Information vs. regulation: a case study of Israeli television*

Ittai Shacham†

27th May 2019

Preliminary and incomplete, please do not circulate

Abstract

Broadcast television channels fund free-to-view content via commercials. The uncertainty surrounding the timing and length of commercial breaks generates an informational friction hindering viewers' choices. Regulators have traditionally approached this issue by administering quantity restrictions. A-priori, the effect of this friction on the equilibrium amount of commercials under alternative regulatory schemes in unclear. To provide insight into the efficacy of different regulatory frameworks, I develop a novel model that explicitly characterizes how choices and beliefs interact, which I estimate using high-frequency data from Israel. Combined with a model of platforms' commercial behavior in the presence of quantity constraints, I simulate the equilibrium amount of commercials under alternative regulatory frameworks. I find that quantity restrictions restrain the average amount of commercials by 23% and 67% on the two commercial channels. Providing viewers with information regarding the length of commercial breaks, attenuates the effect partially, leading to an average increase of 16% and 56%, respectively. The results suggest that an alternative regulation framework, providing viewers with information and partially relaxing the quantity restriction constitutes a welfare improving intervention.

*Special thanks to my advisors Jaap Abbring, Tobias Klein and Jeffrey Campbell for support and guidance. I am grateful to Robert Clark, Ulrich Doraszelski, Alon Eizenberg, Richard Friberg, Martin Hackmann, Aureo de Paula, Niccola Pavanini, Mark Roberts, Mark Schankerman, Jörgen Weibull and participants of the Tilburg Structural Econometrics Group and Law & Economics group for comments and suggestions.

†Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: i.shacham@tilburguniversity.edu

1 Introduction

The past decade has seen the rise of many alternatives to broadcast television, with that, it remains one of the primary forms of entertainment and advertising. In the second quarter of 2017, average adults in the US watched approximately 4.5 hours of TV per day, four of which was live TV (Nielsen (2017)). In Israel, the numbers are similar, in 2011, the average Israeli adult watched 3.9 hours of television per day[1]. Moreover, many of the alternatives to traditional television are variants of business models prevalent in TV, e.g. Hulu and YouTube allocate advertisements within their content. Similarly, Netflix has announced of late that it will test unskippable ads between shows, increasing its similarity to broadcast television.[2] Broadcast TV also remains the largest advertising platform, accounting for 33% of global ad expenditure in 2018[3]. The main advantage of online platforms is their targeting ability. Even along this dimension, the disparity between traditional content providers and online ones is shrinking. In the Netherlands, a major TV provider has begun displaying targeted advertising to viewers within the same commercial break.[4] A stark contrast between broadcast television and new forms of entertainment lies in the regulation they face[5]. Most countries, among them the EU countries, Australia and Israel, regulate the permitted amount of commercials and in some cases, also their temporal distribution, i.e. the permitted timing and length of the commercial breaks. While in many countries broadcast television faces substantial quantity restrictions, their online equivalents remain unregulated in this aspect, creating scope for regulatory arbitrage.

This paper focuses on broadcast television, trying to address the positive and negative welfare effects quantity restrictions may induce. Broadcast TV has several distinct characteristics, namely quantity restrictions and informational frictions. TV and radio provision (including their online equivalents) are alike and differ from other forms of media in that one of the key factors affecting consumer behavior is an *informational friction*: The uncertainty shrouding the timing and length of commercial breaks requires viewers' to base their viewing choice on their assessment of the commercial probability, constraining their switching behavior relative to the full information scenario, thereby increasing the platforms' market power. To avoid an excessive amount of advertisements, many countries impose regulatory restrictions on broadcast television channels. With that, quantity regulation[6] comes at an efficiency cost. Specifically, regulation is never perfect, it may set an upper bound that is too high in which case it has no effect on the service providers, as shown by Zhang (2017) for the case of television in France. Alternatively, the bar may be set too low, driving down industry standards (Wright (1994)) and decreasing product diversity (Gabszewicz et al. (2000)). Moreover, price regulation doesn't allow channels the flexibility required to contend with varying demand environments[7].

Alternatively, quantity regulation may act as a commitment mechanism for the channels, restricting their ability to allocate an excessive number of ads within a specific day, thereby hindering their future revenues. Within a dynamic context, the channels' chosen amount of commercials will influence both the

[1] Gili Izikovich. Study: Israelis Spend Nearly Four Hours a Day Watching TV. Haaretz. December 20, 2011. `https://www.haaretz.com/1.5221105/1.5221105`

[2] Luke Bouma. Updated: Netflix is Testing Ad Between Episodes. Cord Cutters News. August 17, 2018. `https://www.cordcuttersnews.com/netflix-is-testing-unskippable-ads-between-episodes/`

[3] `https://www.statista.com/`

[4] Je buurman krijgt bij KPN straks andere tv-reclames te zien dan jij. NOS. September 26, 2018. `https://nos.nl/artikel/2252164-je-buurman-krijgt-bij-kpn-straks-andere-tv-reclames-te-zien-dan-jij.html`

[5] For example, the European Audiovisual Media Services Directive specifies that broadcasters cannot exceed 12 minutes of advertising per hour. Australia imposes similar restrictions as specified in the Commercial Television Industry Code of Practice.

[6] Regulation of the amount of commercials may be more intuitively seen as quantity regulation although Gabszewicz et al. (2000, 2004) show its equivalence to price regulation.

[7] For example, regulations restricting the amount of advertisements on TV cannot account for seasonal fluctuations in the advertisers demand thereby creating distortions in the product markets (Gal-Or and Dukes (2003); Dukes and Gal-Or (2003)).

intensive and the extensive demand margins they face. At the intensive margin, the amount of ads will influence the viewers' behavior within the day. The extent to which they are averse to advertisements along with the quality of their outside options (all other viewing alternatives as well as non-TV alternatives), will determine their switching behavior. At the extensive margin, the viewers' anticipation of the number of ads will guide their decision whether to view broadcast TV as a viable alternative, i.e. participate in the market. When viewers' update their beliefs in a Bayesian manner, a revenue maximizing amount of ads within a day may be detrimental to their revenue in the future as well as to societal welfare. When the channels are not sufficiently patient or lack the necessary capabilities to optimize their long run profits, the number of advertisements may be inefficient. In the future I aim to use the framework developed herein to address the potential positive effects of regulation on the broadcast TV market.

An alternative to behavioral regulation is alleviation of the friction generating the enhanced market power of the service providers. Providing consumers with the necessary information accompanied by uninhibited pricing allows firms to price their products according to their characteristics, e.g. accounting for quality differences, while enabling consumers to curb excessive pricing through switching behavior. The main drawback of transitioning to this form of regulation lies in the uncertainty shrouding the equilibrium amount of commercials. Depending on the elasticity of viewing demand, the resulting equilibrium may constitute an excessive amount of advertisements from a policy perspective. Furthermore, as shown in Anderson and Coate (2005), such an equilibrium may incorporate an excessive amount of ads also from a social welfare perspective. Alternatively, if information provision generates a substantial increase in switching behavior, the equilibrium amount of commercials may be too low from a societal perspective and the platforms may be unable to operate profitably. A-priori, the effect of this sort of policy intervention is unclear, requiring an empirical investigation of its potential effects.

This paper provides a framework to conduct an empirical regulation assessment in television, comparing the welfare effects of the commonplace quantity regulation to alternative ones. To this end, I propose a viewer demand model incorporating informational frictions regarding the true state of the world (commercial state). I formulate the viewers' belief structure pertaining to the probability of a commercial, conditional on their viewing history. The commercial probability acts as a dynamic component by which an individual's viewing history is summarized. I estimate the model using high-frequency data from Israel on each of the three broadcast channels on each minute of prime-time (20:00-22:00) throughout 2002. Two issues confound demand estimation in TV markets. The first arises from the endogoneity of commercials, stemming from the correlation between the commercial status and unobserved program quality (Trajtenberg (1989)). I overcome this concern using the high-frequency nature of the data. Within high frequency data, the ad status is an indicator as opposed to a continuous value in lower frequency data. Since ads are broadcast on all types of shows - more and less popular - the correlation between ad status and show popularity diminishes. Moreover, I utilize an amendment to the commercial regulation in Israel that transitioned from a per hour quantity restriction to a per prime-time (two hours) quantity restriction without affecting the aggregate amount of ads permitted. Finally I include Berry et al. (1995) instruments. Controlling for endogeneity, without accounting for the informational friction, inflates the parameter estimate for the sensitivity to commercials by approximately 22%, from -.195 to -0.237.

The second challenge to demand estimation within TV market lies in the interplay between the informational friction - requiring the viewers' to make their viewing decisions based on individual assessments of commercial probability - and preferences. Disregarding the informational aspect of viewers' decision process will attenuate their commercial sensitivity, frustrating estimation of the households' underlying preferences.

I account for the asymmetric information by incorporating structure on the viewers' beliefs and the manner by which they affect decisions. Initially, I provide evidence from the raw data pertaining to the existence and form of the informational frictions, motivating the specification of the expectation formation. Specifically, viewers' leaving behavior is proportional to the length of the commercial break, implying rational expectations. Furthermore, the delayed leaving of viewers at the beginning of a commercial break and their gradual return imply a learning process. The magnitude of the parameter estimate for the average disutility from commercials is greater than not accounting for the informational friction by roughly a factor of 2, decreasing from -2.225 to -4.727.

To predict the equilibrium level of commercials under alternative regulatory policies, I follow the formulation of Berry et al. (1995) and update it to account for the specific characteristics of the television industry. I formulate the platforms as competing á la Nash-Bertrand and concentrate on the platforms' long run ad propensities for each minute, thereby abstracting away from the within day dynamics. These commercial propensities constitute the platforms' mixed-strategies within a static equilibrium in which the viewers' beliefs are consistent with the platforms' strategies. Using the results from the demand model together with the supply formulation and equilibrium concept, I conduct policy experiments for alternative regulation schemes. I find that the quantity restrictions substantially curb the two commercial channels' advertising strategies. Absent any regulation restrictions, the average (median) amount of commercials would increase by 23% (11%) and 67% (67%) on each of the two channels. Supplying viewers with a timer informing them of the duration of a commercial break and relaxing the quantity restrictions would generate a more subtle increase in the equilibrium amount of commercials, leading to an average (median) increase of 16% (4%) and 56% (54%), respectively. These simulations imply scope for a welfare improving policy intervention that supplies viewers' with the information required to overcome the informational friction, accompanied by a relaxation of the quantity restrictions.

Relation to the literature

This paper relates to several strands of literature. Methodologically, it conducts a structural demand estimation in television markets. There is a large body of literature which estimated demand in television and radio markets. With that, none have accounted for the informational frictions in the households' choices. Waldfogel and Berry (1999); Berry et al. (2016a,b) estimate the socially optimal amount of radio stations accounting for both horizontal and vertical differentiation, and provide evidence of excessive entry. Berry and Waldfogel (2001) documents that the excessive entry in the radio industry enhanced programming variety relative to the number of stations, in line with theoretical predictions. Berry and Waldfogel (1999) estimate the degree of substitutability between public and commercial radio stations while Berry et al. (2016a) document asymmetric listening behavior of different populations leading radio stations un-proportionally cater to the majority taste preference in the US. Baker and George (2015) estimate a multi-stage game in which television channels choose the type of programming and afterwards the amount of commercials. They find that local news programming is under-supplied in equilibrium relative to the viewer optimum. Zhang (2017) extends the viewer demand estimation by accounting for heterogeneous viewer preferences. He finds that under competition among the channels, market forces are a sufficient restraint on the channels' incentives, making regulation unnecessary. Alternatively, within a collusive environment the regulation improves viewer welfare. Shcherbakov (2016) considered the effect of switching costs by estimating a dynamic demand model for cable TV consumption and finds evidence of significant economic switching costs, affecting the cable providers' pricing strategy.

