The Impact of Time-shift Television on TV Viewership Behavior [*] [†]

Rodrigo Belo[4,a], Pedro Ferreira[1,2,b], Miguel Godinho de Matos[3,c], and Filipa Reis[1,3,d]

[1]Heinz College, Carnegie Mellon University
[2]Department of Engineering and Public Policy, Carnegie Mellon University
[3]Católica-Lisbon, School of Business and Economics
[4]Rotterdam School of Management, Erasmus University Rotterdam
[a]rbelo@rsm.nl
[b]pedrof@cmu.edu
[c]miguel.godinhomatos@ucp.pt
[d] frodrigu@cmu.edu

July 9, 2016

Abstract

We collaborate with a large telecommunications provider to study the impact of time-shift television (TSTV) on TV consumption. TSTV automatically records the contents broadcasted on TV and makes them available for a given number of days from the cloud. This technology significantly increases the amount and variety of content available to viewers and the level of flexibility with which it can be consumed at any point in time. For now, TSTV does not include additional dynamic ad placement (e.g. pre-rolls) and, in many cases, allows for unrestricted ad skipping. In this paper, we use both observational and experimental micro level data from two distinct stages of the life cycle of TSTV - a natural experiment around the initial introduction of TSTV by our industrial partner, and a randomized experiment deployed by this provider 3 years later. Our analyses of both these datasets provide consistent results showing that TSTV increases total TV time without cannibalizing live TV time. We also find that entertainment captures a disproportionate amount of the time-shifted viewership relative to live viewership and thus TSTV induces a change in the preferences of users.

[*]This work in part of Filipa Reis's PhD dissertation. Authors are ordered in alphabetical order.

[†]This work has been partially supported by PhD doctoral grant SFRH/BD/51568/2011, by the grant SFRH/BPD/94212/2013, and by the iLab and the Living Analytics Research Center at the Heinz College. We also thank our Industrial Partner for their support.

Finally, we find evidence of a positive impact of TSTV on the concentration of TV viewership consistent with a super start effect. Our findings have useful implications for advertising. We show that advertisers need to take into account that the introduction of TSTV displaces viewership across channels and that the sub-population of users who adjust consumption is different from the average user. Cable TV providers stand in a privileged position to assemble data on the characteristics and preferences of these users and can use this information to negotiate ad placement on top of TSTV in ways similar to what is done today in online markets using cookies.

1 Introduction

Traditional live television imposes two main constraints upon users. First, users interested in a specific program must accommodate programming grids within their own schedule, possibly compromising on their convenience; second, users with a fixed time slot available for leisure and media consumption will be restricted to the contents being televised during that interval of time, possibly compromising on their content choices.

Technological innovation has evolved towards providing increased convenience to TV viewers, first with the videocassette recorder (VCR), later with the DVD recorder, and more recently with Digital Video Recording (DVR) technology and Time-Shift Television (TSTV). Both DVR and TSTV allow users to go back in time and watch past programs but the former typically does not allow for recording concurrent programs, while requiring consumers to explicitly schedule recordings. TSTV has all the features of DVR with the advantage that one does not need to decide in advance which programs to record - it automatically records the contents broadcast on TV and makes them available to users for a given period of time (e.g. one week) from the cloud (Grece et al., 2015).

Time-Shift technology disrupts both the time and content constraints imposed by programming schedules upon viewers. It increases the variety and the amount of TV content available at all times while offsetting the restrictions of programming grids — viewers may select which programs to watch, pause, rewind, fast-forward and skip commercials. With TSTV, users are empowered at the expense of content providers and advertisers. As audiences shift in their choices of

what content to consume and when to consume it, accurate audience prediction and measurement becomes not only more difficult (Anderson and Gans, 2008) However, the exact way in which TSTV will impact TV consumption behavior is challenging to anticipate.

TSTV may trigger several concurrent changes to the way people watch television — it may impact the amount of TV consumed, the type of content consumed, and the time of consumption — and therefore, its impact on TV consumption is essentially an empirical question. The interplay of changes in these three aspects of TV consumption has critical implications for the industry. Namely, the shifts in audience size for specific time windows and content will necessarily affect advertisement audiences, both in size and demographic characteristics. Additionally, any reduction in live TV audiences in favor of Time-Shift consumption may negatively impact advertisement consumption if viewers engage in ad-skipping behavior similarly to what has been observed in DVR usage (Wilbur, 2008). Findings on these issues may call for a revision of current TV advertisement practices and entail the generation of new business models relying on the collection and exchange of household profile information for better advertisement targeting in close resemblance to current online advertisement practices.

In this work, we collaborate with a large telecommunications provider to investigate how TSTV impacts households' TV viewership behavior. Our analysis relies on two distinct data sources — an observational dataset from a natural experiment, and a dataset from a randomized experiment. The natural experiment was set off by the introduction of time-shift technology for over 50 TV channels to a select group of households served by our industrial partner. We use state of the art methodologies for causal measurement in observational data to infer how the introduction of TSTV impacted TV viewership behavior. We find that after the initial introduction of TSTV, households' daily TV time increased by 11 minutes ($p < 0.01$) (a 5% increase), and also find evidence consistent with the super-star effect described by Rosen (1981) — TSTV further skews the distribution of content viewership towards the most popular programs increasing

viewership concentration. Importantly, we found that the introduction of TSTV did not cause a reduction in the consumption of live television, which may lift some of the more immediate concerns of advertisers.

We rely in the experimental dataset to better grasp the mechanisms at work in regard to the distinct impact of increased content and increased flexibility on TV consumption. The randomized experiment was conducted by the same telecommunications provider 3 years after the introduction of TSTV. In this experiment, a random sample of subscribers was split among three experimental conditions: the first group received a set of 10 premium entertainment channels with TSTV for free; the second group received the same set of channels for free but without TSTV; and a third was assigned to a control condition in which subscribers were held out from any intervention. Note that while in the initial introduction of TSTV first studied, TSTV was associated to channels that were already available to subscribers, in this experiment, treated subscribers received a new (not previously available to them) set of channels with or without TSTV.

Comparing TV usage across households in the three experimental groups, we found that receiving the set of premium channels without TSTV led users to substitute content from the set of previously available channels for the new set of channels — on average, the new premium channels were viewed for 12 minutes per day ($p < 0.01$) while viewership of the set of previously available channels decreased in the same amount (corresponding to, approximately, a 4.6% decrease ($p < 0.01$) relative to the baseline of 263 minutes per day) — resulting in no significant changes in total TV view time. Similarly to what happened right after the introduction of TSTV, we found that receiving the set of premium channels with TSTV, caused an increase in total TV view time: on average, viewers watched the new premium channels for 19 minutes per day ($p < 0.01$) and decreased their viewership of the set of previously available channels by 14 minutes (corresponding to, approximately, a 5.3% percent decrease relative to the baseline of 263 minutes per day ($p < 0.01$)), resulting in a net increase of total TV time of 5 minutes per day ($p < 0.01$), which corresponds to a 1.9 percent increase rel-

ative to the baseline level of 263 minutes of daily TV time. Total live viewership time showed no significant alterations as a result of receiving the set of premium channels with TSTV. Consistent with the results from our analysis of the observational dataset, we find that receiving the new set of channels with TSTV led to a super-star effect Rosen (1981) — the most popular programs further increased their audience share and overall viewership concentration increased.

Despite having no impact on live viewership, receiving the set of free premium entertainment channels (with and without TSTV) stirred viewers attention away from channels with advertisement (mostly away from other entertainment channels) towards channels without advertising (the new premium entertainment channels), impacting households' advertisement view time. Using information on TV ads sponsored by our industrial partner and broadcast during the period of the experiment, we compared the viewership of these advertisements in the three experimental groups. Households in the control condition viewed an average of 820 seconds/month of these ads while those who received the set of premium channels without and with TSTV viewed 35 and 53 seconds less (corresponding to a decrease of 4.3% and 6.5%), respectively ($p < 0.01$). This analysis also allowed us to isolate the effect brought about by the technology alone — a difference of 18 seconds/month ($p < 0.05$) between those who received the premium channels with and without TSTV. Therefore, we do not find evidence that TSTV induced additional strategic avoidance of live advertisements as the observed reduction in live advertisement viewership is proportional to the overall reduction in view time of the channels in which such advertisement was broadcast. Households without and wth TSTV watched 6 min and 9.3 min less of TV in the original channels, and thus the shift in exposure to advertising is similar without and with TSTV ($35/53 \sim 6/9.3$). Our dataset does not allows us to observe ad consumption during time-shifted viewership hence we are unable to evaluate the existence and extent of ad-skipping during time-shifted viewership.

Finally, we found statistically significant differences between the profiles of households who consumed the same contents in live and time-shift, namely in

their use of TSTV prior to the experiment and in their preferences for specific types of TV content (e.g. entertainment, children, news).

In short, we find that TSTV increases the size of the "TV pie" (which could potentially be a positive effect for the industry) without hurting live viewership time (easing some of the most pressing concerns of advertisers). However, TSTV is also responsible for a shift in the contents consumed. Viewers tend to prefer entertainment content during time-shifted consumption, thereby shifting their attention from general content to entertainment, with consequences for advertisement viewership. Popular content becomes even more popular when TSTV is available, and viewership concentration increases. Audiences become fragmented between live and time-shifted consumption and the demographic characteristics of the two types of audience differ. These findings suggest that part of the value derived from advertising shifts to the content itself and the time at which it is consumed (not necessarily the same as the time of broadcast). By measuring the impact of TSTV on TV consumption and unraveling some of the mechanisms behind households' choices in the face of TSTV and their implications for advertising, this work provides valuable insights to researchers and stakeholders in the media industry — telecom providers, networks, and advertisers.

This paper is organized as follows. In section 2 we review the literature related to our research topic and put forth our research hypotheses. In section 3, we briefly describe the empirical context of our work and the datasets used. Section 4 presents the methods used in our analyses and section 5 presents the relevant descriptive statistics. Section 6 presents our results on the user level impacts of TSTV - TV view time, type of content watched, time of of viewership, advertisement viewership, and demographics of live and TSTV audiences. Section 7 presents our results on the program level impacts of TSTV, namely, its impact on viewership concentration. Finally, in section 8 we discuss our findings and conclude.

2 Related Work and Hypotheses

In this work we do an in-depth analysis of how TSTV impacts TV consumption behavior . Our hypotheses draw from several streams of literature that inform us on the likely impact of TSTV on TV consumption.

First, TSTV may impact the amount of TV consumed. According to Hinz and Eckert (2010) all markets aimed at leisure must be considered unsaturated markets — markets in which additional consumption is possible as opposed to saturated markets in which additional consumption in only possible through within-market substitution — as consumption of one leisure activity can generally increase at the expense of other leisure activities (cross-market substitution) (Hinz and Eckert, 2010). TSTV effectively increases product variety for consumers with respect to television programming and it does so with no costs for consumers. Previous research studying the consumer welfare effects of increased product variety on the Internet found that greater variety resulted in meaningful gains in consumers' welfare as consumers were able to located products better matched to their preferences (Brynjolfsson et al., 2003). We hypothesize that by allowing viewers to find contents that are better matched to their preferences, TSTV will lead households to increase their TV consumption.

Hypothesis 1: TSTV has a positive impact in total TV time.

Second, TSTV may impact the content choices of users. More variety may either trigger long-tail-like phenomena (Brynjolfsson et al., 2006) causing the distribution of content shares to drift away from prime time and popular content towards the tail of the programs viewership distribution or, alternatively, a superstar effect (Rosen, 1981; Elberse and Oberholzer-Gee, 2006) leading consumers to allocate even more time to prime time and popular content. By increasing the amount and variety of contents available to consumers TSTV may also impact viewers' search costs — recent studies show that search costs increase with the number of products that consumers need to scan. For example, (Boatwright and Nunes, 2001; Kuksov and Villas-Boas, 2010) report situations in which increased

variety reduced sales. Thus, although classic economic theory predicts that more variety increases consumer welfare (Hotelling, 1929; Dixit and Stiglitz, 1977; Salop, 1979), with too much variety consumers may be unable to internalize the potential benefits of increased choice (Sawhney and Eliashberg, 1996).

In the particular case that we study, even though TSTV increases the amount and variety of content available to users at any point in time, it does not provide any additional recommendation system or search filters to users. This means that search costs will be different for different types of contents - blockbuster movies will have lower search costs on TSTV as they have easier recall (resulting from advertising, word of mouth and such factors) (Frank and Cook, 1995). In the case studied, the TSTV menu does include a search box where the user may search for a specific movie title. However, previous research on the effect of search costs and search tools on product concentration, has shown that this sort of directed search has no significant impact on the consumption of niche products (Brynjolfsson et al., 2011a). For these reasons, we expect that TSTV will further increase the audience share of popular contents and, as a result, increase the concentration of TV viewership.

Hypothesis 2: TSTV increases TV viewership concentration.

Thirdly, with TSTV programs that viewers would have otherwise forgone watching due to an inconvenient schedule can easily be viewed at a different time. Similarly, programs that viewers would otherwise watch due to their availability may be easily replaced by programs that better match the viewer's preferences. We therefore expect viewers to take advantage of the increase flexibility and, to some extent, reorganize their TV viewing schedule at their convenience. As a side effect of each viewer's shifts in both content choices and time of consumption, we expect TSTV to impact the demographic characteristics of live audiences. People who prior to TSTV would consume certain contents will now opt for other contents, and people who would watch TV at a specific time may now do so at a different time.

Hypothesis 3: TSTV changes households' TV viewership schedule.

Hypothesis 4: TSTV changes the demographic profile of households who consume live TV.

Additionally, when several functionally similar media channels are available for satisfying a certain need, users will chose the one that provides the better fit to her needs (Rubin, 2002). An improvement in quality of a given media channel should lead to substitution between that and the other channels satisfying the same set of needs (Rubin, 2002; Ferguson and Perse, 2000). As such, following our initial hypothesis that TSTV will have a positive impact on TV time, we hypothesize that this increase in TV viewership will take place at the expense of other functionally similar leisure activities, in particular, live television.

Hypothesis 5: TSTV negatively impacts live TV time.

Finally, we hypothesize that as audiences attention is stirred away from the consumption of live TV (which generally includes commercial breaks) towards the consumption of time-shifted TV (during which commercial breaks can be skipped as has been found to be the case with TiVo and DVR usage (Downey, 2007; Pearson and Barwise, 2007)), TSTV will have cause a reduction in advertisement view time.

Hypothesis 6: TSTV negatively impacts advertisement view time.

The implications of our set of hypotheses are most relevant for the advertisement industry, which still plays a dominant role on TV network's profitability[1].

TV advertising has traditionally been a two-sided model in which media channels use content to attract viewers and then sell the access to those viewers to advertisers (Evans, 2008). The prices of advertisement spots are negotiated between advertisement agencies and broadcasters or channel networks based on the estimated size and characteristics of the audiences that specific contents are able to attract (which are generally measured by external rating agencies) (Evans,

[1]By 2010 TV advertising amounted 0.47% of the US gross domestic product (GDP) and revenues from direct pay subscriptions reached 0.6% of GDP (Waterman et al., 2012). Four decades earlier, advertising revenues represented only 0.35% of GDP while direct pay contributions amounted to 0.03% of GDP (Waterman et al., 2012).

2008). The ability of specific contents to attract the desired audiences will determine which content is televised and which is not (Evans, 2008).

TV advertising revenues have been facing growing competition from online advertising, which has been consistently gaining ground due to its ability to provide advertisers with a largely more efficient way for targeting the desired consumers (Evans, 2009). Online advertising not online deviates ad money from TV to the Internet but also exerts downward pressure on the value of TV advertisement spots with consequences for the network's ability to broadcast quality content (Evans, 2009). The expansion of digital TV and modern TV delivery methods such as IPTV allows TV providers to implement online advertisement placement methods on TV (Evans, 2009) and imitate the online advertisement business models.

Though TV targeting is likely to never reach the level of refinement made possible online (through keyword search and tracking of one's browsing history) much can nonetheless be done in what concerns the characterization of the viewer's location, day and time of viewership, content preferences, viewership format (live, TSTV, DVR), and device used. Digitization makes this characterization increasingly more cost effective surpassing the limitations previously imposed by analogue systems. Time-shift technologies such as TSTV and DVR provide pay-TV providers with further opportunities for assimilating online advertisement practices such as pre, post, or mid-roll ads associated to specific content and targeted to specific demographics and contexts of viewership and approximating the near real-time targeting process of the online world. Additionally, time-shift technologies allow for more accurate audience measurement (Wilbur, 2008) while also openning the door for the sort of ad-effectiveness measurement experimentats that are currently done online (Johnson et al., 2015) to be done on TV as well.

