

The Impact of DNS Blocking on Digital Piracy Activity

Filipa Reis

Católica Lisbon School of Business and Economics and Heinz College of Carnegie Mellon University
flipareis@clsbe.lisboa.ucp.pt

Miguel Godinho de Matos

Católica Lisbon School of Business and Economics, miguel.godinhomatos@clsbe.lisboa.ucp.pt

Pedro Ferreira

Heinz College and Department of Engineering and Public Policy, Carnegie Mellon University, pedrof@cmu.edu

We use a unique household level dataset to measure the impact of DNS blocking on digital piracy and on the consumption of several legal alternatives not covered before in the literature. Our results show that DNS blocking reduces Internet traffic, including upload traffic, which proxies piracy. These blocks increase TV viewership, in particular across channels devoted to movies and TV shows, but they do not change the use of paid legal alternatives such as video-on-demand rentals or the subscription of premium TV channels. Furthermore, we observe that DNS blocking increases Google searches for tools to bypass them. We find evidence that more BitTorrent users remained active after the blocks in regions where this search behavior was more prevalent. Finally, combining the household level data with survey data from a smaller subscriber sample, we find that the presence of teenagers in the household and the head of household's computer literacy level and willingness to pay for content are significant moderators of the household's response to the DNS blocking policy.

Key words: Piracy, Internet, TV, DNS blocks, natural experiment

1. Introduction

Increased digitization and the dissemination of broadband Internet led to digital piracy whereby copyrighted content is illegally distributed and accessed on the web. This practice is perceived as a serious business threat by a number of large industries such as music, media, software, and publishing. As a consequence, it did not take long for major industry players to take legal action against providers that allowed infringement (epitomized by the shutdown of Napster in 2001) and exert pressure on legislators to limit digital piracy. In parallel, digital piracy sparked significant interest of academics who sought to analyze the claims of the industry and evaluate how piracy impacted businesses, from sales (e.g. Peitz and Waelbroeck (2006)) to product quality (e.g. ?).

Theoretical work on the impact of digital piracy on sales yielded mixed predictions (Peitz and Waelbroeck 2006, Chellappa and Shivendu 2005, Reavis Conner and Rumelt 1991), leading most researchers to address the topic empirically. Most of such work focuses on the music and movie

industries and supports the general belief that piracy hurts firms (Rob and Waldfogel 2006, Zentner 2006, Waldfogel 2010, Liebowitz 2008, Bounie et al. 2006, Rob and Waldfogel 2007, Danaher et al. 2010). As a natural follow-up, researchers investigated the effectiveness of a number of piracy control strategies including litigation threats (Blackburn 2004, Bhattacharjee et al. 2006), digital rights management (Sinha et al. 2010, Zhang 2014), graduated response laws (Danaher et al. 2014), pollution and poisoning of file-sharing networks (Christin et al. 2005), versioning (?), channel competition (Godinho de Matos et al. 2015, Belleflamme et al. 2014, Waldfogel 2012, Danaher et al. 2010), increasing search costs and reducing the quality of piracy options (Geng and Lee 2013, Johar et al. 2012), pricing and sampling (Sundararajan 2004, Chellappa and Shivendu 2005, Belleflamme and Picard 2007), website shutdown (Danaher and Smith 2014, Aguiar et al. 2015), and website blocking (Poort et al. 2014, Danaher et al. 2015, 2016).

Our paper contributes to the latter stream of research on the effectiveness of piracy control measures. More specifically, we study the role of public policy in mitigating the negative effects of digital piracy and promoting the use of legal alternatives. In this smaller body of research, Danaher and Smith (2014) found that some users turned to legal digital channels as a result of the shutdown of Megaupload Aguiar et al. (2015) found that the shutdown of a popular illegal streaming website in Germany reduced piracy activity but the effect was short-lived and followed by the emergence of several smaller similar sites. Danaher et al. (2015, 2016) studied the impact of website blocking in the UK and found that blocking The Pirate Bay in April 2012 did not increase the use of legal sites. The same authors found that a batch block of infringing websites in November 2013 in the UK increased the use of paid legal streaming and ad-supported services such as Netflix. Danaher et al. (2014) studied the effect of the HADOPI law in France and found that it led to a 20-25% increase in music sales, and Adermon and Liang (2014) studied the impact of a Swedish law that increased the chances of being caught and punished for illegal file sharing and found that it led to a 16% drop in Internet traffic and a 36% increase in music sales.

Despite the evidence that a number of these measures effectively reduce digital piracy and promote the use of paid and unpaid legal alternatives, the costs and benefits of each approach can vary widely. Specifically, strategies such as graduated response laws, such as HADOPI, can be extremely unpopular among the public (Breindl and Briatte 2013, De Filippi and Bourcier 2016). Website shutdown entails lengthy and costly legal disputes (Aguiar et al. 2015), and website blocking via IP address can result in over-blocking (blocking of legal content) (Ofcom 2010). Consequently, in recent years, new approaches to limit digital piracy have become increasingly popular. In particular, DNS-based website blocking was adopted by a number of European countries such as Belgium, Denmark, Italy, Portugal, Russia, Spain and the UK, while older approaches, such as the HADOPI

law in France and the six-strikes law in the US¹ have been mostly revoked. In this paper, we measure the effectiveness of batch DNS website blocking in reducing digital piracy. We measure the extent to which these blocks led users to turn to legal channels, and we investigate the heterogeneity in how households responded to them. We analyze a natural experiment whereby a country implemented DNS-based batch website blocking as a piracy control measure. We collaborate with a large telecommunications provider to collect a unique large-scale household level dataset on media consumption. To the best of our knowledge, this is the first analysis of the effectiveness of a DNS based batch website blocking policy and also the first to use disaggregated observational household level data to study the effectiveness of a website blocking policy.

We use a difference-in-differences estimator with household fixed effects to compare the behavior of households that, prior to the blocks, were BitTorrent² users (the “treated” group by the website blocks) to the behavior of households that did not use BitTorrent prior to the blocks (the control group). We estimate the impact of this policy on digital piracy, proxied by downstream and upstream Internet traffic, and on the use of a set of legal channels not previously covered by the literature, namely TV viewership, video-on-demand expenditure, and the subscription of paid TV channels. Our results show that DNS blocks led to a significant reduction in the amount of Internet traffic of BitTorrent users, namely 24.5% ($p < 0.01$) and 26.8% ($p < 0.01$) for download and upload traffic, respectively, and to a modest increase in TV viewership time, namely 2.5 minutes per day ($p < 0.01$), or 1.2%. This increase is essentially driven by viewership in TV channels devoted to movies and TV shows. However, these blocks did not change expenditures in paid legal alternatives, such as video-on-demand or premium TV channels. Therefore, and in our setting, DNS blocks reduced Internet traffic in favor of free legal alternatives to consume content³.

Like other website blocking techniques (e.g. IP address based, URL based), DNS blocks can also be circumvented. We find that, despite the effectiveness of this policy in reducing Internet traffic, and thus reducing piracy, Google searches for keywords related to circumventing DNS blocks, such as “DNS”, “VPN” and “Proxy”, spiked when these blocks were launched and remained at a higher

¹ The Copyright Alert System (CAS) (six-strikes system) was a graduated response system established in 2013 by a group of American ISPs and copyright owners in which participating ISPs would notify subscribers engaging in copyright infringement. Subscribers that exceeded 5 warnings could be subject to sanctions such as limited web access or bandwidth throttling.

² Although the BitTorrent protocol can be used for both legal and illegal purposes, previous work suggests that the vast majority of the content exchanged via this protocol is illegal (Watters et al. 2011, House 2011).

³ Note that households’ Internet navigation activity other than BitTorrent activity is invisible to us; we can only observe daily download and upload traffic in Mbps. We also do not observe subscriptions of products and services other than those provided by our Industry Partner (e.g. Netflix). Thus, it is possible that our control group includes illegal streamers alongside with BitTorrent non-users, meaning that some control households may have also been affected by the policy. Our results, therefore, do not provide us with the average treatment effect but with a conservative estimate of the effect of treatment.

level after that. Combining data on these Google searches and BitTorrent activity, we find that more households remained active users of BitTorrent in regions where the former were more prevalent. Therefore, this result suggests that at least some Internet users responded to the blocks by searching for (and likely learning) ways to bypass them.

Finally, we investigate which household characteristics relate to a stronger or weaker response to the DNS blocks. To this end, we complemented our observational dataset on household Internet activity with data from an online survey conducted by our Industry Partner to a large sample of its subscribers after the DNS blocks were implemented. This survey inquired subscribers about their media consumption habits and preferences, and also allowed us to collect demographic data on them. We find that DNS blocks were less effective in deterring piracy across households with teenagers and households with higher computer literacy levels but more effective in households with a higher willingness to pay for video content. Identifying the characteristics of households that relate to better policy responses allows for designing better policies to curb piracy. For instance, and once households with teenagers seem to respond less to DNS blocks, policy makers may be interested in complementing these blocks with educational campaigns in schools and/or high-schools or media campaigns targeting this specific population segment.

This paper is structured as follows. Section 2 describes how DNS blocks work. Section 3 introduces our research hypotheses, and section 4 describes the data and methods that we use in our analyses. Section 5 provides summary statistics. Section 7 and section 8 present the main results of our work. Finally, section 9 concludes.

2. DNS Website Blocking and Circumvention Tools

This paper focuses on a specific supply-side piracy control strategy - DNS website blocking. This anti-piracy tool is based on the premise that hampering access to illegal media content leads users to turn to legal distribution channels to satisfy their entertainment needs. There are several techniques to block websites that can be applied alone or in tandem. These include Internet Protocol (IP) address blocking; Domain Name System (DNS) filtering; Uniform Resource Locator (URL) blocking; and Shallow or Deep Packet Inspection (SPI and DPI, respectively). These techniques differ in their ease, speed, and cost of implementation as well as in their effectiveness. Ofcom (2010) evaluated them against a set of seven criteria, namely speed of implementation, cost, blocking effectiveness, difficulty of circumvention, ease of administrative or judicial process, integrity of network performance, and impact on legitimate services/law abiding consumers. Ofcom found that while no single technique achieved a perfect score, DNS blocking was identified as a quick, effective, and low cost approach.

In the case studied in this paper, DNS blocks were implemented through a legal court order requiring Internet Service Providers (ISPs) to deny subscribers access to websites marked as illegally providing access to copyrighted content. Our industry partner used DNS website blocking to do so. The Domain Name System (DNS) works as a phone book for IP addresses so that by typing a user-friendly and easy to remember name into one's browser (that is, typing, for example, `www.youtube.com` instead of `208.65.153.238`) redirects the user to the correct IP address. ISPs keep their own DNS "phone books" and can edit them by adding and removing entries or by associating a different IP address to a given domain, a process called DNS filtering or blocking. When a subscriber tries to access a website that has been removed from the DNS server, that website will not load. If the domain name has been associated to a different IP address, the subscriber is redirected to that website, which in the case of piracy may be used to show a warning page notifying the user that the website that she tried to access is blocked due to copyright violation. However, note that with DNS website blocking, the website providing illegal access to copyright content will still be up and running and thus can be accessed by users using circumvention tools.