Thematically, this paper concentrates on the equilibrium commercial behavior of the channels under alternative regulatory frameworks. Gabszewicz et al. (2000, 2004) show, in a theoretical model, that setting a constraining upper bound on the level of advertising will result in less differentiated programing. Anderson and Coate (2005) clarify the two-sided nautre of TV platforms, whose payoff is a function of their being able to get both the advertisers and the viewers on board, while the preferences of these two groups are opposing. They show that the amount of advertising can be over or under provided in regards to the social optimum, in both a monopolistic and duopolistic setting, depending on the viewers' distaste for commercials and the level of competition between the platforms. Sweeting (2010) finds that mergers in the radio industry lead the merged stations to increase the differentiation among themselves. With that, he also finds that the merged stations locate closer to their competitors, thereby not necessarily increasing variety or decreasing duplication. As opposed to the evidence of duplication, Goettler and Shachar (2001) find that under free entry, the locational choices of television broadcasters are close to the social optimum. Sweeting (2006, 2009) investigate radio stations incentive to coordinate commercial timing. These papers produce evidence of stations actively trying to coordinate commercial timing, although the coordination is not perfect. Epstein (1998) finds similar patterns within the broadcasting time of four national US television broadcasters during prime time. Crawford and Yurukoglu (2012) model and estimate the effects of channel bundling in cable TV. They find that the welfare effects of á la carte pricing are offset by the increased costs associated with the alternative pricing scheme.

This paper contributes to the literature by incorporating an informational friction inherent in media markets and show that the presence of such frictions have a significant economic impact on viewer choices. It provides evidence of the equilibrium effects of changes in regulation in the presence of such frictions and proposes an information based welfare improving policy intervention. Although many of the digital content providers are unregulated, they at times choose to provide their viewers with the necessary information to diminish the informational friction, as in the case of YouTube. Moreover, the similarity in the type of product as well as in the business models between the traditional and digital platforms imply that the viewers' preferences do not differ substantially across these different outlets. Hence, any results and conclusions derived in this paper should apply, with the appropriate caveats, also to online entertainment outlets.

The paper is organized as follows: Section 2 presents an institutional background of the television industry in Israel, overviews the data and produces evidence of frictions; sections 3 & 4 present the viewer demand model, its estimation procedure and results. Section 5 presents a model of platform behavior and equilibrium; section 6 presents the results from alternative policy experiments. Finally, section 7 concludes.

2 Regulation and informational frictions in television

A short history of Israeli television

The first Israeli television channel was the Israel national channel also referred to as *Channel 1*, which began broadcast in 1968. 25 years later, in 1993, the second Israeli television channel was introduced, commonly referred to as *Channel 2*. In contrast to channel 1 which is a governmentally financed non-profit channel, channel 2 is a commercial for-profit channel. In an attempt to introduce a degree to competitiveness in the programming, the regulator divided the broadcasting days of channel 2 among three networks: Keshet, Telad and Reshet. Each of the three networks broadcast on different days on the same channel, e.g. in 2003 Telad broadcasted on Sunday and Wednesday; Reshet on Tuesday, Friday and Saturday (until April); and

Keshet broadcasted on Monday, Thursday and Saturday (from April onwards)[8]. In 2002 a second commercial channel was introduced, *Channel 10*.

The regulation of commercial television is explicit on the number and amount of commercials allowed to be broadcasted. The initial law regulating commercial activity of the commercial channels dates back to 1992, several changes have been implemented in the regulations throughout the years, where the final change was implemented in 2009. Table 1 presents the main points in the regulation of commercial placement during the time frame of the data used in the empirical analysis.

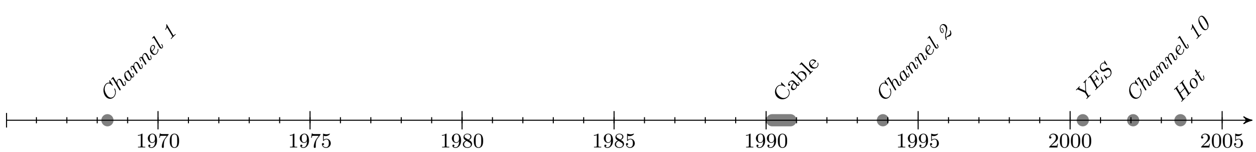
Table 1: Commercial placement regulation

Criterion	Until 22.5.2002	After 22.5.2002
General		
Length of commercials[1]	12 min per hour	24 min per 2 hours
Length of commercial break	5 min	5 min
Length of single commercial	1.5 min	No limit
Number of commercial breaks	3 per hour	4 per hour
Broadcast on memorial days[2]	disallowed	disallowed
News program		
Length of commercial break	2 min	3 min
Number of commercial breaks	$\ell \leq 10$: no break;	$\ell \leq 10$: no break;
	$10 < \ell \leq 40$: 1;	$10 < \ell \leq 40$: 1;
	$40 < \ell \leq 60$: 2;	$40 < \ell \leq 50$: 2; $50 < \ell \leq 60$: 3;
	$\ell > 60$: 3	$\ell > 60$: 4 per hour

Notes:
[1] Within the two hours of prime-time. Outside of prime-time the channels are permitted 12 minutes per hour.
[2] There are three memorial days - Yom Hashoa, Yom Hazikaron and Tisha'a BeAv - defined by the Jewish calendar and are therefore on different dates of the Georgian calendar on each year.

Parallel to the development of broadcast commercial television, multi-channel broadcast - cable and satellite - began operating circa 1990. Three operators were chosen to provide cable television throughout Israel, each in a different geographical area. In 2003 the three cable companies were unified under one company, Hot Telecommunication Systems. 2000 saw the introduction of a direct broadcast satellite operator - YES. By the mid 2000's, roughly 66% of households had access to multichannel broadcast, consisting of roughly 74% of households owning a television. Figure 1 presents a timeline of the major events in the Israeli TV industry up to 2005.

Figure 1: Israeli TV industry timeline



[8]Telad lost concessionaires broadcasting bid for channel 2 and ceased broadcast on 30th of October 2005, leaving only Reshet and Keshet.

Data overview

I utilize two unique dataset regarding viewer shares and characteristics on three broadcast television channels in Israel in 2002. The data was acquired from Kantar Media, a private firm that collects data on television viewer shares. The *Programs* dataset details the start time, end time and genre[9] of every program and commercial break on channels 1, 2 and 10 that aired during prime-time (20:00-21:59). The *Dayparts* dataset details the viewer market share across various segments of the population[10] for every minute on every one of the three channels throughout the aforementioned time period. The data was aggregated at the minute level. Whenever more than one program was broadcast on a given minute, the program for which the largest amount of time was allocated was associated with the specific minute of broadcast.

The broadcasting behavior of the two commercial channels - 2 & 10 - is very similar and stands in contrast to that of the public channel - 1. Table 2 presents summary statistics regarding the programming of the three channels. The similarity between the commercial channels is apparent in both the commercial placement behavior as well as the program length. While the average amount of in-program commercials on channels 2 and 10 is roughly 9.5 minutes, the average length of in-program commercials on channel 1 is only around 20 seconds. Additionally, there is an average of two commercial breaks per program in the commercial channels, while breaks in programming of channel 1 are uncommon as implied by an average of 0.07 breaks per program. Furthermore, while the average program length on channels 2 and 10 is roughly 50 minutes, the average program length on channel 1 is significantly shorter, roughly 40 minutes.

Table 2: Program summary statistics

	Mean	SD	Max	Min	Obs.
Channel 1					
Program length[1]	40.56	19.03	1	120	
In-program commercial length	.29	1.34	0	13	
Between-program commercial length	4.04	1.42	1	10	215,813
Commercial breaks per program	.15	.66	0	7	
Length of commercial break	2.78	1.81	0	10	
Channel 2					
Program length	50.17	19.47	1	120	
In-program commercial length	9.96	5.89	0	29	
Between-program commercial length	6.37	2.09	1	15	214,138
Commercial breaks per program	2.02	1.05	0	7	
Length of commercial break	5.53	1.72	0	16	
Channel 10					
Program length	51.72	19.91	1	120	
In-program commercial length	9.05	4.25	0	21	
Between-program commercial length	4.29	1.87	1	9	168,337
Commercial breaks per program	2.37	1.17	0	10	
Length of commercial break	4.25	1.21	0	14	

Notes:
[1] Programs were censored at both their beginning and end to the time frame of the data, e.g. if a program began prior to 20:00, the start time used to calculate the program duration is 20:00.
[2] All time variables, i.e. *Program length*, *In-program commercial length*, *Between-program commercial length* and *Length of commercial break* are denoted in minutes.

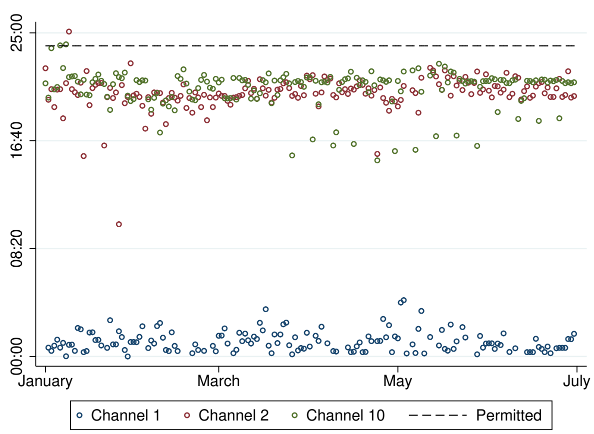
[9] The genres were defined by the data collection agency and are: (1) children; (2) cinema; (3) commercial; (4) culture, leisure & education; (5) documentary; (6) entertainment; (7) news & current events; (8) sports; (9) television drama; and (10) other.

[10] The subpopulation are: (1) all households; (2) Jewish households; (3) households by income (a lot below average, below average, average, above average, a lot above average); and (4) households by age (4-24, 25-44, 45-64, 65+). Note that the households self-classified themselves into income groups and the age denotes the age of the head of the household.

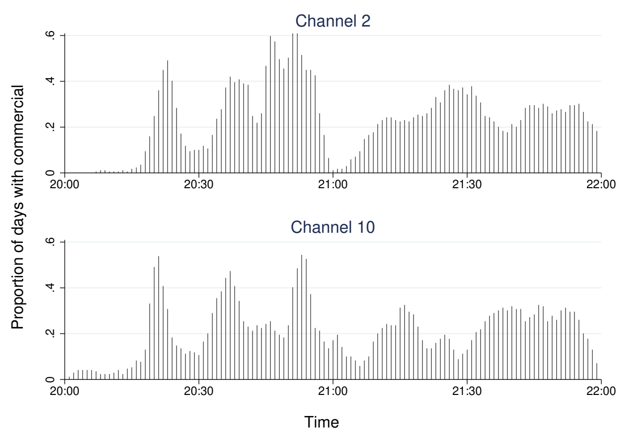
The commercial placement behavior of the channels exhibits a large heterogeneity in the daily amount of commercials between days for each channel[11] as well as large within day variation in the temporal allocation of these commercials. Figure 2a shows the daily number of ads across the channels. Several interesting facts stand out. The similarity between the two commercial channels and difference with the governmental channel is clear from the figure. While on most days, both commercial channels broadcast an amount that is very close to the permitted amount, the public channel almost always broadcasts an amount of commercials substantially less. Moreover, the figure displays the entry behavior of channel 10, with low amounts in the first year and a half, after which the amount of commercials aligns with that of the incumbent channel 2. Figure 2b displays the propensity of a commercial in each minute of the two-hour prime time window on the two commercial channels. Beyond the variation in commercial placement, the figure also shows that some minutes are more prone for commercials. Moreover, there is a similarity in the commercial propensities among the two channels. This similarity is much more pronounced in the first half of the prime-time than in the latter half.

Figure 2: Commercial distributions

(a) Daily commercial amount



(b) Commercial timing patterns



Notes: Panel 2a displays the distribution of the daily amount of commercials on each channel during prime time (20:00-21:59) between 2002 and 2005. The data used to generate this graph is the *spots* dataset detailing the precise length (at the second level) of each commercial. Memorial days are excluded as well as 10 days in which the amount of commercials exceeded 1.5 times the permitted amount (36 minutes).
Panel 2b presents the commercial timing patterns across the two commercial channels. The horizontal axis presents the minutes within the day and the vertical axis presents the proportion of days within the data in which a channel broadcasted a commercial.