Unlike the web, where advertisers (websites) collect digital cookies themselves and can then exchange these cookies between them, TV channels cannot do so and thus the ability to perform this kind of data collection and processing lies solely on the hands of the TV service provider. Hence, TV providers are in an

unique position to collect and leverage upon the demographic and TV viewership activities of their subscribers largely expanding upon current audience metrics, demographics, and behavior characterization. By imitating the online advertising business model TV may be able to stop further losses in its advertising revenues. Such recovery would have economic implications not only for channel networks and pay TV providers but also for the other advertising industry stakeholders relying primarily on traditional media (e.g. ad agencies)

Up until now, much of the previous literature on the impact of time shifted TV (DVR or TSTV) on media consumption has focused on the consumers behavior to avoid advertisement (Wilbur, 2006a,b; Anderson and Gans, 2008). One should note that even in the absence of time-shifting technologies such as DVR or TSTV, advertisement avoidance is a common behavior among viewers who either change channels, deviate their attention or engage in other activities during commercial breaks (Van Meurs, 1998; Krugman et al., 1995; Tse and Lee, 2001). Ad avoidance varies according to ad content — type of product or service advertised, advertiser, and content elements of the ad (Liaukonyte et al., 2015).

Existing evidence on the impact of time-shift technologies on advertisement viewership is mixed. Previous work has investigated how technologies such as DVR impact advertising revenues by allowing users to skip TV commercials (Wilbur, 2008) and TiVo and DVR users have been shown to skip through most of commercials when consuming recorded contents (Downey, 2007; Pearson and Barwise, 2007). Contrary evidence was provided by Bronnenberg et al. (2010) who analyzed data from a three year field study carried in partnership with five firms in which a sample of fourteen thousand households were solicited to accept TiVo for free[2]. Using propensity score matching, the authors found that skipping advertising occurred relatively infrequently and were unable to reject the null hypothesis that DVR did not change consumer purchase behavior.

[2]Solicitation was non-random, and was based on households that had IRI's BehaviorScan service, which is a testing service that allows quantifying the ROI of advertising by showing different commercials to different households and monitoring the purchasing behavior of willing panelist participants.

This work contributes to the stream of literature on the impact of time-shift technology on households media consumption behavior. We use micro level data on households' TV viewership and TSTV use that covers both the period of the initial introduction of TSTV by a large telecom provider and a later stage when the technology is mature. TSTV provides viewers with even more content and flexibility than those provided by other time-shift technologies (such as TiVo or DVR) that have previously been studied in the literature (Wilbur, 2008; Anderson and Gans, 2008; Zigmond et al., 2009). Unlike previous work that relied on observational data alone (Zigmond et al., 2009; Downey, 2007; Pearson and Barwise, 2007), we use both observational data and data from a randomized field experiment — the gold standard for inferring causal relationships in social sciences (Bapna and Umyarov, 2012). We find that TSTV increases total TV view time while causing no changes in live view time. We also find that entertainment content captures a disproportionate amount of time-shift viewership relative to live viewership and that TSTV stirs viewers attention away from general purpose channels to entertainment channels. We expand on previous work on the impact of time-shift technologies on advertisement viewership (Wilbur, 2008; Anderson and Gans, 2008; Zigmond et al., 2009), which has mostly focused on ad-skipping behavior, by showing that, TSTV also impacts advertisement viewership by leading consumers to watch different contents on TV. Finally, we find evidence of a positive impact of TSTV on TV viewership concentration consistent with the super start effect (Rosen, 1981; Elberse and Oberholzer-Gee, 2006).

3 Data

3.1 Empirical Context

This work was developed in collaboration with a multinational telecommunications provider. Our analysis will focus on one region where our industrial partner is the market leader in the Pay-TV segment with over one million sub-

scriber households. The firm's services include Pay-TV, Video-on-Demand, Digital Video Recording, Automated Cloud Recording, online TV, broadband and mobile Internet, and fixed and mobile telephony.

Subscribers can opt for either the standard or premium service, which differ in the number of TV channels and on the set of complementary features available. Our study focuses on premium subscribers who, in 2012 accounted for roughly 40% of the households subscriber base having grown to around 90% in 2015. Premium subscribers have at least one Set-Top-Box (STB) with over one hundred TV channels, a high-speed internet connection and unlimited fixed telephone.

On the summer of 2012, our industrial partner introduced TSTV to premium subscribers at no additional cost. This extended the previously available DVR service. With TSTV, subscribers became able to watch time-shifted TV with rewind capabilities that went as far as one week. TSTV was initially available for about half of the channels available from this provider and was available for the majority of its channel offering by 2015.

The TV channel offering from our industrial partner's basic service bundle can be complemented with additional thematic channel packages (e.g. children, music, sports, documentaries, and movies and TV shows), which can be purchased separately for a fixed monthly fee. The *movie bundle* is a set of 10 premium movie and series channels that can be purchased for an additional fee of 13 dollars/month. This pack of channels includes 8 movie channels and 2 TV-series channels. Most contents broadcast are recent and popular movies and series. In particular, the TV-series broadcast on these channels air only a couple of days after their U.S. broadcast.

3.2 Data

3.2.1 Media Consumption Dataset

Our observational dataset focuses on a random sample of 10,000 premium subscribers who were given TSTV for over 50 TV channels at the time of its in-

troduction. This dataset also includes a sample of 50,000 randomly selected standard subscribers who were not given access to TSTV. This dataset spans the period comprehended between July 15 and October 15 of 2012. Our experimental dataset focuses on 40,500 subscribers who were part of a randomized experiment run for a period of 6 weeks between May 13 and June 30, 2015 by our industrial partner.

Both our observational and experimental dataset were built using anonymized click-stream data from households subscribing to the triple-play services provided by our industrial partner. This was made possible due to a specific API feature whereby Set-top Boxes request information about the content they start playing. This feature allows the provider to infer how much time each user spends watching a given program and whether viewership was live or time-shift. Events are coded *live* if all viewing time is within the original program airing time slot; and coded *TSTV* if viewing either starts before the end of the program and ends after the program finishes or starts after the end of a program. This results in the following data: Live TV viewership time aggregated hourly by channel and channel type (aggregated in categories such as entertainment, children, sport, and news) for each household in our sample. TSTV viewership time aggregated hourly and structured similarly to the live TV data.

Additionally we were given access to an anonymized customer purchase history — a monthly snapshot of the services purchased by each subscriber including information on the subscriber's service tenure and monthly service fee. Finally, we were also given access to the aggregate daily fixed Internet download and upload traffic in MB for each household in our sample for the duration of the randomized experiment. We were not able to access the Internet activity data from 2012.

3.2.2 Advertisement Dataset

Our industrial partner granted us access to a dataset including all TV advertisement spots that it purchased and that were broadcast between June 1 and June

30, 2005. For each ad that was broadcast, the dataset includes all viewership instances by the subscribers in our experimental sample: each entry includes an anonymized subscriber identifier, the corresponding advertisement viewed (uniquely identified by an identifier code, timestamp of transmission, and the channel in which it was broadcast) and the duration in seconds of both the advertisement and the viewership event.

We note that by focusing only on advertisements sponsored by one specific company, this dataset may not be representative of the overall advertisement distribution in what regards the channels or the schedule selected for these advertisements to be broadcast or the advertisement characteristics[3].

4 Methods

We analyze the impact of TSTV at two main levels — user level and program level. Our user level analysis relies on difference-in-differences (diff-in-diff) for estimating the impact of TSTV on a number of user level outcomes, while our product level analysis follows the same methodology used by Brynjolfsson et al. (2011b) — the Pareto curve — for estimating the effect of search costs on the concentration of product sales to analyze the impact of TSTV on TV viewership concentration. The variables used in our analyses originate from two different datasets — a natural experiment (observational) and randomized experiment dataset. The natural experiment resulted from the introduction of TSTV by our industrial partner to all premium consumers in 2012. The randomized experiment was conducted by our industrial partner in 2015. On the one hand, the natural experiment dataset has the advantage of allowing for the analysis of the impact of TSTV when it is introduced to most of the TV channels relevant to viewers and not only to a limited set of entertainment channels. On the other hand, the experimental dataset though focusing only on a specific set of channels, allows us

[3]We are currently working towards obtaining a similarly structured dataset for advertisements by other organizations which we hope to be able to include in a future version of this work.

to address the endogeneity concerns associated to the use of observational data. Hence, most of our analyses use the experimental dataset as it is both the most complete dataset and the one that allows for unequivocal causal interpretation of results. The user level analyses conducted using the observational dataset rely on propensity scores for creating an appropriate control group from the subset of households who did not received TSTV in 2012 (McCaffrey et al., 2004). The main outcome variables considered in this work and the datasets and estimation procedures used are summarized in table 1.

4.1 User Level Analysis

The goal of this analysis is to compare the behavior of consumers with TSTV — *Treated* — with that of consumers who do not have TSTV — *Control*. We use differences-in-differences to perform this analysis. The equation below presents our general empirical specification for measuring the effect of TSTV:

$$Y_{it} = \beta_1 After_t + \beta_2 Treated_i + \beta_3 Treated_i * After_t + \mu_{it} \quad (1)$$

The subscript i corresponds to a subscriber (household) and t to the time-period (before or after the experiment). β_3 gives us the impact of receiving TSTV on the outcome variable considered[4].

4.1.1 Natural Experiment

In our analysis using the observational dataset, the control group was artificially created using propensity score matching. We used inverse probability of treatment weighting (IPTW), i.e., we weight observations of standard consumers by

[4]As a robustness check, the analyses of the impact of treatment on TV consumption using the experimental dataset were also conducted using subscriber level fixed effects and the full day-level panel data set. The same empirical approaches were used with the addition of day-level fixed-effects in the analysis using the day-level data. These analyses yielded consistent results with the ones here presented and are available upon request.

Table 1: Data and Methods Overview

	User Level Used for estimating the effect of TSTV on TV time, time of viewership, and advertisement viewership.		**Program Level** Used for estimating the effect of TSTV of TV viewership concentration.	
	2012	**2015**	**2012**	**2015**
Data Source:	Natural Experiment	Randomized Experiment	Natural Experiment	Randomized Experiment
Period of Analysis:	July 16 to October 15	May 1 to June 30	July 16 to October 15	May 1 to June 30
Estimation:	PSM and Diff-in-Diff	Diff-in-Diff	Diff-in-Diff	Diff-in-Diff
Dependent Variables:				
TV time Live	✓ (Tables 2 and 11)	✓ (Table 2)	✓	✓
TV time TSTV	✓ (Table 2)	✓ (Table 2)	✓	✓
TV time Total	✓ (Table 2)	✓ (Table 2)	✓ (Table 8)	✓ (Table 9)
TV time by Content	-	✓ (Tables 13 and 14)	-	-
TV time by Hour	-	✓ (Table available upon request)	-	-
Advertisement View Time	-	✓ (Tables 4 to 7)	-	-
Internet Traffic (Mb)	-	✓ (Table 12)	-	-
N	60K households: 10K with TSTV 50K without TSTV (for matching)	40.5K households: 13.5 with content and TSTV 13.5 with content without TSTV 13.5 without content or TSTV	306K programs: 3372 programs/day for 91 days	2K programs

their propensity to be treated, to ensure that both groups of consumers are similar on average in observed covariates that determine whether they had the TSTV feature (Rosenbaum, 1987). Propensity scores were calculated using boosted regression modelling (McCaffrey et al., 2004; Guo and Fraser, 2014). This strategy reduces the probability that standard and premium consumers are different in unobserved covariates related to the propensity of being treated, which in turn reduces the selection bias.

The success of propensity score weighting as an empirical strategy to study the effects of natural experiments relies on finding all key relevant covariates that may determine the likelihood of treatment. In our case, we need to control for the factors that determine whether a consumer is standard or premium. We calculated propensity scores based on covariates that proxy consumers' interest on TV and on their availability to watch TV. We proxied their interest in TV with the number of days consumers watched TV per month and the number of programs they watched per month. We also proxied their availability to watch TV by the time they spent watching TV and by the hour of the day they turned on the TV. Appendix B provides more details on the matching procedure.

We aggregate our observational dataset at the week level, resulting in 7 time periods before the introduction of TSTV — *Before* — and 6 after its introduction — *After*. The difference-in-differences specification is run over the weighted sample with those who received TSTV being the *Treated* group and those who did not receive this technology being the *Control* group.

4.1.2 Randomized Experiment

The randomized experiment consisted in splitting a random sample of 40,566 subscribers who had not previously subscribed to the *movie bundle* among three experimental conditions of equal size[5]. Subscribers in the first experimental con-

[5]The experiment's design is closely related to that presented in previous work by Godinho de Matos et al. (2015). This work studied the substitution patterns of TV and Internet in a population of households that engaged in digital piracy. The authors analyzed the impact of giving premium TV channels with TSTV on

dition, *Treated TS* received the *movie bundle* channels with full access to the time-shift feature for free for a period of 6 weeks. Subscribers in the second experimental condition, *Treated No TS*, received the *movie bundle* channels without the time-shift feature during the same time period. Subscribers in the third experimental group, *Control*, were held out from any intervention for the duration of the experiment.

In this experiment, a random sample of subscribers was split among three experimental conditions: the first group received a set of 10 premium entertainment channels with TSTV for free; the second group received the same set of channels for free but without TSTV; and a third was assigned to a control condition in which subscribers were held out from any intervention. Note that while in the initial introduction of TSTV first studied, TSTV was associated to channels that were already available to subscribers, in this experiment, treated subscribers received a new (not previously available to them) set of channels with or without TSTV.

Subscribers were notified about the temporary offer of the *movie bundle* (with and without TS) by both text message and email [6]. No setup action was needed on their part to access the channels as they became automatically available

whether digital pirates stopped pirating. The authors found that consumers who received the premium TV channels increased their consumption of TV and decreased their download traffic. However, as all treated households received the premium TV channels with TSTV, the authors were unable to tease out whether TSTV was responsible for the effects found. The results from this work were used to test if the selected sample size would yield enough statistical power for econometric identification - the calculations indicated that for a significance level of 5 percent and 80 percent of statistical power, it sufficed to identify effects of a magnitude smaller or equal to those identified in Godinho de Matos et al. (2015).

[6]Subscribers in the control condition were not locked-out from voluntarily subscribing to the *movie bundle* during the treatment period. As such, it is possible that some of those allocated to the control group have defied their treatment assignment by opting into treatment. Similarly, nothing forces a household that received the *movie bundle* to actually view these channels, which means that we may have non compliance with the treatment as well. Consequently, our difference-in-differences estimations will give us a conservative estimate of the effect of TSTV (Angrist et al., 1996) known as the effect of the intention to treat (Hollis and Campbell, 1999).

through remote activation by our industrial partner. For technical reasons, the remote activation of the *movie bundle* for all users in our sample took nearly one week to complete, having followed a random order in regard to the treatment conditions. This period of data was removed from our analysis. Additional details on the experiment design and descriptive statistics can be found in the appendix, section A.

We aggregate our experimental dataset in two periods — before the experiment (May 1st to 12th, 2015 — *Before*) and during the experiment (May 18th to June 30th, 2015 — *After*). In our main analyses, we compare the behavior of those who were given access to the *movie bundle* with TSTV — *Treated* — to those who received this set of channels without TSTV — *Control*. When relevant, we also compare those who received the *movie bundle* with and without TSTV against the set of households that were held out from intervention — *Control*.

4.2 Program Level Analysis

The goal of this analysis is to measure the impact of TSTV on TV viewership concentration. We follow the methodology of Brynjolfsson et al. (2011b), who use the Pareto distribution to model product sales as a function of its rank. The cumulative density function for a random variable in $[1, +\infty]$ with a Pareto distribution with shape parameter $\beta > 0$ is given by $f(x) = 1 - x^{-\beta}$. The shape parameter measures, for example, the decay in sales across rank. A larger shape parameter will place more density towards the head of the distribution, that is, the better ranks, for a higher concentration.

Using our observational dataset, we study how the introduction of TSTV affected the shape parameter of the distribution of viewership time per program. In this analysis each observation pertains to a program in a viewership type, either live or total. This means that there are two observations per program. We pool all observations together and use an indicator variable, $Total_i$ to identify

which type of viewership each observation pertains to. Therefore, our model is:

$$log(Time_i) = \alpha_0 + \alpha_1 log(Rank_i) + \alpha_2 TotalTV_i + \alpha_3 log(Rank_i) \times Total_i + \varepsilon_i$$

where $Time_i$ represents the viewership time for observation i, $Rank_i$ represents the rank associated to that observation.

The parameter of interest in this case is α_3, which measures the differences in the decay between total and live viewership. In this setting identification relies on the assumption that live viewership did not change from the pre-TSTV period to the post-TSTV period. If, for example, concentration in live viewership decreased with the introduction of TSTV, the decay of live viewership would have become flatter, which would lead to an over-estimation of differences between live and total viewership time and, consequently, to an upper bound for the increase in concentration caused by TSTV.