There are three main ways to circumvent DNS blocks - using a DNS server that has not been altered or a proxy or VPN connection. By simply editing their network preferences, users can opt for a different DNS server from their ISPs such as Google's DNS server, a very quick, costless, and safe solution. Users can also resort to proxy servers which work as intermediaries between their computers and the wider network. By connecting to a proxy server, a user can then connect to websites that are restricted to users in her actual location but not in the proxy server's location. Finally, users can use a Virtual Private Network (VPN). Both proxies and VPNs allow users to hide their location by appearing to be connected to the internet from a different location. However, while a proxy server works as middleman for a single application (e.g. web browser), a VPN captures all internet traffic (e.g. web browser, applications, etc.), replacing the ISP and allowing for total encryption of one's internet traffic. The quality, cost, and safety of proxy and VPN services will vary as there are numerous paid and free options available online.

3. Research Hypotheses

We aim at measuring the effectiveness of DNS blocking and start by noting that the potential of these blocks to curb piracy is not given. First, owners of copyright infringing websites may respond to DNS blocks by registering their websites at multiple domains in order to escape the blocks. In other words, the effectiveness of DNS blocks may be low if new domains appear at a faster rate than new block batches are implemented. This would be similar to a phenomena known as the Hydra effect, observed by Aguiar et al. (2015), whereby the shutdown of a large piracy website in Germany led users to redistribute themselves across a number of new smaller websites. Likewise,

and similarly to what was observed by Danaher et al. (2015) when Pirate Bay was blocked in the UK, blocking large websites may simply lead their users to redistribute themselves across existing smaller websites.

Second, users can resort to a number of tools to bypass DNS blocks including changing their DNS settings, using a proxy server, or using a VPN connection. Though these circumvention strategies may initially be limited to more technically savvy users, their popularity increases with the number of blocked websites. The availability of free, safe, quick, and easy to implement circumvention tools, which can also be readily found by any user searching online for ways to circumvent website blocks, makes block circumvention a real threat to the effectiveness of DNS website blocking as a policy to curb piracy. Finally, public disclosures or announcements with lists of blocked websites may translate themselves into unintended advertisement for such websites as observed by Clemente (2015)'s study on website blocking in Italy. Combined with the availability of information on block circumvention techniques, this may further contribute to reduce the effectiveness of these measures.

For the reasons above, it is important to evaluate the effectiveness of this policy empirically. Following, Adermon and Liang (2014) we start by using Internet traffic as a proxy for digital piracy. Using Internet traffic as a measure for digital piracy should allow us to capture piracy activity as peer-to-peer sharing of media files generates significant upload and download traffic while also allowing us to keep track of hidden piracy activities (BitTorrent traffic hidden through VPN or proxy servers). Our preliminary hypotheses are thus that DNS blocks reduce Internet traffic from pirates, which as we describe below in section 6.1, are identified as BitTorrent users in our setting:

H1a: DNS blocks reduce download traffic of BitTorrent users.

H1b: DNS blocks reduce upload traffic of BitTorrent users.

We then analyze the effectiveness of DNS blocking in promoting the use of legal alternatives to consume media. As argued in Danaher and Smith (2014) "a necessary condition for supply-side anti-piracy policies to be worthwhile is that we must see a causal gain in media sales and revenues resulting from the reduction in piracy." Our unique setting allows us to test the impact of DNS blocks on a set of previously unexplored legal alternatives, namely the consumption of Video-on-Demand (VoD), the subscription of paid entertainment TV channels and overall TV viewership. Theoretically, these channels to consume media seem likely alternatives to piracy. Users are familiar with them, our industry partner actively promotes them and, in the case of TV, usage is free. Therefore, our hypotheses are that DNS blocks increase the use of these legal alternatives:

H2a: DNS blocks increase the expenditure in Video on Demand of BitTorrent users.

H2b: DNS blocks increase the subscription of paid entertainment channels by BitTorrent users.

H2c: DNS blocks increase TV viewership of BitTorrent users.

Our second research question concerns which household characteristics relate to a stronger response to DNS blocks. Previous research has looked into the association between economic factors (e.g. income, price) and piracy. In general, income is negatively related to piracy (??). At a national or state level, studies have shown that richer countries present lower piracy rates (???), and similarly, in the context of music piracy, previous work has shown a negative relationship between income and piracy at the individual level (?). In addition to income, a higher willingness to pay for copyright contents is generally assumed to be associated to lower piracy (?). We thus hypothesize that the economic characteristics of households, namely their income and willingness to pay⁴ for video content will moderate the effectiveness of the DNS blocking policy on curbing households' digital piracy activity and promoting the use of legal alternatives to piracy:

Hypothesis 3a: The effectiveness of DNS blocks is moderated by households' income level.

Hypothesis 3b: The effectiveness of DNS blocks is moderated by the head of household's willingness to pay for video content.

In addition to economic factors, demographic factors have also been shown to be a determinant of piracy behavior. In the context of software piracy, Al-Rafee and Cronan (2006) study the influencing factors of one's attitude towards digital piracy and find that, along with several cognitive factors, the individual's age played a significant role. In the context of music piracy, ? develop a behavioral model of digital piracy also finding that age was a significant determinant of piracy behavior. These findings are aligned with the extant ethics literature that suggests that ethical standards are higher among older individuals (e.g. ???). Mateus and Peha (2011) characterized and quantified worldwide BitTorrent traffic and found that contents "appealing to the teenager and young adult demographics have disproportionately higher ratios of BitTorrent transfers to sales than titles that appeal to an older segment of the population" (Mateus and Peha (2011), pp. 37). Gunter et al. (2010) studied the prevalence of digital piracy among adolescents and found that more than half of 8th graders had pirated before, 44% had pirated in the previous year, and 35.1% had pirated in the previous month. These statistics increased to 72.3%, 63.8%, and 52.8% among 11th graders, respectively (Gunter et al. 2010).

The fact that digital piracy is a prevalent behavior among young people has also motivated a number of studies on digital piracy across diverse fields - information systems, economics, law, criminology, and ethics - to focus exclusively on college student samples (e.g. Rob and Waldfogel (2006), Al-Rafee and Cronan (2006), Solomon and O'Brien (1990), Higgins et al. (2008), Yoon

⁴ In our analysis, we use the households monthly bill value before the policy implementation as a proxy for household income. Willingness to pay for video content is measured using a survey question in which respondents were asked how much they were willing to pay to have their favourite TV show immediately available in its entirety.

(2011)). However, no study to date has actually measured how individuals from different age groups respond to anti-piracy policies.

In addition to online piracy, younger demographics also seem more familiar with the tools that may help circumvent blocks. These tools, e.g. proxy networks and VPNs, are often promoted by the pirate websites themselves. The Pirate Bay is an instance of such a website that has explicitly advised their users to do so (BBCNews 2012). According to the GlobalWebIndex (Q4, 2015), about a quarter of adult Internet users have used a VPN service. When inquired about the motivations to do so 30% mentioned access to better entertainment content and 27% mentioned access to restricted content and websites in their country (Young 2016). Breaking down VPN use by generation, Statista reports that in the first quarter of 2014, 32% of millennials, 27% of generation X, and 13% of baby boomers from the worldwide population of Internet users used VPN or proxy servers (Statista 2014). Millennials used these services in order to "access restricted content and social networks or to use download sites", while generation X used them to "keep in contact with those abroad and access restricted sites."

In this regard, we hypothesize the following:

Hypothesis 3c: The effectiveness of DNS blocks is moderated by the presence of teenagers in a household.

Hypothesis 3d: The effectiveness of DNS blocks is moderated by the presence of young adults in a household.

Finally, one other aspect considered is the role of individuals' computer and Internet proficiency as a determinant of digital piracy behavior. In our analysis we proxy households' computer literacy level with the indicator variable of whether the household opted for electronic billing, an indicator that at least the head of household is proficient with email. Although we consider it as a separate factor, computer literacy is related to individuals' age, as exemplified by the above statistics on the use of VPN by different demographic groups. Our final hypothesis is the following:

Hypothesis 3e: The effectiveness of DNS blocks is moderated by households' computer literacy level.

4. Empirical Context and Data

We study the effectiveness of an anti-piracy protocol resulting from a collaborative effort between local copyright owners, telecommunication providers, and government. The protocol requires that Internet providers implement DNS blocks of websites listed by an external auditing entity as engaging in copyright infringement. These lists are periodically and continually sent to Internet providers and the blocks must be made effective within a short time window.

We collaborate with a multinational telecommunications provider. This firm provides pay-TV, broadband and mobile Internet, fixed and mobile telephony to over 1 million households. We use

data from this firm to study the impact of DNS website blocking. ISPs in this country were required to block websites listed by an external auditing entity as engaging in copyright infringement. Lists with these website were periodically sent to ISPs and the blocks had to be made effective within a short time window. Our analyses focus on triple and four play households at this firm. We collected data from a random sample of 100k of these households among those who were active (i.e. had a working 3P or 4P service) during our period of analysis, that is from 4 months preceding the start of website blocking up to 10 months after the blocks were first introduced. This sample includes subscriber’s of both our IP’s premium and standard service, which differ in the number of TV channels offered, Internet speed, and VoD credits offered.

Each household in our dataset is uniquely identified with an anonymized identifier. For each month of our period of analysis, we have information on each households’ aggregate download and upload Internet traffic (in MB), paid VoD expenditure (in cents), aggregate TV viewership time (in minutes) and BitTorrent activity. BitTorrent data were obtained from a third party provider that routinely scans popular active online swarms (the peers downloading and uploading a given torrent) keeping track of the IP addresses illegally sharing copyright content and the identifiers of the content associated to each IP address’s torrent activity. These data were matched to households in our dataset through their IP addresses by a separate provider providing data escrow services in order to ensure the full anonymity of subscribers at our industry partner. Using the BitTorrent data, we assign the households in our sample into one of two groups – BitTorrent users and non-users, corresponding to households that used BiTtorrent at least once in the 4 month-period preceding the blocks and those who did not. Our main dataset also includes information on the additional services purchased by each household including the subscription of Video-on-Demand (VoD) service and of paid TV channels (e.g. sports and entertainment), and on the household’s monthly bill value and whether it opted for electronic billing. Appendix A provides a detailed description of all household level variables used in our analysis.

In order to explore some of the possible mechanisms determining the effectiveness of DNS blocks, we collected data on online searches for a set of keywords that may be used by Internet users that attempt to bypass DNS blocks, namely “DNS”, “Proxy” and “VPN”. These data were obtained from Google Trends and covers 14 regions from the country of interest and 15 months (4 months before the blocks and 11 months after). Google Trends’ data is normalized. The search interest for a specific term is the share of searches for that term relative to all Google searches that took place in the same region during the same time period. The search peak for a given term in a given location during a specified time period is indexed to 100. These data were combined with the BitTorrent activity of households in our dataset at the regional level, that is, the country analyzed comprises

several administrative sub-regions and the Google trends' data in each such region were associated to households whose address belonged in that region.

Finally, and to gain a better understanding of the characteristics of the households that engage in digital piracy, and how these characteristics might relate to heterogeneity in their responses to DNS blocks, we complement our observational dataset described above with data from an online survey ran by our Industry Partner to a separate sample of households. The survey inquired subscribers about their media preferences, media consumption habits, and demographics, age group and gender for the individuals in the household plus the education level of the survey respondent. The survey data covers 1.4K households. Although not perfectly representative of the population of subscribers at our Industry partner, these survey responses allow us to derive valuable insights into household behavior, namely we use the demographic data to find how different household characteristics moderate their response to DNS blocks. The survey was distributed to households after the DNS blocks started. Therefore, we use only the household demographic data from the survey in our analysis because everything else in the survey could have been affected by the block and thus lead us into problems of reverse causality. However, this concern is unwarranted with respect to demographic data. For example, it is unreasonable to think that the DNS blocks caused households to have more teenagers.