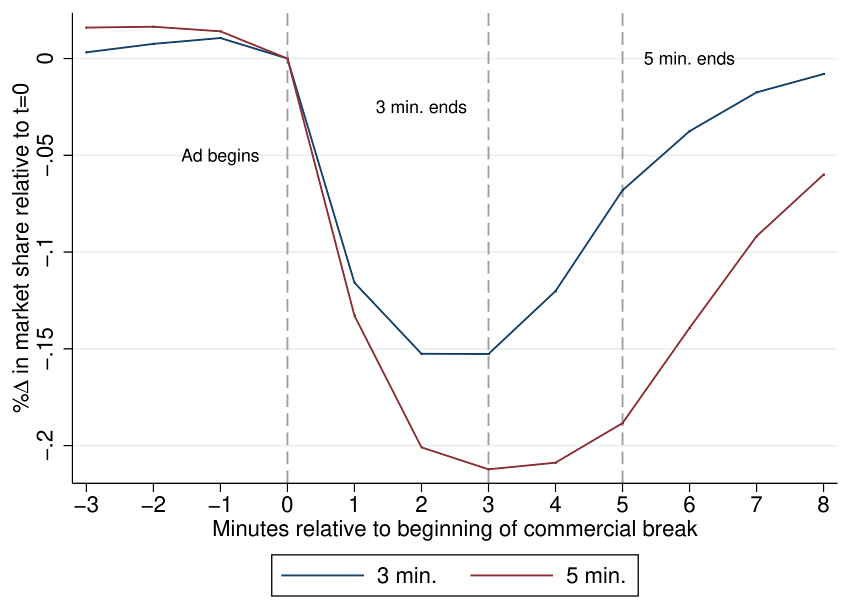
Information frictions

Figure 3 provides initial insight into viewer behavior, presenting the percentage change in the market share of a channel before and after the beginning of a commercial break, broken down by commercial break length. The first conclusion from the figure is that coordination frictions exist. For every commercial length, there are viewers which return to the channel broadcasting a commercial too early, while the channel is still broadcasting commercials, whereas other viewers return too late, after the program has resumed. The second conclusion derived from this simple figure is that viewers' expectations pertaining to the length

[11] Industry professionals stressed that during high ad-demand periods, the channels at times choose to broadcast more ads than permitted by the regulation, consequently receiving fines and sanctions. Moreover, during low demand periods, the channels choose to broadcast less commercials at a relatively high price than 'fill the quota' and let the ad prices decrease.

of a commercial break are rational. The number of minutes in which viewers leave a channel is directly proportional to the true length of the commercial break, e.g. within three minute commercial breaks, channels lose market share in the five minutes after the beginning of a commercial break, while for one minute commercial breaks the loss is noticeable for only two minutes after the beginning of the break. The rationality of the viewers is in their ability to correctly form expectations regarding the length of the commercial break. Finally, the shape of the viewership loss provides insight into the informational frictions. Specifically, the viewers leave 'en mass at the beginning of the commercial break, but it is important to notice that this occurs with a lag. Afterwards, the viewers gradually return. These observations will guide the formulation of the coordination frictions within the model.

Figure 3: Percent change in market share, by commercial break length



Notes: The horizontal axis presents the minutes before and after the beginning of a commercial break and the vertical axis presents the percentage change in the channels' market share. The reference point by which the percentage change was calculated is the minute before the beginning of the commercial break, denoted by $t = 0$.

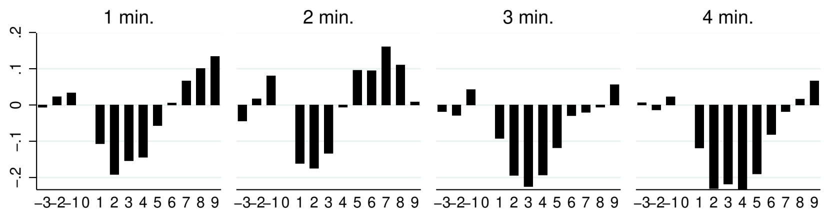
The entry of channel 10 within the period covered by the data provides a case study to examine the development of the viewers' beliefs pertaining to the length of a commercial break via their switching behavior. Panel 4a presents the percentage change in the market share of channel 10 in the first month of the channel's operation, while panel 4b displays the switching behavior from the second month onwards. It is clear form the figure that the lower panel coincides with figure 3 yielding similar conclusions. The upper panel, on the other hand, does not present a clear distinction in the switching behavior between different commercial length breaks. Therefore, between these two time periods, the viewers learned the commercial broadasting behavior of the channel and generated (rational) beliefs accordingly.

3 Viewing demand model

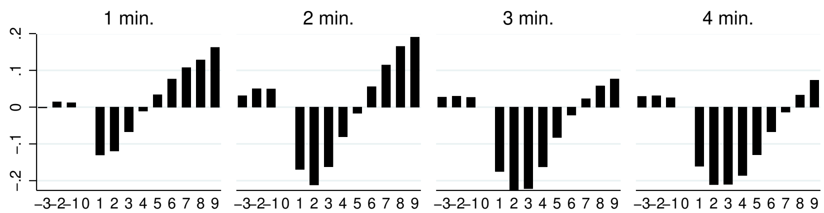
In the following section I develop a model for TV viewership demand following the characteristics-based demand literature in line with previous research in the field (e.g. Berry et al. (2016b); Baker and George (2015); Zhang (2017)). In this model, potential viewers face a two stage process: initially, viewers choose

Figure 4: Percent change in market share, by commercial break length

(a) First month



(b) Months 2+



Notes: The horizontal axis presents the minutes before and after the beginning of a commercial break with the minute before the beginning of the commercial break denoted by $t = 0$. The vertical axis presents the percentage change in channel 10's market share. The reference point by which the percentage change was calculated is $t = 0$.

whether to opt-in to television watching within prime-time of a specific day or participate in different activities, e.g. reading a book, going out or browsing the web. The TV watching decision is decided upon ex-ante before the day, requiring the viewers to consider their expected utilities. Upon making their leisure decision, the market size for the day is determined. Within the prime-time and conditional on deciding to watch television, the households' viewing decisions are myopic in that they choose, at each period, the alternative that maximizes their contemporaneous utility without consideration of the future. I augment the model by incorporating uncertainty pertaining to the commercial status of the channels, resulting in informational frictions. The informational friction embeds a dynamic component in the model: viewer specific commercial probabilities are dependent on the viewer's information set, making their viewing choice conditional on their viewing history. The gradual and continuous updating of individuals' information sets incorporates a learning process. In this process, viewers update their individual assessment of the commercial probabilities based on the new information they acquire.

Let time within the day be discrete (and finite), denoted by $t = 1, ..., T$ and the day by $d = 1, ...D$ (also discrete and finite). Potential viewers have several options from which they can choose: watch one of the commercial broadcast channels, $\mathcal{J}_c$; one of the broadcast public channels, $\mathcal{J}_p$; one of the multi-channel options, $\mathcal{J}_m$; or not watch television. Within each period viewers are assumed to choose one of the $j \in \{\mathcal{J}_c, \mathcal{J}_p, \mathcal{J}_m\}$ viewing alternatives or the outside option, denoted by $j = 0$. Furthermore, viewers are assumed to choose only one alternative in each period. Within the data used in the empirical analysis, a time period is a minute, making this assumption rather straightforward. Hence, the set of alternatives in each time period is given by $\mathcal{J} = \{0, 1, ..., J\}$ where J is the sum of the viewing alternatives in $\{\mathcal{J}_c, \mathcal{J}_p, \mathcal{J}_m\}$.

Stage 2: viewing decision

Viewers make their decisions within an incomplete information framework. The *informational structure* of the model is that viewers possess a "viewing guide" knowledge, meaning the only uncertainty pertains to the timing of intra-program commercial breaks on the commercial channels. [12]. The timing of the model is as follows:

0. At the outset of each period individual taste shocks are realized;
1. Viewers make their viewing choice for the upcoming period, $y_{it} \in \mathcal{J}$;
2. Viewers learn the commercial status of their chosen channel, $a_{y_{it}t} \in \{0, 1\}$.

As implied by the timing of the model, viewers possess full knowledge of all relevant characteristics except for the commercial state of each channel. The discrepancy in the available information to the viewers at the time of making their viewing choice and the realized commercial state of the channels, creates a divide between the viewers' utility guiding their choices (ex-ante utility) and their realized utility (ex-post utility). The ex-post indirect utility of viewer i from watching channel j in period t and day d is separably additive in the channel characteristics at time t:

$$u_{ijtd} = \alpha a_{jtd} + x'_{jtd}\beta + \zeta_{jtd} + \varepsilon_{ijtd} \quad \forall \;\; i \in \mathcal{I},\; j \in \mathcal{J},\; t = 1, ..., T,\; d = 1, ..., D \tag{3.1}$$

For convenience, I will omit the day subscript for the remainder of the section. a_{jt} is an indicator of whether an ad is being broadcast on channel j at time t; x_{jt} is a K-dimensional vector of observed characteristics of channel j at time t; ζ_{jt} summarizes the unobserved (by the econometrician) characteristics of channel j at time t; and finally ε_{ijt} is an idiosyncratic taste shock iid across individuals, alternatives and time. I will assume ε_{ijt} follow a type 1 extreme value distribution (henceforth TIEV). Within the parameter vector $\theta = (\alpha, \beta)$, α is a scalar representing the average viewer's (dis)utility from a commercial and β is a K-dimensional vector of average taste coefficients for a channel's characteristics.

The viewers may decide not to watch any of the channels, opting for the outside option, in which case the indirect utility is normalized to zero, yielding the viewer a reservation utility of $u_{i0t} = \varepsilon_{i0t}$.

The ambiguity surrounding the commercial state at the time of making their viewing choice propels viewers to consider their expected contemporaneous utilities, conditional on their information set at time t, i.e their ex-ante indirect utilities:

$$U_{ijt} \equiv \mathbb{E}(u_{ijt}|\mathcal{H}_{it}) = \alpha\mu_{ijt} + x'_{jt}\beta + \zeta_{jt} + \varepsilon_{ijt} \tag{3.2}$$

where $\mu_{ijt} \equiv \Pr(a_{jt} = 1|\mathcal{H}_{it})$ and $\mathcal{H}_{it} = (\mathcal{H}_{ijt})_{j=1}^{J}$ is the information set of individual i regarding all channels. A detailed exposition of the development and evolution of the viewers' ad probabilities is presented in section 3. Aggregation of the viewers' choices at each minute, together with market size denoted by M, generate the channels' market shares:

$$\mathfrak{s}_{jt} = \int_{i \in \mathcal{I}_d} y_{ijt} dF_\varepsilon / M \tag{3.3}$$

where F_ε denotes the population distribution of the taste shocks.

[12]Concisely: 1. all show attributes are known; 2. commercial break times of public channels are known; 3. timing of inter-program commercial breaks are known; 4. timing of intra-program commercial breaks are unknown.

Stage 1: participation decision

Households' decision whether to watch television within a certain day's prime time determines the market size within that day. The choice whether to do so relies on their assessment of the potential utility they would derive from participation in the market, in turn relying on the broadcast characteristics and commercial placement decisions of the channels. The informational structure implies that the characteristics of the programs are known in advance. With that, the commercial placement behavior of the channels remains unknown. In line with the viewers' rational expectations (the assumption will be more thoroughly detailed in section 3), the expected advertising behavior of the channels' is known to the viewers. As such, the maximal utility viewer i derives from participating in the market at time t would be simply $\max_j(U_{ijt})$. Using the TIEV distribution of the viewers' taste shocks, this simplifies to:

$$\max_j(U_{ijt}) = \log\left[\sum_{j\in\mathcal{J}} \exp\left(\alpha\lambda_{jt} + x'_{jt}\beta + \xi_{jt}\right)\right] \tag{3.4}$$

where $\lambda_{jt} = \Pr\left(a_{jt} = 1|x_{jt}\right)$. The viewers' expected utility from participation in a specific day is given by the total expected utility from viewing along with a day effect:

$$\tilde{U}_{ijd} = \Omega_{jd} + \gamma_d + \psi_{ijd} \tag{3.5}$$

where $\Omega_{jd} = \sum_{t=1}^{T} \log\left[\sum_{j\in\mathcal{J}} \exp\left(\alpha\lambda_{jt} + x'_{jt}\beta + \xi_{jt}\right)\right]$ is the total expected utility from participation in the market, γ_d is a day effect and ψ_{ijd} is an idiosyncratic taste shock, independent of the within day taste shocks ε_{ijt}. Normalization of the coefficient of Ω_{jd} implies that the day effects are estimated in relation to it. The utility from alternative activities, i.e. not participating in the market, is normalized such that $\tilde{U}_{i0d} = \psi_{i0d}$.