Our analysis using the experimental dataset allows us to overcome this potential overestimation problem. We follow a similar approach to that just described with a few differences. Each observation pertains to a program in the *movie bundle* channels during the experimental period. For each program we computed the total viewership time by households who received this set of channels with and without TSTV and its corresponding rank in the two groups. We pool all observations together and use an indicator variable, $TreatedTSTV_i$ to identify which treatment condition the observation pertains to. Any differences in the viewership time of these programs by the two groups of households can be causally attributed to TSTV. Our model for this analysis is the following:

$$log(Time_i) = \alpha_0 + \alpha_1 log(Rank_i) + \alpha_2 TreatedTSTV_i + \alpha_3 log(Rank_i) \times TreatedTSTV_i + \varepsilon_i$$

where $Time_i$ represents total view time of program i, and $Rank_i$ represents the rank associated to that program .

5 Descriptive Statistics

5.1 Live and Time-Shifted TV Viewership

In this section we characterize the consumption of live and time-shifted television in both the period immediately following the introduction of TSTV (2012) and the later period during which the randomized experiment took place (2015). Note, that the household samples from each period of analysis are distinct — 2012's data concerns the group of premium subscribers who were initially given this technology and 2015's data concerns the experimental group that was left out from any intervention (the control condition).

On average, households in our 2012's sample, watched around 3.4 hours of television per day prior to the introduction of TSTV. In our 2015's household sample the average total TV time (including both live and time-shifted consumption) amounts to 4.4 hours per day[7]. General content channels (such as general purpose free to air channels) attract most of audiences' attention (40%), followed by entertainment channels (15%), children channels (9%), news channels (7.6%) and sport channels (3.8%). People are the least likely to watch live television during night-time and dawn hours and become increasingly more likely to watch TV as the day evolves with prime-time (the time interval between 8 pm and midnight) being the period of the day during which people are the most likely to watch live television. People are most likely to watch time-shifted television during the late night (between 10 pm and 1 or 2 am), likely due to the lack of quality programming during this time. Unlike live TV viewership, time-shifted viewership drops at around 7 to 9 pm, the time when most news shows are transmitted on the free to air channels.

Figures 1 shows the overall use of TSTV as a fraction of total TV time during the two periods of analysis considered - 2012 and 2015. On the lefthand side fig-

[7]These values are consistent with those reported by official audience tracking entities for the geography considered. For privacy reasons we are unable to disclose this sources to readers but will make them available to reviewers upon request.

ure we can observe that after the initial introduction of TSTV, TSTV viewership amounted, on average, to 7.5% of the total TV viewership time of the households who received this feature. This statistic increased slightly to nearly 8 % three years later, visible on the righthand side figure, in the period preceding May 15. The righthand side figure also shows that once the randomized experiment was in place households in the different experimental groups altered their consumption of time-shift TV, in particular, those who received the set of new premium entertainment channels with TSTV show a slightly increased use of this feature relative to users in the two other experimental conditions.

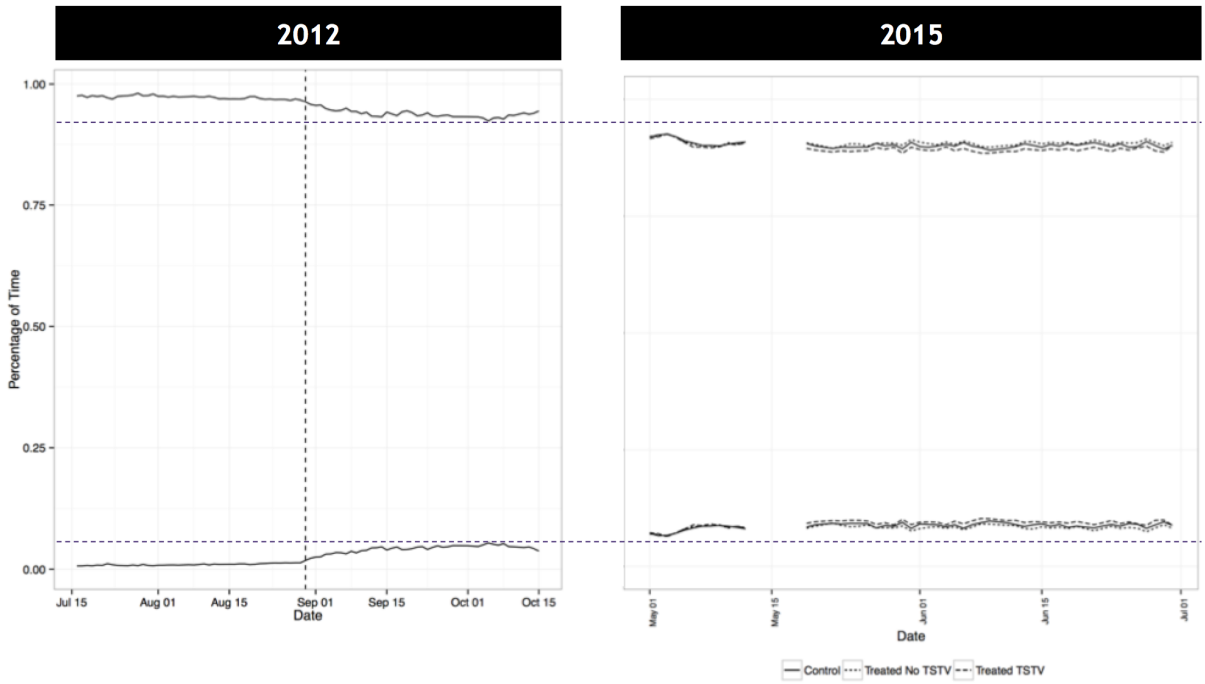

Figure 1: Percentage of time in Live and TSTV over time (observational dataset, 2012)

Figures 2 and 3 present the market share of each channel content category for live and time-shifted viewership for 2012 and 2015, respectively. The figures show that for both live and TSTV viewership and in both time periods considered, general purpose channels attract most of viewers attention, followed by entertainment, and children content. However, while the share of general content channels is similar in live and time-shift viewership, entertainment channels have

a noticeable larger viewership share in TSTV than in live. The opposite occurs for news content, which though capturing a significant portion of live audiences is not much consumed in time-shift (which can be explained by the fact that the value of the contents broadcast in these types of channels is usually highly time-dependent).

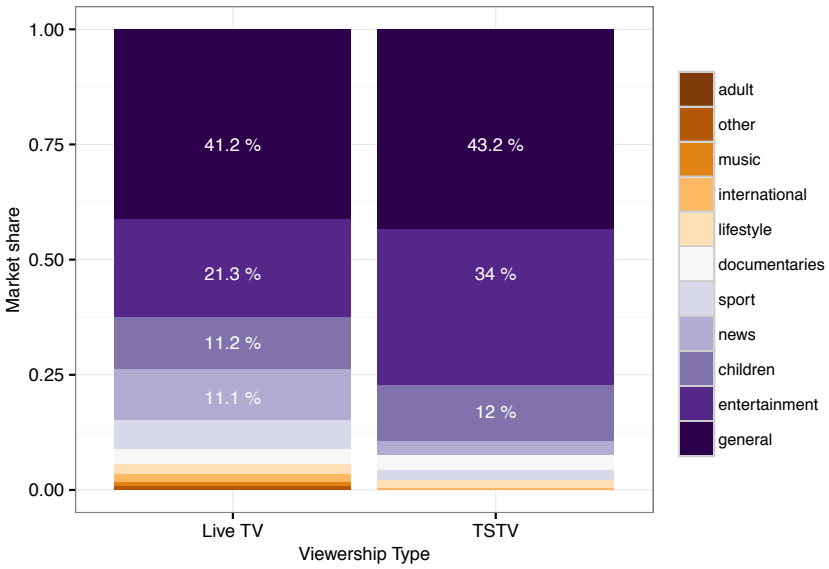

Figure 2: Market Share of Live TV and TSTV time by channel category in 2012

Figures 4 and 5 show the cumulative distribution of TSTV viewership by the number of days that passed from the moment the program originally aired to the moment it was watched. In figure 4, the earlier period of analysis (depicting behavior right after the introduction of the technology) corresponds to the line in the lighter shade, while the darker shaded line depicts behavior three years past. In both time periods, roughly 80% of TSTV viewership was for content that had aired in the previous couple of days. This shows a clear preference for recent content or an indication that search costs are higher for older content. This behavior seems to increase slightly from 2012 to 2015.When the cumulative distribution of TSTV viewership is broken down by the content type in figure 5, we can observe clear differences in how time sensitive certain types of content are. In particular, we find that for sports and news contents nearly 90% of view-

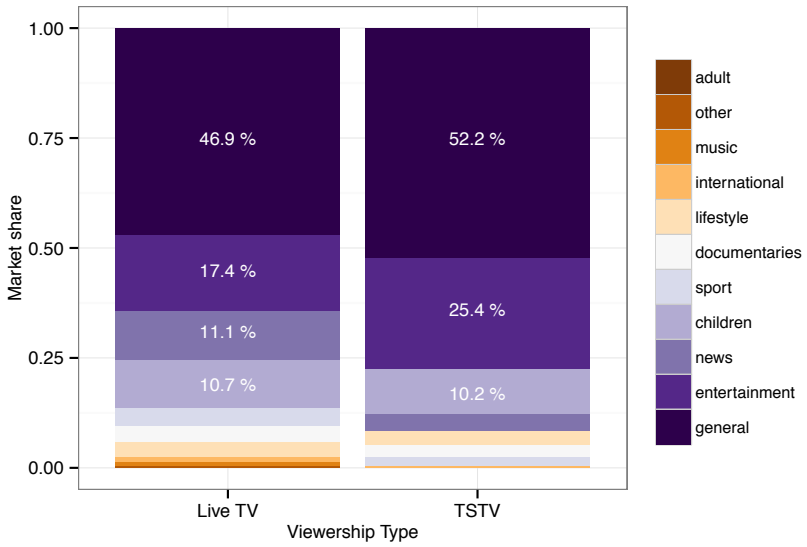

Figure 3: Market Share of Live TV and TSTV time by channel category in 2015

ership concerns programs that aired on the same day while, in the other extreme, less than 70% of TSTV viewership of entertainment content and documentaries concerns contents aired on the same day, while the remaining content categories fall in between this interval.

Figures 6 and 7 show the distribution of time spent watching live and TSTV broken down by the time of the day the content originally aired. Prime-time content (content aired between 8pm and midnight) is the most popular content in both time periods and viewership formats. In 2012, prime-time content attracted around 37% and 38% of live viewership and TSTV, respectively. If we take into account the fact that networks optimize their programming schedules so that the content of higher quality and that is the most relevant to viewers is aired throughout the period of the day during which people are the most likely to be at home and available to watch TV, it makes sense that the content televised during this time is also the content that is most consumed in time-shift viewership. Contents broadcast during the night and dawn periods attract the least amount of live or time-shifted attention.

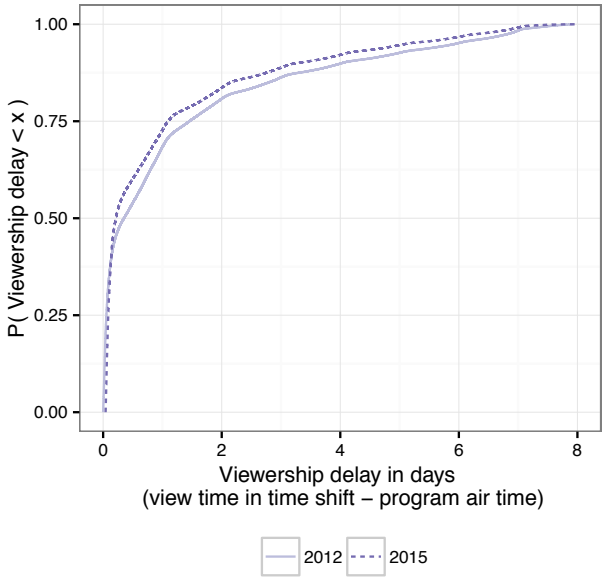

Figure 4: Cumulative distribution of TSTV viewership delay (delay: difference between TSTV event and original air time).

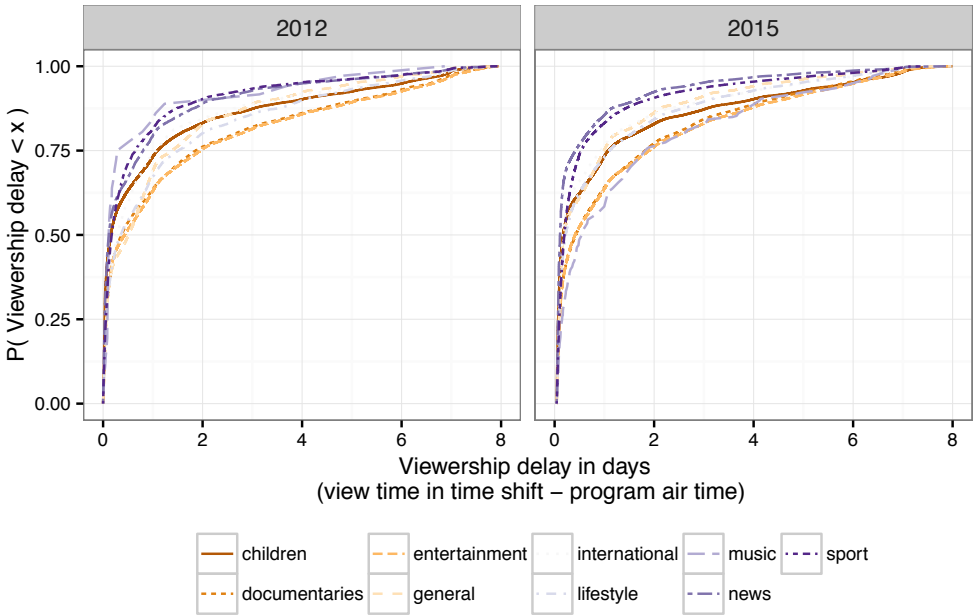

Figure 5: Cumulative distribution of TSTV viewership delay by channel type (delay: difference between TSTV event and original air time).

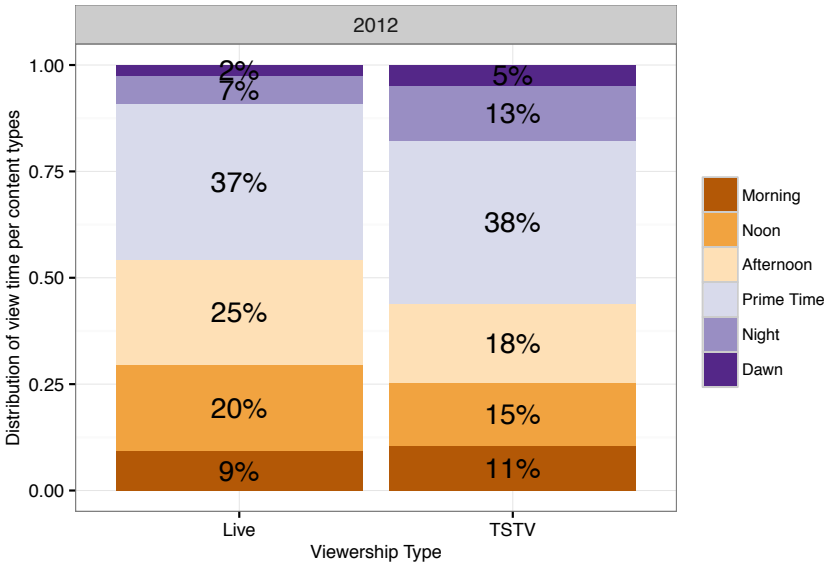

Figure 6: Distribution of View Time per Content (2012)

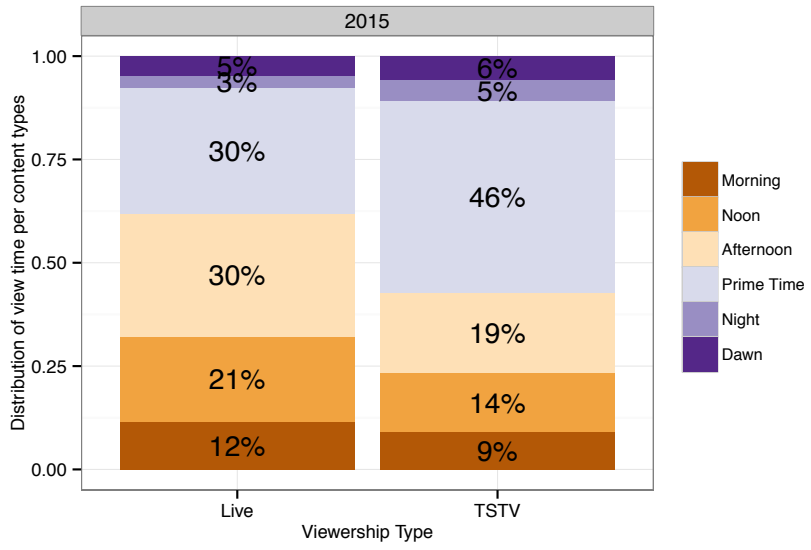

Figure 7: Distribution of View Time per Content (2015)

5.2 Advertisement Viewership

In this section we characterize both the advertisement investment by our industrial partner and how its advertisement was consumed by the household sample that was held out from any intervention during 2015's experimental period.