5. Descriptive Statistics

Tables 1 and 2, provide summary statistics for the main variables used in our analysis for both BitTorrent users and non-users, respectively, for the month just before the implementation of website blocking. From our random sample of 100K households, about 9% were observed using BitTorrent at least once in the 4 months before the blocks. From this subset of BitTorrent users, 78% subscribed to our IP's premium service and 21% subscribed to paid entertainment channels. On average, BitTorrent users consumed 2.3 GB of download traffic and 0.6 GB of upload traffic per day. These users spent an average of 14 cents/month on VoD and watched an average of 4.7 hours of TV per day, of which 70 minutes were dedicated to entertainment. As for non-users of BitTorrent, 66% of them subscribed to the premium service and 14% to paid entertainment channels. On average, non-BitTorrent users download and upload only 0.7 GB and 0.1 GB of Internet traffic per day, respectively. These users spent an average of 16 cents/months on VoD and watched an average of 4.7 hours of TV per day, of which 55 minutes were devoted to entertainment. These statistics show clear differences in media consumption between BitTorrent and non-BitTorrent users, particularly in what concerns Internet use. Appendix B provides descriptive statistics for the households that answered the media survey.

Figures 1 and 2, depict the monthly BitTorrent activity for households who used BitTorrent before the blocks and households who did not during 15 months – 4 months before the start of the

Table 1 Summary Statistics for BitTorrent users.

Statistic	N	Mean	St. Dev.	Min	Max
N Torrent (torrents /month)	8,847	2.920	17.018	0.000	742.000
Download (Mb/day)	8,847	2,337.603	2,402.189	0.000	48,991.090
Upload (Mb/Day)	8,847	576.038	1,295.393	0.000	42,442.120
VoD (cents/month)	8,847	14.484	116.433	0.000	4,130.400
TV Total (min/day)	5,021	283.451	167.775	0.400	1,124.500
TV Entertainment (min/day)	5,021	70.978	75.228	0.000	618.367
TV Video-on-Demand (min/day)	5,021	2.051	7.790	0.000	118.467
Month Bill (USD)	8,819	64.998	25.208	14.300	699.985
TV Service Tenure (month)	8,847	94.529	60.873	3.258	249.161
Internet Service Tenure (month)	8,847	68.635	37.062	0.000	190.129
Set-Top-Box Tenure (month)	8,847	46.755	27.142	0.000	88.387
Flag Premium	8,847	0.782	0.413		
Flag Paid Entertainment Channels	8,847	0.215	0.411		
Flag Sports-Bundle	8,847	0.156	0.363		
Flag Electronic Receipt	8,847	0.388	0.487		
Flag Bank Transfer	8,847	0.323	0.468		
Flag Active Contract	8,847	0.783	0.412		

TV time only available for households with certain set-top-box models

Table 2 Summary Statistics for Non-BitTorrent users.

Statistic	N	Mean	St. Dev.	Min	Max
N Torrent (torrents /month)	91,153	0.000	0.000	0.000	0.000
Download (Mb/day)	91,153	731.308	1,345.599	0.000	47,257.730
Upload (Mb/Day)	91,153	110.692	467.151	0.000	22,730.020
VoD (cents/month)	91,153	16.639	185.401	0.000	26,908.800
TV Total (min/day)	41,586	282.160	171.032	0.033	1,157.000
TV Entertainment (min/day)	41,586	55.978	68.672	0.000	940.800
TV Video-on-Demand (min/day)	41,586	1.540	6.718	0.000	190.800
Month Bill (USD)	90,007	55.303	24.338	13.702	172.484
TV Service Tenure (month)	91,153	82.094	65.576	0.000	250.161
Internet Service Tenure (month)	91,153	35.881	39.329	0.000	188.871
Set-Top-Box Tenure (month)	91,153	29.152	27.132	0.000	88.387
Flag Premium	91,153	0.656	0.475		
Flag Paid Entertainment Channels	91,153	0.135	0.342		
Flag Sports-Bundle	91,153	0.149	0.356		
Flag Electronic Receipt	91,153	0.259	0.438		
Flag Bank Transfer	91,153	0.330	0.470		
Flag Active Contract	91,153	0.692	0.462		

TV time only available for households with certain set-top-box models

DNS blocks and the 11 months after. Figure 1 plots the average number of BitTorrent downloads per household and Figure 2 plots the fraction of households in each group that were observed using BitTorrent. The vertical lines mark the start of the DNS blocks. Both figures present a clear drop in BitTorrent activity immediately after the start of the first wave of DNS blocks. There are some periods of recovery but essentially BitTorrent activity keeps dropping after the blocks. Figure 3 plots the average Google Search Index for weekly web searches for the set of blocked websites. This

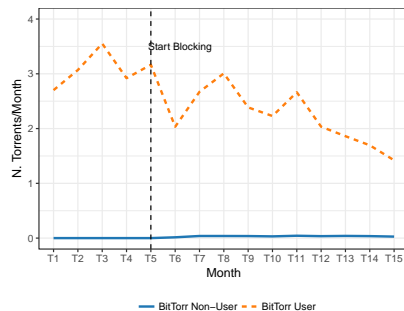


Figure 1 Average number of BitTorrent downloads per household per month.

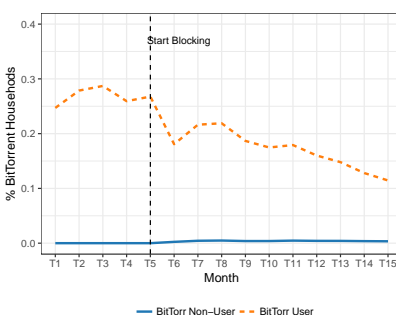


Figure 2 Fraction of households using BitTorrent per month.

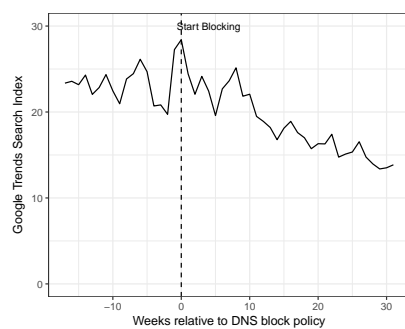


Figure 3 Average Google search index for the set of blocked websites.

figure shows a steady decline in search activity for these websites. The patterns observed in these figures suggest that website blocking was at least partially effective in leading some households away from using BitTorrent and accessing the blocked torrent sites. These figure also suggest that some households knew or learnt how to circumvent the DNS blocks and continued downloading content from the blocked websites.

Figure 4, depicts the average number of Torrent downloads per month for households that were using BitTorrent in each month. The Figure shows that the intensity of BitTorrent usage was stable across these households before the blocks started and increased right after. Therefore, fewer households used BitTorrent to illegally download copyright content after the blocks but the households that kept doing so seem to have increased their piracy activity. It appears that the households that stopped using BitTorrent to illegally download copyright content were the non-heavy users of BitTorrent while the heavy users sought out ways to circumvent the DNS blocks and kept downloading. Figure 5 provides additional evidence that may backup this hypothesis. This figure shows that Google searches for keywords related to block circumvention, such as “DNS”, “VPN” and “Proxy”, spiked right when the blocks were first issued and remained roughly twice as large after the blocks compared to before. This provides some evidence that the blocks may have had the side effect of leading some Internet users to learn how to circumvent them.

6. Empirical Strategy

6.1. Household Level Analysis

The goal of our household level analysis is to compare the media consumption activity of subscribers who used BitTorrent before the website blocks, our *Treated* group of households, to those who did not, our *Control* group of households. Our expectation is that blocking copyright infringing websites impacts the behavior of BitTorrent users but not of that of non-BitTorrent users. To measure the effectiveness of the website blocking policy in deterring digital piracy, we use difference-in-differences to estimate the change in Internet download and upload traffic that resulted from

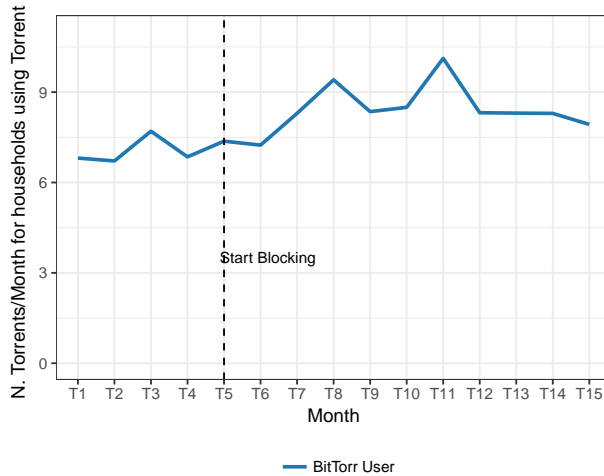


Figure 4 Average number of BitTorrent downloads per household using BitTorrent per month.

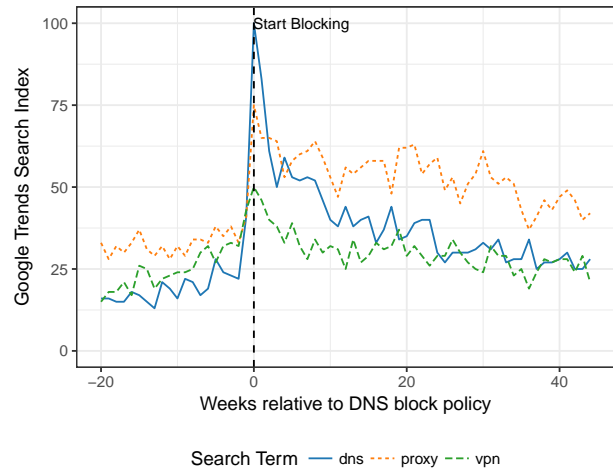


Figure 5 Google search index for keywords DNS, PROXY and VPN.

the blocks. Internet traffic is used as a proxy for piracy activity as discussed above in section 3. In addition, we note the following average Internet traffic profile at our industry partner: 37% is peer-to-peer file sharing (with BitTorrent making up 20% of total identifiable download traffic and 68% of total identifiable upload traffic), 35% is web browsing, 19% is video consumption, and 9% is associated to other services including gaming and voice over IP⁵. Thus, we believe that any significant change in piracy activity should be necessarily reflected in a significant change in Internet traffic. Furthermore, a fundamental reason to use the amount of Internet traffic as a proxy for piracy is that while BitTorrent traffic can be hidden using VPN or proxy servers, the amount of traffic it generates cannot. Thus, using Internet traffic as a measure for digital piracy, allows us not to lose sight of 'hidden' piracy activities. We also measure the effect that the blocks had in promoting the use of legal alternatives for media consumption and thus we study how they changed VoD expenditure, subscriptions of paid entertainment channels, and TV viewership.

