Advertising Quantity Regulation and Content Quality Distortion

Ittai Shacham*

2nd December 2020

Job Market Paper

Link to latest version

Abstract

This paper studies the effects of advertising quantity regulation in television on channels' content quality decisions. Many countries impose quantity restrictions on the amount of advertising in broadcast television. Such restrictions suppress advertising levels; simultaneously, their influence on channels' incentives may manifest itself through distortions to other attributes of television programming, such as content quality. I develop a structural model of a television market that I estimate using high frequency data from Israel. The results illustrate that regulatory policies that do not account for equilibrium quality responses may be detrimental from a welfare perspective. Counterfactual experiments show that the current regulation is too restrictive and that relaxing the ad quantity constraint can lead to substantial welfare gains.

*Tilburg University, P.O. Box 90153, 5000LE Tilburg, The Netherlands. E-mail: i.shacham@tilburguniversity.edu
I extend my gratitude to my supervisors, Jaap Abbring, Jeff Campbell, and Tobias Klein for continual guidance and support. I am grateful to Alon Eizenberg, Richard Friberg, Martin Hackmann, Aureo de Paula, Niccola Pavanini, Jörgen Weibull and participants of the Tilburg Structural Econometrics Group, Tilburg Law & Economics group, CEPR IO 2019, and EEA 2020 for comments and suggestions.

1 Introduction

The effects of regulatory price caps on welfare are a-priori ambiguous. In scenarios where firms only set prices, such restrictions are generally welfare enhancing. However, when firms set multiple product attributes, price restrictions may influence non-price attributes as well. Neglecting this aspect of firm decisions may lead to biased welfare calculations and incorrect policy conclusions. This paper illustrates these effects in the broadcast television industry.

Broadcast television provides households with content at no direct cost, commonly funded through advertising (*commercial broadcast channels*) or governmental transfers (*public broadcast channels*).[1] The commercial broadcast television industry is subject to advertising quantity constraints in many countries.[2] While ad quantity restrictions suppress advertising levels, they also influence channels' programming incentives (Wright (1994); Gabszewicz et al. (2000)). The goal of this paper is to study the effects of ad quantity restrictions on the content quality decisions of commercial broadcast channels and their ensuing welfare implications.

Broadcast television continues to be one of the primary forms of entertainment worldwide. In the first quarter of 2020, adults in the US on average watched 4.25 hours of TV per day, 3.75 of which was live TV.[3,4] Broadcast TV also constitutes a leading advertising platform, accounting for 31% of global ad expenditure in 2019.[5] 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. There are two stark contrasts between broadcast television and new forms of entertainment. The first lies in the regulation they face. While broadcast television is subject to ad quantity restrictions in many countries, their online equivalents remain unregulated in this aspect. Furthermore, online content providers have the technological capacity to require viewers to view an ad. Consequently, viewer ad avoidance is operates differently in these settings.

This paper presents a structural model of a television market consisting of viewer demand, advertiser demand, and channels' supply. A factor affecting viewer behavior

[1] There are several additional business models in the industry, e.g. broadcast networks charging cable providers directly to be included in their packages. For the goals of this paper, I will concentrate on the traditional two financing forms.

[2] 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. 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.

[3] The Nielsen total audience report August 2020

[4] In Israel, the numbers are similar, in 2011, the average Israeli adult watched 4 hours of television per day.
Gili Izikovich. Study: Israelis Spend Nearly Four Hours a Day Watching TV. Haaretz. December 20, 2011.

[5] Based on Statista measurement.

is *incomplete information* surrounding the timing of commercial breaks: viewers' uncertainty pertaining the end of an ad break would induce some viewers to return while ads are still broadcast and others to return after the program has resumed. Alternatively, perfect foresight on the side of the viewers would imply a sharp and constant decrease in viewer shares during an ad break, that recuperates fully at the end of the ad break. The observed trends in the data are consistent with incomplete information (as shown in Figure 2). Consequently, viewers' ad avoidance is inherently intertwined with the informational framework. I incorporate a learning model regarding the ad dynamics, resulting in *channel viewing persistence* consistent with an empirical regularity that has been documented thoroughly (Kinjo and Ebina (2015); Esteves-Sorenson and Perretti (2012) and especially Anand and Shachar (2004) and the references therein).[6]