Regarding advertisement investment, during the period considered, a total of 708,354 seconds of this single advertiser's ads were broadcast across a number of channels. This investment was split among different types of channels - 38% in entertainment channels, 36% in general content channels (such as free to air channels), 12% in news channels, 10% in documentaries channels and the remaining in children channels. Overall, 44% of this advertisement time was broadcast during prime-time.

Figure 8 shows the breakdown of this advertiser's investment by channel type and by prime-time and non-prime time broadcast. The split between prime and non-prime time broadcast was similar across the different channel types. Figure 9 shows the advertisement time consumed by households from the control group of the experimental dataset. Strikingly, most of the consumed advertisement time came from general content channels which captured 64% of advertisement view time. The remaining time was distributed in the following way: news channels captured 16% of view time, entertainment channels 14%, documentary channels 5%, and children channels less than 1%. Over 40% of advertisement view time took place during prime-time for general content, news, and entertainment channels. The distribution of advertisement consumption by channel type and prime and non-prime time is presented in figure 9.

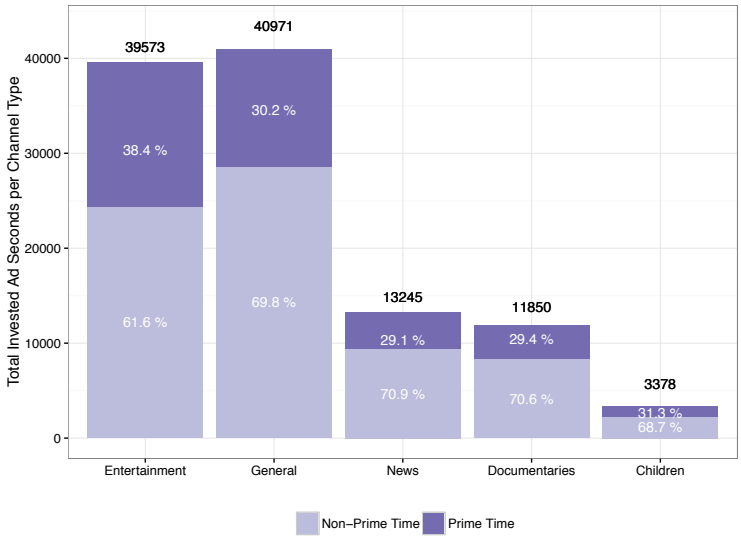

Figure 8: Advertisement Seconds Invested

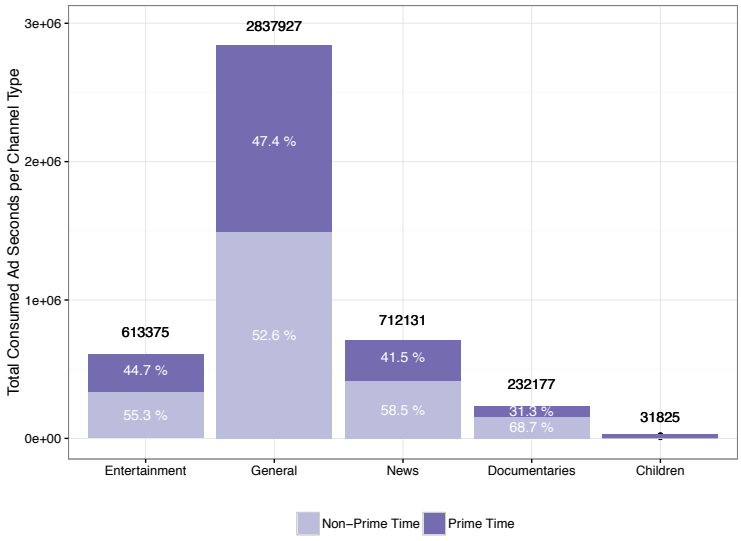

Figure 9: Advertisement Seconds Consumed (Experimental Control Group)

6 User Level Impacts of TSTV

6.1 Impact of TSTV on TV Viewership Time

Table 2, shows our difference-in-differences regression results for the analyses using the observational dataset (columns 1 to 3) and the experimental dataset (columns 4 to 6). Columns 1 and 4 pertain to total TV viewership, columns 2 and 5 to live viewership and columns 3 and 6 to TSTV. Our variable of interest is the interaction between *After* (corresponding to the period after the introduction of TSTV in the observational dataset and to the period during which the *movie bundle* was available in the experimental dataset) and *Treated* (corresponding to the *premium* consumers in the observational dataset and to the subset of subscribers who received the *movie bundle* with TSTV in the experimental data). In short, the results presented in the first three columns contrast the behavior of *premium* subscribers against that of *standard* subscribers after the initial introduction of TSTV in the summer of 2012, while the results presented in the last three columns compare the behavior of households who received the *movie bundle* with TSTV to that of households who received the same set of channels without TSTV in the summer of 2015, three years after the introduction of TSTV. Our variable of interest is the interaction *Treated TSTV * After*.

In column 1 of table 2, we can see that the initial introduction of TSTV caused an increase in total TV time of 0.182 hours per day ($p < 0.01$), corresponding to a 5.35% increase relative to the baseline daily total TV view time of 3.39 hours. Column 2, shows us that live TV time remained unchanged after the introduction of TSTV and column 3, shows that TSTV view time increased from 0 to 0.2 hours per day ($p < 0.01$) in the period after its introduction. Hence, the increase in total TV time was made up solely of the increase in time-shifted viewership. In column 3, we find that receiving the *movie bundle* with TSTV led to an increase in total daily TV time of 0.075 hours ($p < 0.01$) relatively to receiving the same set of channels without the time-shift feature. This corresponds to an increase of around 1.7% relative to the baseline of 4.4 hours of total daily TV time. Similarly

to what is observed in the analysis using observational data, columns 4 and 5 show that receiving the *movie bundle* with TSTV caused no significant changes to live viewership time while significantly increasing TSTV time by 0.053 hours per day ($p < 0.01$), a 15.5% increase relative to the baseline of 0.346 hours. The results of both analyses are highly consistent - both show that giving the TSTV feature to users leads them to spend more time watching TV and that live viewership does not decrease as a result of being able to consume contents in time-shift.

Table 2: View Time as a function of TSTV

	Dependent variable:					
	Total TV	Live TV	Time-Shift TV	Total TV	Live TV	Time-Shift TV
	survey-weighted	*survey-weighted*	*survey-weighted*	*experiment*	*experiment*	*experiment*
	(1)	(2)	(3)	(4)	(5)	(6)
Treated TSTV	0.104***	−0.073***	0.063***	0.028	0.015	0.008
	(0.024)	(0.024)	(0.001)	(0.038)	(0.036)	(0.008)
After	−0.068**	−0.068**	−0.000***	−0.452***	−0.487***	0.035***
	(0.028)	(0.028)	(0.000)	(0.022)	(0.020)	(0.004)
Treated TSTV * After	0.182***	−0.034	0.200***	0.075***	0.014	0.053***
	(0.033)	(0.033)	(0.002)	(0.021)	(0.020)	(0.005)
Constant	3.386***	3.386***	0.000***	4.980***	4.591***	0.338***
	(0.020)	(0.020)	(0.000)	(0.030)	(0.028)	(0.006)
week	Yes	Yes	Yes	Yes	Yes	Yes
Observations	652,799	652,799	652,799	182,507	182,507	182,507
Akaike Inf. Crit.	4,170,118.000	4,128,326.000	1,077,946.000			
R^2				0.005	0.006	0.003

Note: *p<0.1; **p<0.05; ***p<0.01

Cluster robust standard errors in parentheses. Observations clustered at the household level.

6.2 Impact of TSTV on TV Content Viewed

Figures 10 and 11 provide a graphical summary of how the changes in total TV time were distributed by the different channel categories (tables 14 and 15 in appendix F present the regression results corresponding to each figure, respectively). Each point in the figures depicts the magnitude of each coefficient accompanied by the corresponding 95 percent confidence interval. The analyses on which the figures are based rely on our experimental dataset. The considered

channels categories are presented in alphabetical order in both figures and are the following: adult (all adult channels are premium content that must be separately purchased), children (includes both premium and non-premium channels), documentaries, entertainment (movies and series channels), generalist (including the free to air channels), international, lifestyle (including travel, cooking, and fashion channels), the *movie bundle* channels (abbreviated as *M. Bundle*), music, news, sport (includes premium and non-premium channels), and other (includes a mix of channels that do not fit the considered categories).

Figure 10 contrasts the behavior of households who received the *movie bundle* with and without TSTV to that of households who were held out from any intervention, and table 11 contrasts households who received the *movie bundle* with TSTV to those who received the same set of channels without this feature. Both analyses follow the empirical strategy described in 4.1 and inform us on the distinct effects of simply increasing the content available to users without providing them with added flexibility in the way that content can be consumed and the effect of simultaneously increasing both content and flexibility.

On the lefthand side plot of figure 10, we can observe that households who received the *movie bundle* without the time-shift feature viewed this set of channels, on average, for 0.208 hours per day (12.48 minutes) ($p < 0.01$). About 40% ($p < 0.01$) of this increase was deducted from the time spent on other entertainment channels, about 31% ($p < 0.01$) from general purpose channels, and the remaining from children, documentaries, lifestyle, and to lesser extent from news, and sport channels. Therefore, gaining access to new entertainment content did attract viewers attention but the time that was dedicated to the *movie bundle* channels resulted from a direct substitution from time spent on the set of channels that was previously available, resulting on a zero net impact on total TV time as identified in section 6. On the righthand side plot, we find that households that received the *movie bundle* with TSTV viewed this set of channels for an average of 0.322 hours per day (19 minutes) ($p < 0.01$). Similarly to the behavior observed among households who received the *movie bundle* without TSTV, those receiving

TSTV also decreased the time spent on the set of previously available channels: entertainment channels lost 0.124 hours (7.4 minutes) ($p < 0.01$), corresponding to 39% of the increase in the *movie bundle channels*; and general content channels lost 0.064 hours (3.8 minutes) ($p < 0.01$), corresponding to about 20% of the increase in view time of the *movie bundle*. Documentaries, lifestyle, news, and sport channels also present statistically significant reductions in their view time though of smaller magnitudes. In total, the reductions in view time of the set of previously available channels by households who received TSTV was smaller than the increase in view time of the *movie bundle*, resulting in a significant net increase of total TV time identified in section 6.

Figure 11 contrasts only the behavior of households who received the *movie bundle* with and without TSTV. We find that the only significant differences between the two groups are the larger increase in view time of the *movie bundle* channels by those treated with TSTV (about 7 minutes) and also the larger reduction in time spent watching the set of previously available entertainment channels by the same household group (about 2.3 minutes). Both differences are statistically significant at the one percent level and indicate that TSTV induces an even larger substitution away from the previously available entertainment channels than that caused by simply receiving new content without TSTV.

6.3 Impact of TSTV on TV Viewership Hours

Figure 12 , presents a graphical summary of our analyses of the impact of receiving the movie bundle with and without TSTV on the time of TV viewership (broken down hourly). The left-hand side figure compares the behavior of households who received the set of premium channels without TSTV to that of households in the control condition and the right-hand side figure compares the behavior of households who received the movie bundle with TSTV to households in the control condition. For each hour of the day, the figures depict the differences in total TV view time between the two groups (measured in seconds) and the corresponding 95 percent confidence interval. We find that receiving the movie

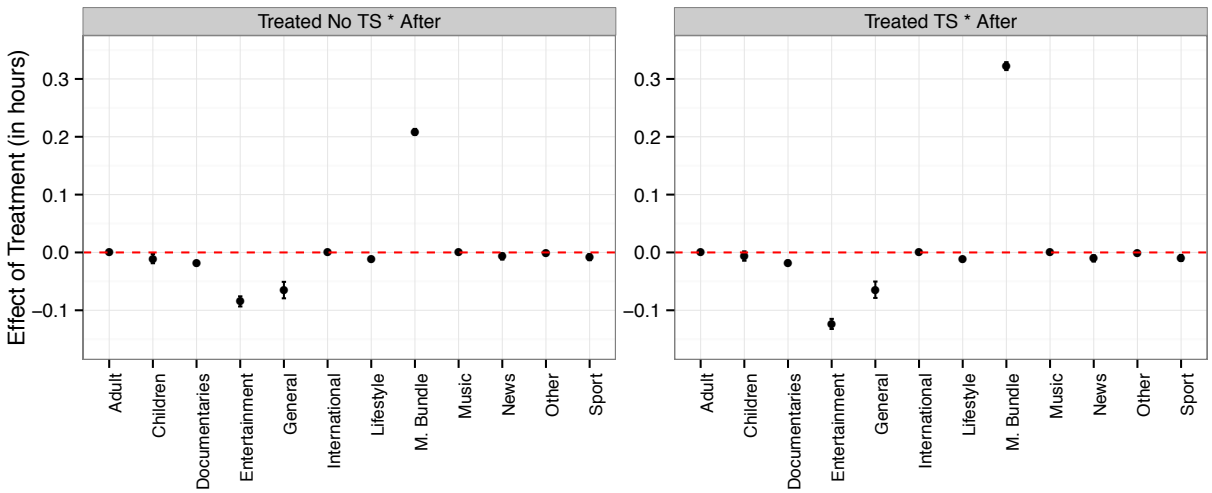

Figure 10: Impact of TSTV on TV Viewership by Channel Type (All Groups)

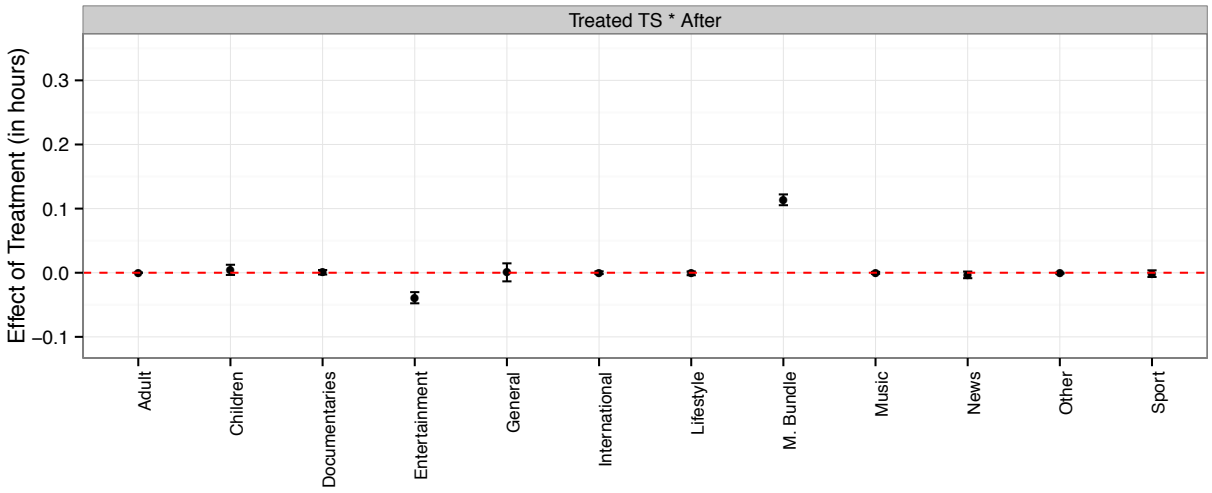

Figure 11: Impact of TSTV on TV Viewership by Channel Type (Treated Groups)

bundle without TSTV caused no significant changes in the TV viewership schedule of households all confidence intervals intersect the horizontal line set at 0. The increase in view time of those treated with time-shift was distributed during the afternoon period (between 2 and 5 pm) and also in the evening (between 8 and 9 pm) but overall we observe no major shifts in TV viewing schedules.