It is important to note that the participation decision does not restrict the viewers' choice set within the prime time to viewing television. Within each period, not viewing remains an option, but non-participation implies that television watching is excluded from the individual's choice set for that day.

The market size for a given day is determined by integration over all individuals' choices. Specifically, the distribution of the taste shocks (ψ_{ijd}, ψ_{i0d}) will determine the form of the aggregation. I will continue with the assumption that the taste shocks follow a TIEV distribution implying that the market size af a specific day is given by:

$$M_d/M \equiv \mathscr{s}_d = \Pr(\tilde{U}_{ijd} \geq \tilde{U}_{i0d}) = \frac{\exp(\Omega_{jd} + \gamma_d)}{1 + \exp(\Omega_{jd} + \gamma_d)} \tag{3.6}$$

Commercial probabilities

The viewer demand model is completed by specifying a belief structure guiding the viewers' expectations regarding ad probabilities. I will begin by specifying the objective ad probabilities, common to all viewers. I will then specify the viewers' information set and define the form by which individual expectations are formed. In the end, I discuss the commercial probability formation within a learning framework.

Objective probabilities

The objective commercial probabilities, common to all viewers, specify the instantaneous probability of an ad at each period, i.e. the hazard rates. These expectations are assumed to be rational in the sense

that they adhere to the programming behavior of the channels. Figure 3 provided a motivation for this assumption, showing that aggregate viewing behavior coincides with the length of the commercial breaks. The viewers' expectations together with the commercial utility parameter jointly determine the switching behavior of households during commercial broadcast. Consequently, this assumption is used to identify each of the components in isolation[13]. The rational expectations assumption could condition on all the possible information on each channel-time, with that, I restrict the informational state space to a subset of the channel-time characteristics, namely, the genre and position within the program.

The objective hazard rates follow a semi-Markov process with two states: program and commercial, where the amount of time spent in a state is denoted by τ_{jt}. These hazard rates, while dependent on the duration within a state, are independent of previous state transitions. Underlying this assumption is limited recall on the side of the viewers. Commercial regulation permits channels to broadcast up to 24 minutes of commercials within the two hours of prime time at the channel's discretion. The large amount of commercials inhibit the viewers' ability to incorporate the past number of commercials into relevant information related to the transition probability in the current state. Furthermore, the hazard rates are assumed to be independent across channels and programs. Hence, knowledge of the state and duration on one channel provides no further information regarding the objective ad probability on the other channels. The independence from previous state transitions and across channels are stated formally in assumption 1. The independence across programs is captured through an initial and terminal condition within each program. Specifically, the first and last period of a program are known to be of a program state. This is stated formally in assumption 2.

Assumption 1. *Let the set of random variables (A_{nj}, T_{nj}) denote the state and time in channel $j \in \mathcal{J} \setminus 0$ at the n^{th} transition with the associated duration $\tau_{nj} = T_{nj} - T_{n-1j}$. The sequence (A_{nj}, T_{nj}) is characterized by:*

$$\Pr\left(\tau_{nj} \leq \tau | (A_{0k}, T_{0k})_{\forall k \in \mathcal{J}}, ..., (A_{n-1k}, T_{n-1k})_{\forall k \in \mathcal{J}}\right) = \Pr(\tau_{nj} \leq \tau | (A_{n-1j}, T_{n-1j}))$$

Assumption 2. *Let the random variable $A_{jt} \in \{0,1\}$ denote the ad state in channel $j \in \mathcal{J} \setminus 0$ in period t and the first and last periods of a program on channel j by $\underline{t}$ and $\bar{t}$ respectively. The sequence of A_{jt} is characterized by:*

$$\Pr\left(A_{j\underline{t}} = 1\right) = \Pr\left(A_{j\bar{t}} = 1\right) = 0$$

The transition probability process follows a logistic distribution and is independent across channels and states. I.e. the parameters shifting the hazard rates are channel-state specific while the covariates include the genre of the program, the percentage of the program and a duration indicator. Specifically, the hazard rate for channel j, at time t, with characteristics w_{jt}, from state a_{jt-1}, after τ_{jt} periods within the state is:

$$\lambda(a_{jt-1}, \tau_{jt}, w_{jt}) = \Pr\left(a_{jt} = 1 | w_{jt}, a_{jt-1}, \tau \geq \tau_{jt}\right) = \frac{\exp(w_{jt}' \gamma_{ja_{jt-1}} + \gamma_{ja_{jt-1}\tau_{jt}})}{1 + \exp(w_{jt}' \gamma_{ja_{jt-1}} + \gamma_{ja_{jt-1}\tau_{jt}})}$$

The relevant parameters are $\Gamma_{ja_{jt-1}} = (\gamma_{ja_{jt-1}}, \gamma_{ja_{jt-1}\tau})$ where the former is a vector of the effects of the relevant state variables in channel j in state a_{jt-1}; and $\gamma_{ja_{jt-1}\tau_{jt}}$ is shorthand for $\gamma_{ja_{jt-1}\tau} \cdot \mathbb{1}\{\tau_{jt} = \tau\}$. This framework allows me to generate hazard rates for all channels, states, times and durations. It is important to note that I cap the maximal possible duration within a state to $\bar{\tau}_a$, assuming that the hazard rates for later durations remain constant. This allows for a larger applicable state space while keeping the parametric underlays constrained.

[13] Note that this is true for any assumption regarding the viewers' expectations.

The data used for the empirical analysis details the minutes from 20-22, hence I do not observe the duration within a state at the beginning of prime-time. As such, the probability of being in any of the two states in the beginning of the evening follows a Markovian process without accounting for durations within a state. I proceed similarly as detailed above, where the instantaneous transition probabilities follow a logistic distribution, with channel specific parameters, γ_j^{init}:

$$\Pr(a_{j1} = 1|w_{j1}) = \frac{\exp(w'_{j1}\gamma_j^{init})}{1 + \exp(w'_{j1}\lambda_j^{init})}$$

The set of parameters guiding the objective transition probabilities are given by $\Gamma = (\gamma_{ja_{jt-1}}, \gamma_{ja_{jt-1}\tau}, \gamma_j^{init})$ for all $j \in \mathcal{J} \setminus 0$. Upon specifying the objective transition probabilities, I will characterize the viewers' information set at each period and how these evolve. Afterwards, I will characterize the way the objective probabilities and the individual information sets interact to generate individual expectations regarding ad probabilities.

Information set

A viewer's conditional commercial probability may depend on the information they acquired throughout their entire viewing history up to period t. With that, assumptions 1 and 2 restrict the state space of the viewers' information set. It is important to note that the state variables create individual knowledge above and beyond the common knowledge structure, given by the viewing guide knowledge and the objective transition probabilities. The additional information of viewer i pertaining to channel j is comprised of the time elapsed since last known state, m_{ijt}; the state of channel j in the last viewed time, $a_{ijt} \equiv a_{jt-m_{ijt}}$; and the amount of time viewer i has observed the last state of channel j, τ_{ijt}. Hence, the *information set* of viewer i on channel j at time t is defined by:

$$\mathcal{H}_{ijt} = \{\lambda_{jt}, m_{ijt}, a_{ijt}, \tau_{ijt}\}$$

where $\lambda_{jt} = \lambda(a_{jt-1}, \tau_{jt}, w_{jt})$ and viewer i's full information set is $(\mathcal{H}_{ijt})_{\forall j \in \mathcal{J} \setminus 0}$. The viewing guide knowledge implies that all program characteristics are known, whereas the factors affecting the objective transition probabilities are a subset of these characteristics, $w_{jt} \subset x_{jt}$. Since these are known, I simplify the definition of the information set to knowledge of the objective transition probabilities instead of knowledge of their underlying factors.

Times in which viewers have not watched a channel throughout a transition between programs, assumption 2 implies that although the viewer has not watched the channel, she posses full knowledge of the state implied by the initial and terminal conditions. Furthermore, the 'viewing guide' knowledge states that within a planned commercial break, the viewers possess full knowledge of the state within all the periods in which the commercial is broadcast.

Individual probabilities

Generating individual probabilities requires specifying the manner by which the individuals' information interacts with the objective transition probabilities. Viewers update their expectations according to a simple Bayesian procedure, by which they take into account all the possible paths leading to an ad in the upcoming period. This is simply illustrated using linear algebra.

To generate channel-time objective transition probability matrices that incorporate all possible states and durations, I stack the hazard rates for all possible states and durations of a channel-time pair into a $(\bar{\tau}_{prog} + \bar{\tau}_{comm}) \times (\bar{\tau}_{prog} + \bar{\tau}_{comm})$ matrix, Q_{jt}. The rows of Q_{jt} define the outgoing state-duration and the columns define the incoming state-duration. Equation 3.7 exemplifies the structure of the objective transition probability matrix of channel j at time t for $\bar{\tau}_{prog} = \bar{\tau}_{comm} = 4$. As noted above, the two last durations within each state are defined as including all further durations (4+). The probabilities within the matrix are the aforementioned hazard rates.

$$
Q_{jt} = \begin{array}{cc}
 & \begin{array}{cccc|cccc} \multicolumn{4}{c|}{\overbrace{\hphantom{xxxxxxxxxxxxxxxxxxxx}}^{\textit{Program}}} & \multicolumn{4}{c}{\overbrace{\hphantom{xxxxxxxxxxxxxxxx}}^{\textit{Commercial}}} \\ 1 & 2 & 3 & 4+ & 1 & 2 & 3 & 4+ \end{array} \\
\begin{array}{r} \textit{Program}\left\{\begin{array}{l}1\\2\\3\\4+\end{array}\right. \\ \textit{Commercial}\left\{\begin{array}{l}1\\2\\3\\4+\end{array}\right. \end{array} &
\left(\begin{array}{cccc|cccc}
0 & 1-\lambda_{jt0}(1) & 0 & 0 & \lambda_{jt0}(1) & 0 & 0 & 0 \\
0 & 0 & 1-\lambda_{jt0}(2) & 0 & \lambda_{jt0}(2) & 0 & 0 & 0 \\
0 & 0 & 0 & 1-\lambda_{jt0}(3) & \lambda_{jt0}(3) & 0 & 0 & 0 \\
0 & 0 & 0 & 1-\lambda_{jt0}(4) & \lambda_{jt0}(4) & 0 & 0 & 0 \\
\hline
1-\lambda_{jt1}(1) & 0 & 0 & 0 & 0 & \lambda_{jt1}(1) & 0 & 0 \\
1-\lambda_{jt1}(2) & 0 & 0 & 0 & 0 & 0 & \lambda_{jt1}(2) & 0 \\
1-\lambda_{jt1}(3) & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{jt1}(3) \\
1-\lambda_{jt1}(4) & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{jt1}(4)
\end{array}\right)
\end{array}
\tag{3.7}
$$

As opposed to transitions between periods, that are represented by semi-Markov hazard rates, the transition to any of the two states in the beginning of the evening is modeled through a Markov process represented by a 1×2 vector (from one initial unknown state to two potential states - program or commercial). Compatibility with later periods' state spaces requires making this transition vector comparable in size to the entire state space. To do so, I define equal transition probabilities to each of the $\bar{\tau}_{prog} + \bar{\tau}_{comm}$ durations within each state. Specifically, the initial transition probability matrix is represented in equation 3.8.

