessary) with results that offer legal options to rent/purchase the movie. Additionally, no infringing links appeared on the first page of the search results and infringing links that would have appeared in these

positions were replaced with legal options such as Amazon.com or iTunes. Neutral results in positions 4-10 (such as IMDB.com, Wikipedia.com, etc.) were left unchanged. See Figure 3 for sample search results in the legal treatment condition.

- Condition 3 – Piracy content manipulation: In the piracy treatment condition, the first 100 search results were retrieved from the major search engine and were displayed to the searcher such that the first 3 results on each of the 10 pages are replaced (if necessary) with results that offer infringing (piracy) options to download/stream the movie. Additionally, no legal options appeared on the first page of the search results and any legal links that would have appeared in these positions were replaced with piracy options such as piratebay.com and torrentz.com. Neutral results in positions 4-10 (such as IMDB.com, Wikipedia.com, etc.) were left unchanged.

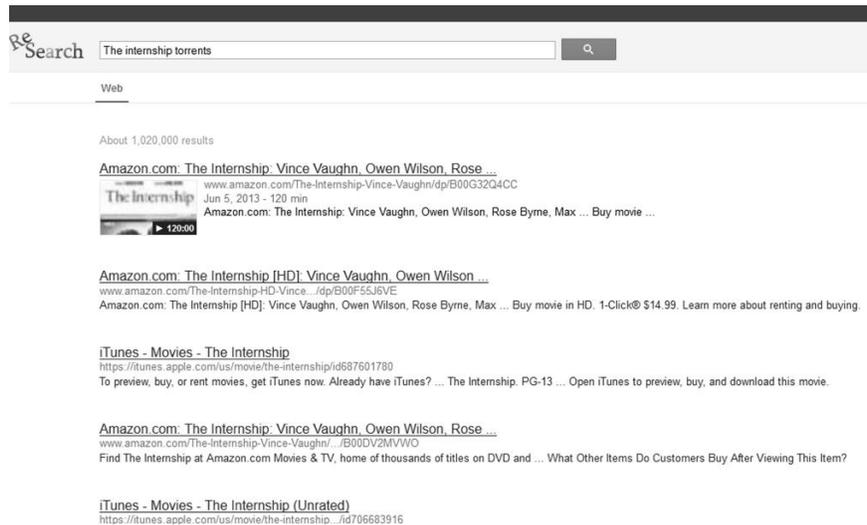


Figure 3: Example Search Results for the Legal Content Manipulation Condition.

After the participants completed the experiment task, they were required to complete a post-experiment survey. In the questionnaire, they were asked about the source from which they acquired the movie, its price (\$0 if the movie was acquired from an infringing source), and they were asked several demographic questions and questions about their media consumption preferences (see Figure 4 for sample post-experiment questions).

4. What website have you used to download/stream/purchase the movie?

5. What was the price of the movie (write \$0 if you've downloaded it for free)?

6. Were you concerned about being monitored?

- No, it had no effect on my choices or behavior.
- I thought about it, but I believe it had no effect on my behavior or choices.
- I was concerned about it and it affected my behavior and choices.

Demographics:

7. What is your gender?

- Male
- Female

8. How old are you?

- 18-21
- 22-25
- 26-30
- 31-40

Figure 4: Post-Experiment Questionnaire.

It is important to note that we can only directly observe user behavior while they are on our search engine: we do not observe their behavior outside of the search engine. Thus, our observation of where the user obtained their content is based on their survey answers. While this may introduce some recall bias, the degree of recall bias should not vary across the control and treatment conditions and the survey occurs immediately after the participant obtains the content, significantly reducing the overall possibility of recall bias. It is also possible, based on prior studies that found that survey respondents are likely to under-report socially undesirable activities, that participants who choose a pirated option could be less likely to reveal that in a survey than other users are (Cannell et al. 1965; Warner 1978; Wyner 1980; and Means et al. 1992). However, again the propensity to misreport should not vary across the control and treatment conditions and any bias would lead to an underestimate of the degree that infringing links induce more piracy. We also note that we can use observed search and clicking behavior to validate a user's survey answers (including in some cases verifying price). In the results section we show that the users' survey responses were very consistent with their observed behavior.

In order to make sure that participants understood the task correctly, we included in our analysis only those users who claimed the \$20 virtual Visa card, used our custom-built search engine, and reported the source from which they acquired the movie and its price. Approximately 1,000 participants were invited to participate in the study. Of these participants, 770 met the initial qualifications for participating in the study by expressing an interest in watching one of the movies on our list. These participants were invited to participate in our

study and 632 of them logged into the system and were presented with the task details. Out of these participants, 235 completed the task as instructed, and 196 (66 men, 130 women) qualified for our study by also completing the post-experiment questionnaire. The dropout rates across each of the different stages described above are not statistically different across experimental conditions.