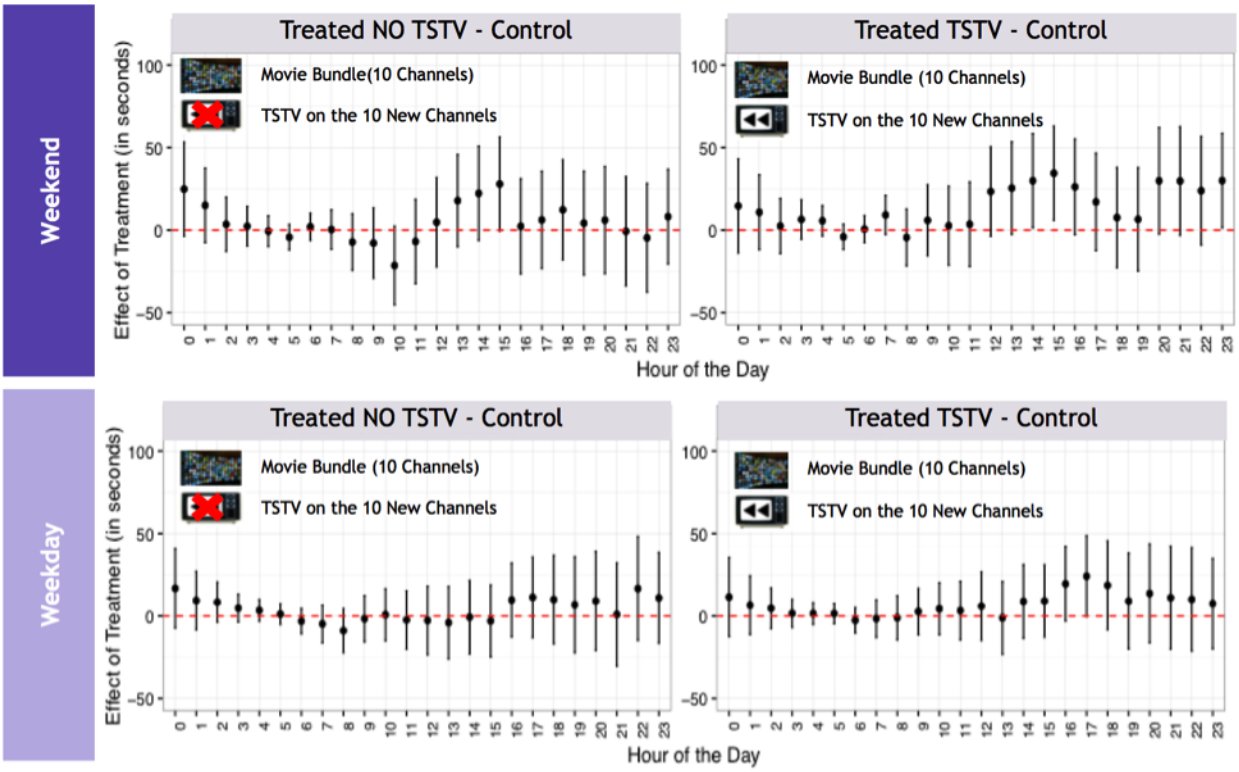

Figure 12: Changes in Viewership per Hour of the Day

6.4 Demographics of Live and Time-Shift Audiences

In this section we use the data from the randomized experiment to compare the demographic profiles of households who do not use TSTV to households that do so, and also compare the profiles of who heavy TSTV users to that of regular users. For this purpose, we consider all households from the randomized experiment during the period before the start of the experiment/ The set of household traits considered includes their weekly TV time (total TV time and by channel type), their Internet activity, their monthly expenditure, the subscriber's

age, whether or not the subscriber opted for an electronic receipt, the subscriber's service tenure, and finally whether or not the household is located in one of the main cities in the country considered.

Figure 13 compares the profile of households who did not watch TSTV (N = 14K) to that of households who did so (N = 20K). We can observe that the profiles of the two types of households differ in a number of features - households who did not use TSTV during this period watched on average less TV, less content from general purpose an entertainment channels but more from news and sports channels than households who watched contents in TSTV in this period. Households who did not use the time-shift feature also had lower Internet traffic (both downloads and uploads), a higher subscriber age, and a longer tenure with the TV service.

Figure 14 compares the profile of households whose TV time was made up of 50% (N = 500) or more of TSTV time to households who watched TSTV but in a smaller proportion to their total TV time. We can see that heavy TSTV users watch, on average, less TV than regular users, while also watching more general content channels and entertainment and less news and sports than regular users. Download traffic of heavy TSTV users is significantly lower than that of regular users, and their monthly expenditure with the company also seems to be slightly lower. Finally, the proportion of households located in the main cities is higher among heavy TSTV users than regular users.

These findings may have important implications for advertising. Advertisements are placed in specific channels and in between specific contents at a particular time of the day or week in order to better target the desired demographic group. However, if the format of consumption - live and time-shifted - attracts viewers with different characteristics, advertisement content and placement should take this differences into account. Additionally, the platform provider stands in a privileged position to characterize the two populations and possibly have an arbitrage opportunity to negotiate advertisement slots associated to content, consumption format, and viewers demographics.

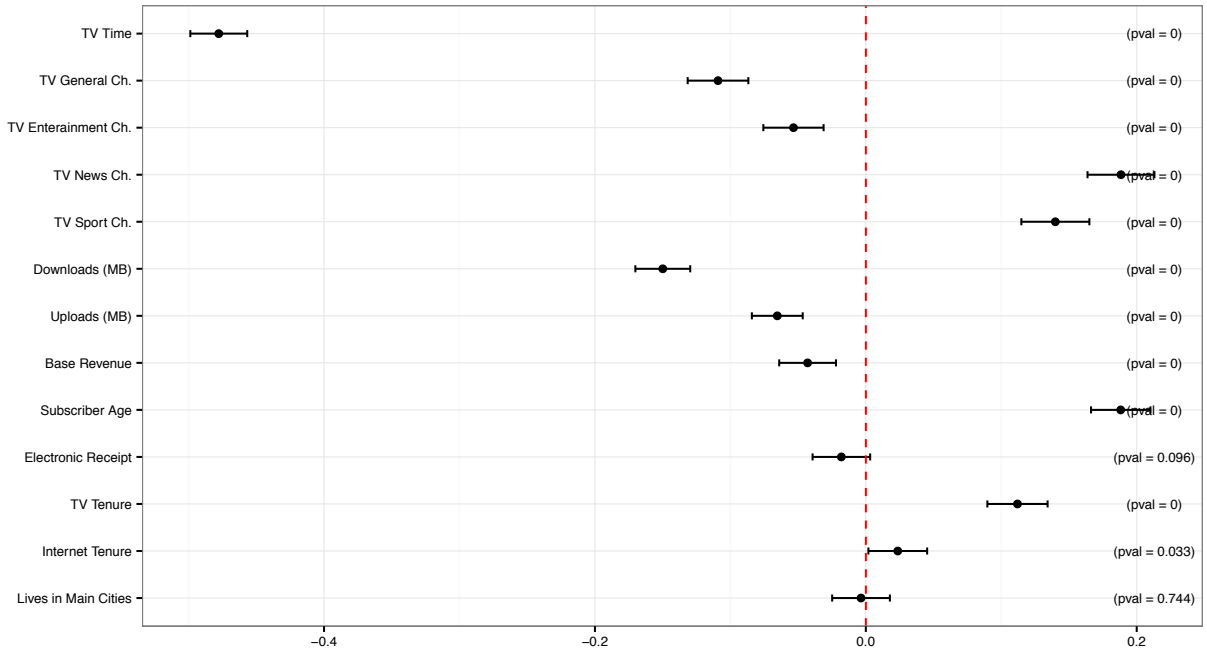

Figure 13: Households who did not use TSTV in the two weeks prior to the experiment are different from those who did so

6.5 Impact of TSTV on Advertisement Viewership

In this section we causally identify the impact of receiving the *movie bundle* with and without TSTV on advertisement view time following the same empirical strategy as presented in Section 4.1. Table 3 shows the impact of receiving the *movie bundle* with and without TSTV on viewership of advertisements sponsored by our industrial partner. Column 1 compares both treated conditions (with and without TSTV) to the control group while column 2 contrasts users who received TSTV to those who did not. The constant from the model presented in column 1, tell us that households who were held out from any intervention watched an average of 708 seconds per month of advertisements by this company. Households that received the *movie bundle* without time-shift watched less 29 seconds/month ($p < 0.01$) than the control group, corresponding to 4.1% reduction in advertisement view time. Those receiving the *movie bundle* presented a larger reduction - a 42 seconds ($p < 0.01$) decline in monthly advertisement view time relative

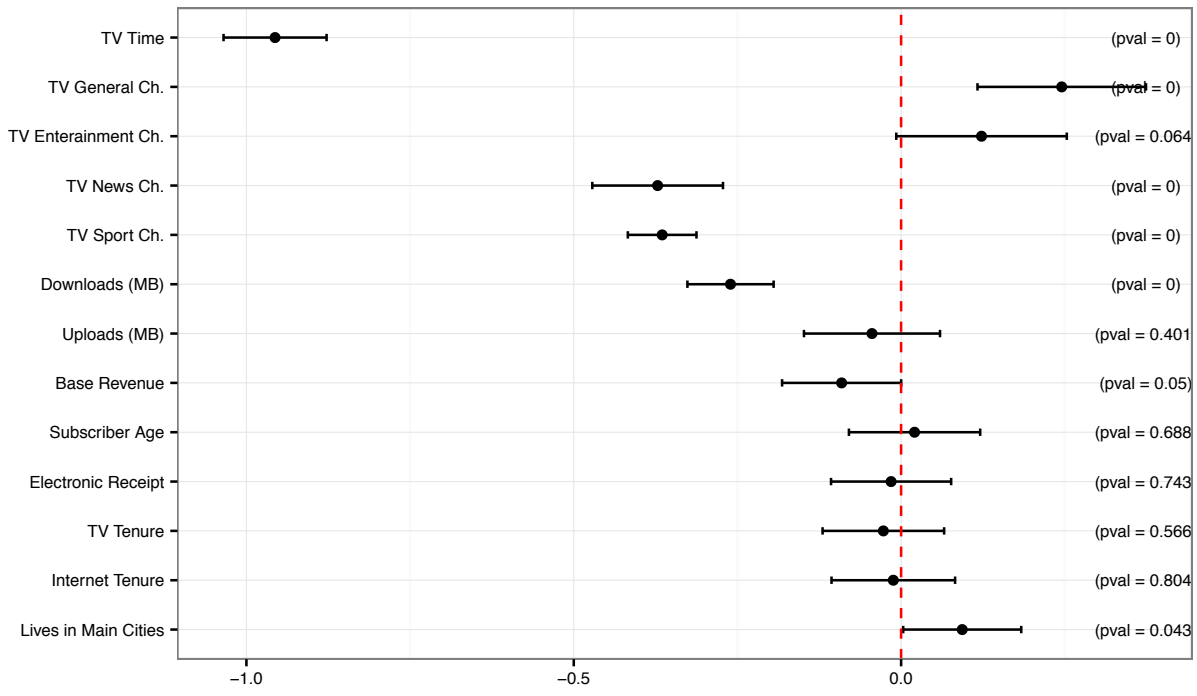

Figure 14: Households who use TSTV the most are different from the other households who also use TSTV

to the control condition (a 5.9% reduction relative to the baseline). Column 2 shows us that the 13 second difference between those treated with and without time-shift is statistically significant at the one percent level, isolating the impact of the time-shift feature on advertisement view time.

Table 3: Impact of TSTV on Ad View time

	Dependent variable:	
	Commercial View Seconds/Month	
	(1)	(2)
Treated No TSTV	-28.791***	
	(6.955)	
Treated TSTV	-42.249***	-13.458**
	(6.858)	(6.763)
Constant	708.354***	679.563***
	(4.983)	(4.852)
Observations	35,007	23,414
R^2	0.001	0.0002
Note:		*p<0.1; **p<0.05; ***p<0.01
	Cluster robust standard errors in parentheses	

Taking into account that not all advertisement slots are priced alike, and that premium time slots are usually significantly more valuable, we repeat our analysis breaking down total advertisement view time in prime-time and non-prime-time viewership. These results are presented in table 4 where columns 1 and 2 compare the results of the *TSTV* and *No TSTV* conditions to the *Control* condition and columns 3 and 4 compare those who received the *movie bundle* with TSTV to those who received it without this feature. The constants in the first and second columns of table 4 tells us that of the total of 708 seconds of advertisement viewed per month, 44% (313.143 seconds) was viewed during the four hour prime-time slot (8 pm to midnight) and 56% (395.211 seconds) was viewed during the

remaining hours of the day. Relative to those in the control condition, prime-time viewership of those receiving the *movie bundle* without TSTV declined by 9.1 seconds/month ($p < 0.01$) while that of households receiving this set of channels with TSTV declined by 13.7 seconds/month (column 1). Both these effects are significant at the one percent level and correspond to a 2.9% and a 4.4% reduction relative to the baseline level. The 4.6 seconds/month difference between the two treated groups is not statistically significant. The decline in advertisement viewership during non-prime time suffered a greater decline as a result of treatment - those who did not receive TSTV present 19.6 seconds/month ($p < 0.01$) decline (approximately a 4.9% drop relative to the control group), and those who received TSTV present a reduction 28.5 seconds/month ($p < 0.01$) (approximately, a 7.2% drop relative to the control condition).The 8.9 seconds difference between the two groups is statistically significant ($p < 0.05$). In sum, although prime-time advertisement viewership was negatively affected by the introduction of the *movie bundle* with and without TSTV, the negative effect was stronger during non-prime time hours.

Next, we explore the heterogeneity of treatment across the different channel categories in which the advertiser invests in. Table 5 contrasts users who received the *movie bundle* with and without time-shift to the those who were held out from intervention and table 6 compares the activity of the two treated groups. For both treated groups, the largest reduction in advertisement view time came from entertainment channels - 13.7 and 20.6 seconds/month ($p < 0.01$) for the *No TSTV* and the *TSTV* conditions, respectively. These values correspond to a 15.3% reduction by the *No TSTV* condition and a 23% reduction by the *TSTV* condition relative to the control group. This result is in line with our findings on the impact of treatment on the content consumed presented in sections 6.2, which indicated that entertainment channels were the channel category that suffered the largest negative impact as a result of the introduction of the *movie bundle*. Those who did not receive TSTV also decrease their advertisement viewership on general content channels by 11.8 seconds/month ($p < 0.1$) (a 2% decline

Table 4: Impact of TSTV on Ad View time by Prime-Time

	Dependent variable:			
	Ad View Prime Time	Ad View Non-Prime Time	Ad View Prime Time	Ad View Non-Prime Time
	(1)	(2)	(3)	(4)
Treated No TSTV	-9.133***	-19.659***		
	(3.208)	(4.631)		
Treated TSTV	-13.727***	-28.522***	-4.594	-8.864**
	(3.170)	(4.564)	(3.136)	(4.490)
Constant	313.143***	395.211***	304.011***	375.552***
	(2.292)	(3.325)	(2.244)	(3.223)
Observations	35,007	35,007	23,414	23,414
R^2	0.001	0.001	0.0001	0.0002

Note: *p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in parentheses

relative to the baseline). Those who received TSTV also significantly decreased advertisement view time on generalist channels - a 17.6 seconds/month (3.2%) ($p < 0.01$) decline. Table 6 shows that the behavior of the two groups did not present any statistically significant differences other than the 96.8seconds difference in advertisement view time in entertainment channels ($p < 0.01$). Again, these results are in line with those presented in section 6.2 of this paper regarding the changes in total TV viewership in the different channel categories. This analysis suggests that other than the simple substitution of viewership in the set of previously available channels for the *movie bundle* channels, treatment did not induce additional advertisement avoidance from users - the reductions in advertisement viewership in the different channel categories are generally proportional to the decline in total viewership in each category.

Finally, we use the full dataset of ad-viewership instances, which covers all viewership events of advertisement sponsored by our industry partner that took

Table 5: Impact of TSTV on Ad View time by Channel Type

	Dependent variable:				
	Children	Documentaries	Entertainment	General	News
	(1)	(2)	(3)	(4)	(5)
Treated No TS	-0.285	-1.582*	-13.789***	-11.775**	-1.361
	(0.216)	(0.838)	(2.071)	(5.413)	(2.026)
Treated TS	-0.236	-1.547*	-20.603***	-17.648***	-2.215
	(0.216)	(0.856)	(2.007)	(5.351)	(1.997)
Constant	6.841***	26.605***	89.348***	502.855***	82.705***
	(0.156)	(0.614)	(1.597)	(3.838)	(1.433)
Observations	35,007	35,007	35,007	35,007	35,007
R^2	0.0001	0.0001	0.003	0.0003	0.00004

Note: *p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in parentheses

Table 6: Impact of TSTV on Ad View time by Channel Type

	Dependent variable:				
	Children	Documentaries	Entertainment	General	News
	(1)	(2)	(3)	(4)	(5)
Treated TSTV	0.048	0.035	-6.814***	-5.873	-0.854
	(0.212)	(0.824)	(1.794)	(5.337)	(1.996)
Constant	6.557***	25.023***	75.559***	491.081***	81.344***
	(0.150)	(0.569)	(1.319)	(3.818)	(1.431)
Observations	23,414	23,414	23,414	23,414	23,414
R^2	0.00000	0.00000	0.001	0.0001	0.00001

Note: *p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in parentheses

place between June 1st and June 30th 2015, by households in the 2015s experimental sample (N=917,054), to study the impact of treatment on ad exiting probability. For each observation, we create an indicator variable that takes the value of 1 if the viewer did not watch the commercial until its end. We then run our usual difference-in-differences specification on this data. Table 4, shows the results of our estimation of the impact of treatment on the probability of the viewer exiting an ad he/she started watching. The baseline probability of exiting an ad is 7.2 percent. We find that receiving the movie bundle without TSTV raises the probability of ad exit by 0.3% ($p < 0.01$) while receiving this set of channels with TSTV raises the probability of ad exiting by 0.2% ($p < 0.05$). There is no statistically significant difference between the two results suggesting that the time-shift feature does not induce additional ad avoidant behavior.