$$
Q_{j1} = \begin{array}{c}
\begin{array}{ccc|ccc} \multicolumn{3}{c|}{\overbrace{\hphantom{xxxxxxxxxxxx}}^{\textit{Program}}} & \multicolumn{3}{c}{\overbrace{\hphantom{xxxxxxxxxxxx}}^{\textit{Commercial}}} \\ 1 & \dots & \bar{\tau}_{prog} & 1 & \dots & \bar{\tau}_{comm} \end{array} \\
\left(\begin{array}{ccc|ccc} \frac{1-\lambda_{j1}}{\bar{\tau}_{prog}} & \dots & \frac{1-\lambda_{j1}}{\bar{\tau}_{prog}} & \frac{\lambda_{j1}}{\bar{\tau}_{comm}} & \dots & \frac{\lambda_{j1}}{\bar{\tau}_{comm}} \end{array}\right)
\end{array}
\tag{3.8}
$$

The first of the three state variables - amount of time since last viewed - defines the individual specific transition probability matrix. Matrix multiplication of the objective transition probability matrices creates probabilities to reach any state-duration from all initial state-durations via all possible routes. E.g. a viewer who has viewed a channel in the previous period has only the last period's transition probability matrix, allowing one route to observe an ad from the outgoing state. Alternatively, a viewer who hasn't viewed a channel for ℓ periods will face greater uncertainty, since the probability of finding a commercial after ℓ periods could result from several broadcasting routes. This uncertainty is captured by the more complex transition

probabilities encapsulated in the multiplied matrix. Formally, define individual i's transition probability matrix of channel j after not viewing the channel for m_{ijt} periods as:

$$Q_{ijt} = \Pi_{\ell=0}^{m_{ijt}} Q_{j(t-\ell)} \tag{3.9}$$

The two latter state variables, namely the commercial state in the last viewed time and the amount of time viewer i has observed the last state, determine the relevant distribution of states conditional on the viewing history. Particularly, relating the two state variables to the matrix structure outlined above and exemplified in figure 3.7, the last observed state and the amount of time observing the last state jointly determine the row within the matrix, $q_{ijt}(a_{ijt}, \tau_{ijt}, \cdot)$, where the row is given by $(a_{ijt}\bar{\tau}_{prog} + \tau_{ijt})$ and $\cdot$ implies all columns within the row. The corresponding commercial probability is given by the sum of the elements in columns $[\bar{\tau}_{prog} + 1, \bar{\tau}_{prog} + \bar{\tau}_{comm}]$ of the relevant row. Formally:

$$\mu_{ijt} = \sum_{\ell=\bar{\tau}_{prog}+1}^{\bar{\tau}_{prog}+\bar{\tau}_{comm}} q_{ijt}(a_{ijt}, \tau_{ijt}, \ell) \tag{3.10}$$

Alternatively, when a viewer does not know the duration within a state, τ_{ijt} includes all possible durations within the observed state. E.g. imagine a 'zapper', that is a viewer who has watched channel j for a while, left to an alternative option and returned to channel j afterwards. The period after returning to channel j, the viewer knows the state, but not the duration within the state. This viewer must therefore construct beliefs concerning the likelihood of each possible duration. Viewers update their expectations according to a Bayesian updating procedure, using previous period's probabilities (the columns of Q_{ijt-1}) as weights for the likelihood of being in any of the current period's durations (the rows of Q_{ijt}). In the example of the zapper above, the individual knows the current state while the duration is unknown, as such, the weights need to be inflated accordingly. Formally, define the weight given to duration τ as:

$$w_{ijt}(\tau) = \frac{q_{ijt-1}(a_{ijt-1}, \tau_{ijt-1}, a_{ijt} \cdot \bar{\tau}_{prog} + \tau)}{\sum_{\ell=1}^{a_{ijt}\cdot\bar{\tau}} q_{ijt-1}(a_{ijt-1}, \tau_{ijt-1}, a_{ijt} \cdot \bar{\tau}_{prog} + \ell)} \tag{3.11}$$

The corresponding commercial probability in this scenario is given by:

$$\mu_{ijt} = \sum_{\tau=a_{ijt}\cdot\bar{\tau}_{prog}+1}^{\bar{\tau}_{prog}+a_{ijt}\bar{\tau}_{comm}} \left[w_{ijt\tau} \cdot \sum_{k=\bar{\tau}_{prog}+1}^{\bar{\tau}_{prog}+\bar{\tau}_{comm}} q_{ijt}(\tau, k) \right] = \sum_{\tau=a_{ijt}\cdot\bar{\tau}_{prog}+1}^{\bar{\tau}_{prog}+a_{ijt}\bar{\tau}_{comm}} w_{ijt\tau} \cdot \mu_{ijt}(\tau) \tag{3.12}$$

where $q_{ijt}(\tau, k)$ is shorthand for $q_{ijt}(a_{ijt}, \tau_{ijt}, k)$.

The learning process

The commercial probability function maps the individuals' knowledge derived from past viewing choices, unto a commercial probability. Whenever a viewer has been constantly viewing a channel and therefore pertains full knowledge regarding the current state, the commercial probability distribution is degenerate and the individual expectation is aligned with the objective probability. The learning process is pronounced whenever this is not the case. At the extreme, let's return to the example of the zapper, who watches channel j, leaves for a while, returns and leaves again. In this scenario, when the viewer contemplates returning to channel j, he knows the commercial state of channel j when he left but not the duration. For illustrative purposes, let the last known state be a program. For this viewer, the commercial probability

distribution is given by all elements within the upper-right quadrant of his commercial probability matrix, Q_{ijt}. Upon returning to channel j and learning the current state, the commercial probability distribution becomes more confined but still a non-degenerate distribution spanning multiple potential durations. As the viewer continuously watches channel j, the distribution becomes more refined, since with the passing of each period, another duration can be ruled out. After viewing for (at most) $\bar{\tau} - 1$ periods, the probability distribution again becomes degenerate and is summarized by a single element.

Upon specifying the individual ad expectations the model is fully specified. Using the ex-ante utility of viewer i of watching channel j outlined in equation 3.2, together with the commercial probability derived above, we can rewrite the viewers' ex-ante utility as:

$$U_{ijt} = \alpha\mu_{ijt}(\mathcal{H}_{ijt}; \Gamma) + \delta_{jt}(x_{jt}, \zeta_{jt}; \beta) + \varepsilon_{ijt} \quad \forall \;\; i \in \mathcal{I}, \; j \in \mathcal{J}, \; t = 1, ..., T \tag{3.13}$$

where $\delta_{jt} = x'_{jt}\beta + \zeta_{jt}$ and the parameters of the model are (Γ, θ): $\Gamma = (\gamma, \gamma_\tau, \gamma^{init})$ govern the commercial probability formation and $\theta = (\alpha, \beta)$ are the utility parameters.

4 Viewership estimation

Estimation of the parameters follows a two-step procedure: in the first step the parameters Γ are estimated from the channels' broadcasting behavior; in the second step, the probability formation parameters are taken as given and the utility parameters, θ, are estimated using a method of simulated moments using the unobserved characteristics as the basis of the moments as proposed in Berry (1994); Berry et al. (1995).

Commercial probabilities

Underlying the viewers' individual expectations are the objective hazard rates that follow a logistic distribution and are derived from the observed programming decisions of the channels. In the empirical implementation, the durations were capped at 23 minutes for the program state[14] and the durations for the commercial state were capped at 7 minutes. Detailed results of the parameters can be found in appendix A.

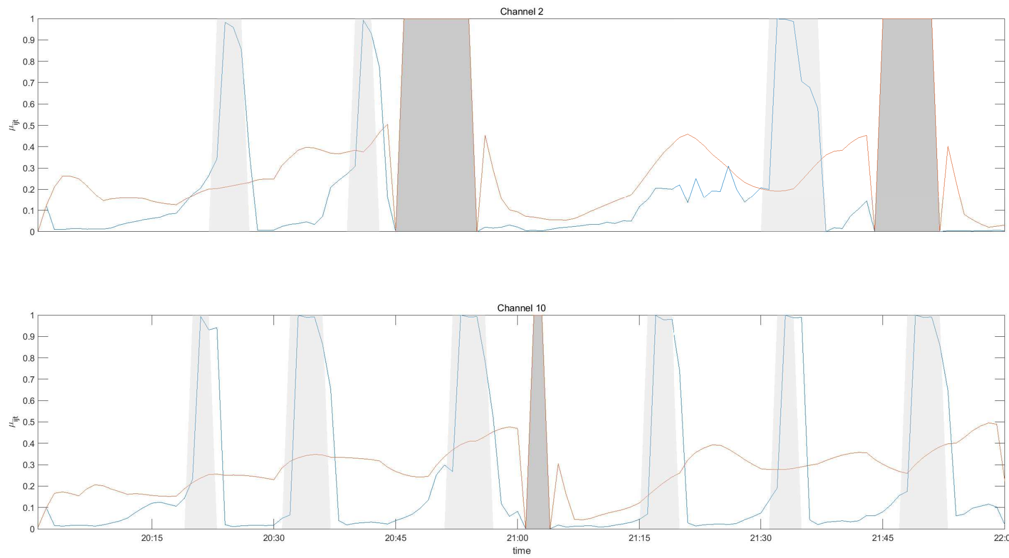
Figure 5 presents the way by which the estimated commercial probabilities are translated into varying beliefs. The figure compares the predicted commercial probabilities between two extremes - an individual who views the channel continuously and an individual who never views the channel. All commercial probabilities will lie in between the two extremes. The difference between the two graphs represent the informational advantage resulting from past viewing. In the figure, the probabilities are similar at the start of the evening resulting from the initial Markovian probabilities. As time progresses, the informational gap between the individuals widens. Moreover, as clarified in the model, there is no informational gap between the individuals during planned transitions, that is, at the first and last periods of a program and during inter-program commercial breaks. Comparison of the figures highlights the large informational advantage resulting from viewing a channel regarding the correct perception of commercial probabilities.

Viewing choice model

The utility parameters are estimated via the method of simulated moments where the basis of the moments is the structural term (ζ_{jt}) as proposed by Berry (1994). While the viewing choice model has no random

[14] After which the hazard rates are assumed to remain constant.

Figure 5: Informational advantage of persistent viewing



Notes: The figure displays the perceived commercial probabilities at each minute of the prime-time on 28$^{\text{th}}$ April, 2005. The horizontal axis displays the minutes within the day and the vertical axis the perceived probability of a commercial. The shaded areas are commercial breaks and the dark shading confers to a planned commercial break (between programs).

coefficients, the individual expectations form a viewer specific component ($\alpha\mu_{ijt}$) while the rest of the utility components are common to all viewers (δ_{jt}). Integrating over the individual expectations and isolating the structural error term is done using a contraction mapping proposed by Berry et al. (1995) (henceforth **BLP**), while accounting for the dynamic aspect in the choices.

The individuals' expectations form a dynamic component in the viewers' choices, requiring the model be solved via forward simulation. That is, for a fixed $\hat{\Gamma}$ and a generic parameter vector θ, we begin by solving the utilities for the first period. Lacking viewing histories results in a commercial probability that equates across all simulated individuals (and equals the objective probabilities). Hence, the initial choices are purely static in the sense that all heterogeneity across individuals is driven by the taste shocks, $\varepsilon_{\cdot j1}$. Upon generating the first period's choices, individuals begin to gather a viewing history, thereby creating differing information sets. The utilities at $t = 1$ and $t > 1$ take the form:

$$\begin{aligned} U_{ij1} &= \alpha\hat{\lambda}_{j1} + \delta_{j1} + \varepsilon_{ij1} \quad \forall\ j \in \mathcal{J},\ i \in \mathcal{I} \\ U_{ijt} &= \alpha\hat{\mu}_{ijt} + \delta_{jt} + \varepsilon_{ijt} \quad \forall\ j \in \mathcal{J},\ i \in \mathcal{I},\ t = 2, ..., T \end{aligned} \tag{4.1}$$

where $\hat{\lambda}_{jt} = \lambda_{jt}(\hat{\Gamma})$ and $\hat{\mu}_{ijt} = \mu_{ijt}(\mathcal{H}_{ijt}; \hat{\Gamma})$. The viewing choice of simulated viewer i in period t is determined by:

$$y_{it} = \underset{j \in \mathcal{J}}{\arg\max}\{U_{i0t}, ..., U_{iJt}\} \tag{4.2}$$

The informational homogeneity across viewers in the first period makes these utilities substantially different from those in later periods. Consequently, these observations were not included in the estimation, only in constructing the viewing choices for future periods.