A second noteworthy feature of the viewing demand model is the *endogenous market size*. At the outset of each day —or prime-time —households decide whether to watch television on that day or participate in alternative leisure activities, i.e. make their *leisure decisions*.[7] Households make their leisure decision under an additional dimension of uncertainty, in addition to advertising uncertainty —the quality of the daily programming. Households consider the expected benefit from watching television in making their daily leisure decision accounting for expected program quality, advertising and the associated learning process. The endogenous market size generates *television viewing persistence*, a phenomenon that has also been documented empirically (Rust and Alpert (1984); Shachar and Emerson (2000)). Advertisers are modeled as a continuum of monopolistic firms, where impressions on all media outlets constitute perfect substitutes.

Channels choose their *advertising strategies* and *content quality strategies* with the goal of maximizing expected profits. Channels' profits are based on the expected advertising revenue and the average variable cost associated with content provision. Quality choices determine the mean quality of a program, whereas episodes are realizations from the ensuing quality distribution. Under a constrained advertising level, i.e. aggregate ad quantity constraint, channels choose the distribution of their permitted ads among the programs broadcast throughout the day. Allocation of ads across episodes is made after observing the realized qualities of the episodes.

The endogenous market size implies that channels' strategies induce non-trivial externalities among each other: content quality increases a channels' market share both by diverting viewers from other alternatives and by increasing the market size. The increased market size enhances competitors' incentives to provide quality. Higher quality provision

[6] Dubé et al. (2010) stress the distinction between "structural state dependence" and "spurious state dependence". While the former results from past choices directly affecting current choices, i.e. persistence, the latter results from unobserved heterogeneity. As shown in Moshkin and Shachar (2002) and ensuing work, television viewing behavior is characterized almost exclusively by the former.

[7] An alternative interpretation of households' leisure decision is households' deciding how much of their leisure time to devote to viewing television each day.

by competitors in turn induces higher switching among the viewers, further intensifying competition.

I consider two regulatory instruments: *ad quantities restrictions* and *information provision*. Both regulatory interventions have countervailing welfare effects. On the one hand they benefit viewers —by exposing them to less ads in the case of ad quantity constraints or allowing them to better allocate their viewing decisions during ads in the case of information provision. On the other hand, restricting advertisements or enhancing ad avoidance, limits channels' ability to transform content into revenue. As a result, channels' incentives to invest in content quality is diminished, which may lead to degradation in content quality.[8] Equilibrium quality effects may prove to offset welfare gains from regulatory interventions.

I estimate the model using high-frequency data from Israel on each minute of prime-time (20:00-22:00) throughout 2004-2005. Identification of the viewer demand relies on the advertising constraint. As in many countries that regulate advertising quantities, also in Israel, ad quantities are restricted for a given time frame, e.g. in the setting I utilize, channels' are allowed to broadcast up to 24 minutes of advertising throughout the two hours of prime-time. Under an aggregate ad quantity constraint, advertising on each episode is related to the quality of all other episodes; a high quality episode will decrease the amount of ads on other episodes, ceteris paribus. Variation in the usage of the allotted advertising and episode qualities allows me to construct a *shadow cost shifter* that identifies viewers' sensitivity to ads.

Identification of advertiser demand elasticities is based on a sample selection of days with a binding advertising constraint. On these days, variation in impressions results from quality differences that are pre-determined. Finally, marginal costs associated with content quality are estimated through equilibrium conditions.

Uncertainty plays a substantial role in viewer behavior: decreasing viewers' uncertainty by 10%,[9] decreases a channel's impressions by 36% on average —via viewers' enhanced ad avoidance —while increasing the average viewership on all minutes by 7%. As a result, equilibrium quality is sensitive to information provision, exhibiting up to 72% decrease in quality expenditure in the case of full information. This point exemplifies the importance of accounting for equilibrium quality in welfare analyses. The viewer surplus implications of a transition to full information without accounting for quality responses has an associated 7% mean increase in households' welfare. Accounting for equilibrium

[8] Viewers' ad avoidance constitutes an important distinction between broadcast television and multi-channel television such as cable and direct broadcast satellite (**DBS**). In the latter, viewers pre-commit to payment for quality via a subscription fee, while in the former, viewer ad avoidance affects channels' quality provision incentives.