Prior to analyzing the results of the experiment we confirmed that the distribution of user characteristics (demographic characteristics, attitudes towards piracy) is similar across the control and treatment groups. The average values for each experimental condition are presented in Table 1. Chi squared tests show that there are no statistically significant differences in demographic characteristics or attitudes toward piracy between the experimental conditions (p -value > 0.05). This confirms that our randomization worked as intended.

	% Women	Average age group	Average household size	Average household income group	Average attitude against piracy (Likert scale)	% Downloads infringing
Condition 1: No manipulation	70.0%	4.20	2.85	2.82	4.60	33.3%
Condition 2: Legal content manipulation	59.2%	4.30	2.85	2.77	4.75	39.4%
Condition 3: Infringing content manipulation	64.6%	3.97	2.68	3.02	4.58	44.6%

Table 1: Between-Conditions Comparison of demographic characteristics and attitude towards piracy. There are 7 possible age group values in the questionnaire: 1 (18-21), 2 (22-25), ..., 7 (61 and over); 10 possible household size values: 1, 2, ..., 9, 10 or more; and 6 possible household income group values: 1 (less than \$30,000), 2 (\$30,000 - \$50,000), ..., 6 (over \$150,000).

To further verify similarity in pre-existing attitudes toward piracy, we tested whether the participants' initial intent for pirate or legal content is similar across the three experimental conditions. We did this by classifying their initial search terms reflected neutral, legal, or infringing intent. We did this classification by analyzing the degree to which pirate or legal links were present in the (unmodified) search results for queries commonly issued by our users. We discovered that search terms using only the movie's name contained almost exclusively "neutral" results (i.e., results that neither promote legal or pirate sources), and thus we classify these searches as "neutral." However, when search terms included the words "buy," "rent," or

“purchase” the search results contained 38% more legal links than pirate links, and when the search term contained a legal domain name (e.g., amazon), the search results contained 78% more legal links than pirate links. Thus, we classify these search terms as representing “legal” intent.

Conversely, when search terms included the words “download,” “stream,” or “full movie” there were 33% more pirate links in the search results than legal links, and including the domain name of an infringing site (e.g., piratebay) resulted in search results that included 65% more pirate links than legal links. Because of this, we classify these search terms as representing “infringing” intent.

We then classify intent based on the initial intent reflected in each user’s search terms (or “neutral” if the user did not express intent in their searches). As above, reassuringly, the distributions are not statistically different across the control and treatment conditions (see Table 2 for frequencies across conditions).

	N	Neutral Searches	Legal Intent	Infringing Intent
Condition 1: No manipulation	60	17	31	12
Condition 2: Legal content manipulation	68	28	23	17
Condition 3: Infringing content manipulation	65	21	33	11

Table 2: Between-Conditions Comparison of the initial intent (based on the first keyword each user entered).

Finally, we note that while the characteristics of users who participated in the experiment were similar across the control and treatment conditions, it is possible that our participant pool skews toward more media or tech savviness than the general population. As such, in the results below one should focus on the relative levels of sales/piracy between the control and treatment conditions as opposed to the absolute levels within any particular condition.

Results

Table 3 compares the proportion of legal purchases (and the average price of a legal purchase) across the different treatment conditions. This table shows that in the control condition, where the search results are not manipulated, 80% of our participants chose to acquire the movie through a legal channel (and the remaining 20% consumed through a pirated channel). Relative

to this baseline, participants in the legal content treatment condition were significantly more likely to acquire the movie from a legal channel ($P = 94.37\%$, $SD = 23.22\%$) than were participants in the baseline condition ($P = 80\%$, $SD = 40.34\%$). A t-test ($t(121) = 2.85$, $p < .01$) confirms that these differences are statistically significant.

Conversely, the results show that participants in the infringing content treatment condition were significantly less likely to acquire content from a legal channel ($P = 56.92\%$, $SD = 49.90\%$) relative to the baseline condition ($P = 80\%$, $SD = 40.34\%$). Again, a t-test confirms that the differences are statistically significant ($t(91) = 2.43$, $p < .01$).

	n	# Legal option	% Legal option	Average price (conditional on acquiring a legal copy)
Condition 1: No manipulation	60	48	80.0%	\$9.80
Condition 2: Legal content manipulation	71	67	94.4%	\$9.89
Condition 3: Infringing content manipulation	65	37	56.9%	\$9.93

Table 3: *Between-Conditions Comparison of the Proportion of Purchases Made from Non-Infringing Websites.*

These results, and the statistically significant differences across groups, strongly suggest that reducing the prominence of pirated content in search results can have a significant impact on user behavior, and thus is a viable anti-piracy strategy. This result may be initially surprising given that our treatment doesn't remove all pirated or legal links, just the ones appearing on the first page and on the first three positions on subsequent pages. However, the fact that our users are willing to substitute between legal and infringing options with even minor changes in the rank of these options suggests that users in our sample are relatively willing to substitute between legal and infringing consumption channels.