Table 7: Impact of TSTV on Ad Exit Probability

	Dependent variable:	
	Ad Exit	
	(1)	(2)
Treated No TSTV	0.003***	
	(0.001)	
Treated TSTV	0.002**	-0.001
	(0.001)	(0.001)
Constant	0.072***	0.077***
	(0.002)	(0.002)
day	Yes	Yes
Observations	917,054	602,633
R^2	0.001	0.001

Note: *p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in parentheses
Errors clustered at the household level

7 Program Level Impacts of TSTV

In this section we investigate the impact of TSTV on TV viewership concentration following the methodology described in section 4.2 of this paper.

Tables 8 and 9 present the results obtained using the observational and experimental datsets, respectively. As expected, in both cases, viewership time decreases with rank, shown by the coefficient of *log(rank)* on the first row of each table. Our variable of interest is depicted in the third row of each table and pertains to the interaction between $log(rank)$ and $totalTV$ in table 8 and the interaction between $log(rank)$ and $TreatedTS$ in table 9.

The coefficient of $log(rank) * TotalTV$ in table 8 captures how the shape parameter of the Pareto distribution of total viewership time differs from that of the live viewership time. The sign of this term is negative and the coefficient is statistically significant ($p < 0.01$). This provides evidence that the concentration of total viewership is higher than that of live viewership as total viewership decays faster with rank than live viewership.

The coefficient of $log(rank) * TreatedTSTV$ in table 9 captures how the shape parameter of the Pareto distribution of total viewership time differs between households who received the *movie bundle* with TSTV and households who received this set of channels without TSTV. The coefficient for this interaction term is also negative and statistically significant ($p < 0.01$) providing evidence that the viewership concentration of households treated with TSTV is higher than that of households who did not receive TSTV, as viewership decays faster with rank for those who received the time-shift feature.

Both analysis provide consistent evidence that TSTV leads to increased concentration of TV viewership - popular content becomes even more popular with time-shift. This result is consistent with the super start effect identified in previous literature (Rosen, 1981; Elberse and Oberholzer-Gee, 2006).

Table 8: View time as a function of rank - Observational Data

	Dependent variable:
	log(viewership hours)
log(rank)	-2.474^{***}
	(0.004)
Total TV	0.401^{***}
	(0.059)
log(rank) * Total TV	-0.030^{***}
	(0.005)
Constant	35.070^{***}
	(0.042)
Observations	414,472
R^2	0.684
Note:	*p<0.1; **p<0.05; ***p<0.01
	Cluster robust standard errors in parentheses

Table 9: View time as a function of rank - Experiment Data

	Dependent variable:
	log(viewership hours)
log(rank)	-0.960^{***}
	(0.035)
Treated TSTV	0.860^{***}
	(0.130)
log(rank) * Treated TSTV	-0.084^{***}
	(0.022)
Constant	9.889^{***}
	(0.216)
Observations	2,178
R^2	0.788
Note:	*p<0.1; **p<0.05; ***p<0.01
	Cluster robust standard errors in parentheses

8 Discussion and Conclusion

In this work we collaborate with a large telecommunications provider to study the impact of time-shift television (TSTV) on household's television consumption behavior. We use data from both a natural experiment setting resulting from the initial introduction of the TSTV by our industrial partner to a select household sample, and data from a randomized experiment conducted by this provider three years after the introduction of this technology. Our analysis focuses on the impact of TSTV on TV time, TV contents watched and viewership concentration, time of viewership, and advertisement viewership. We also compare the characteristics of households who consume contents in time-shift to those of households who consume the same contents live.

We start by characterizing the live and TSTV viewership patterns of our two household samples. We find that prime-time content is the most popular content in both live and time-shifted viewership and that most time-shifted viewership consists of programs aired in the two days prior to the time-shift event, suggesting that the relevancy of content to the user is highly time-dependent. Also, different types of content are significantly more likely to be consumed in time-shift - e.g.

entertainment - than others - e.g. news and sports.

Our results regarding the impact of TSTV on TV time and concentration are highly consistent between our analyses of the observational and experimental datasets - TSTV leads households to increase their total TV time without reducing live TV time and also to increase their viewership concentration, that is, popular content becomes even more popular with time-shift. Thus, even though TSTV may allow some consumers to find niche content, on average, its dominant effect seems to be that of allowing consumers to watch even more popular programs. Using the experimental dataset we find that TSTV also impacts the time of the day viewers watch television, leading to a significant increase in viewership during the afternoon hours. These results suggest that prior to TSTV, viewers were compromising on the contents they watched, likely due to scheduling issues - individuals with similar content preferences differed in their preferred time to watch such contents and consequently dispersed their program viewership without TSTV. TSTV brings value to the consumer by decoupling content from airtime. Moreover, our results show that households who consumed the same content live and in time-shift differ in a number of characteristics, indicating that shared content preferences are not necessary synonymous to a similar subscriber profile.

Combining our experimental dataset with data on viewership of all advertisements sponsored by our industrial partner that aired during the experiment's period, we analyze the impact of TSTV on viewership of TV commercials isolating the effect of added content of that of increased flexibility in the way that content can be consumed. This was done by offering a set of premium entertainment channels with and without TSTV to a sample of users and comparing their TV viewership patterns to that of a control group that was held out from any intervention. The bundle of channels offered contained no advertisement. We find that receiving channels without TSTV (an increase in the amount and variety of content available to users but without added flexibility) lead viewers to substitute viewership of previously available content for the new content and thus

to significantly decrease their advertisement consumption in a similar proportion. When in addition to content, households also received increased flexibility (TSTV), the substitution effect was significantly larger and accompanied by an also proportionally larger decrease in advertisement viewership. By leading users away from channels with advertisements and towards contents without associated commercials, TSTV negatively impacts advertisement viewership. Other than this substitution effect, we did not find evidence of any additional strategic behavior by users regarding ad-skipping.

One limitation of this work concerns our analysis of the impact of TSTV on advertisement viewership and arises from the fact that our advertisement datasets only includes ads by a single firm. As a result, the advertisements studied may not be representative of the overall advertisement distribution in terms of the channels or the schedule selected for these their transmission or the advertisement characteristics. Our advertisement dataset further limits the reach of our analysis by not including advertisement viewership events in time-shift which prevents us from investigating possible ad-skipping behavior during time-shifted viewership.

Our results have useful implications for advertising, which has traditionally been tied to predetermined content and time slots in order to better target the desired demographic groups. With TSTV, advertisement remains associated to the content but becomes displaced in time. Our work provides some evidence the demographic profiles of live and time-shift audiences of the same content differ in their demographic profiles. Platform providers stand in a privileged position to characterize the two populations and possibly have an arbitrage opportunity to negotiate advertisement slots associated to content, consumption format, and viewers demographics. In the not so distant future, TV advertisement will likely resemble that we now observe on popular video streaming websites such as YouTube, non skippable pre-rolls or mid-rolls that can (potentially) be adapted to each specific viewers.

References

Anderson, S. and Gans, J. (2008). TiVoed: the effects of ad-avoidance technologies on broadcaster behaviour. SSRN Scholarly Paper ID 1295046, Social Science Research Network, Rochester, NY.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, 91(434):444–455.

Bapna, R. and Umyarov, A. (2012). Do your online friends make you pay? a randomized field experiment in an online music social network. NBER.

Boatwright, P. and Nunes, J. C. (2001). Reducing assortment: An attribute-based approach. *Journal of marketing*, 65(3):50–63.

Bronnenberg, B. J., Dube, J.-P., and Mela, C. F. (2010). Do digital video recorders influence sales? *Journal of Marketing Research*, 47(6):998–1010.

Brynjolfsson, E., Hu, Y., and Simester, D. (2011a). Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Science*, 57(8):1373–1386.

Brynjolfsson, E., Hu, Y., and Smith, M. D. (2003). Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management Science*, 49(11):1580–1596.

Brynjolfsson, E., Hu, Y. J., and Simester, D. (2011b). Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Science*, 57(8):1373–1386.

Brynjolfsson, E., Hu, Y. J., and Smith, M. D. (2006). From niches to riches: Anatomy of the long tail. *Sloan Management Review*, 47(3):67–71.

Dixit and Stiglitz (1977). Monopolistic competition and optimum product diversity. *The American Economic Review*, Vol. 67(No. 3):297–308.

Downey, K. (2007). Study: First ad does get more eyeballs. *Media Life Magazine*, 21.

Elberse, A. and Oberholzer-Gee, F. (2006). Superstars and underdogs: An examination of the long tail phenomenon in video sales. *Harvard Business School Working Paper Series*.

Evans, D. S. (2008). The economics of the online advertising industry. *Review of network economics*, 7(3).

Evans, D. S. (2009). The online advertising industry: Economics, evolution, and privacy. *Journal of Economic Perspectives, Forthcoming*.

Ferguson, D. A. and Perse, E. M. (2000). The world wide web as a functional alternative to television. *Journal of Broadcasting & Electronic Media*, 44(2):155–174.

Frank, R. H. and Cook, P. J. (1995). The winner-take-all society: How more and more americans compete for fewer and bigger prizes, encouraging economic waste, income inequality, and an impoverished cultural life.

Godinho de Matos, M., Ferreira, P., and Smith, M. D. (2015). The effect of television content on pirates online activity: Evidence from a randomized experiment. *SI 2015 Economics of IT and Digitization Workshop, National Bureau of Economic Research*.

Grece, C., Lange, A., Schneeberger, A., and Valais, S. (2015). The development of the european market for on-demand audiovisual services. Technical report.

Guo, S. and Fraser, M. W. (2014). *Propensity score analysis: Statistical methods and applications*, volume 11. Sage Publications.

Hinz, O. and Eckert, J. (2010). The impact of search and recommendation systems on sales in electronic commerce. *Business & Information Systems Engineering*, 2(2):67–77.

Hollis, S. and Campbell, F. (1999). What is meant by intention to treat analysis? survey of published randomised controlled trials. *Bmj*, 319(7211):670–674.

Hotelling, H. (1929). Stability in competition. *The Economic Journal*, 39, No 53:41–57.

Johnson, G. A., Lewis, R. A., and Nubbemeyer, E. I. (2015). Ghost ads: A revolution in measuring ad effectiveness. *Available at SSRN*.

Kernan, W. N., Viscoli, C. M., Makuch, R. W., Brass, L. M., and Horwitz, R. I. (1999). Stratified randomization for clinical trials. *Journal of clinical epidemiology*, 52(1):19–26.

Krugman, D. M., Cameron, G. T., and White, C. M. (1995). Visual attention to programming and commercials: The use of in-home observations. *Journal of Advertising*, 24(1):1–12.

Kuksov, D. and Villas-Boas, J. M. (2010). When more alternatives lead to less choice. *Marketing Science*, 29(3):507–524.

Liaukonyte, J., Teixeira, T., and Wilbur, K. C. (2015). Television advertising and online shopping. *Marketing Science*, 34(3):311–330.

McCaffrey, D. F., Ridgeway, G., and Morral, A. R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological methods*, 9(4):403.

Pearson, S. and Barwise, P. (2007). Pvrs and advertising exposure: a video ethnographic study. *International Journal of Internet Marketing and Advertising*, 4(1):93–113.

Rosen, S. (1981). The economics of superstars. *The American economic review*, 71(5):845–858.

Rosenbaum, P. R. (1987). Model-based direct adjustment. *Journal of the American Statistical Association*, 82(398):387–394.

Rubin, A. M. (2002). *The uses-and-gratifications perspective of media effects.* Lawrence Erlbaum Associates Publishers.

Salop, S. C. (1979). Monopolistic competition with outside goods. *The Bell Journal of Economics*, 10(1):141.

Sawhney, M. S. and Eliashberg, J. (1996). A parsimonious model for forecasting gross box-office revenues of motion pictures. *Marketing Science*, 15(2):113–131.

Tse, A. C. B. and Lee, R. P. (2001). Zapping behavior during commercial breaks. *Journal of Advertising Research*, 41(3):25–29.

Van Meurs, L. (1998). Zapp! a study on switching behavior during commercial breaks. *Journal of Advertising Research*, 38(1):43–44.

Waterman, D., Sherman, R., and Ji, S. W. (2012). The economics of online television: Revenue models, aggregation, and 'TV everywhere'. SSRN Scholarly Paper ID 2032828, Social Science Research Network, Rochester, NY.

Wilbur, K. (2006a). How the digital video recorder changes traditional television advertising. SSRN Scholarly Paper ID 918088, Social Science Research Network, Rochester, NY.

Wilbur, K. (2006b). A two-sided, empirical model of television advertising and viewing markets. SSRN Scholarly Paper ID 885465, Social Science Research Network, Rochester, NY.

Wilbur, K. C. (2008). How the digital video recorder (DVR) changes traditional television advertising. *Journal of Advertising*, 37(1):143–149.

Zigmond, D., Interian, Y., Lanning, S., Hawkins, J., Mirisola, R., Rowe, S., and Volovich, Y. (2009). When viewers control the schedule: Measuring the impact of digital video recording on tv viewership. In *Key Issues Forums at ARF Audience Measurement Conference*.

A Additional Descriptive Statistics

The distribution of audiences attention at the different hours of the day is depicted in figure 15, which shows how likely users are to watch TV at any point in time during the day, conditional on the viewership format - live or time-shift. We can observe that people are the least likely to watch live television during night-time and dawn hours and become increasingly more likely to watch TV as the day evolves with prime-time (the time interval between 8 pm and midnight) being the period of the day during which people are the most likely to watch live television. Time-shift viewership follows a somewhat similar distribution with a few key differences - people are most likely to watch time-shifted television during the late night (between 10 pm and 1 or 2 am), likely due to the lack of quality programming during this time; and, unlike live TV viewership, time-shifted viewership drops at around 7 to 9 pm, the time when most news shows are transmitted on the free to air channels.

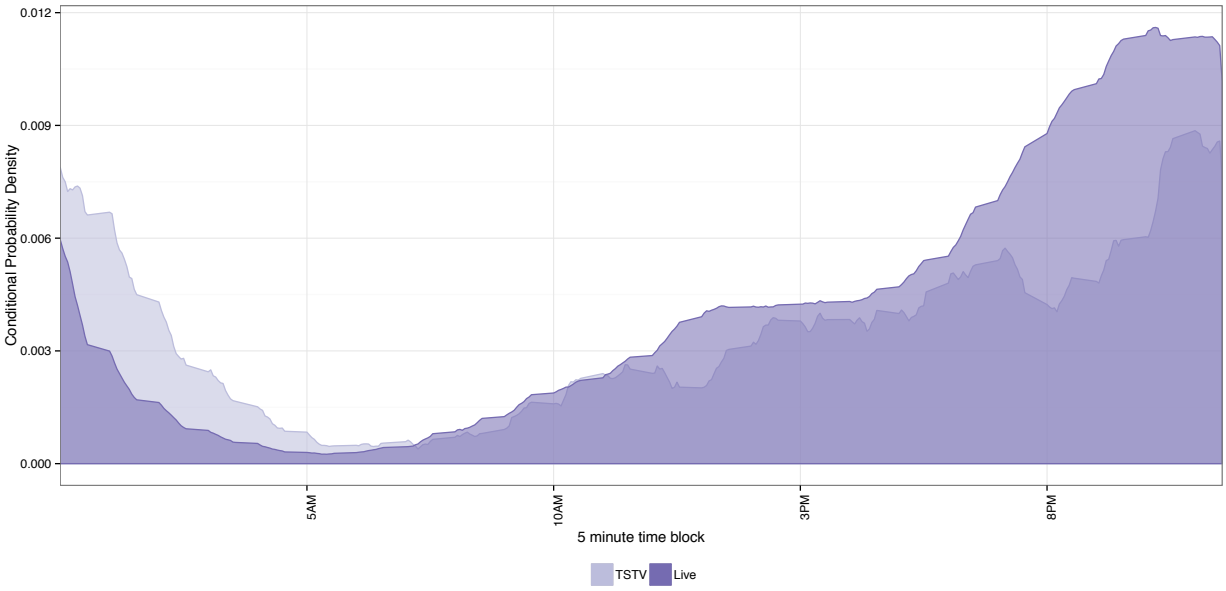

Figure 15: Audiences per Hour

B Propensity Score Weighting

We calculate propensity scores based on covariates that proxy consumers' interest on TV and on their availability to watch TV. We proxy their interest in TV with the number of days consumers watch TV per month and the number of programs they watch per month. We also proxy their availability to watch TV by the time they spend watching TV and by the hour of the day they turn on the TV. We use consumers' average and standard deviation of view time, average and standard deviation of watched programs, TV, Internet and voice service tenure, as well number of days, time they turn on the TV, and average change in total view time in the period preceding TSTV availability.