While the choice of each simulated viewer is given by 4.2, their choice probabilities for each alternative follow a logistic distribution, resulting from the type 1 extreme value distribution of the taste shocks. Accordingly, the channels' simulated market shares in each period are given by integrating over the choice probabilities of the ns simulated viewers[15]:

$$\begin{aligned} \mathcal{s}_{ijt}(\delta_{\cdot t}, \hat{\mu}_{i \cdot t}; \alpha) &= \frac{\exp(\delta_{jt} + \alpha \hat{\mu}_{ijt})}{\sum_{k \in \mathcal{J}} \exp(\delta_{kt} + \alpha \hat{\mu}_{ikt})} \\ \mathcal{s}_{jt}(\delta_{\cdot t}) &= \frac{1}{ns} \sum_{i=1}^{N} \mathcal{s}_{ijt}(\delta_{\cdot t}, \hat{\mu}_{i \cdot t}; \alpha) \end{aligned} \tag{4.3}$$

The contraction mapping proposed by BLP allows us to isolate the ζ_{jt} by solving the system of equations $\mathcal{s}_{\cdot t}(\delta_{\cdot t}) = S_{\cdot t} \ \forall \ t = 2, ..., T$, where $S_{\cdot t}$ is a vector of the observed market share of all channels at time t. Lacking a closed form representation of the structural error term, the contraction mapping requires iterating over the series:

$$\delta_{\cdot t}^{\ell+1} = \delta_{\cdot t}^{\ell} + \ln(S_{\cdot t}) - \ln\left[\mathcal{s}_{\cdot t}(\delta_{\cdot t}^{\ell})\right] \quad \forall \ t = 2, ..., T, \ \ell = 0, ..., L \tag{4.4}$$

Until the norm between two successive iterations is smaller than a pre-defined tolerance level, $||\delta_{\cdot t}^{L} - \delta_{\cdot t}^{L-1}|| \leq \epsilon$. Dubé et al. (2012) emphasize the importance of a tight tolerance level as a stopping criterion for the contraction mapping to prevent propagation of simulation errors to the estimation[16]. Finally, the structural error term is defined as:

$$\zeta_{jt}(\delta_{jt}, x_{jt}, S_{\cdot t}; \beta) = \delta_{jt}(S_{\cdot t}) - x_{jt}'\beta \tag{4.5}$$

The structural shocks were stacked to a $D \cdot (T-1) \cdot J \times 1$ vector ζ[17]. The identification assumption is that the structural error term is independent of a set of exogenous variables (Z). Formally, define the set of moments corresponding to the orthogonality of the structural error term and the instruments as $m(Z, \zeta; \theta) = E[Z'\zeta(\theta)]$. The identification assumption is that $m(Z, \zeta; \theta^0) = 0$ at the true parameter vector, θ^0. The resulting MSM estimator is the solution to:

$$\hat{\theta} = \arg\min_{\theta} \bar{m}(Z, \zeta; \theta)' W \bar{m}(Z, \zeta; \theta)$$

where $\bar{m}(\cdot)$ is the sample counterpart of $m(\cdot)$. The weighting matrix was computed in a two stage procedure, the initial weighting matrix is the identity matrix, upon generating initial estimates for the model, the updated weighting matrix is given by $W = \left(N^{-1}\bar{m}\bar{m}'\right)^{-1}$ where $N = D \cdot (T-1) \cdot J$. Standard errors were computed with the standard methods shown in Newey and McFadden (1994).

Instruments

The commercial probability may be endogenous due to the correlation between the commercial status and unobserved program quality captured in ζ_{jt}. The claim regarding simultaneity bias resulting from unobserved quality is similar to that in other works, commencing with Trajtenberg (1989). Specifically, higher quality

[15] The initial stage in the estimation procedure requires simulation of $ns \times D \cdot T$ taste shocks ε_{ijt}. This was done using Halton draws to reduce simulation error for $ns = 1000$ households as explained by Train (2009).

[16] In the presence of dynamic choices, there is another role for the tolerance level, preventing propagation simulation errors to later periods' mean valuation. If the tolerance level is not sufficiently tight, a minuscule decrease in the tolerance level may result in a discrete change in choices as opposed to a smooth change in the choice probabilities. While this change may have a sufficiently small effect on the choice probabilities, the effect of imprecision on choices and therefore on future mean valuations may be large. The importance of a tight tolerance level is therefore magnified in the presence of dynamic choices. The dynamic simulation error is also mitigated through a sufficiently large simulated sample ensuring that the simulated choices will also be sufficiently responsive to changes in the mean value.

[17] The $T - 1$ results from the exclusion of the first time period within each day.

programs may attract more viewers which in turn incentivize the channel to broadcast more ads. I overcome this concern using the high-frequency nature of the data: Within high frequency data the ad status is an indicator as opposed to a continuous value in lower frequency data. Since ads are broadcasted on all types of shows, not only on high quality ones, the correlation between ad status and show popularity diminishes. Simultaneously, from the demand perspective, viewers may be more averse to missing out on the program within high quality programs, making them less sensitive to ad broadcast. This concern is mitigated by incorporating program dummies. The differential effect between programs, which is captured in the program dummies, will capture this unobserved quality. Furthermore, I include regulation based variables[18]. The regulation based variables incorporate exogenous variation shifting the amount of commercials. Specifically, as shown in table 1, the regulation affects the overall amount of commercials as well as the amount of commercials within a break. The exogenous variation across genres and across times of the day create exogenous variation affecting the commercial placement behavior and will help identify the commercial (dis)utility parameter. Finally, I include Berry et al. (1995) instruments[19].

5 Channels' supply & equilibrium

Setup

Assume that a subset $\mathcal{J}_c \subset \mathcal{J}$ of channels are strategic in their commercial placement behavior, while the rest place advertisements only between programs. In the empirical implementation, channels 2 and 10 are strategic while channel 1 isn't. Channel j's strategy vector is the couple $\mathcal{S}_j = (\rho_j, \phi_j) \in \mathbb{R}^2$ where the former denotes the change in the baseline commercial probability and the latter the change in the commercial persistence. The channel's strategy vector and objective ad probabilities map to updated objective transition probabilities, $\lambda_{jt}(\rho_j, \phi_j; \tau) : \ \mathcal{S}_j \times \lambda_{jt}(\tau) \to [0,1]$. Specifically, the objective transition probabilities are constructed for each commercial channel, at each minute, pertaining to all possible durations, $\lambda_{jta}(\tau) \ \forall \ a \in \{0,1\}, \ j \in \mathcal{J}_c, t = 1, ..., T, \tau$. The mapping updates the probabilities according to:

$$\lambda(a_{jt-1}, \tau_{jt}, w_{jt}, \phi_j, \rho_j) = \frac{\exp(w'_{jt}\gamma_{ja_{jt-1}} + \gamma_{ja_{jt-1}\tau_{jt}} + (1-a_{jt-1}) \cdot \phi_j + a_{jt-1} \cdot \rho_j)}{1 + \exp(w'_{jt}\gamma_{ja_{jt-1}} + \gamma_{ja_{jt-1}\tau_{jt}} + (1-a_{jt-1}) \cdot \phi_j + a_{jt-1} \cdot \rho_j)}$$

While the probabilities corresponding to all potential durations enter the objective transition probability matrix on which the viewers' base their decisions, $Q_{jt}(\phi_j, \rho_j)$, the probability relating to the true state (a_{jt-1}, τ_{jt}) determines the ad probability in each minute, $\sigma_{jt}(\rho_j, \phi_j)$. The rational expectations assumption results in that the viewers' beliefs correspond to the channels' programming behavior. Changes to the channels' strategies will affect the viewers' beliefs and switching behavior accordingly. The equilibrium condition I impose is that the viewers' expectations are aligned with the channels' strategies. Finally, I assume the channels choose their commercials at the beginning of the day implying a static equilibrium. This is formally stated below:

Definition 1 (Rational expectations static equilibrium)**.** *An equilibrium is defined by a set of strategies* $(\rho_j, \phi_j) \ \forall \ j \in \mathcal{J}_c$ *such that:*

[18]Specifically, the variables which were used are: (a) the remaining amount of commercials left for each channel to broadcast within the hour; and (a) the remaining amount of commercials left for each channel to broadcast within each commercial break.

[19]The variables are: (a) the sum of the commercial state on the other channels at each time; (a) the sum of the percentage of the program on the other channels at each time; and (a) the number of other channels broadcasting the same genre as the pivotal channel at each time.

1. *The strategies are best responses to each other;*

2. *Viewers' beliefs are consistent with the channels' resulting ad probabilities,* $\lambda_{jt}(\rho_j^*, \phi_j^*, \tau) = \sigma_{jt}(\rho_j^*, \phi_j^*; \tau)\ \forall\ t, j, \tau$.

Payoffs

Channel j's payoff from advertising an ad at minute t is given by:

$$\pi_{jt}(x_{\cdot t}, \lambda^t) = p_j \cdot \boldsymbol{\mathscr{s}}_{jt}(x_{\cdot t}, \lambda^t; \theta) \tag{5.1}$$

Equation 5.1 specifies channel j's revenue according to the per viewer price of an ad (p_j) and the channel's expected market share ($\boldsymbol{\mathscr{s}}_{jt}$). Channel j's expected viewer share is determined by the characteristics of all channels at the specific time ($x_{\cdot t} = (x_{jt})_{\forall j \in \mathcal{J}}$), the viewers' objective commercial probability expectation ($\lambda^t = (\lambda_{jm})_{\forall j \in \mathcal{J}_c, m=1,\ldots,t}$) and preferences ($\theta$). The per viewer ad price is taken as exogenous. Conditional on the viewer side beliefs and advertising strategies, channel j's expected daily payoff is given by:

$$\tilde{\Pi}_j(\phi, \rho, \lambda, x) = \sum_{t=1}^{T} \sigma_{jt}(\rho_j, \phi_j) \cdot \pi_{jt}(x_{\cdot t}, \lambda^t) \tag{5.2}$$

Equilibrium strategies

The channels' strategies maximize their expected payoff subject to the regulation constraint and the competing channel's equilibrium ad probabilities, $\sigma_k(\rho_k^*, \phi_k^*)$. Channel j's maximization problem is given by:

$$\begin{aligned} (\rho_j^*, \phi_j^*) = & \arg\max_{\rho_j, \phi_j} \Pi_j(\phi_j, \rho_j, \sigma_k^*, \lambda, x) \\ s.t. \quad & \sum_{t=1}^{T} \sigma_{jt}(\rho_j, \phi_j) \leq \kappa \ (\textit{regulation constraint}) \end{aligned} \tag{5.3}$$

where κ is the amount of time permitted for advertisements by the regulator. It is important to note that the, in this setting, regulation constraint restricts the expected number of minutes devoted to commercials. Furthermore, whenever the regulation constraint is binding, the constraint implicitly determines the the value of one of the strategy components subject to the other. Explicitly, define the expected number of ads within a day as:

$$G(\rho_j, \phi_j; \lambda) = \sum_{t=1}^{T} \sigma_{jt}(\rho_j, \phi_j) \tag{5.4}$$

Any of the two strategic components can therefore be isolated and denoted by:

$$\phi_j = G^{-1}(\kappa, \rho_j; \lambda) \tag{5.5}$$

Incorporating equation 5.5 in the channel's maximization problem 5.3, yields a straightforward one variable maximization problem:

$$\rho_j^* = \arg\max_{\rho_j} \Pi_j(\rho_j, \sigma_k^*, \lambda, x; \kappa) \tag{5.6}$$

The first order condition of channel j for the static equilibrium strategy implicitly defining the equilibrium strategy is given by:

$$\rho_j^* : \quad \sum_{t=1}^{T} \frac{\partial \sigma_{jt}}{\partial \rho_j} \left[\boldsymbol{s}_{jt} + \sum_{\tau=t}^{T} \sigma_{j\tau} \cdot \frac{\partial \boldsymbol{s}_{j\tau}}{\partial \sigma_{jt}} \cdot \frac{\partial \sigma_{jt}}{\partial \rho_j} \right] = 0 \tag{5.7}$$

Using the implicit function theorem we can analyze the strategic interaction between the commercial channels.