To further explore user's willingness to substitute between channels, we analyze whether changing the prominence of pirated links can affect the behavior of users who state an initial preference for legal or pirated content. This is important because these users might be considered among the most "committed" to consume through legal or pirated channels, and thus among the least likely to change their behavior.

We do this by following the methodology described above (see Table 2 and surrounding text) to classify each user’s “intent” based on whether their initial searches express a preference for legal or illegal content. In Table 4, we compare how the purchase likelihood for users with pirated or legal intent varies across the different treatment groups.

Each cell in the table reports the number of users observed in a particular experimental and “initial intent” condition, the number of those users who made a legal purchase, and the proportion of legal purchases. For example, Table 4 shows that 33% (4 of the 12) “control group” users who initially expressed intent to consume through infringing channels, ultimately purchased through a legal channel.

First search term	Control group	Legal content manipulation	Infringing content manipulation
Legal intent	31/31 (100%)	22/23 (96%)	24/33 (73%)
Infringing intent	4/12 (33%)	15/17 (88%)	5/11 (45%)

Table 4: Legal purchase rates across treatment conditions and initial intent

This table shows that users who initially express intent to consume legally are significantly *less* likely to do so when placed in the infringing content manipulation (73%) than in either the legal (96%) or control (100%) conditions (t-tests: $t(48) = 2.55, p < 0.01$; $t(32) = 3.46, p < 0.01$, respectively). Likewise, users who initially express intent to consume infringing (pirated) content are significantly *more* likely to purchase legally in the legal content manipulation (88%) than in the infringing (45%) or control (33%) conditions (t-tests: $t(15) = 2.42, p < 0.05$; $t(18) = -3.36, p < 0.01$, respectively). These results suggest that even users with an initial preference for pirate or legal channels have a relatively high substitutability between channels based on the ranking of search results.

While the tests reported above are sufficient to determine differences between the control and treatment groups based on our experimental manipulations, we can also use a logistic regression model to control for and analyze differences between groups based on observed characteristics. Specifically, we use the following logistic regression model to control for observable participant characteristics:

$$(3) \quad \log \frac{PR(Legal_i)}{1 - PR(Legal_i)} = \alpha + \beta_1 \cdot NI_i + \beta_2 \cdot I_i + \sum_{j=3}^5 \beta_j \cdot DC_i^j + \sum_{j=6}^8 \beta_j \cdot MCP_i^j + \sum_{j=9}^{10} \beta_j \cdot ATP_i^j + \beta_{11} \cdot \text{intent} + \varepsilon_i$$

where, $Legal_i$ denotes whether the movie was acquired from a non-infringing source; NI_i is an indicator variable denoting whether participant i was treated with the non-infringing search condition; I_i is an indicator variable denoting whether participant i was treated with the infringing search condition; $\sum_{j=3}^5 \beta_j \cdot DC_i^j$ includes the following demographic characteristics: Gender (an indicator variable for whether the participant was a woman), Age ≤ 40 , Household size, and Income; $\sum_{j=6}^8 \beta_j \cdot MCP_i^j$ includes the following media consumption preferences: Time online (the average hours spent online per day), Acquires movies online (whether the participant ever downloaded or streamed a movie, including pirated movies), and Movies per year (the number of movies the participant watched in the last 12 months); $\sum_{j=9}^{10} \beta_j \cdot ATP_i^j$ includes the following attitude towards piracy variables: Against piracy (on a 6-point Likert Scale ranging from (1) 'There is nothing wrong with it' to (6) 'It is the same as stealing'), Downloads infringing (whether the participant indicated that that s/he uses torrents or other free online downloads/streaming/file-sharing); and *Non – infringing intent* _{i} (whether the first search term that the user entered indicate that his/her intent is to acquire a legal copy of the). We present the results of this model in Table 5 below.

These results are consistent with our means comparison results above in that the treatment variables are statistically different from the control condition and that they have the expected sign (the legal treatment condition increases the likelihood of purchasing legally and the infringing treatment condition decreases the likelihood of purchasing legally). These results also confirm, as one would expect, that participants who consumed infringing content in the past are less likely to purchase the movie legally and that those who use a search term that implies legal intent are more likely to purchase legally.

In the next section, we report the results of a similar experiment that uses college-aged participants and varies the intensity of the search manipulations.