Note that in our setting the DNS blocks concern all sorts of websites providing copyright infringing materials in any format. This includes websites providing BitTorrent content but also websites providing, for example, streaming services. We are only able to observe the first type of piracy activity and therefore our classification of households as BitTorrent users is a conservative definition of “digital pirates” which leaves out households who engage in all other types of illegal consumption of copyrighted content that we are unable to measure. This means that our control group might include, alongside BitTorrent non-users, digital pirates whose pirating behavior we are unable to

⁵ Note that Internet traffic is dependent on both use and file size. For this reason, consumption of movies, specially HD movies, via streaming or BitTorrent will necessarily entail a large amount of traffic. This explains why P2P (often used for large files) and video streaming capture such a large portion of traffic.

observe (e.g. streamers). The implication of this fact for our analysis is that some households in our control group may have been affected by the treatment, and thus our difference-in-differences estimators do not provide us with the average treatment effect but instead with a conservative estimate of the effect of treatment because households in the control group that are affected by the policy will also reduce the amount of internet traffic they consume.

The equation presented below summarizes our general difference-in-differences empirical specification with fixed-effects to estimate the effect of website blocking:

$$Y_{it} = \beta_0 + \beta_1 \text{BitTorrentUser}_i + \beta_2 \text{BitTorrentUser}_i \times \text{After}_t + \alpha_i + \tau_t^T + \epsilon_{it} \quad (1)$$

where BitTorrentUser_i is a dummy variable indicating whether household i used BitTorrent prior to the blocks and After_t indicates a time period after the blocks started. τ^T are time dummies and α_i the household level fixed effects. β_2 measures the impact of website blocking on the outcome variable of interest. Outcome variables considered in our study include two proxies for piracy activity – $\text{Log}(\text{Downloads Mb/Day})$ and $\text{Log}(\text{Uploads Mb/Day})$ – and the consumption of three legal alternatives to piracy – VoD Expenditure , $\text{Paid Entertainment Channels}$ and TV min/Day . We also look at the total time people spend watching movie channels on TV, represented by TV Mov. min/Day . This specification is used to test hypotheses H1a, H1b, H2a, H2b and H3b. We estimate this specification using OLS, adjusting standard errors for heteroskedasticity, serial correlation and clustering them at the household level. We test hypotheses H3a and H3b using the same model with two additional terms, namely an indicator for the presence of teenagers in the household and a triple interaction term between this indicator, BitTorrentUser_i and After_t .

We also test if internet usage trends were similar across BitTorrent and Non-BitTorrent users before the implementation of the DNS blocking. This is achieved using the following model:

$$Y_{it} = \beta_0 + \beta_1 \text{BitTorrentUser}_i + \gamma_2^T \text{BitTorrentUser}_i \times \tau_t^T + \alpha_i + \tau_t^T + \epsilon_{it} \quad (2)$$

where the vector of coefficients γ_2^T captures the difference in behavior across groups (BitTorrent users and non-users) over time. This specification also allows us to study how the impact of the policy evolved over time.

6.2. Region Level Analysis

The goal of our region level analysis is to explore heterogeneity in the responses to website blocking associated to learning and increased sophistication across BitTorrent users in what regards their ability to access blocked content. The outcome of interest in this case is the fraction of active BitTorrent users (those who were identified as having used BitTorrent at least once in that month)

in each region relative to the number of subscribers in the region per month (in log form). Our empirical model is the following:

$$\%BitTorrentUsers_{jt} = \beta_0 + \beta_1 Search_{jt} + \beta_2 Search_{jt} \times After_t + \tau_t + region_j + \mu_{it} \quad (3)$$

Subscript j corresponds to a region and subscript t to a month. $Search_{jt}$ represents the Google search index for a specific keyword in region j and month t . We analyze three different keywords – “DNS”, “proxy”, and “VPN”. The coefficient of interest in this specification is β_2 , which gives us the region level effect in percentage of active BitTorrent users of the search intensity for each of these keywords. All our models here include month and region level fixed effects and time trend controls. Coefficients are estimated using a panel linear model with fixed effects. Standard errors are estimated using a robust covariance matrix accounting for serial correlation with observations clustered at the region level.

7. Results

7.1. Impact of Website Blocking on Household Behaviour

Table 3 presents our main difference-in-differences estimates. Columns (1) and (2) show the impact of treatment on download and upload traffic, respectively. The results show that after the start of DNS blocks, daily download traffic of BitTorrent users decreased by $100 \times (exp(-0.275) - 1) = 24.5\%$ ($p < 0.01$) and upload traffic decreased by roughly $100 \times (exp(-0.312) - 1) = 26.8\%$ ($p < 0.01$). This evidence suggests that DNS blocks effectively decreased the households’ piracy activity thus supporting hypotheses H1a and H1b. Columns (3), (4), (5) and (6) measure the impact of the blocks on VoD expenditure, subscription of paid entertainment channels, unpaid viewership of TV channels devoted to movies and TV shows and unpaid overall TV consumption. We find that the treatment did not change VoD expenditure (column 3) but increased the subscription of paid entertainment channels by 4.5% ($p < 0.01$) (column 4, the baseline subscription rate of these channels is 13.4%⁶). In addition, the treatment increased the consumption of TV by 2.5 minutes/day ($p < 0.01$) (column 5) but this represents only an increase of 1.2% in TV viewership. This increase is essentially associated to the increase in the viewership of TV channels broadcasting movies and TV shows (+6.1 minutes/day).

Our preliminary results support hypotheses H2b and H2c but not H2a. The DNS blocks seem to have led BitTorrent users to slightly increase their use of legal alternatives but the increase in the use of paid alternatives may be modest. A possible explanation for the lack of take up in transactional VoD consumption is that there might be a mismatch between VoD catalogs and the

⁶ Baseline references are calculated using the constant of the model interacting the *BitTorrent User* indicator and the *After* indicator with time trend controls as presented in the main results but without household level fixed effects.

Table 3 Effect of DNS blocks on Internet traffic and use of legal alternatives.

	<i>Dependent variable:</i>					
	Log(Download) FE	Log(Upload) FE	VoD Rev. FE	Paid Entertain. Ch. FE	TV FE	TV Mov. FE
	(1)	(2)	(3)	(4)	(5)	(6)
After × BitTorrent User	−0.275*** (0.009)	−0.312*** (0.010)	−0.008 (0.778)	0.006*** (0.002)	2.488*** (0.755)	6.085*** (0.446)
Month Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,500,000	1,500,000	1,500,000	1,500,000	962,580	962,580
F Statistic	896.832***	1,230.551***	33.260***	1,098.312***	1,371.978***	14,487.110***

Note:

*p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in ()
Clustering at individual level
FE is the fixed effects estimator

preferences of pirates as observed in Godinho de Matos et al. (2015). Another explanation might be that BitTorrent users are heavy media consumers and prefer subscription-based services rather than the pay-per view alternatives captured by our estimate of VoD expenditure. Finally, we see a similar reduction in download and upload traffic across BitTorrent users, which suggests that these households are not entirely replacing their torrent traffic by alternative streaming websites or paid legal alternatives provided by other firms over the Internet, such as Netflix or Hulu. If this was the case then we would expect upload traffic to decline as households abandon BitTorrent and download traffic to remain roughly constant given that all other legal over the top VoD alternatives require as much download bandwidth as BitTorrent transfers do.

Figures 6 depict the γ_2^T of equation 2 for all dependent variables of interest. The top plots in the first row of this figure show that the trend of downloads and uploads was similar across BitTorrent and Non-BitTorrent users before the DNS blocks were introduced but that after they came into play BitTorrent users reduced their Internet usage substantially. Furthermore, these plots highlight that the effect of the DNS blocks was not one off but rather continued to increase over time. This is consistent with the fact that more and more websites were blocked every month since the policy came into effect. The second row of plots in this figure provides insight into the effects of the policy over time on the take up of VoD and premium TV channels. We observe no change in the consumption of paid VoD and a modest increase in the subscription of premium channels devoted to movies and TV shows. These plots also show that BitTorrent and non-BitTorrent users exhibited similar VoD expenditures and similar subscription rates for the premium channels before the blocks were introduced. Finally, the last two panels in the bottom row of this figure show that BitTorrent users spend more time watching TV, in particular watching TV channels devoted to movies and TV shows. However, this figure also shows that this was already the case before the DNS blocks came into play.

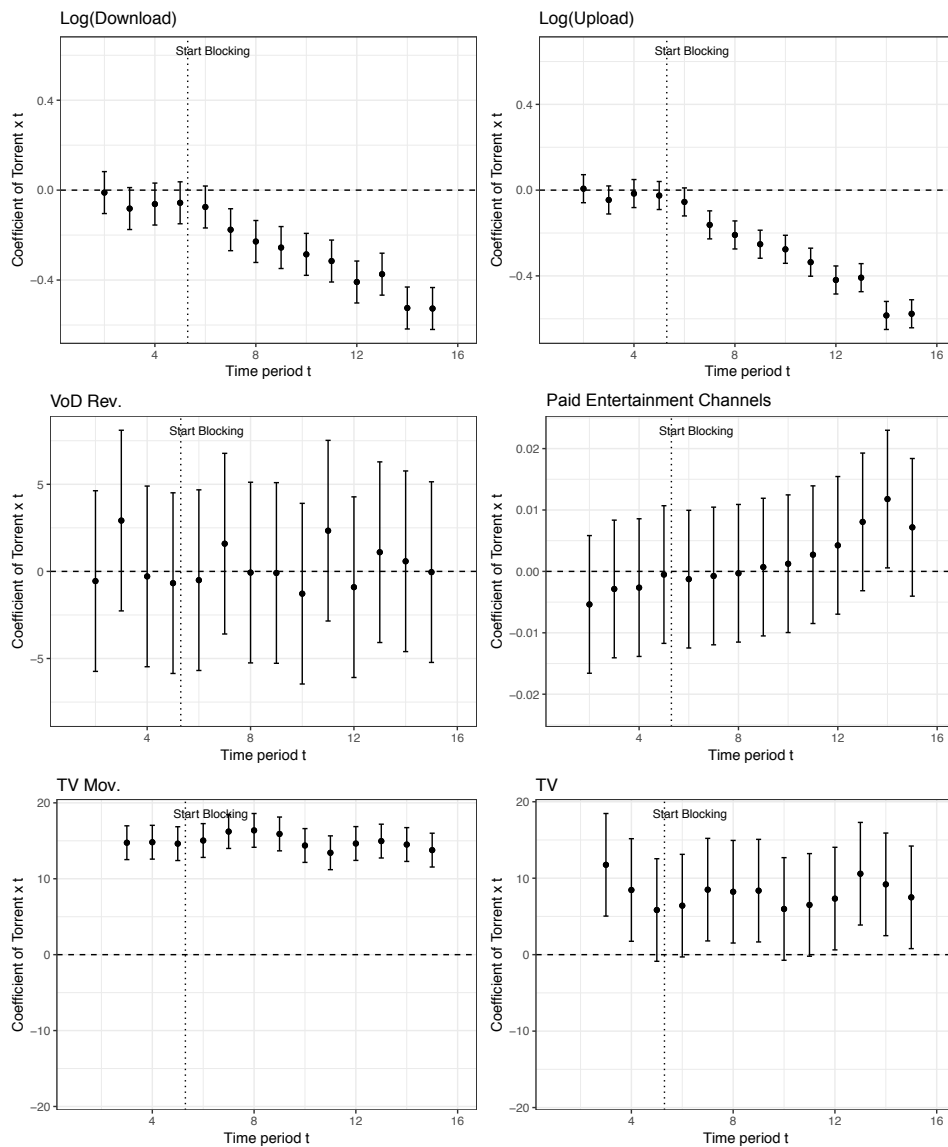


Figure 6 Impact of DNS blocks on downloads, uploads, video-on-demand expenditure, subscription of premium TV channels and TV viewership over time. Error bars are the 95% confidence intervals.