Equilibrium simulation

Changes to the channels' choice variables will affect the amount of commercials through their underlying ad probabilities. With that, while the strategies determine the ad probabilities, the payoffs of the channels are determined by their realizations. Each realization of ad probabilities will affect the viewing behavior of the households along with the ad prices, resulting in different payoffs to the channels. To resolve this tension, I use the implied ad probabilities from a strategy vector to simulate a daily ad realization. Once a realization is determined, so are the expected viewer shares and ad prices and hence, the respective payoff. This is done numerous times where the averaged payoff is an approximation to the expected payoff resulting from the strategy vector within a specific day.

More specifically, a strategy vector implies a Bernoulli probability distribution for an ad at each time, conditional on the previous periods' realizations (through the duration dependence). For a sufficient number of simulations, the empirical propensity of ads will equal the ad probability. Equivalently, for this sufficient number of simulations, the averaged payoff will be a close approximation to the expected payoff resulting from the strategy vector. The simulation process follows a grid search over the quadruple strategy elements $\mathcal{S} = (\rho_j, \phi_j)_{j \in \mathcal{J}_c}$ according to the process formulated below. First, notice that the strategies are bounded by the upper bound of the smallest probability and the lower bound implied by the highest probability. I generate an evenly spaced 4-dimensional grid within the bounds and for each point simulate the best response of each channel with respect to the grid points of the competing channel. This generates focal points of the channels' empirical best response functions. I interpolate the empirical best response function between grid points using a cubic spline. An equilibrium is defined by intersections of the best responses of the two commercial channels' strategies.

Equilibrium simulation procedure

Algorithm 1: Equilibrium simulation

input : Strategy grid $\mathcal{S}_G = \{\mathcal{S}^1, ..., \mathcal{S}^N\}$, programming characteristics (x, ζ), baseline ad probabilities λ and demand parameters θ

output: Set of equilibrium strategies $\mathcal{S}^* = (\rho^*, \phi^*)$

for i=1,...,N **do**

  Ad probabilities are given by: $\sigma^i \equiv \sigma(\mathcal{S}^i) = \big(\sigma_j(\rho^i_j, \phi^i_j), \sigma_k(\rho^i_k, \phi^i_k)\big)$;

  Viewers' have rational beliefs: $\lambda^i = \sigma^i$;

  for s=1,2,... **do**

    Simulate daily ad distributions according to ad probabilities: $\hat{A}^{is} \equiv \hat{A}^s(\sigma^i) = \Big(\hat{A}^s_j(\sigma^i_j), \hat{A}^s_k(\sigma^i_k)\Big)$;

    Simulate viewer shares: $\hat{\jmath}^{is} \equiv \hat{\jmath}(x, \zeta, \hat{A}^{is}, \lambda^i; \theta_1)$;

    Simulate ad prices: $\hat{P}^{is} \equiv \hat{P}(\hat{A}^{is}, \hat{\jmath}^{is}; \theta_2)$;

    Calculate simulated daily profit: $\hat{\Pi}^{is}_j \equiv \hat{\Pi}_j(\hat{A}^{is}, \hat{\jmath}^{is}, \hat{P}^{is}) = \sum_{t=1}^{T} \hat{a}^{is}_{jt} \cdot \hat{P}^{is}_{jt} \cdot \hat{\jmath}^{is}_{jt}$;

    Calculate simulated ad probabilities $\hat{\sigma}^{is}$: $\hat{\sigma}^{is}_{jt} = s^{-1} \sum_{\ell=1}^{s} a^{\ell}_{jt}$;

    Stop when simulated ad probabilities are sufficiently close to implied ad probabilities: $|\hat{\sigma}^{is} - \sigma^i| < \epsilon$;

  end

  Calculate average simulated payoff $\bar{\Pi}^i = (\bar{\Pi}^i_j, \bar{\Pi}^i_k)$: $\bar{\Pi}^i_j \equiv \bar{\Pi}_j(\mathcal{S}^i) = S^{-1} \sum_{i=1}^{S} \hat{\Pi}^{is}_j$;

end

Expand strategy-payoff grid using cubic spline: $(\mathcal{S}, \bar{\Pi})_G \to (\mathcal{S}, \bar{\Pi})_F$;

Find best response correspondence for each point along F grid: $\mathcal{S}^{BR}_j(\mathcal{S}_k) = \arg\max_{\mathcal{S}_j} \bar{\Pi}_j(\mathcal{S}_j, \mathcal{S}_k) \; \forall \; j \in \mathcal{J}_c$;

for $i \in$ *BR-grid* **do**

  if $\mathcal{S}^i_j(\mathcal{S}^i_k) = \mathcal{S}^i_k(\mathcal{S}^i_j)^{-1}$ **then**

    $\mathcal{S}^i \in \mathcal{S}^*$ is an equilibrium point;

  end

end

6 Policy analysis

Viewers' uncertainty pertaining to the timing and length of commercial breaks generates a friction that diminishes switching behavior and increases the platforms' market power. Regulators have approached this market failure by restricting the channels' permitted amount of commercials. An alternative approach would relax the friction underlying the market failure and allowing the viewers' switching behavior to act as the disciplining power of the market. It is not clear a-priori the effects of the regulation, on the one hand whether the quantity restriction is necessary, and on the other, whether information provision would sufficiently decrease the market power documented via the small semi-elasticities (in absolute terms). The agglomeration of total commercial timing around the 24 minute permitted amount in figure 2 implies that in the current informational environment, the regulation curbs the platforms' equilibrium strategies. In this section I try to answer two questions: (a) *to what extent does the current quantity restriction curb the*

platforms' advertising strategies? and (b) *would alleviation of the informational friction sufficiently contain the advertising strategies of the platforms?*

I will use the results from the viewer demand estimation together with the platforms' optimal response equation and equilibrium concept to analyze the distribution of commercial strategies under two policy interventions. In the first, the quantity restriction is relaxed, allowing the two commercial channels to choose their commercial strategies according to their optimal allocation. In this experiment, the viewers' informational structure remains unchanged, namely their underlying perceptions, driving their viewing choices, are consistent with the channels' strategies. Likewise, I simulate the equilibrium amount of commercials without quantity restrictions, but with supplying the viewers with information regarding the duration of commercial breaks. This policy experiment can be thought of as providing viewers with a timer on the screen, informing them how long until the end of the commercial break. As a result, the channels will still use mixed strategies between days but the viewers' make their viewing decision during a commercial break based on the true commercial state as opposed to the channels' strategies. A caveat regarding these policy analyses is that the advertising firms' demand is not explicitly modeled and hence, not incorporated in the platforms' strategies. The two sided market effect, if binding, would act as another force curbing the channels' commercial strategies. As such, the results from these experiments should be seen as an upper bound on the channels' behavioral responses.

To conduct these policy experiments I simulate the alternative distribution of commercial strategies from the equilibrium condition laid out in equation **??**. Formally, I simulate $\sigma_d^i = \Delta_d^{-1} \boldsymbol{s}_d^i$ for $i \in \{NR, IP\}$ for all days, where *NR* stands for *No-Regulation* and *IP* for *Information Provision*. In the information provision scenario I simulate the market shares conditional on the observed programming behavior of the channels during commercial broadcast, while retaining uncertainty pertaining to when the commercial break will commence. In the case of no regulation, to allow the strategy vector underlying the market shares to condition on the relevant state variables, i.e. the channel-day-time characteristics, I derived the vector of strategies from a channel specific logit model with time dummies and controls for percentage of the program and weekday dummies. Furthermore, a similar smoothing was implemented on the channel-day-minute transition matrix, Δ_d. Finally, these experiments were conducted while keeping the advertising strategies of channel 1 unchanged.

The distribution of commercial strategies and market shares with regulation restrictions, without them and with information provision are presented in figure 6 and summary statistics in table 3. Several interesting insights can be drawn from these results. First is that commercial strategies substantially increase without regulation, increasing by 23% and 67% on average on channels 2 and 10, respectively. More interestingly, although information provision restrains the increase in the commercial strategies, the increase remains significant, by 16% and 56% on average. Notice that the median percentage increase is noticeable lower, 4% and 54%. Comparing the distribution of strategies with regulation and without on figure 6a, we see that the new equilibrium distribution is not characterized by stochastic dominance, there are cases in which the probabilities decrease. More specifically, the distribution of commercial strategies becomes smoother in that there is less agglomeration at the extremes and most strategies are non-degenerate. Comparing the strategy distribution between the case in which all regulation is relaxed and the case with information provision, we notice that the two distributions are very similar, with that, absent regulation the equilibrium amount of commercials is higher than with information provision on average by 6% and 7% on channels 2 and 10 respectively. As such, information provision curbs the amount of commercials allocated by the channels, but not significantly. This result is in line with the elasticities documented in section 4, by which TV viewers

Table 3: Policy experiment 1 strategy summary statistics

	Strategies					**Market shares**				
Channel	Mean	S.D.	Median	25^{th} ptile	75^{th} ptile	Mean	S.D.	Median	25^{th} ptile	75^{th} ptile
(a) Current										
2	0.238	0.140	0.247	0.139	0.321	0.032	0.009	0.033	0.028	0.038
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
10	0.208	0.112	0.201	0.134	0.275	0.007	0.003	0.007	0.006	0.009
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
(b) No regulation										
2	0.291	0.136	0.275	0.200	0.370	0.031	0.007	0.031	0.028	0.035
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
10	0.347	0.153	0.336	0.248	0.426	0.007	0.002	0.007	0.005	0.008
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
(c) Information provision										
2	0.274	0.145	0.256	0.177	0.358	0.032	0.007	0.032	0.028	0.036
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
10	0.324	0.162	0.311	0.217	0.401	0.007	0.002	0.007	0.006	0.008
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: The table details summary statistics of the results from simulations of two policy experiments. The statistics in panel (a) pertain to the predicted strategies and market shares from the current regulation scheme. Standard errors were computed using a bootstrap procedure with 2,000 repetitions. All reported statistics are significant at the 1% level.

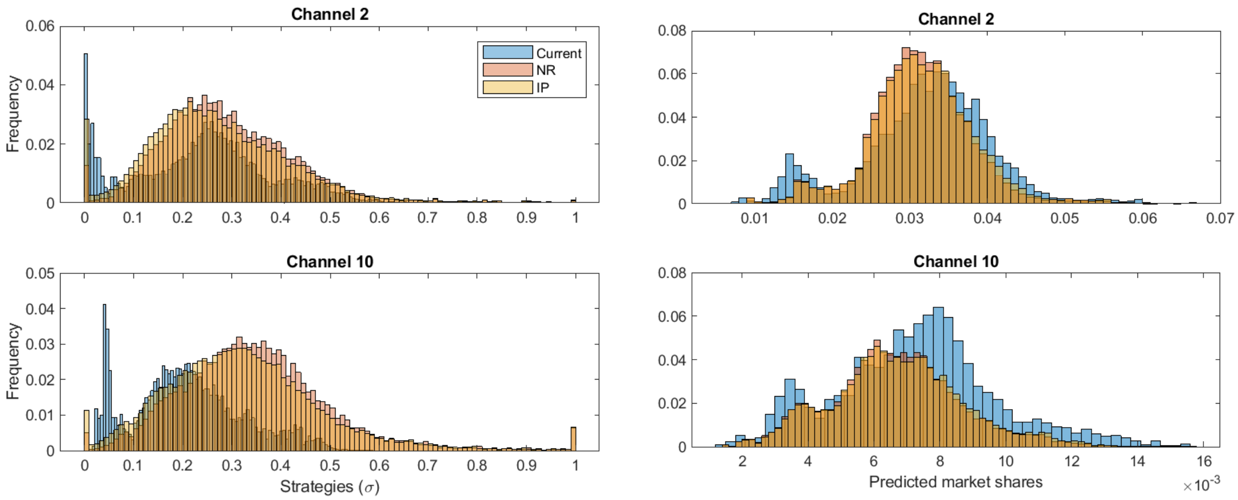
exhibit very inelastic demand. This inelasticity also results in the small changes to the market shares of the channels after the regulation changes and the new equilibria commercial strategies. The market shares exhibit a slight decrease.