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Basic	Including	Including	Including	Including
Acquired legally	Model	Demographic	Media	Attitude	Intent to
		Characteristics	Consumption	Towards	Acquire
			Preferences	Piracy	Legally
Constant	1.386*** (0.323)	1.637** (0.663)	-0.333 (1.288)	0.148 (1.511)	0.0575 (1.533)
Non-infringing mode	1.432** (0.608)	1.458** (0.613)	1.606** (0.634)	1.855*** (0.673)	1.905*** (0.685)
Infringing mode	-1.108*** (0.409)	-1.152*** (0.422)	-1.232*** (0.438)	-1.235** (0.507)	-1.261** (0.516)
Woman		-0.00896 (0.406)	0.0163 (0.416)	0.0167 (0.482)	0.0421 (0.491)
Age <= 40		-0.603 (0.398)	-0.744* (0.436)	-0.669 (0.491)	-0.492 (0.500)
Household size		-0.134 (0.146)	-0.216 (0.151)	-0.0772 (0.186)	-0.106 (0.193)
Income		0.126 (0.142)	0.132 (0.147)	-0.0719 (0.166)	-0.0333 (0.173)
Time online			0.0287 (0.213)	0.235 (0.249)	0.169 (0.250)
Acquires movies online			0.191 (0.231)	0.453* (0.275)	0.412 (0.282)
Movies per year			0.409** (0.178)	0.229 (0.210)	0.255 (0.213)
Against piracy				0.554 (0.486)	0.404 (0.504)
Downloads infringing				-2.424*** (0.516)	-2.410*** (0.523)
Non-infringing intent					1.266** (0.621)
Number of obs.	196	196	196	196	196
Pseudo R2	0.1392	0.1611	0.1933	0.3529	0.3758

Table 5: Logistic regressions results
Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

4. Experiment 2 — Younger Audience, 4 Treatment Conditions

Experimental Design

We obtained participants for the second experiment from the same independent survey panel company used in Experiment 1. However, in experiment 2, we limited participation to users who were 18-25 years old and who were college students at the time of the experiment or who had

at least 2 years of college education. We did this to analyze the degree to which our results extend to a college-aged population, a population that tends to be disproportionately more likely to pirate than other demographic segments are (e.g., Rob and Waldfogel 2007, Vandiver et al 2012).

The second difference between our first and second experiments is that our second experiment adds two additional treatment conditions. We retain our original “legal” and “infringing” manipulations that, respectively, remove all pirated or legal options from the first page of results and from the first three results displayed on subsequent pages. However, in experiment 2 we add two additional manipulations that soften the impact of these manipulations by only changing the top three results on the first page and making no other changes to the displayed results. Specifically, the two additional treatment conditions are:

- Condition 2a – Mild legal content manipulation: The first 100 search results were retrieved from a major search engine and were displayed to the searcher such that the first 3 results on the first page (and only the first page) were replaced (if necessary) with results that offer legal options to rent/purchase the movie.
- Condition 3a – Mild piracy content manipulation: The first 100 search results were retrieved from the major search engine and were displayed to the searcher such that the first 3 results on the first page (and only the first page) were replaced (if necessary) with results that offer infringing (piracy) options to download/stream the movie.

Data and Results

Following our approach in experiment 1, in our second experiment we included in our analysis only those users who claimed the \$20 virtual Visa card, used our custom-built search engine for at least one search, and reported the source from which they acquired the movie and how much they paid for it. A total of 234 participants (86 men, 148 women) qualified for our study.⁹ In Table 6 we report average statistics for these participants across the manipulation conditions.

⁹ Approximately 650 participants that were invited to participate in the study and expressed an interest in watching one of the movies in the study. A total of 550 participants logged into the system and were presented with the task details. Out of these participants 270 completed the experimental task as instructed, and of these participants 234 (86 men, 148 women) qualified for our study by also completing the post-experiment questionnaire. The dropout rates in each of the different stages described above are not statistically different across experimental conditions.

These statistics show that the distribution of demographic characteristics and attitudes towards piracy are similar across the control and treatment groups, as one would expect given the experimental assignment.

	% Women	Average age group	Average household size	Average household income group	Average attitude against piracy (Likert scale)	% Downloads infringing
Condition 1: No manipulation	57.14%	1.69	3.02	2.60	3.40	80.95%
Condition 2a: Mild legal content manipulation	73.47%	1.71	3.18	2.35	3.73	69.39%
Condition 2: Legal content manipulation	62.50%	1.75	2.79	2.13	3.85	62.50%
Condition 3a: Mild infringing content manipulation	54.55%	1.68	3.41	2.18	3.66	75.00%
Condition 3: Infringing content manipulation	66.67%	1.73	3.29	2.31	3.76	64.71%

Table 6: *Between-Conditions Comparison of demographic characteristics and attitude towards piracy. There are 7 possible age group values in the questionnaire: 1 (18-21), 2 (22-25), ..., 7 (61 and over); 10 possible household size values: 1, 2, ..., 9, 10 or more; and 6 possible household income group values: 1 (less than \$30,000), 2 (\$30,000 - \$50,000), ..., 6 (over \$150,000).*

As in experiment 1, we also compare the initial search intent expressed by users across the different treatment conditions (Table 7), finding no significant differences in expressed intent across conditions.