This significant difference between BitTorrent and non-BitTorrent users before the blocks came into play hampers the interpretation of our results. To mitigate this concern and to provide a robustness check for our findings, we use Coarsened Exact Matching (CEM) and redo the previous analysis on a matched sample of households. CEM is a monotonic imbalance bounding matching

method that has been shown to be superior to propensity score matching for estimating causal effects (??). CEM is typically used in k-to-k matches such that different numbers of treated and control units are matched to each other. CEM uses weights to compensate for the differential strata sizes that result from this procedure. For additional details into how CEM works see ?.

Figure 7 provides balance statistics for the covariates that we used to match BitTorrent and non-BitTorrent users before and after the match. It is clear from this figure that the matched households are highly similar on all these covariates. Figure 8 replicates Figure 6 for the matched households only. This figure shows that after matching, BitTorrent and non-BitTorrent users exhibited similar trends in all the dependent variables of interest. The disadvantage of matching is that we are only left with 4,564 households (of which 3,466 are controls and 1,188 are treated) out of the initial 100,000 used in the previous analysis. Table 4 replicates the analysis of table 3 with the matched dataset. Columns (1) and (2) show again that download and upload traffic reduced quite significantly after the DNS blocks. Columns (4) and (5) suggest that there was no take-up in paid legal alternatives. Columns (5) and (6) show that households using BitTorrent before the DNS blocks increased their daily usage of TV and of entertainment content in particular. In sum, our results after using CEM come all in line with our previous results thus increasing the robustness of our findings. In short, the introduction of DNS blocks reduced Internet traffic significantly but did not increase the use of paid legal alternatives to consume media. At best, there is a modest increase in TV viewership driven by an increase in the consumption of TV channels devoted to movies and TV shows.

Table 4 Effect of DNS blocks on Internet traffic and use of legal alternatives (matched sample).

	<i>Dependent variable:</i>					
	Log(Download) FE	Log(Upload) FE	VoD Rev. FE	Paid Entertain. Ch. FE	TV Mov. FE	TV FE
	(1)	(2)	(3)	(4)	(5)	(6)
After × BitTorrent User	−0.179*** (0.033)	−0.238*** (0.035)	1.289 (1.508)	−0.012 (0.007)	2.149** (0.914)	7.824** (3.456)
Month Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	69,810	69,810	69,810	69,810	69,810	69,810
Residual Std. Error	0.907	0.923	68.842	0.160	25.610	87.623

Note:

*p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in ()
Clustering at individual level
FE is the fixed effects estimator

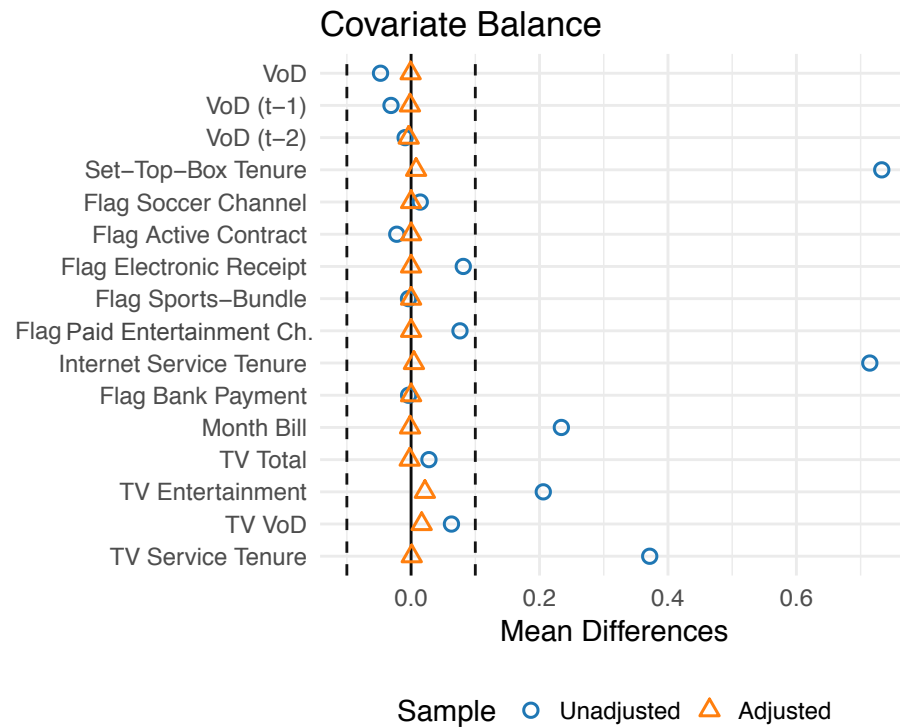


Figure 7 Covariate balance before and after Coarsened Exact Matching.

7.2. Avoidance of DNS Blocking Policy

The previous section showed robust evidence that DNS blocking reduced Internet traffic – both uploads and downloads – and that households that used BitTorrent prior to the DNS blocks

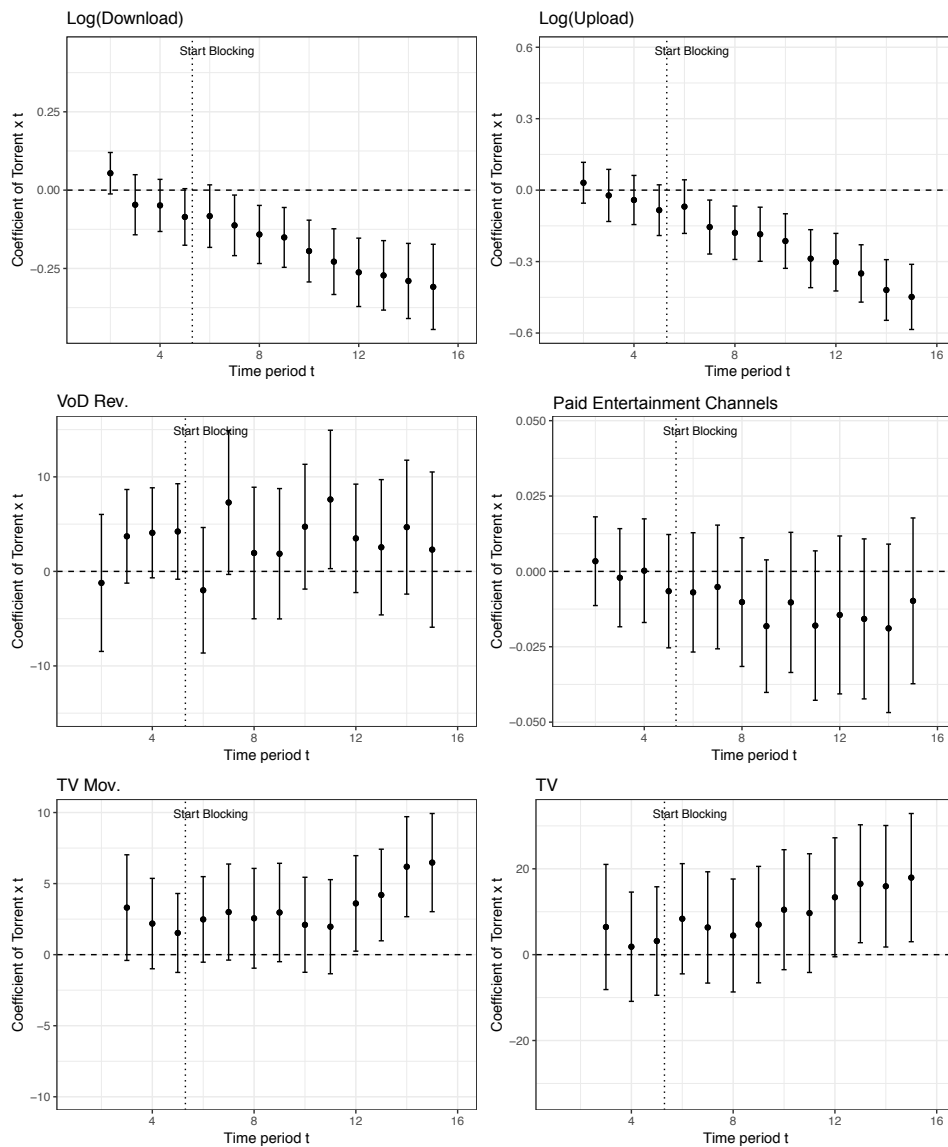


Figure 8 Impact of DNS blocks on downloads, uploads, video-on-demand expenditure, purchase of premium channels and usage of TV over time for households in the matched sample. Error bars are the 95% confidence interval.

increased their usage of free and legal entertainment content. In this section we study whether looking for ways to circumvent these blocks affects these results.

Table 5 presents the results of our region level analysis for the relationship between the search intensity for the keywords “DNS”, “Proxy”, and “VPN” and the fraction of active BitTorrent users in each region. The interaction term with the search index for DNS is positive and statistically significant ($p < 0.05$) in four of the five models we present below. This result suggests that, after the DNS blocks, country regions with greater search intensity for the keyword “DNS” had a statistically significantly higher fraction of active BitTorrent users than regions with smaller search intensity for this keyword. None of the interaction terms with “Proxy” or “VPN” are statistically significant. This may be explained by the fact that VPN and Proxy services entail financial costs and households that are tech savvy enough to buy these services could also probably change their DNS provider at no cost and thus opt for the latter. We also note that our results are robust to different polynomial time trends but we lose statistical significance when we use both time and region fixed effects in column (5). In fact, the whole regression in this column is itself not statistically significant ($F = 0.759$). This happens because we only have a small number of observations for this analysis resulting in only a very few degrees of freedom after the two way fixed effects are introduced.

Despite the correlational nature of this aggregated results, they provide some evidence that part of the BitTorrent users learned how to bypass the DNS blocks. The regions where this behavior was more prevalent present a higher fraction of active BitTorrent users after the blocks. Policy makers and industry players interested in preventing block circumvention may need to consider a combination of different website blocking methods in order to further raise the cost and difficulty of circumvention and increase the effectiveness of this anti-piracy policy.

8. Heterogeneous Effects of DNS Blocks

In this section, we investigate which household characteristics moderate their response to the DNS blocking policy. For this purpose, in addition to the main dataset on 100K households, we use data from a media survey that collected information on household demographics for 1.4K households. The descriptive statistics in Appendix B show that this selected sample of households are more interested in TV, entertainment and VoD than the average household in the original sample. Therefore, we start by replicating the analyses provided in section 7.1 over this sub-population. The results obtained are presented in table 6. As before, we find a significant decline in Internet usage across BitTorrent users, no change in VoD expenditure or in the subscription of premium TV channels, and an increase in the time devoted to watch TV channels that broadcast movies and TV shows. However, and unlike the results in section 7.1, we see that for these households the reduction in upload traffic is much larger than that in download traffic. This may suggest that some of these households substituted illegal BitTorrent activities by other over the top illegal or legal streaming services, such as Netflix. Content streaming takes up almost no upload traffic but keeps

Table 5 Number of households using BitTorrent after the blocks per region and as a function of Google searches for terms related to circumventing DNS blocks.