The results from the above experiments illuminate the effects of two different policy interventions. In both cases regulation plays an important role in curbing the advertising strategies of the channels. As these results together with the semi-elasticity estimates imply, television viewership demand in Israel during the years 2002-2005 was inelastic which alloted a substantial degree of market power to the channels. Abstaining from quantity restrictions would substantially increase the amount of commercials on most days. Furthermore, the inelastic demand also results in that information provision would curb the channel strategies only partially, still resulting in an increase in the commercial strategies, albeit a more subtle increase. These results provide scope for interesting policy interventions which constitute a welfare improvement. The fact that information provision doesn't lead the viewers to leave én-mass allows for information provision accompanied by a ease of the quantity restriction. Information provision allows the most sensitive viewers, who are most aversely affected by ads to avoid them more easily without early or late return. Simultaneously, a partial relaxation of the quantity restriction would suffice for the channels not to incur a loss form the new policy.

Figure 6: Effects of regulation

(a) Distribution of commercial strategies

(b) Distribution of market shares



Note: The figure displays the distributions of commercial strategies and market shares in equilibrium with quantity restrictions and in their absence. The strategies depicted in figure 6a are minute level strategies conditional on the state variables: genre, % program and date & time controls. The market shares in figure 6b are derived from the Logit model outlined in equation **??** based on the two vectors of commercial strategies: with regulation and without.

7 Conclusion

This paper emphasized the role of informational frictions within the households' decision process in television markets. Based on observed patterns in viewing data from Israel, the paper developed and estimated a model incorporating the interaction between beliefs and choices. I was able to overcome the main difficulty within demand estimation, that is the endogeneity of the commercial status, using a regulatory change as well as the high frequency nature of the data, which in turn yielded consistent and reliable estimates. The informational frictions in the viewer decision process were shown to substantially impact the parameter estimates. Ignoring these frictions attenuated the sensitivity to commercials in a magnitude of more than a factor of 3. These estimates contribute to the literature on demand estimation in media markets by using high-frequency data to more finely estimate the viewers' sensitivity to commercials. Moreover, the proposed method creates a framework for demand estimation in the presence of informational frictions within a learning environment.

Together with a formulation of the channels' supply behavior and an equilibrium concept, this paper provided a framework to conduct an empirical regulation analysis. This framework was used to simulate the equilibrium commercial strategies under two hypothetical regulation schemes: absent quantity restrictions while keeping the informational framework fixed and relaxing the quantity restrictions while providing viewers information pertaining to the duration of commercial breaks. The model predictions show that the regulation substantially suppresses the equilibrium amount of commercials broadcast by the channels. Absent the restriction, the amount of commercials would increase on average (median) by 23% (11%) and 67% (67%) on each of the two commercial channels. Information provision restrained the increase leading to an average (median) increase of 16% (4%) and 56% (54%) respectively. These results are in line with the estimated semi-elasticities that showed the television viewer demand in Israel in between 2002 exhibited inelasticity. These results provide a basis for a welfare improving policy intervention by which information regarding the duration of commercial breaks is provided to viewers and the quantity restrictions are partially eased.

The changing environment in content provision, makes this an opportune time for a policy assessment. The increasing similarity between traditional content providers and online ones, together with the difference in the regulation they face, requires a re-assessment of the current regulatory policy. This work took a first step in conducting a data-driven policy evaluation aimed at providing policy makers and regulators information required to advance alternative policy regulations. Further work is required to refine the predicted effects of regulation changes. A main drawback of the results in this paper is the lack of heterogeneity in the demand specification. TV viewers are constructed of various groups within the population, with differing preferences as well as different advertising value. Accounting for these differences across the population would allow to refine the results and provide additional insight into optimal regulation.

References

Anderson, S. P. and S. Coate (2005). Market provision of broadcasting: A welfare analysis. *The review of Economic studies 72*(4), 947–972.

Baker, M. and L. M. George (2015). The news hour: Welfare estimation in the market for local tv news.

Berry, S., A. Eizenberg, and J. Waldfogel (2016a). Fixed costs and the product market treatment of preference minorities. *The Journal of Industrial Economics 64*(3), 466–493.

Berry, S., A. Eizenberg, and J. Waldfogel (2016b). Optimal product variety in radio markets. *The RAND Journal of Economics 47*(3), 463–497.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841–890.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 242–262.

Berry, S. T. and J. Waldfogel (1999). Public radio in the united states: does it correct market failure or cannibalize commercial stations? *Journal of Public Economics 71*(2), 189–211.

Berry, S. T. and J. Waldfogel (2001). Do mergers increase product variety? evidence from radio broadcasting. *The Quarterly Journal of Economics 116*(3), 1009–1025.

Crawford, G. S. and A. Yurukoglu (2012). The welfare effects of bundling in multichannel television markets. *American Economic Review 102*(2), 643–85.

Dubé, J.-P., J. T. Fox, and C.-L. Su (2012). Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation. *Econometrica 80*(5), 2231–2267.

Dubois, P., L. Lasio, et al. (2018). Identifying industry margins with price constraints: Structural estimation on pharmaceuticals. *American Economic Review 108*(12), 3685–3724.

Dukes, A. and E. Gal-Or (2003). Negotiations and exclusivity contracts for advertising. *Marketing Science 22*(2), 222–245.

Epstein, G. S. (1998). Network competition and the timing of commercials. *Management Science 44*(3), 370–387.

Gabszewicz, J. J., D. Laussel, and N. Sonnac (2004). Programming and advertising competition in the broadcasting industry. *Journal of Economics & Management Strategy 13*(4), 657–669.

Gabszewicz, J. J., D. Laussel, N. Sonnac, et al. (2000). *TV-Broadcasting competition and advertising*. Center for Operations Research & Econometrics. Université catholique de Louvain.

Gal-Or, E. and A. Dukes (2003). Minimum differentiation in commercial media markets. *Journal of Economics & Management Strategy 12*(3), 291–325.

Goettler, R. L. and R. Shachar (2001). Spatial competition in the network television industry. *RAND Journal of Economics 32*(4), 624–656.

Newey, W. K. and D. McFadden (1994). Large sample estimation and hypothesis testing. *Handbook of econometrics 4*, 2111–2245.

Nielsen (2017). *The Nielsen total audience report Q2 2017.* The Nielsen total audience series.

Shcherbakov, O. (2016). Measuring consumer switching costs in the television industry. *The RAND Journal of Economics 47*(2), 366–393.

Sweeting, A. (2006). Coordination, differentiation, and the timing of radio commercials. *Journal of Economics & Management Strategy 15*(4), 909–942.

Sweeting, A. (2009). The strategic timing incentives of commercial radio stations: An empirical analysis using multiple equilibria. *The RAND Journal of Economics 40*(4), 710–742.

Sweeting, A. (2010). The effects of mergers on product positioning: evidence from the music radio industry. *The RAND Journal of Economics 41*(2), 372–397.

Train, K. E. (2009). *Discrete choice methods with simulation.* Cambridge university press.

Trajtenberg, M. (1989). The welfare analysis of product innovations, with an application to computed tomography scanners. *Journal of political Economy 97*(2), 444–479.

Waldfogel, J. and S. T. Berry (1999). Free entry and social inefficiency in radio broadcasting. *RAND Journal of Economics 30*(3), 397–420.

Wright, D. J. (1994). Television advertising regulation and program quality. *Economic Record 70*(211), 361–367.

Zhang, J. (2017). Regulating advertising time: Evidence from french tv.

Appendices

A Objective transition probability estimation

The data used to generate the objective transition probabilities is the minute level broadcasting data spanning 2004-2005. The public channel, i.e. channel 1, has significantly different commercial broadcast behavior than the two commercial channels, notably, the channel scarcely broadcasts ads within a program, mainly between programs. This is evident from the summary statistics presented in table 2. As such, I will assume perfect foresight of the viewers with respect to channel 1 and concentrate on the implications of informational frictions on the two commercial channels alone. Table A.1 presents the number of observations and censoring statistics for each of the durations within the data used to estimate the objective transition probabilities.

Table A.1: Commercial state durations

	Channel 2			Channel 10		
Duration (min.)	Number	Left	Right	Number	Left	Right
Commercial state						
1	127	1	22	275	1	22
2	255	·	13	146	1	18
3	1,054	·	16	907	1	31
4	556	·	12	1,127	1	20
5	396	·	12	991	1	6
6	536	·	6	358	·	·
7	390	·	3	77	·	1
8+	175	·	1	1	·	·
Program state						
1-4	645	·	211	707	26	176
5-9	1,201	4	195	1,084	31	221
10-14	1,355	13	112	1,675	43	107
15-19	701	76	46	794	173	68
20-24	776	465	23	566	321	17
25-29	211	112	14	173	83	·
30-34	96	27	5	70	6	1
35-39	30	9	·	7	1	·
40+	37	7	6	85	10	7

Notes: The table displays the number of spells of each duration observed in the data independently for a commercial state (excluding between program commercial breaks) and a program state. The columns denoted *Left* refer to the number of spells that were left-censored within the data, i.e. began prior to 20:00, and equivalently for *Right*, i.e. ended after 22:00.

Figure A.1 and table A.2 together present the factors affecting the transition probabilities. Figure A.1 displays the average marginal effect of each of the durations on the transition probability while table A.2 shows the parameter estimates for the genre effects and the program progress effects.

Table A.2: Objective transition probability estimates

	Channel 2		Channel 10	
	Commercial	Program	Commercial	Program
Percentage program				
10-19	-0.849***	-0.256*	-0.875***	0.375***
	(0.21)	(0.10)	(0.22)	(0.10)
20-29	-0.986***	-0.205	-1.562***	0.811***
	(0.20)	(0.11)	(0.21)	(0.10)
30-39	-0.714***	-0.442***	-1.468***	0.689***
	(0.20)	(0.12)	(0.21)	(0.12)
40-49	-0.723***	-0.183	-1.843***	0.835***
	(0.20)	(0.12)	(0.22)	(0.11)
50-59	-0.283	-0.147	-2.070***	0.961***
	(0.19)	(0.11)	(0.22)	(0.10)
60-69	-0.853***	-0.063	-2.187***	0.966***
	(0.20)	(0.11)	(0.21)	(0.10)
70-79	-1.355***	0.392***	-1.784***	0.785***
	(0.20)	(0.10)	(0.22)	(0.10)
80-89	-0.162	1.000***	-1.912***	1.344***
	(0.19)	(0.10)	(0.21)	(0.10)
90-100	0.891***	2.158***	-1.258***	2.362***
	(0.21)	(0.09)	(0.22)	(0.10)
Genre				
Cinema	-0.568	-3.201***	-1.171***	-2.965***
	(0.46)	(0.38)	(0.23)	(0.20)
Culture	0.007	-0.246	0.859*	-0.357
	(0.29)	(0.19)	(0.35)	(0.22)
Documentary	-0.125	-0.535***	0.522*	-0.470***
	(0.21)	(0.13)	(0.21)	(0.12)
Drama	0.270	-0.428**	0.952***	-0.329**
	(0.23)	(0.15)	(0.20)	(0.12)
Entertainment	-0.035	-0.683***	0.294	-0.634***
	(0.20)	(0.13)	(0.20)	(0.11)
News	1.111***	-0.122	0.761***	-0.524***
	(0.20)	(0.13)	(0.19)	(0.11)
Sports	-0.744*	-0.496*	-0.647**	-0.403***
	(0.33)	(0.25)	(0.21)	(0.12)
Obs.	15132	47948	15374	49141

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure A.1: Duration effect on transition probability