	N	Neutral Searches	Legal Intent	Infringing Intent
Condition 1: No manipulation	42	6	23	13
Condition 2a: Mild legal content manipulation	49	16	19	14
Condition 2: Legal content manipulation	48	14	20	14
Condition 3a: Mild infringing content manipulation	44	11	18	15
Condition 3: Infringing content manipulation	51	15	27	9

Table 7: *Between-Conditions Comparison of the initial intent (based on the first keyword each user entered).*

Next, in Table 8 we compare the proportion of legal purchases (and the resulting average price of a legal purchase) made by participants in each of the experimental conditions. This table shows that in the control condition, where the search results were not manipulated, 61.9% of the participants chose to acquire the movie through a legal channel (and the remaining 38.1% consumed through a pirated channel). Note that in the general population sample, we found that 80% of the participants in the control condition acquired the movie through a legal channel. Comparing this figure to the 61.9% figure above suggests that, as expected, younger users are less likely to acquire content through legal channels than are users in the general population.

Relative to the proportion of legal purchases in the control condition, Table 8 shows that participants who were assigned to the mild legal content treatment condition were more likely to acquire the movie from a legal channel ($P = 75.5\%$, $SD = 43.4\%$) than were participants in the baseline condition ($P = 61.9\%$, $SD = 43.4\%$), a difference that is statistically significant ($t(83) = -1.39$, $p < 0.1$). We also see that the effect of the more intense legal content manipulation (which is the same the legal content manipulation used in study 1) resulted in a statistically higher proportion of legal purchases compared to the milder legal manipulation condition ($P = 91.7\%$, $SD = 27.9\%$, t-test: $t(63) = -2.18$, $p < .05$).

Conversely, the results show that participants in the mild infringing condition were significantly less likely to acquire content from a legal channel ($P = 47.7\%$, $SD = 50.5\%$) relative to the baseline condition ($P = 61.9\%$, $SD = 43.4\%$; t-test: $t(84) = 1.32$, $p < 0.1$). Likewise, the effect of the more intense infringing content manipulation (which is the same as the infringing content manipulation used in study 1) is stronger than that of the mild treatment, resulting in a directionally lower

proportion of legal purchases compared to the baseline condition ($P = 38.2\%$, $SD = 49\%$). However, this difference is not statistically significant (t-test: $t(90) = 0.83$, $p > 0.1$), possibly due to the relatively small sample size.

	n	# Legal option	% Legal option	Average price (conditional on acquiring a legal copy)
Condition 1: No manipulation	42	26	61.9%	\$12.26
Condition 2a: Mild legal content manipulation	49	37	75.5%	\$9.57
Condition 2: Legal content manipulation	48	44	91.7%	\$11.84
Condition 3a: Mild infringing content manipulation	44	21	47.7%	\$12.76
Condition 3: Infringing content manipulation	51	20	39.2%	\$14.75

Table 8: Between-Conditions Comparison of the Proportion of Purchases Made from Non-Infringing Websites in Study 2.

These results not only suggest that increased search costs associated with finding pirated content can have a significant impact on user behavior even for younger users; they also suggest that the effect varies with the intensity of the treatment, and is present even for relatively minor reductions in search costs.

First search term	Legal content manipulation			Infringing content manipulation	
	Control	Mild	Intense	Mild	Intense
Legal intent	21/23 (91%)	18/19 (95%)	20/20 (100%)	16/18 (89%)	17/27 (63%)
Infringing intent	0/13 (0%)	4/14 (29%)	10/14 (71%)	7/15 (47%)	2/9 (22%)

Table 9: Legal purchase rates across treatment conditions and initial intent

Our analysis of the initial intent expressed by participants is also consistent with that seen in experiment 1. Table 9 suggests that users who initially express intent to consume legally are significantly less likely to do so when placed in the infringing content manipulation (63%) than in either the other treatment conditions ($p < 0.05$ in all four t-tests performed). Likewise, users who initially express intent to consume infringing (pirated) content are significantly more likely to consume legally in the legal content manipulation (71%) than in the other treatment

conditions ($p < 0.1$ in the t-test that compares the mean of this condition with the mean of the intense legal manipulation, and $p < 0.05$ in the other three t-tests performed).

Finally, we use a logistic regression model to control for and analyze differences between groups based on observed characteristics. This model is similar to the model from Experiment 1, except that it includes dummy variables for the additional treatment conditions.