From the original 10,000 premium consumers that we selected at random for our study only 8,348 were active, i.e., had watched TV at least one day in the period before the introduction of TSTV. We aggregate the data at the week level and select at random a set of 50,000 standard consumers, who do not have TSTV, to add to our random sample of 10,000 premium consumers, who were all given TSTV. Table 10 shows the average of these covariates for standard and premium consumers before the introduction of TSTV. On average and before the introduction of TSTV, standard subscribers watched TV less frequently, and in particular watched fewer programs than premium consumers both during weekdays and during the weekend. They also tend to have a lower tenure on voice and Internet subscription and to turn on the TV slightly earlier than the latter.

Table 10 also shows weighting reduces significantly the bias in these covariates from comparing standard and premium consumers. After weighting, all these differences decrease and all estimated effect sizes decrease below the reference threshold of 0.2 (McCaffrey et al., 2004). This is accomplished by increasing the weights of standard subscribers that watch more days and more programs, that is, the standard subscribers that are more likely to like TV and digital content.

Figure 16 shows the relative influence of each of the variables for the weighting

Table 10: Information on pre-treatment variables before and after weighting. Chosen weights minimize the mean standardized effect size (ATT) in the pre-treatment variables. Standardized effect size of a variable is defined as the difference between the mean of the variable in premium and standard consumers divided by the standard deviation of the variable over the premium consumers.

	Premium	Premium (s.d.)	Standard	Standard (s.d.)	Std. Eff. Size	Weight. Standard	Weight. Standard (s.d.)	W. Std. Eff. Size
view_time_total_hr	3.452	2.705	2.959	2.441	0.182	3.330	2.534	0.045
sd_view_time_total_hr	2.750	1.227	2.692	1.372	0.048	2.703	1.227	0.038
n_days_weekend	8.174	3.496	7.009	3.609	0.333	7.993	3.538	0.052
n_days_weekday	21.810	8.814	18.980	9.216	0.321	21.260	9.013	0.063
n_programs_weekend	39.780	36.270	17.950	18.080	0.602	37.870	34.010	0.053
n_programs_weekday	32.070	29.020	15.120	14.450	0.584	30.670	27.730	0.048
sd_n_programs	31.830	20.120	15.270	10.220	0.823	29.930	18.270	0.094
mode_first_event_start_hour	14.270	5.441	13.810	5.418	0.086	14.120	5.434	0.028
d_view_time_total_hr	-0.013	0.148	-0.019	0.175	0.040	-0.015	0.155	0.016
tv_tenure	68.380	59.140	68.430	56.820	-0.001	70.840	60.080	-0.042
tv_tenure:<NA>	0	0	0.010	0.100	-0.110	0	0	0
net_tenure	41.790	36.360	34.470	31.650	0.201	42.100	35.970	-0.008
net_tenure:<NA>	0	0	0.010	0.100	-0.110	0	0	0
voice_tenure	29.390	18.590	27.680	17.240	0.092	29.780	18.290	-0.021
voice_tenure:<NA>	0	0	0.010	0.100	-0.110	0	0	0

task. The most important discriminant variables are the standard deviation of the number of programs watched, the standard deviation of the total view time, and tenure with the Internet service.

Figure 17 shows how the absolute standard differences for each of the variables. Weighting decreased all the biases, as expected.

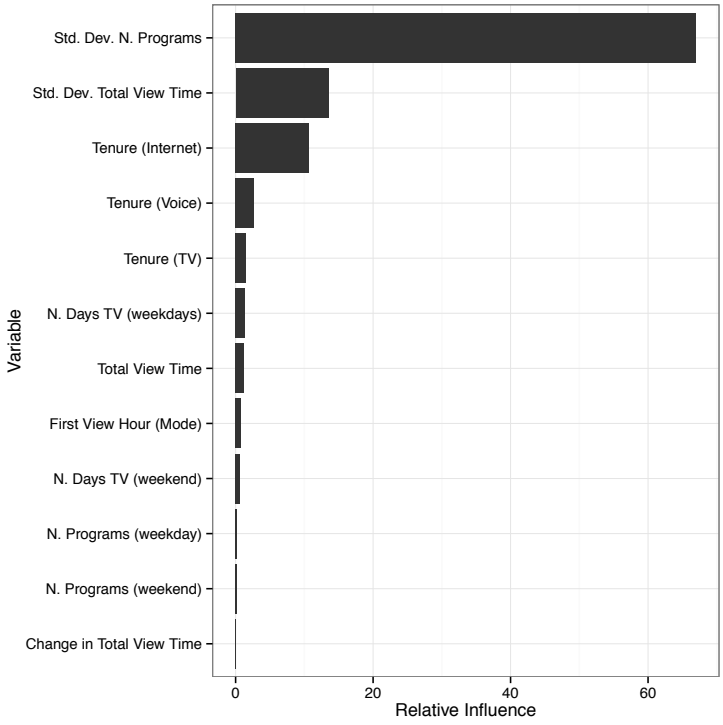

Figure 16

Table 11: View Time as a function of TSTV for unweighed and weighted sample.

	Dependent variable:					
	Total	Live	TS	Total	Live	TS
	OLS	*OLS*	*OLS*	*survey-weighted normal*	*survey-weighted normal*	*survey-weighted normal*
	(1)	(2)	(3)	(4)	(5)	(6)
TSTV	−0.035***	−0.035***	−0.000***	−0.068**	−0.068**	−0.000***
	(0.008)	(0.008)	(0.000)	(0.028)	(0.028)	(0.000)
Premium	0.455***	0.278***	0.063***	0.104***	−0.073***	0.063***
	(0.032)	(0.031)	(0.002)	(0.024)	(0.024)	(0.001)
TSTV*Premium	0.149***	−0.067***	0.200***	0.182***	−0.034	0.200***
	(0.021)	(0.020)	(0.004)	(0.033)	(0.033)	(0.002)
Constant	3.035***	3.035***	0.000***	3.386***	3.386***	0.000***
	(0.012)	(0.012)	(0.000)	(0.020)	(0.020)	(0.000)
Observations	652,799	652,799	652,799	652,799	652,799	652,799
R^2	0.005	0.001	0.176			
Akaike Inf. Crit.				4,170,118.000	4,128,326.000	1,077,946.000

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors clustered at the user level.

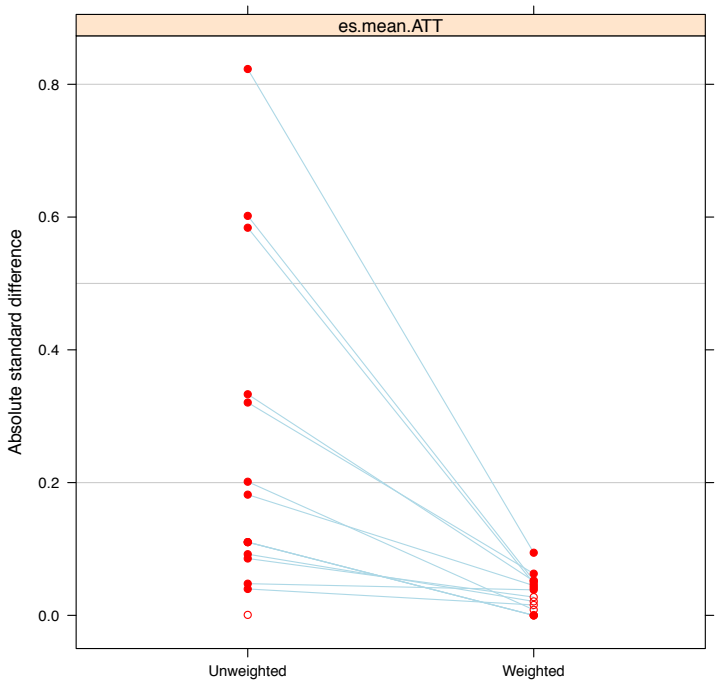
es.mean.ATT
Absolute standard difference
0.0
0.2
0.4
0.6
0.8
Unweighted
Weighted

Figure 17

C Experiment Data Details

We collected data between May 1^{st} and June 30^{th} 2015. This period covers the 12 days before the start of the experiment and the six weeks of the experiment from which we exclude the first 6 days (May 13^{th} to 18^{th}) during which the *movie bundle* activation took place.

Our initial sample of 40,566 households suffered some attrition mostly resulting from lack of recorded TV or Internet activity by these households during the period of analysis. In total, 5,459 were dropped from the sample, 30 percent of which due to not showing any TV or Internet activity during the entire period of analysis and 67 percent due to not being active in both the considered time periods - before and during the experiment. Lack of registered activity by subscribers may be due to them having legacy set-top-boxes which failed to accurately track their TV activity, due to churn, or due to actual lack of activity. The remaining households were removed from our final sample due to problems regarding household contact information and opt-outs from proactive marketing campaigns.

Our final sample has 35,107 households distributed across the three treatment conditions as depicted in figure 18. As shown in this figure, sample attrition affected all treatment conditions equally (13% to 14% of households in each group were lost), having no relationship to treatment assignment. Sample attrition, thus did not harm our ability to causally interpret the impact of treatment on household activity. However, we should note that the representativeness of our sample regarding the customer base of our industrial partner is not complete but should, nonetheless, represent a relevant majority group including active triple and quad-play subscribers with up-to-date equipment and who have not opted out from the provider's marketing campaigns.

Despite randomization, our ability to causally interpret the effects of treatment on the outcomes of interest can be hindered if the households in each treatment group differ on key characteristics liable to impact the effect of treatment

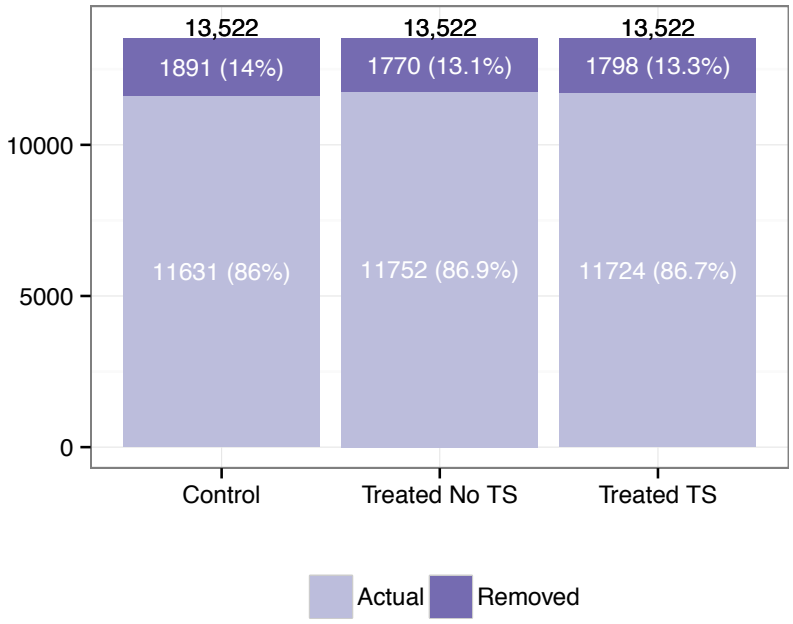

Figure 18: Randomized Experiment Sample Summary

(Kernan et al., 1999). We use the pre-experiment data (from May 1^{st} to the 12^{th}) to ensure that all experimental groups were equivalent on the relevant observable metrics prior to treatment. Table 12, presents summary statistics of the main variables of interest for the three experimental groups in the pre-treatment period. The table includes the p-value for the t-test of the differences in means between treatment conditions. Table 12 shows that, despite attrition, our final sample is balanced - the t-tests indicate that, at a 95% confidence level, there are no significant differences between the three treatment conditions. In other words, the random selection of households into treatment conditions did result in three equivalent groups.

From table 12 we also learn that the subscribers in our sample viewed an average of around 264 minutes of television per day during this period, 21 minutes of which were time-shifted viewership (about 8 percent). The average daily time spent viewing the *movie bundle* channels was under one minute indicating that most subscribers in our sample had not voluntarily subscribed to these chan-

nels prior to treatment. The average daily Internet download traffic level was approximately 1.25 Gb.

Table 12: Balance Table (before treatment)

	Treatment Group	Variable	Control Avg	Treatment Avg	T Test P Value
1	M. Bundle No TSTV	Total TV time (min)	263.743	264.227	0.832
2	M. Bundle No TSTV	Total CPack time (min)	0.603	0.593	0.770
3	M. Bundle No TSTV	TSTV TV time (min)	20.608	20.755	0.758
4	M. Bundle No TSTV	Downloads (Mb)	1,248.078	1,228.666	0.454
5	M. Bundle No TSTV	VoD expenditure (cents)	31.315	27.118	0.162
6	M. Bundle No TSTV	Prior Customer Retention	1.102	1.112	0.591
7	M. Bundle TSTV	Total TV time (min)	263.743	265.418	0.461
8	M. Bundle TSTV	Total CPack time (min)	0.603	0.578	0.408
9	M. Bundle TSTV	TSTV TV time (min)	20.608	21.124	0.280
10	M. Bundle TSTV	Downloads (Mb)	1,248.078	1,244.956	0.911
11	M. Bundle TSTV	VoD expenditure (cents)	31.315	37.404	0.082
12	M. Bundle TSTV	Prior Customer Retention	1.102	1.109	0.701

Figure 19 shows the average time spent watching the *movie bundle* channels in each day of our period of analysis by the users in each treatment condition. We can see that all groups had identical viewing times prior to treatment but, once treatment was implemented, a significant gap appeared between the vieweing times of the treated and control groups. Particularly, those who received the *movie bundle* with TS spent significantly more time viewing these channels than those who received the *movie bundle* without TS. This plot provides us with some visual evidence that the set of channels offered did succeed in capturing viewers attention and that the TS feature did attract users to consume more of the *movie bundle*.

Figure 20 shows a similar graph for the time spent viewing the *movie bundle* channels in time-shift. We can observe a similar pattern to that of figure 19 whereby the treatment and control groups have an identical activity level prior to treatment but a large gap grows between the group that received the *movie bundle* with TS and the two other groups once treatment is in place. This plot provides additional evidence that subscribers who received the *movie bundle* with the TS feature did use TS to consume the contents of this set of channels.

Figure 21 shows the percentage of subscribers in each group who viewed the

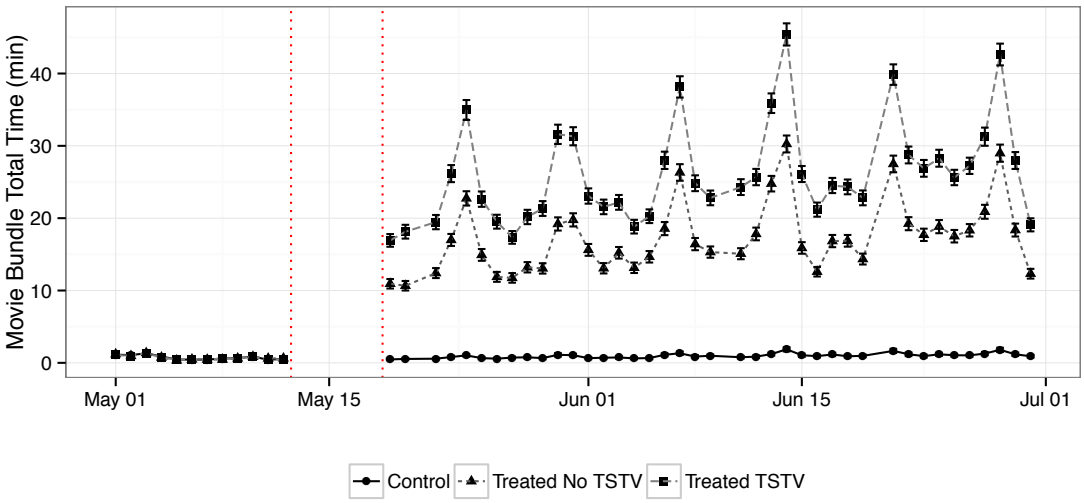

Figure 19: Average daily minutes viewed in Movie Bundle channels

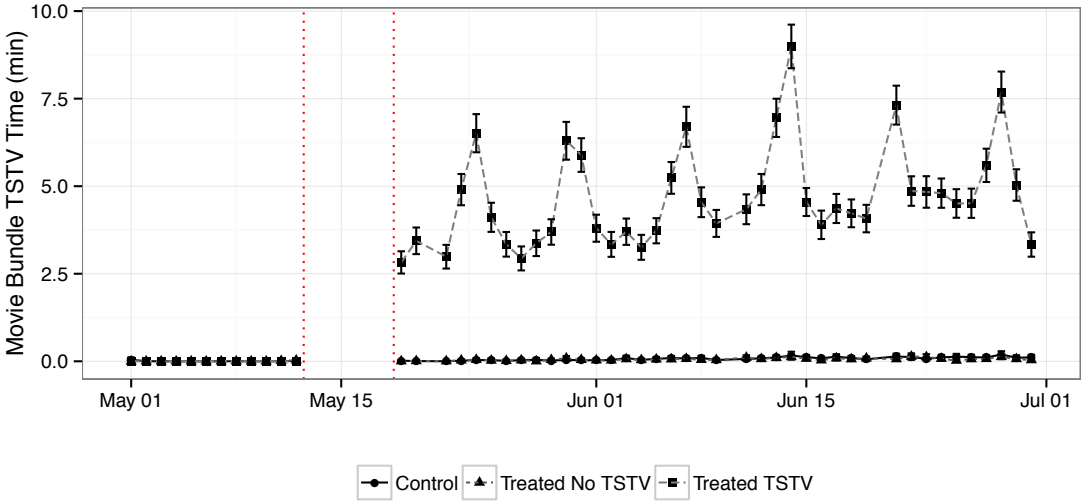

Figure 20: Average daily time-shift minutes viewed in Movie Bundle channels

movie bundle channels for 90 minutes at least once during the treatment period. Here we can observe the lack of perfect compliance with treatment - about 10 percent of users in the control group subscribed to and viewed these channels while a large fraction of those in the treatment conditions did not view these channels despite having received free access to them (50% and 40% in the no time-shift and time-shift conditions, respectively).