[9] Although this experiment is hard to imagine in reality, information provision also has different levels. E.g. one could display a ticker on screen during ads disclosing the amount of time left for an ad break; Alternatively, the ticker may present the number of ads left, without specifying distinctly the amount of time remaining.

quality produces an opposite effect —viewer surplus decreases by 13% on average.

More generally, the quality deterioration effect of information provision implies that information provision always reduces viewer surplus. Namely, viewers are willing to trade-off more ads with higher content quality, with that, viewers lack of commitment to view ads induces an equilibrium that is detrimental to all participants. This can be seen as a 'missing market' whereby channels are unable to compensate viewers directly for watching ads. Online equivalents of broadcast television, among them YouTube and Hulu, have the technological capacities to resolve this issue by conditioning content provision on ad broadcast for each viewer. I investigate the value derived from these technologies by measuring the equilibrium effects of requiring viewers to watch ads in my setting. Canceling viewers' ad avoidance substantially affects channels' quality provision incentives, leading to a 35.6% decrease in equilibrium quality expenditure. Consequently, viewers' surplus diminishes by $317 mil., while channels' profits increase by $607.1 mil. Hence, technologies that provide a solution to the missing market lead to substantial societal welfare gains —34.5% increase equivalent to $290.1 mil. in my setting. With that, the distribution of the surplus may be uneven, benefiting some at the expense of others.

Optimal regulation differs substantially between the households and channels. While the channels' unconstrained equilibrium advertising levels are approximately 22 minutes of ads per hour, the households' surplus is maximized at roughly 10 minutes of advertising per hour. Social welfare is maximized under a regulation consisting of no information provision and 13 minutes of ads per hour. The socially optimal advertising constitutes 9% increase over the current regulation, leading to welfare gains of up to $23 million throughout 2005.

The estimated model also provides a framework to analyze the mechanics driving quality competition. Simulations show that the advertiser side of the market induces strategic substitutability in quality competition among the channels. Simulations of a scenario characterized by infinite elasticity of advertiser demand - i.e. no ad price setting power - while holding the viewer demand characteristics fixed, show that quality provision increases by 113%. Both viewer demand characteristics —viewers' ad avoidance and the market expansion effect —induce channels to provide higher quality. The market expansion effect increases quality provision by 10%; while the ad avoidance increases quality provision by 23%. Market structure is shown to also have important implications for quality. Transition from a multi-channel monopoly to symmetric duopoly entails a 44% increase in quality investment.

Related Literature. This paper relates to several strands of literature. First it provides a welfare assessment of regulation in media markets. Past research on media markets emphasized several important aspects relating to welfare and *market structure* (Berry and Waldfogel (2001); Goettler and Shachar (2001); Berry et al. (2016b)); *mergers* (George (2007); Chandra and Collard-Wexler (2009); Sweeting (2010)); and *incentives in* *two-sided markets* (Berry et al. (2016a); Baker and George (2019); Durante et al. (2020)). This paper is most similar to Fan (2013) who showed that mergers in the newspaper industry induce price effects as well as changes to other product characteristics. Similarly, this paper also addresses competitive interaction with endogenous quality. I contribute to this literature by quantifying the extent to which content quality declines as a result of advertising quantity restrictions. I extend the findings in Zhang (2017), showing that quality responses have substantial effects on welfare, in line with the findings of Fan (2013). Specifically, I show that not accounting for quality responses leads to biased welfare measurement and may lead to incorrect policy recommendations.

This paper emphasizes scenarios in which information provision may be detrimental to consumers. Recent empirical research has analyzed the effects of price transparency in several settings, among them gas stations (Rossi and Chintagunta (2016); Luco (2019); Montag and Winter (2019)); retail (Ater and Rigbi (2019)); and health care (Brown (2019); Whaley (2019); Grennan and Swanson (2020)). The main finding of this research is that price transparency is generally effective in decreasing prices. Alternatively, several theoretical papers have emphasized that information provision may actually be detrimental to consumers, e.g. Levin (2001); Belleflamme and Peitz (2014). This paper reconciles the two strands of research by showing how information provision may lead to quality degradation, that is eventually detrimental for consumers. The mechanism underlying this finding is similar to that in Matsa (2011) who showed how enhanced competition is associated with lower quality in the retail industry.