Specifically, we use the following logistic regression model to control for observable participant characteristics:

$$(5) \quad \log \frac{PR(legal_i)}{1 - PR(legal_i)} = \alpha + \beta_1 \cdot MNI_i + \beta_2 \cdot NI_i + \beta_3 \cdot MI_i + \beta_4 \cdot I_i + \sum_{j=5}^7 \beta_j \cdot DC_i^j + \sum_{j=8}^{10} \beta_j \cdot MCP_i^j + \sum_{j=11}^{12} \beta_j \cdot ATP_i^j + \beta_{13} \cdot intent + \varepsilon_i$$

where, MNI_i is an indicator variable denoting whether participant i was treated with the mild non-infringing search condition; MI_i is an indicator variable denoting whether participant i was treated with the infringing search condition; and the other variables are the same as before. We present the results of this model in Table 10 below.

The results in Table 10 are consistent with our means comparison results above in that the intense treatment variables are statistically different from the control condition and that all treatment variables have the expected signs (the non-infringing treatment condition increases the likelihood of purchasing legally and the infringing treatment condition decreases the likelihood of purchasing legally). As with experiment 1, the results show that participants who consumed infringing content in the past are less likely to purchase the movie legally, and that those who use a search term that reveals a clear non-infringing intent are more likely to purchase legally.

In summary, the results from experiment 2 are consistent with those in experiment 1 in showing that search results can have a significant impact on whether consumers choose to acquire legal or infringing media content. In addition, experiment 2 further shows that these results hold for younger users and that stronger treatments yield stronger responses.

	(1)	(2)	(3)	(4)	(5)
Dependent variable: Acquired legally	Basic Model	Including Demographic Characteristics	Including Media Consumption Preferences	Including Attitude Towards Piracy	Including Intent to Acquire Legally
Constant	0.486 (0.318)	0.772 (0.529)	3.077** (1.215)	3.265** (1.295)	2.932** (1.430)
Mild non-infringing mode	0.641 (0.460)	0.741 (0.470)	0.733 (0.479)	0.694 (0.484)	1.046** (0.529)
Intense non-infringing mode	1.912*** (0.611)	1.975*** (0.620)	1.977*** (0.628)	1.942*** (0.641)	2.321*** (0.671)
Mild infringing mode	-0.576 (0.438)	-0.644 (0.452)	-0.700 (0.462)	-0.732 (0.469)	-0.727 (0.528)
Intense infringing mode	-0.924** (0.428)	-0.932** (0.438)	-1.152** (0.469)	-1.228** (0.477)	-1.280** (0.530)
Woman		-0.626* (0.321)	-0.824** (0.341)	-0.940*** (0.350)	-1.178*** (0.390)
Younger (Age < 22)		-0.462 (0.338)	-0.593* (0.349)	-0.518 (0.356)	-0.383 (0.394)
Household size		0.0843 (0.0977)	0.0831 (0.1000)	0.0409 (0.102)	-0.00624 (0.109)
Income		-0.0100 (0.118)	-0.0267 (0.122)	-0.0174 (0.123)	-0.119 (0.136)
Time online			-0.317* (0.187)	-0.319* (0.190)	-0.229 (0.208)
Acquires movies online			-0.442* (0.264)	-0.209 (0.284)	-0.324 (0.302)
Movies per year			0.0780 (0.146)	0.0928 (0.150)	0.132 (0.165)
Against piracy				0.148 (0.410)	0.425 (0.448)
Downloads infringing				-0.968** (0.416)	-0.889* (0.458)
Non-infringing intent					3.127*** (0.693)
Number of obs.	234	234	234	234	234
Pseudo R2	0.1321	0.1519	0.1755	0.1987	0.3107

Table 10: Logistic regressions results
Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Discussion

As search engines become an increasingly important source of public information, they are being asked to play a stronger role in regulating what information is made available to the public. This role is seen in a variety of contexts including court ordered enforcement of an individual's "right to be forgotten" in a variety of countries, the United States Federal Drug Administration and Department of Justice requirements that search engines ban sponsored search listings from non-FDA approved pharmacies, and efforts by Google to inform its consumers about hacked sites,¹⁰ insecure sites,¹¹ or sites suspected of malware and phishing. Recently, many in the content industries have added pirated content to this list by asking whether changing the availability of pirated content in search results will change consumers' choices for piracy versus legal consumption channels.

The goal of this study is to assess whether search results can influence consumers choices for piracy versus legal consumption channels. To address this question, we used a unique experimental design to simulate consumers' online movie search and consumption processes. Our custom search engine allowed us to experimentally manipulate the rank and positioning of pirate and legal links in the search results.