	<i>Dependent variable:</i>				
	Active BitTorrent Users per 1,000 households				
	FE (1)	FE (2)	FE (3)	FE (4)	FE (5)
After x DNS Google Search Index	7.315*** (1.513)	4.851** (1.992)	5.858*** (2.026)	5.359*** (2.074)	-0.603 (2.148)
After x VPN Google Search Index	-0.916 (2.439)	0.454 (2.084)	0.290 (1.948)	0.322 (1.889)	0.976 (1.662)
After x Proxy Google Search Index	-0.930 (3.454)	-0.806 (2.284)	-0.307 (2.211)	-0.482 (2.275)	-3.246 (2.406)
After	-18.062*** (1.860)	-15.711*** (1.504)	-15.925*** (1.734)	-14.113*** (2.408)	
DNS Google Search Index	-1.497 (1.175)	-5.223*** (1.533)	-6.817*** (1.726)	-6.170*** (1.816)	0.787 (1.548)
VPN Google Search Index	0.371 (1.684)	-1.005 (1.514)	-1.267 (1.315)	-1.237 (1.234)	-1.157 (1.173)
Proxy Google Search Index	1.064 (2.171)	-0.114 (1.816)	-0.355 (1.751)	0.067 (1.874)	2.642 (1.857)
T		2.169*** (0.395)	4.015*** (0.780)	6.006*** (0.940)	
T^2		-0.162*** (0.020)	-0.437*** (0.083)	-1.046*** (0.351)	
T^3			0.011*** (0.003)	0.070* (0.037)	
T^4				-0.002 (0.001)	
Month Dummies	No	No	No	No	Yes
Observations	210	210	210	210	210
F Statistic	91.615***	187.931***	179.378***	164.772***	0.759

Note:

*p<0.1; **p<0.05; ***p<0.01
FE is the within estimator; Cluster Robust Errors in ()

the download link significantly busy and thus this would explain the asymmetry in the reduction between upload and download traffic across these households. For completeness, we present figure 9 in appendix B where we show that the pre-block trends in download traffic, upload traffic, VoD expenditure, subscription of premium TV channels, TV viewership and the viewership of channels broadcasting movies and TV shows are similar for BitTorrent and non-BitTorrent users, which allows us to use difference-in-differences to study the effect of the blocks on these outcome variables.

Table 6 shows the results of the analysis of the heterogeneity in the households' response to DNS blocks using the original dataset. In order to test hypothesis 3a and 3e, we use the households' month bill in the first month of our period of analysis as a proxy for income and the indicator of whether the household opted for electronic billing in that same month as a proxy for computer literacy. The coefficient of the triple interaction term *BitTorrent User x After x Month Bill T1* is negative and marginally significant ($p < 0.1$) in column (2) – Uploads – and column (4) – Paid

	<i>Dependent variable:</i>					
	Log(Download) FE	Log(Upload) FE	VoD Rev. FE	Paid Entertain. Ch. FE	TV Mov. FE	TV FE
	(1)	(2)	(3)	(4)	(5)	(6)
After x BitTorrent User	-0.256*** (0.047)	-0.510*** (0.053)	-4.516 (5.031)	-0.015 (0.015)	5.247*** (2.006)	8.129 (5.497)
Month Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21,330	21,330	21,330	21,330	20,205	20,205
R ²	0.033	0.052	0.002	0.028	0.274	0.074
Adjusted R ²	-0.037	-0.016	-0.070	-0.042	0.222	0.008
F Statistic	44.947***	73.338***	2.771***	38.368***	475.201***	101.016***

Note:

*p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in ()
Clustering at individual level
FE is the fixed effects estimator

Table 6 Effect of DNS blocks on Internet traffic and use of legal alternatives (sample of households that answered the media survey).

Entertainment Channels, and not statistically significant in the remaining models. The evidence on the role of income as moderator of the effect of the DNS blocks is thus weak, only suggestive that higher income pirate households reduced slightly more their upload traffic but also slightly reduced their Paid Entertainment Channels subscriptions. The evidence is stronger on the moderator role of computer literacy. The triple interaction term *BitTorrent User x After x Flag Elect. Receipt T1* is positive and statistically significant for downloads and uploads ($p < 0.01$ and $p < 0.1$, respectively), suggesting that more tech-savvy pirate households reduced less their Internet traffic. Possibly, these households were able to circumvent the blocks or started using legal alternatives online. Regarding the legal alternatives to piracy considered, the triple interaction term is only statistically significant ($p < 0.05$) on model (4) – Paid Entertainment Channels – indicating that more tech-savvy pirate households subscribed more to this set of premium channels. Note that, when analysing variables pertaining to customers’ profiles such as the two billing variables considered, it is possible that unobserved household characteristics that correlate to the way households respond to the blocks but also to the way they respond to other events such as marketing campaigns or promotions (for which we are unable to control for) come into play. For instance, households that opt for electronic receipt may be more likely to open emails with targeted marketing and promotions that our Industry Partner regularly sends to its customers and thus change their behavior accordingly. Such factors could possibly explain the significant results of the triple interaction terms on the Paid Entertainment Channels subscriptions. This limitation is specific to our analysis of the heterogeneity of households response to the blocks as in the main analysis any effects are averaged out.

Table 7 expands on the analysis presented in table ?? by using the survey sample to test hypotheses 3a, 3b, 3c, 3d, and 3e. In addition to the variables of interest, all models also include a control

variable indicating whether the head of household had a higher education and for the presence of children in the household. Again, evidence on the moderator role of income (proxied by Month Bill before the blocks) is weak – the triple interaction term is only statistically significant ($p < 0.05$) in model (5), suggesting that higher income pirate households increased the time spent watching movie channels on TV less. Once more, evidence is stronger on the role of computer literacy. Households that opted for electronic billing reduced their Internet traffic less ($p < 0.05$ for downloads and $p < 0.01$ for uploads) and also increased the time spent watching movie TV channels less ($p < 0.1$). Regarding household composition, specifically the moderator role of teenagers, the results show that the coefficient of interest – the triple interaction *BitTorrent User x After x N. Teens* – is positive and statistically significant ($p < 0.05$ and $p < 0.01$, respectively) in columns (1) and (2) and negative and statistically significant in columns (3) and (6) ($p < 0.05$ and $p < 0.1$, respectively). Thus, the presence of teenagers in the household reduced the policy’s effectiveness in limiting piracy activity and promoting the use of legal alternatives to piracy. This results support hypothesis 3c and constitute a valuable insight for policy makers who may be interested in complementing current DNS blocking policies with educational programs targeting schools and/or high schools. Such programs could also highlight to parents the risks and liabilities that they may face if someone in their home is engaging in illegally sharing copyrighted content. Surprisingly, the triple interaction term *BitTorrent User x After x N. Young Adults* is not statistically significant in any of the models suggesting that the policy response of BitTorrent users from this demographic group was no different from that of adults. Finally, we find some support for hypothesis 3b on the moderator role of the household’s willingness to pay for video content. Specifically, we consider whether the households’ willingness to pay for video exceeds the price of the cheapest paid entertainment channels available (6.5 USD). The coefficient of the triple interaction term *BitTorrent User x After x Flag WTP 6.5* is negative and statistically significant in models (1) and (2) ($p < 0.1$ and $p < 0.05$, respectively), indicating that a higher willingness to pay is associated to a greater reduction in internet traffic although no corresponding take-up is observed in the set of legal alternatives considered.

9. Conclusion

We study the effectiveness of DNS blocks as a deterrent of piracy activity and incentive to the use of legal alternatives to consume media. We focus on one specific country where a DNS blocking protocol targeting websites providing copyright infringing content (video, audio, text, etc.) was implemented. We take advantage of a unique dataset on disaggregated household level media consumption to measure the impact of this policy on piracy and on the use of a set of legal alternatives not yet covered by the previous literature, namely TV use, Transactional VoD use,

	<i>Dependent variable:</i>					
	Log(Download)	Log(Upload)	VoD Rev.	Paid Entertain. Ch.	TV Mov.	TV
	FE	FE	FE	FE	FE	FE
	(1)	(2)	(3)	(4)	(5)	(6)
After x BitTorrent User	-0.316*** (0.029)	-0.258*** (0.031)	0.596 (1.941)	0.003 (0.005)	4.016** (1.601)	9.888** (4.722)
After x Month Bill T1	-0.001*** (0.0002)	-0.001*** (0.0001)	-0.021 (0.022)	0.001*** (0.00003)	0.180*** (0.008)	0.003 (0.027)
After x Flag Elect. Receipt T1	-0.060*** (0.008)	-0.029*** (0.006)	-1.311 (0.898)	0.011*** (0.001)	2.285*** (0.263)	6.867*** (0.883)
After x BitTorrent User x Month Bill T1	0.001 (0.001)	-0.001* (0.001)	-0.013 (0.040)	-0.0002* (0.0001)	0.030 (0.027)	-0.126 (0.078)
After x BitTorrent User x Flag Elect. Receipt T1	0.071*** (0.020)	0.040* (0.021)	0.972 (1.745)	0.010** (0.005)	-1.212 (0.894)	-1.183 (2.660)
Month Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,500,000	1,500,000	1,500,000	1,500,000	962,580	962,580
R ²	0.010	0.013	0.0004	0.016	0.196	0.023
Adjusted R ²	-0.061	-0.057	-0.071	-0.054	0.139	-0.047
F Statistic	740.436***	1,003.266***	27.005***	1,197.130***	11,556.430***	1,094.633***

Note:

*p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in ()
Clustering at individual level
FE is the fixed effects estimator

Table 7 Moderators of the effect of DNS blocks on Internet traffic and use of legal alternatives (original 100K sample)