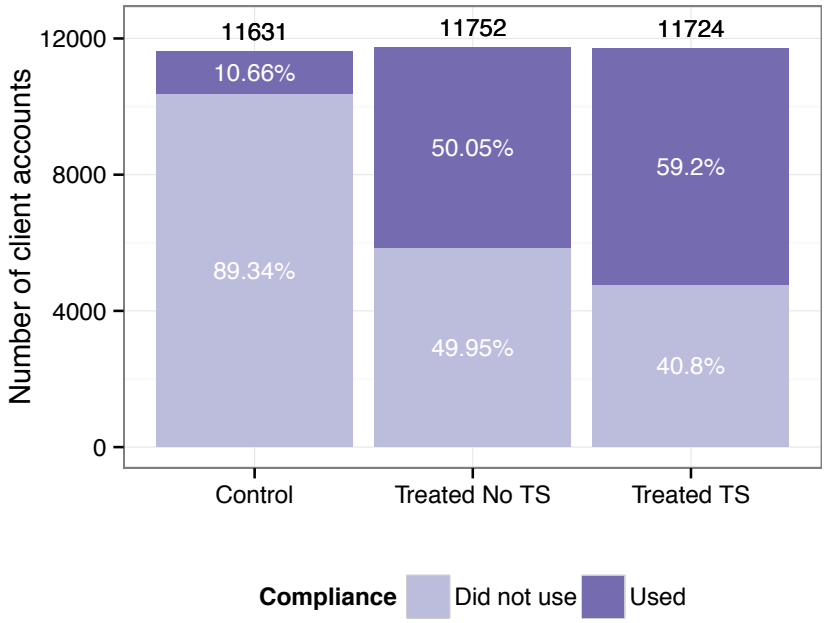

Figure 21: Compliance with Treatment per Treatment Group

D Additional User Level Impacts of TSTV

As hypothesize in section 2 we expect that the observed increase in TV view time was to some extent deducted from other media consumption activity. Given that, similarly to TV, the Internet is a platform commonly used for entertainment and the consumption of media, we use our experimental dataset to test whether Internet traffic was affected by the introduction of the *movie bundle*. Table 13 compares daily download traffic (in Mb) of those who received the *movie bundle* with TSTV against those who received this set of channels without TSTV. We find that those who received the time-shift feature significantly decreased their activity by 31.3 Mb ($p < 0.1$) (a 2.5 percent relative to the baseline of 1227 Mb per day).

Combining the results of the impact of treatment on total TV time and Internet use, we find evidence of substitution between TV and Internet as a result of receiving the *movie bundle* with time-shift - the significant increase in TV time among this group of users is associated to a significant decrease in their download traffic. According to our difference-in-differences estimates, and considering that 5 minutes of YouTube streaming at a 720p resolution (a common high-definition resolution) uses approximately 38 Mb of download traffic, our results can be tentatively translated into a substitution of slightly less than 5 minutes of Internet streaming for 5 minutes of TV as a result of receiving the *movie bundle* with time-shift, which nears a one-to-one substitution effect. Although we are unable to verify whether this was a direct substitution of online video for TV (users may have decreased download traffic by reducing the amount of time spent on social networks or email, for instance) these results suggest some extent of direct substitution.

Table 13: Impact on TSTV on Media Consumption - Difference-in-Differences ITT (OLS)

	Dependent variable:	
	Downloads (Mb)	
	(1)	(2)
Treated No TSTV	−19.411	
	(25.922)	
Treated TSTV	−3.121	16.290
	(27.948)	(26.884)
After	−20.695*	−30.729***
	(11.494)	(11.661)
Treated No TS * After	−10.034	
	(16.373)	
Treated TS * After	−41.326**	−31.292*
	(17.299)	(17.410)
Constant	1,248.078***	1,228.666***
	(19.109)	(17.515)
Observations	70,214	46,952
Adjusted R^2	0.0001	0.0002

Note: *p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in parentheses
Errors clustered at the household level

E Additional Program Level Impacts of TSTV

In this section we analyze the impact of TSTV on viewership concentration using our observational dataset and using the Theil Index and a measure of TV viewership concentration. On average, in 2012, there were 3372 different TV programs offered each day by our industrial partner to consumers with pay-TV service. We computed the share of each program to determine the Theil Index. The share of each program was computed as the total viewership of a program as a fraction of the total viewership for the duration of the program.Therefore, in this analysis each observation pertains to a TV program and contains that program's share of time both for total and live viewership. We bootstraped (1,000 re-samples with replacement) the population of programs to assess the statistical significance of the difference in concentration between total and live viewership after the introduction of TSTV. Therefore, we are comparing live and total viewership concentration within a given bootstrapped program population.

Figure 22 shows the results obtained. Each observation represents the difference in concentration (as measured by the Theil index) between live and total viewership in a bootstrapped program population. The concentration of total viewership is higher than that of live viewership both in the short-term ($p - value = 0.011$) and in the long-term ($p - value = 0.001$). Also, this difference is larger in the longer-term ($p - value = 0.045$). Therefore, it seems that the introduction of TSTV increases the concentration of TV viewership.

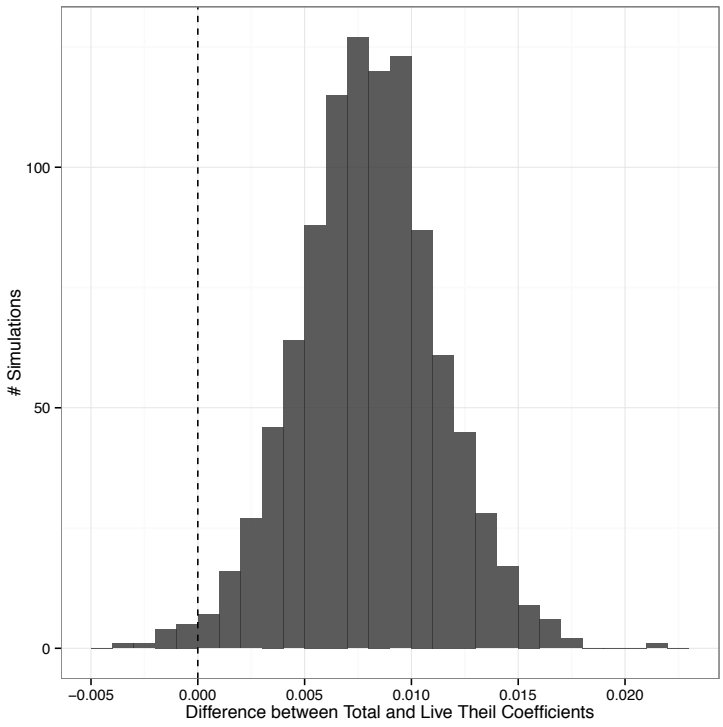

Figure 22: Distributions of the difference in the Theil Index between live and TSTV time.

F Impact of TSTV by Channel Type

Table 14: Impact of TSTV on TV Time by Channel Type (All Groups)

	Dependent variable:											
	Adult	Children	M. Bundle	Documentaries	Entertainment	Generalist	International	Lifestyle	Music	News	Other	Sport
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treated No TS	0.0002	0.0001	−0.0002	0.003	−0.006	0.019	−0.002	0.002	−0.002	−0.008	−0.0004	−0.003
	(0.0003)	(0.009)	(0.001)	(0.004)	(0.011)	(0.019)	(0.002)	(0.003)	(0.002)	(0.007)	(0.001)	(0.007)
After	0.0002	0.143***	0.003***	0.005***	0.080***	0.269***	0.006***	0.024***	0.002*	0.069***	−0.001	−0.051***
	(0.0003)	(0.004)	(0.001)	(0.002)	(0.005)	(0.008)	(0.001)	(0.002)	(0.001)	(0.003)	(0.0005)	(0.003)
Treated TS	0.0002	0.006	−0.0005	0.004	−0.003	0.022	−0.002	−0.003	−0.001	−0.002	−0.0001	0.002
	(0.0003)	(0.009)	(0.0005)	(0.004)	(0.011)	(0.019)	(0.002)	(0.003)	(0.002)	(0.007)	(0.001)	(0.007)
Treated No TS * After	−0.0001	−0.011*	0.208***	−0.018***	−0.085***	−0.065***	0.001	−0.011***	0.00004	−0.007*	−0.0004	−0.008*
	(0.0003)	(0.006)	(0.004)	(0.003)	(0.007)	(0.011)	(0.002)	(0.002)	(0.001)	(0.004)	(0.001)	(0.004)
Treated TS * After	−0.0001	−0.006	0.322***	−0.018***	−0.124***	−0.064***	0.001	−0.012***	−0.0001	−0.010**	−0.001	−0.010**
	(0.0003)	(0.006)	(0.005)	(0.003)	(0.007)	(0.011)	(0.002)	(0.002)	(0.001)	(0.004)	(0.001)	(0.004)
Constant	0.001***	0.327***	0.010***	0.130***	0.622***	1.661***	0.042***	0.115***	0.028***	0.322***	0.013***	0.219***
	(0.0002)	(0.006)	(0.0003)	(0.003)	(0.008)	(0.013)	(0.002)	(0.002)	(0.001)	(0.005)	(0.001)	(0.005)
Observations	70,214	70,214	70,214	70,214	70,214	70,214	70,214	70,214	70,214	70,214	70,214	70,214
R^2	0.00002	0.008	0.161	0.0004	0.002	0.006	0.0003	0.001	0.0001	0.003	0.0002	0.004
Adjusted R^2	0.00002	0.008	0.161	0.0004	0.002	0.006	0.0003	0.001	0.0001	0.003	0.0002	0.004

Note: *p<0.1; **p<0.05; ***p<0.01

Cluster robust standard errors in parentheses

Table 15: Impact of TSTV on TV Time by Channel Type (Treated Groups)

	Dependent variable:											
	Adult	Children	M. Bundle	Documentaries	Entertainment	Generalist	International	Lifestyle	Music	News	Other	Sport
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treated TS	0.0001	0.006	−0.0002	0.001	0.003	0.003	−0.0002	−0.005	0.002	0.005	0.0004	0.005
	(0.0003)	(0.009)	(0.001)	(0.004)	(0.011)	(0.019)	(0.002)	(0.003)	(0.002)	(0.007)	(0.001)	(0.007)
After	0.00005	0.132***	0.211***	−0.013***	−0.004	0.204***	0.007***	0.013***	0.002*	0.062***	−0.001***	−0.059***
	(0.0001)	(0.004)	(0.004)	(0.002)	(0.005)	(0.008)	(0.001)	(0.002)	(0.001)	(0.003)	(0.0004)	(0.003)
Treated TS * After	0.0001	0.004	0.114***	0.001	−0.039***	0.001	0.0001	−0.001	−0.0002	−0.003	−0.001	−0.001
	(0.0002)	(0.006)	(0.007)	(0.003)	(0.007)	(0.011)	(0.002)	(0.002)	(0.001)	(0.004)	(0.001)	(0.004)
Constant	0.001***	0.327***	0.010***	0.133***	0.616***	1.679***	0.040***	0.117***	0.025***	0.314***	0.012***	0.216***
	(0.0002)	(0.006)	(0.0004)	(0.003)	(0.007)	(0.013)	(0.002)	(0.002)	(0.001)	(0.005)	(0.0005)	(0.005)
Observations	46,952	46,952	46,952	46,952	46,952	46,952	46,952	46,952	46,952	46,952	46,952	46,952
R^2	0.00001	0.008	0.131	0.0004	0.001	0.005	0.0004	0.001	0.0001	0.003	0.0002	0.004
Adjusted R^2	0.00001	0.008	0.131	0.0004	0.001	0.005	0.0004	0.001	0.0001	0.003	0.0002	0.004

Note: *p<0.1; **p<0.05; ***p<0.01

Cluster robust standard errors in parentheses

G Week-level Analyses

Table 16: View Time as a function of TSTV - All Groups

	Dependent variable:		
	Total TV	Live TV	Time-Shift TV
	(1)	(2)	(3)
Treated No TS * After	0.006	0.030	-0.022^{***}
	(0.021)	(0.020)	(0.004)
Treated TS * After	0.086^{***}	0.048^{**}	0.032^{***}
	(0.021)	(0.020)	(0.005)
week	Yes	Yes	Yes
Observations	272,849	272,849	272,849
Adjusted R^2	0.018	0.021	0.007

Note: *p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in parentheses
Errors clustered at the household level

Table 17: View Time as a function of TSTV - Treated Groups

	Dependent variable:		
	Total TV	Live TV	Time-Shift TV
	(1)	(2)	(3)
Treated TSTV * After	0.080^{***}	0.017	0.055^{***}
	(0.021)	(0.020)	(0.005)
week	Yes	Yes	Yes
Observations	182,507	182,507	182,507
Adjusted R^2	0.018	0.021	0.007

Note: *p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in parentheses
Errors clustered at the household level

Table 18: Downloads as a function of TSTV - All Groups

	Dependent variable:
	Downloads (Mb)
Treated No TS * After	−20.083
	(17.895)
Treated TS * After	−49.791***
	(19.107)
week	Yes
Observations	272,849
Adjusted R^2	0.004
Note:	*p<0.1; **p<0.05; ***p<0.01
	Cluster robust standard errors in parentheses
	Errors clustered at the household level

Table 19: Downloads as a function of TSTV - Treated Groups

	Dependent variable:
	Downloads (Mb)
Treated TSTV * After	−29.704
	(19.624)
week	Yes
Observations	182,507
Adjusted R^2	0.004
Note:	*p<0.1; **p<0.05; ***p<0.01
	Cluster robust standard errors in parentheses
	Errors clustered at the household level

Table 20: Impact of TSTV by Channel Type - All Groups

	Dependent variable:											
	Adult	Children	M. Bundle	Documentaries	Entertainment	Generalist	International	Lifestyle	Music	News	Other	Sport
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treated No TSTV * After	−0.0001	−0.010	0.216***	−0.019***	−0.086***	−0.018***	0.001	−0.011***	0.00003	−0.001	−0.001	−0.010**
	(0.0003)	(0.006)	(0.004)	(0.003)	(0.007)	(0.004)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)	(0.005)
Treated TSTV * After	−0.0001	−0.006	0.334***	−0.019***	−0.131***	−0.022***	0.001	−0.013***	−0.0001	−0.001	−0.001	−0.010**
	(0.0004)	(0.007)	(0.006)	(0.003)	(0.007)	(0.004)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)	(0.005)
week	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	272,849	272,849	272,849	272,849	272,849	272,849	272,849	272,849	272,849	272,849	272,849	272,849
Adjusted R^2	0.00002	0.033	0.094	0.007	0.023	0.028	0.004	0.009	0.001	0.002	0.001	0.026

Note: *p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in parentheses
Errors clustered at the household level

Table 21: Impact of TSTV by Channel Type - Treated Groups

	Dependent variable:											
	Adult	Children	M. Bundle	Documentaries	Entertainment	Generalist	International	Lifestyle	Music	News	Other	Sport
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treated TS * After	0.00000	0.004	0.117***	−0.0005	−0.044***	−0.004	0.0001	−0.001	−0.0001	0.0004	−0.0003	−0.001
	(0.0003)	(0.006)	(0.007)	(0.003)	(0.007)	(0.004)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)	(0.005)
week	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	182,507	182,507	182,507	182,507	182,507	182,507	182,507	182,507	182,507	182,507	182,507	182,507
Adjusted R^2	0.00003	0.032	0.101	0.007	0.021	0.027	0.004	0.009	0.001	0.002	0.001	0.028

Note: *p<0.1; **p<0.05; ***p<0.01

Cluster robust standard errors in parentheses

Errors clustered at the household level