We ran our experiment in two stages. In the first stage, we ran the experiment on a representative sample of the population recruited through an independent company maintaining large survey panels. We then exposed these users to three randomly assigned search conditions: a control condition which displayed search results from a major search engine, an infringing content treatment condition which artificially promoted infringing sites in the search results, and a legal content treatment condition which artificially promoted legal sites in the search results. In the second stage of the experiment we recruited a separate sample of college-aged students (18-25 year olds) as participants, and added two additional treatment conditions to the experiment: a mild legal treatment condition and a mild infringing treatment condition.

Our results in the first experiment with the general population show that 80% of users in the control condition choose to purchase content through legal channels. Relative to this baseline, we find that 94% of users in the legal treatment and 57% of users in the infringing treatment

¹⁰ <https://support.google.com/websearch/answer/190597?hl=en>

¹¹ <http://googlewebmastercentral.blogspot.com/2014/08/https-as-ranking-signal.html>

purchase content through legal channels. We see similar results for the second experiment with a college-aged audience: 62% of users in the control condition purchase content versus 92% in the legal treatment and 39% in the infringing treatment. The second experiment also shows that stronger treatments lead to stronger outcomes: 76% of users purchase in the mild legal treatment (versus 92% in the legal treatment) and 48% of users purchase in the mild infringing treatment (versus 39% in the infringing treatment).

We also find evidence that search results can even affect the behavior of users with a stated preference for legal or infringing content. When we classify user's intentions based on their initial search terms, we found that users who initially express an intent to consume legally are less likely to purchase legally in the infringing treatment condition than in other conditions, and that users who initially express an intention to consume through pirate channels are more likely to consume legally when they are placed in the legal treatment condition.

Together our results suggest that reducing the prominence of piracy links in search results can have a significant impact on consumer behavior, and may be a viable policy option in the fight against intellectual property theft. However we also note that more research is needed to shape such a policy. Identifying infringing content in real-time search results is different, and potentially more challenging, than identifying malware or phishing content. There also could be significant social costs from false positives, and opportunities for strategic manipulation of which sites and links are designated "infringing." Any policy should take these issues into account, and may benefit from the lessons learned from similar issues in the context of the enforcement of DMCA notices.

Our experiment is, of course, not without limitations. One possible limitation of our study is that while our custom search engine closely replicates the functionality of standard search engines, it still could be perceived differently by the users relative to a non-experimental setting. A second possible limitation is that the participants received a \$20 Visa virtual card and were asked to follow a specific experimental task. While this is standard practice in the experimental economics literature, it may not necessarily reflect their normal movie consumption environment. For these two reasons we emphasize that, in interpreting our results, one should focus on the relative levels of sales/piracy between the control and treatment conditions as opposed to the absolute levels within any particular condition. Finally, we note that our

experiment does not allow us to test potential longer-term changes in user behavior from manipulating search results.

References

- Abhishek, Vibhanshu, Rahul Telang, Yi Zhang. 2013. Gone with the Bits: Digital Release in the Presence of Piracy. Working Paper, Carnegie Mellon University, Pittsburgh, PA.
- Abhishek, Vibhanshu, Kartik Hosanagar and Peter Fader. 2013. Aggregation Bias in Sponsored Search Data: The Curse and the Cure. Working Paper, Carnegie Mellon University, Pittsburgh, PA. (<http://ssrn.com/abstract=1490169>)
- Agarwal, Ashish, Kartik Hosanagar, and Michael D. Smith. 2011. Location, Location, Location: An Analysis of Profitability of Position in Online Advertising Markets. *Journal of Marketing Research*. 48(6), 1057-1073.
- Bhattacharjee, Sudip, Ram D. Gopal, Kaveepan Lertwachara, and James R. Marsden. 2006. Impact of Legal Threats on Online Music Sharing Activity: An Analysis of Music Industry Legal Actions. *Journal of Law and Economics*. 49 91–114.
- Baye, Michael R., Babur De los Santos, and Matthijs R. Wildenbeest. 2012. What's in a Name? Measuring Prominence and Its Impact on Organic Traffic from Search Engines. *Working Paper*.
- Brooks, Nico. 2004. The atlas rank report II: How search engine rank impacts conversions. Atlas Institute, New York.
- Cannell, Charles F., Gordon Fisher, and Thomas Bakker. 1965. Reporting of Hospitalization in the Health Interview Survey. *Vital and Health Statistics*, PHS Pub. No. 1000, Ser. 2, No. 6, Washington, D.C.: U.S. Government Printing Office.
- Chesnes, Matthew, Weijia Dai, and Ginger Z. Jin. 2014. Banning Foreign Pharmacies from Sponsored Search: The Online Consumer Response Jin. *NBER Working Paper Series*.
- Danaher, Brett, Samita Dhanasobhon, Michael D. Smith, Rahul Telang. 2010. Converting Pirates without Cannibalizing Purchasers: The Impact of Digital Distribution on Physical Sales and Internet Piracy. *Marketing Science*, 29(6) 1138–1151.
- Danaher, Brett, Samita Dhanasobhon, Michael D. Smith, Rahul Telang. 2014. Understanding Media Markets in the Digital Age: Economics and Methodology, Greenstein, Goldfarb, and Tucker, eds. *Economics of Digitization: An Agenda*. University of Chicago Press, Chicago, Illinois. (<http://ssrn.com/abstract=2355640>)
- Danaher, Brett, Michael D. Smith, Rahul Telang. 2014. Piracy and Copyright Enforcement Mechanisms, Lerner and Stern, eds. *Innovation Policy and the Economy*, Volume 14. University of Chicago Press, Chicago, Illinois. (<http://ssrn.com/abstract=2273716>)
- Danaher, Brett, Michael D. Smith. 2014. Gone in 60 Seconds: The Impact of the Megaupload Shutdown on Movie Sales. *International Journal of Industrial Organization*. 33 1-8.