	<i>Dependent variable:</i>					
	Log(Download)	Log(Upload)	VoD Rev.	Paid Entertain. Ch.	TV Mov.	TV
	FE	FE	FE	FE	FE	FE
	(1)	(2)	(3)	(4)	(5)	(6)
After x BitTorrent User	-0.156 (0.166)	-0.498*** (0.191)	-8.725 (21.357)	-0.067 (0.056)	20.435*** (7.524)	25.464 (20.897)
After x Month Bill T1	-0.001 (0.002)	0.0005 (0.002)	-0.474** (0.200)	0.001** (0.001)	0.286*** (0.069)	0.110 (0.206)
After x Flag Elect. Receipt T1	-0.214*** (0.075)	-0.238*** (0.071)	1.136 (9.378)	0.026 (0.019)	4.309* (2.433)	-0.672 (7.443)
After x N. Children	0.064* (0.035)	0.039 (0.038)	3.631 (4.442)	0.016 (0.014)	-3.163** (1.326)	6.111 (4.526)
After x N. Teenagers	0.031 (0.057)	-0.019 (0.054)	21.696** (9.577)	-0.014 (0.018)	8.422*** (2.391)	16.163** (6.470)
After x N. Young Adults	0.068 (0.089)	0.058 (0.085)	-4.685 (7.463)	-0.017 (0.021)	4.242 (3.757)	-6.073 (8.097)
After x Flag WTP 6.5	0.044 (0.073)	0.030 (0.067)	2.348 (7.411)	-0.021 (0.019)	1.355 (2.506)	-1.985 (7.544)
After x Flag High Educ.	0.037 (0.067)	0.025 (0.063)	-2.988 (8.087)	-0.026 (0.019)	-2.646 (2.387)	-0.991 (7.075)
After x BitTorrent User x Month Bill T1	-0.003 (0.002)	-0.003 (0.003)	0.259 (0.248)	0.0003 (0.001)	-0.232** (0.113)	-0.126 (0.309)
After x BitTorrent User x Flag Elect. Receipt T1	0.247** (0.102)	0.383*** (0.113)	-3.040 (11.057)	-0.006 (0.029)	-7.359* (4.180)	14.541 (11.652)
After x BitTorrent User x N. Children	0.005 (0.065)	-0.087 (0.070)	1.814 (5.350)	-0.002 (0.021)	-1.541 (2.340)	-11.178 (6.973)
After x BitTorrent User x N. Teenagers	0.165** (0.077)	0.306*** (0.083)	-25.477** (11.868)	0.020 (0.027)	1.681 (3.839)	-16.721* (9.818)
After x BitTorrent User x N. Young Adults	-0.098 (0.122)	-0.087 (0.121)	9.191 (8.629)	0.010 (0.028)	7.696 (5.037)	-7.255 (11.401)
After x BitTorrent User x Flag WTP 6.5	-0.185* (0.111)	-0.289** (0.120)	-7.783 (10.327)	0.026 (0.033)	-4.651 (4.088)	-11.672 (12.244)
After x BitTorrent User x Flag High Educ.	-0.059 (0.093)	0.016 (0.105)	-2.521 (9.925)	0.037 (0.030)	2.255 (4.020)	-3.573 (11.289)
Month Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21,330	21,330	21,330	21,330	20,205	20,205
R ²	0.038	0.059	0.003	0.034	0.283	0.077
Adjusted R ²	-0.032	-0.010	-0.069	-0.036	0.231	0.009
F Statistic	27.304***	42.709***	2.365***	24.227***	256.651***	54.046***

Note:

*p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in ()
Clustering at individual level
FE is the fixed effects estimator

Table 8 Moderators of the effect of DNS blocks on Internet traffic and use of legal alternatives (sample of households that answered the media survey.)

and paid TV channel subscriptions. Our results suggest that the policy was effective at reducing piracy, in our case proxied by Internet usage – both downloads and uploads. However, the take up

of legal alternatives is modest. VoD expenditure and the subscription of premium TV channels does not change. Only TV viewership increases slightly, in particular that of TV channels broadcasting movies and TV shows. We also find that Google searches for terms that are related to circumventing DNS blocks spiked right when the blocks started and remained higher after the blocks compared to pre-block levels. This provides some evidence that some BitTorrent users became more sophisticated in using tools that allow them to bypass these blocks. We find correlational evidence that the fraction of active BitTorrent users after the policy was higher in regions where this behavior was more prevalent. Finally, we find that the effectiveness of DNS blocks in reducing piracy was lower across households with teenagers and more computer literate households but higher across households with a higher willingness to pay for video content.

Our work contributes to the extant literature on digital piracy and on the effectiveness of piracy control strategies by empirically analyzing the effectiveness of an increasingly popular policy intervention - DNS-based website blocking. To the best of our knowledge, this is the first study analyzing the effectiveness of a batch website blocking policy based solely on DNS filtering technology. It is also the first work using disaggregated observational household level data to study the effectiveness of a website blocking policy. In addition to measuring the effectiveness of this policy in reducing piracy activity, we also measure its impact on a set on legal alternatives to consume media that have not been explored before in the literature. Finally, we complement this analysis with insights regarding whether pirates are likely to circumvent blocks and we find household characteristics that moderate their effectiveness, thus informing policy makers about the types of households that they may be willing to target. Our work also informs researchers, policy makers, and industry practitioners about the benefits and limitations of DNS blocking as a piracy control measure. However, our results also show that Internet service providers, who are a key player in implementing website blocking, do not necessarily have much to gain from devoting resources to implement this kind of measures (if anything, Internet traffic reduces). This may impact their predisposition to collaborate in the implementation of this or of similar anti-piracy efforts or else the costs associated to doing so need to be carefully shared across different industry players.

9.1. Limitations

This work is not without its limitations. First, all analyses pertain to a specific geography. This means that attitudes towards piracy and the observed responses to regulatory changes might be specific to local culture. Second, we know that a household in our sample was affected by the DNS blocks if it is marked as using BitTorrent prior to these blocks but we do not know who are the households that used streaming services that were also affected by the blocks. Therefore, some households in the control group may have been affected by the treatment and thus our estimates

are lower bounds for the average effect of treatment. Third, the media preferences survey that we analyze was not answered by all households in the original sample. This introduces limitations, namely the households that responded to this survey self-selected to do so and, therefore, are not a representative sample of the households included in the main analysis. Still, for some households the survey provides valuable demographic information that allows us to identify heterogeneous responses to the policy that prior research had not yet been able to produce. Fourth, the matched sample used in the robustness check drops a very large portion of the original sample. However, this is an inherent trade-off between the quality of the match and the resulting matched sample size. Finally, this work relies on observational data, and even if pre-intervention trends are similar across treated and control households, we can never rule out entirely that such trends could have diverged after the DNS-blocks for reasons unknown to us.

References

- Adermon, A., C.-Y. Liang. 2014. Piracy and music sales: The effects of an anti-piracy law. *Journal of Economic Behavior & Organization* **105** 90–106.
- Aguiar, L., J. Claussen, C. Peukert. 2015. Online copyright enforcement, consumer behavior, and market structure. *Consumer Behavior, and Market Structure (March 30, 2015)* .
- Al-Rafee, S., T. P. Cronan. 2006. Digital piracy: Factors that influence attitude toward behavior. *Journal of Business Ethics* **63**(3) 237–259.
- Angrist, J. D., G. W. Imbens, D. B. Rubin. 1996. Identification of causal effects using instrumental variables. *Journal of the American statistical Association* **91**(434) 444–455.
- BBCNews. 2012. File-sharers look to vpns to overcome pirate bay ban. URL <http://www.bbc.com/news/technology-17922214>. [Online; posted 2-May-2012].
- Belleflamme, P., M. Peitz, et al. 2014. Digital piracy: an update. Tech. rep., UCL.
- Belleflamme, P., P. M. Picard. 2007. Piracy and competition. *Journal of Economics & Management Strategy* **16**(2) 351–383.
- Bhattacharjee, S., R. D. Gopal, K. Lertwachara, J. 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**(1) 91–114.
- Blackburn, D. 2004. On-line piracy and recorded music sales. *Unpublished manuscript* .
- Bounie, D., M. Bourreau, P. Waelbroeck, others. 2006. Piracy and the demand for films: Analysis of piracy behavior in french universities. *Review of Economic Research on Copyright Issues* **3**(2) 15–27.
- Breindl, Y., F. Briatte. 2013. Digital protest skills and online activism against copyright reform in france and the european union. *Policy & Internet* **5**(1) 27–55.
- Chellappa, R. K., S. Shivendu. 2005. Managing piracy: Pricing and sampling strategies for digital experience goods in vertically segmented markets. *Information Systems Research* **16**(4) 400–417.
- Christin, N., A. S. Weigend, J. Chuang. 2005. Content availability, pollution and poisoning in file sharing peer-to-peer networks. *Proceedings of the 6th ACM conference on Electronic commerce*. ACM, 68–77.
- Clemente, G. 2015. Gli effetti sulla pirateria dei provvedimenti in materia di diritto d’autore. Tech. rep., Autorita per le Garanzie nelle Comunicazioni (AgCOM).
- Danaher, B., S. Dhanasobhon, M. D. Smith, R. Telang. 2010. Converting pirates without cannibalizing purchasers: The impact of digital distribution on physical sales and internet piracy. *Marketing Science* **29**(6) 1138–1151. doi:10.1287/mksc.1100.0600.
- Danaher, B., M. D. Smith. 2014. Gone in 60 seconds: The impact of the megaupload shutdown on movie sales. *International Journal of Industrial Organization* **33** 1–8. doi:10.1016/j.ijindorg.2013.12.001.

-
- Danaher, B., M. D. Smith, R. Telang. 2015. The effect of piracy website blocking on consumer behavior. *Available at SSRN 2612063* .
- Danaher, B., M. D. Smith, R. Telang. 2016. Website blocking revisited: The effect of the uk november 2014 blocks on consumer behavior. *Available at SSRN 2766795* .
- Danaher, B., M. D. Smith, R. Telang, S. Chen. 2014. The effect of graduated response anti-piracy laws on music sales: evidence from an event study in france. *The Journal of Industrial Economics* **62**(3) 541–553.
- De Filippi, P., D. Bourcier. 2016. three-strikes response to copyright infringement: The case of hadopi. *The Turn to Infrastructure in Internet Governance*. Springer, 125–152.
- Geng, X., Y.-J. Lee. 2013. Competing with piracy: A multichannel sequential search approach. *Journal of Management Information Systems* **30**(2) 159–184.
- Godinho de Matos, M., P. Ferreira, M. Smith. 2015. The effect of television content on pirates online activity: Evidence from a randomized experiment. si 2015 economics of it and digitization workshop.
- Gunter, W. D., G. E. Higgins, R. E. Gealt. 2010. Pirating youth: Examining the correlates of digital music piracy among adolescents. *International Journal of Cyber Criminology* **4**(1/2) 657.
- Higgins, G. E., S. E. Wolfe, C. D. Marcum. 2008. Digital piracy: An examination of three measurements of self-control. *Deviant Behavior* **29**(5) 440–460.
- Hollis, S., F. Campbell. 1999. What is meant by intention to treat analysis? survey of published randomised controlled trials. *Bmj* **319**(7211) 670–674.
- House, B. 2011. Technical report: An estimate of infringing use of the internet. *Analysis* 1–56.
- Johar, M., N. Kumar, V. Mookerjee. 2012. Content provision strategies in the presence of content piracy. *Information Systems Research* **23**(3-part-2) 960–975.
- Liebowitz, S. J. 2008. Research note-testing file sharing’s impact on music album sales in cities. *Management Science* **54**(4) 852–859.
- Mateus, A. M., J. M. Peha. 2011. Quantifying global transfers of copyrighted content using bittorrent. TPRC.
- Oberholzer-Gee, F., K. Strumpf. 2007. The effect of file sharing on record sales: An empirical analysis. *Journal of political economy* **115**(1) 1–42.
- Ofcom. 2010. site blocking to reduce online copyright infringement - a review of sections 17 and 18 of the digital economy act. Tech. rep., Ofcom.
- Peitz, M., P. Waelbroeck. 2006. Piracy of digital products: A critical review of the theoretical literature. *Information Economics and Policy* **18**(4) 449–476.
- Poort, J., J. Leenheer, J. van der Ham, C. Dumitru. 2014. Baywatch: Two approaches to measure the effects of blocking access to the pirate bay. *Telecommunications Policy* **38**(4) 383–392.