Danaher, Brett, Michael D. Smith, Rahul Telang, Siwen Chen. 2014. The Effect of Graduated Response Anti-Piracy Laws on Music Sales: Evidence from an Event Study in France. *Journal of Industrial Economics*. Forthcoming.

Forrester Research. 2006. *Forrester Research Report*. <http://www.forrester.com>.

Jerath, Kinshuk, Liye Ma, Young-Hoon Park, and Kannan Srinivasan. 2011. A "Position Paradox" in Sponsored Search Auctions. *Marketing Science*, 30 (4), 612-627.

Jeziorski, Przemyslaw and Ilya R. Segal, I. 2012. What Makes Them Click: Empirical Analysis of Consumer Demand for Search Advertising. Working Paper (<http://ssrn.com/abstract=1417625>).

Liebowitz, Stan J. 2008. File-Sharing: Creative Destruction or just Plain Destruction? *Journal of Law and Economics*, 49 1-28.

Means, Barbara, Katherine Habina, Gary Swan, and Lisa Jack. 1992. Cognitive Research on Response Error in Survey Questions on Smoking. *Vital Health & Statistics*. Hyattsville, MD, National Center for Health Statistics. Series 6, Number 5.

Narayanan, Sridhar and Kirthi Kalyanam. 2011. Measuring Position Effects in Search Advertising: A Regression Discontinuity Approach. Working Paper.

Nielsen//NetRatings. 2006. Online Search Hits All-Time High of 5.7 Billion Searches. *Press release*. March 3 (http://www.nielsen-online.com//pr/pr_060302.pdf).

Oberholzer-Gee, Felix, Koleman Strumpf. 2010. File Sharing and Copyright, Lerner and Stern, eds. *Innovation Policy and the Economy*, Volume 10, University of Chicago Press, Chicago, Illinois.

Pan, Bing, Helene Hembrooke, Thorsten Joachims, Lori Lorigo, Geri Gay, and Laura Granka. 2007. In Google we trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication*. 12(3), article 3.

Reimers, Imke. 2014. The Effect of Piracy Protection in Book Publishing. Working Paper, Northeastern University, Boston, MA.

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

Rob, Rafael, Joel Waldfogel. 2007. Piracy on the Silver Screen. *The Journal of Industrial Economics*, 55(3) 379-395.

Rutz, Oliver J. and Michael Trusov. 2011. Zooming In on Paid Search Ads—A Consumer-Level Model Calibrated on Aggregated Data. *Marketing Science*. 30(5), 789-800.

Telang, Rahul and Joel Waldfogel. 2014. Piracy and new product creation. A Bollywood story. Working Paper, Carnegie Mellon University, Pittsburgh, PA.

Vandiver, Donna M., Scott Bowman, Armando Vega. 2012. Music Piracy Among College Students: An Examination of Low Self-Control, Techniques of Neutralization, and Rational Choice. *The Southwest Journal of Criminal Justice*, 8(2) 92-111.

Waldfogel, Joel. 2012. And the Bands Played on: Digital Disintermediation and the Quality of New Recorded Music. Working Paper, University of Minnesota, Minneapolis, Minnesota.

Warner, Kenneth E. 1978. Possible increases in the under-reporting of cigarette consumption. *Journal of the American Statistical Association*, 73, 314-318.

Wyner, Gordon. 1980. Response errors in self-reported number of arrests. *Sociological Methods and Research*, 9: 161-177.

Yang, Sha and Anindya Ghose. 2010. Analyzing the Relationship Between Organic and Sponsored Search Advertising: Positive, Negative, or Zero Interdependence? *Marketing Science*, 29(4), 602-623.

Zentner, Alejandro. 2006. Measuring the Effect of File Sharing on Music Purchases. *The Journal of Law and Economics*, 49 63-90.