- Reavis Conner, K., R. P. Rumelt. 1991. Software piracy: an analysis of protection strategies. *Management Science* **37**(2) 125–139.
- Rob, R., J. Waldfogel. 2006. Piracy on the high cs: Music downloading, sales displacement, and social welfare in a sample of college students. *Journal of Law and Economics* **49**(1) 29–62. doi:10.1086/jle.2006.49.issue-1.
- Rob, R., J. Waldfogel. 2007. Piracy on the silver screen. *The Journal of Industrial Economics* **55**(3) 379–395.
- Sinha, R. K., F. S. Machado, C. Sellman. 2010. Don't think twice, it's all right: music piracy and pricing in a drm-free environment. *Journal of Marketing* **74**(2) 40–54.
- Smith, M. D., R. Telang. 2009. Competing with free: the impact of movie broadcasts on DVD sales and internet piracy 1. *MIS Quarterly* **33**(2) 321338.
- Solomon, S. L., J. A. O'Brien. 1990. The effect of demographic factors on attitudes toward software piracy. *Journal of Computer Information Systems* **30**(3) 40–46.
- Statista. 2014. Use of vpns and proxy servers worldwide in 1st quarter 2014, by generation. URL <https://www.statista.com/statistics/306955/vpn-proxy-server-use-worldwide-by-generation/>. [Online].
- Sundararajan, A. 2004. Managing digital piracy: Pricing and protection. *Information Systems Research* **15**(3) 287–308.
- Waldfogel, J. 2010. Music file sharing and sales displacement in the itunes era. *Information Economics and Policy* **22**(4) 306–314.
- Waldfogel, J. 2012. Copyright research in the digital age: Moving from piracy to the supply of new products. *American Economic Review* **102**(3) 337–342. doi:10.1257/aer.102.3.337.
- Watters, P. A., R. Layton, R. Dazeley. 2011. How much material on bittorrent is infringing content? a case study. *Information Security Technical Report* **16**(2) 79–87.
- Yoon, C. 2011. Theory of planned behavior and ethics theory in digital piracy: An integrated model. *Journal of business ethics* **100**(3) 405–417.
- Young, K. 2016. 4 things to know about vpn users. URL <https://www.globalwebindex.net/blog/4-things-to-know-about-vpn-users>. [Online; posted 2-February-2016].
- Zentner, A. 2006. Measuring the effect of file sharing on music purchases. *Journal of Law and Economics* **49**(1) 63–90. doi:10.1086/jle.2006.49.issue-1.
- Zhang, L. 2014. Intellectual property strategy and the long tail: Evidence from the recorded music industry. Available at SSRN .

Appendix A: Variable Definitions

Variable	Description
N Torrent (torrents/month)	Number of BitTorrent downloads in each month by household
Download (Mb/day)	Household's average daily download traffic (Mbps) in each month
Upload (Mb/Day)	Household's average daily upload traffic (Mbps) in each month
VoD (cents/month)	Household's total expenditure (USD cents) in Video on Demand per month
TV Total (min/day)	Household's average daily television view time (minutes) in each month
TV Entertainment (min/day)	Household's average daily entertainment channels view time (minutes) in each month
TV Video-on-Demand (min/day)	Household's average daily Video on Demand view time (minutes) in each month
Month Bill (USD)	Value (in USD) of household's monthly service bill
TV Service Tenure (month)	Months since the households subscribed to the television service
Internet Service Tenure (month)	Months since the households added an internet connection to its service bundle
Set-Top-Box Tenure (month)	Months since the household added a Set-Top-Box to its service bundle
Flag Premium	Households that have Set-Top-Box that allow tracking TV consumption
Flag Entertainment-Bundle	Indicator of whether the household purchased a set of premium movie and TV-shows channels in each month
Flag Sports-Bundle	Indicator of whether the household purchased a set of premium sport channels in each month
Flag Electronic Receipt	Indicator of whether the household opted for an electronic receipt in each month
Flag Bank Transfer	Indicator of whether the household opted for an automated bank transfer in each month
Flag Active Contract	Indicator of whether the household is locked in a contract in each month

Appendix B: Descriptive statistics for survey respondents

Statistic	N	Mean	St. Dev.	Min	Max
N Torrent (torrents /month)	610	6.905	20.702	0	225
Download (Mb/day)	610	2,794.062	2,566.066	0.000	21,661.460
Upload (Mb/Day)	610	1,164.247	1,901.067	0.000	19,214.270
VoD (cents/month)	610	13.932	87.552	0.000	1,077.600
TV Total (min/day)	478	282.468	154.986	19.000	923.167
TV Entertainment (min/day)	478	81.672	73.882	0.000	423.433
TV Video-on-Demand (min/day)	478	4.087	12.468	0.000	164.933
Month Bill (USD)	610	73.779	22.180	33.670	148.187
TV Service Tenure (month)	610	100.826	54.632	11.871	241.968
Internet Service Tenure (month)	610	78.107	34.428	14.290	188.677
Set-Top-Box Tenure (month)	610	61.303	20.125	0.710	88.387
Flag Premium	610	0.997	0.057	0	1
Flag Paid Entertainment Channels	610	0.389	0.488	0	1
Flag Sports-Bundle	610	0.169	0.375	0	1
Flag Electronic Receipt	610	0.575	0.495	0	1
Flag Bank Transfer	610	0.377	0.485	0	1
Flag Active Contract	610	0.918	0.275	0	1
N. Children in the Household (Ages 0 - 12)	610	0.539	0.836	0	8
N. Teenagers in the Household (Ages 13 - 18)	610	0.252	0.530	0	4
N. Young Adults in the Household (Ages 18-25)	610	0.279	0.559	0	3
Willingness to Pay for Content > 6.5 USD (Cheapest Paid Entertainment Ch.)	610	0.310	0.463	0	1
Flag Higher Education of Head of Household	610	0.562	0.497	0	1

TV time only available for households with certain set-top-box models

Table 9 Summary Statistics BitTorrent users that answered the media survey

Statistic	N	Mean	St. Dev.	Min	Max
N Torrent (torrents /month)	812	0.000	0.000	0	0
Download (Mb/day)	812	1,263.874	1,500.203	0.000	11,697.370
Upload (Mb/Day)	812	258.108	1,151.982	0.000	25,543.800
VoD (cents/month)	812	39.211	161.212	0.000	1,616.400
TV Total (min/day)	593	289.328	159.788	1.433	850.733
TV Entertainment (min/day)	593	71.818	73.972	0.000	493.400
TV Video-on-Demand (min/day)	593	3.747	10.755	0.000	125.867
Month Bill (USD)	812	68.243	21.470	23.387	149.812
TV Service Tenure (month)	812	80.185	59.505	3.129	243.581
Internet Service Tenure (month)	812	59.109	41.090	0.000	184.355
Set-Top-Box Tenure (month)	812	45.371	26.226	0.000	88.355
Flag Premium	812	0.984	0.126	0	1
Flag Paid Entertainment Channels	812	0.325	0.469	0	1
Flag Sports-Bundle	812	0.160	0.367	0	1
Flag Electronic Receipt	812	0.596	0.491	0	1
Flag Bank Transfer	812	0.435	0.496	0	1
Flag Active Contract	812	0.930	0.256	0	1
N. Children in the Household (Ages 0 - 12)	812	0.635	0.828	0	6
N. Teenagers in the Household (Ages 13 - 18)	812	0.250	0.502	0	3
N. Young Adults in the Household (Ages 18-25)	812	0.172	0.438	0	3
Willingness to Pay for Content > 6.5 USD (Cheapest Paid Entertainment Ch.)	812	0.328	0.470	0	1
Flag Higher Education of Head of Household	812	0.550	0.498	0	1

TV time only available for households with certain set-top-box models

Table 10 Summary Statistics Non-BitTorrent users that answered the media survey

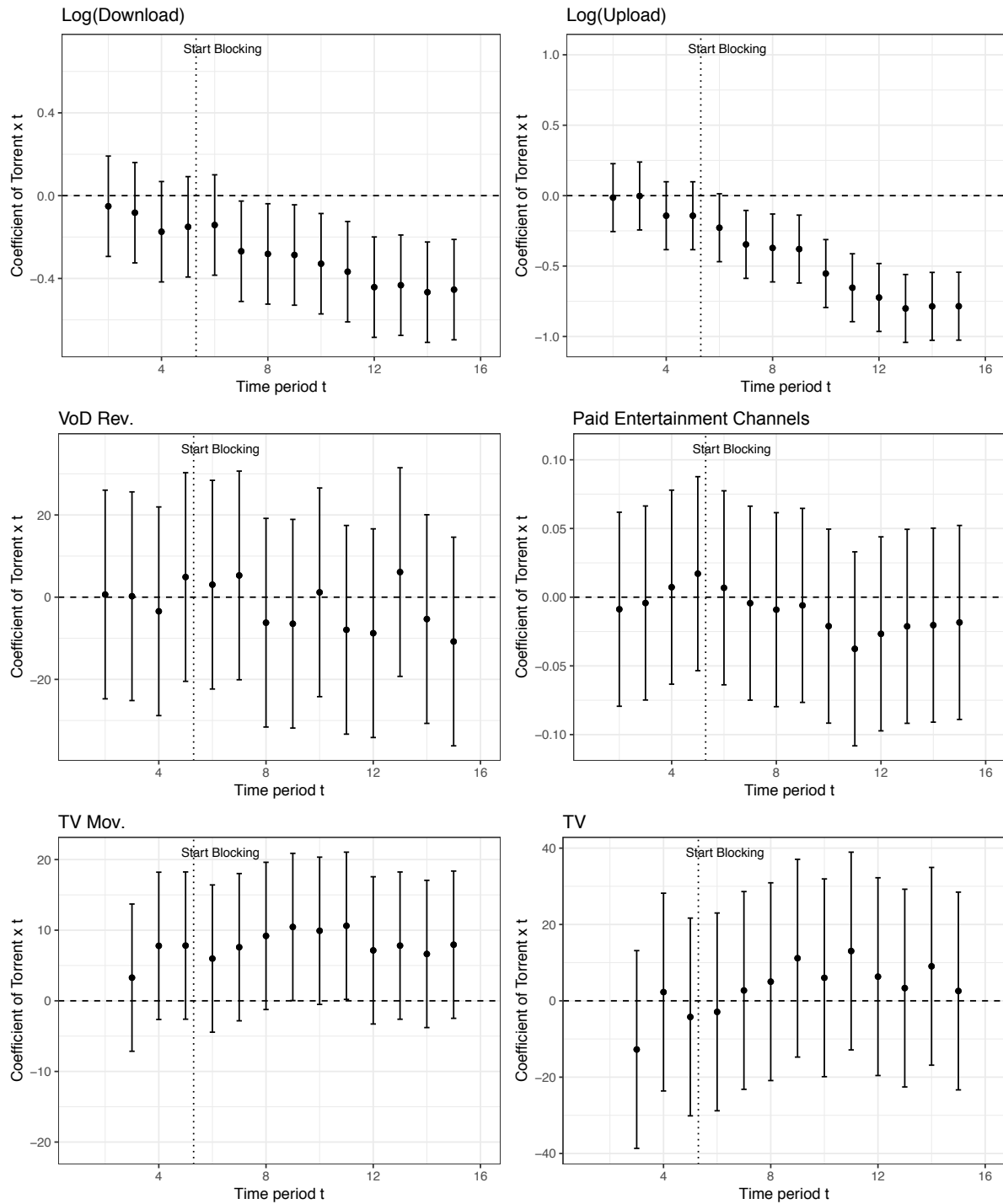


Figure 9 Impact of DNS blocks on downloads, uploads, video-on-demand expenditure, purchase of premium channels and usage of TV over time. Error bars are the 95% confidence interval.