This paper also contributes to the literature on endogenous product choice, among them Sweeting (2013); Fan (2013); Crawford and Shum (2007); Crawford et al. (2019).[10] Sweeting (2013) shows the distortions to radio stations' horizontal product placement incentives —musical genre —resulting from regulation. Crawford and Shum (2007); Crawford et al. (2019) concentrate on the quality choice of a monopolist[11] where the former provides evidence of quality degradation in cable television. Similarly to the approach in Fan (2013), I assume the set of programs offered by each channel remains fixed while their endogenous characteristics may change. I contribute to this literature by analyzing quality investment choice under oligopolistic competition.

From a policy perspective, this paper contributes to the literature on the effects of price restrictions. A summary of the literature is provided by Joskow and Rose (1989); Joskow (2005); Armstrong and Sappington (2007). Effects of price caps on quality provision have been examined empirically in several settings, examples include the *airline industry* (Basso (2008); Anderson and Kraus (1981)), *financial markets* (De Pinho (2000); Cuesta and Sepúlveda (2019); Galenianos and Gavazza (2020); Romero (2020)), and the

[10]Additional papers that considered endogenous product choice not in the media context are for example Mazzeo (2002); Gandhi et al. (2008); Draganska et al. (2009); Eizenberg (2014).

[11]Crawford et al. (2019) analyzes equilibrium quality distortions arising in 2$^{\text{nd}}$ degree price discrimination under competition from a non-strategic outside option.

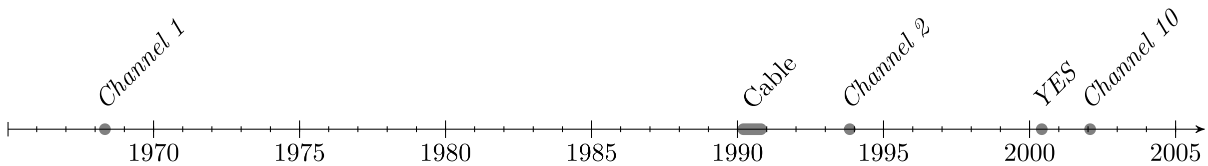
healthcare industry (Gaynor and Town (2011); Gaynor et al. (2015)). This paper provides an empirical example emphasizing the critique made by Bork (1978) whereby regulators focus on salient attributes, disregarding the effects of regulation on less salient attributes. By doing so, welfare calculus is distorted and regulation may be detrimental.

The paper is organized as follows: Section 2 provides an overview of the Israeli television industry; Section 3 describes the data; Section 4 presents the model; Section 5 details the empirical analysis; Sections 6 and 7 present results from policy experiments. Finally, Section 8 concludes.

2 Institutional background

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 public channel, channel 2 is commercial.[12] In 2002 a second commercial channel was introduced, *Channel 10*. Parallel to the development of broadcast commercial television, multi-channel broadcast - cable and DBS - 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 DBS operator - YES. By the mid 2000's, 66% of households had access to multichannel broadcast, consisting of 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



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

[12]In an attempt to introduce a degree to competitiveness in the programming, the broadcasting days of channel 2 were distributed 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 broadcast on Sunday and Wednesday; Reshet on Tuesday, Friday and Saturday (until April); and Keshet broadcast on Monday, Thursday and Saturday (from April onwards). Telad lost concessionaires broadcasting bid for channel 2 and ceased broadcast on 30$^{\text{th}}$ of October 2005, leaving only Reshet and Keshet.

Table 1: Commercial placement regulation

Criterion	Rule
Amount of commercials	24 minutes per two hours (prime-time) 12 minutes per hour (non-prime-time)
Length of commercial break	5 minutes 3 minutes (*News* program)
Number of commercial breaks	4 per hour

Notes: *Prime-time* refers to the two hours from 20:00 - 22:00.

3 Data overview

I utilize three unique datasets regarding broadcast characteristics, television viewership and ad prices in Israel throughout 2004 & 2005.[13] The *broadcast* data details the start time, end time, name and genre[14] of each program and commercial break on channels 1, 2, and 10 that aired during prime-time (20:00-22:00).[15] The *viewer* data details three measures of viewership on each of the channels at each minute: number of households watching television, market share with respect to the number of households owning a television, and market share with respect to the number of households watching television. Finally, the *pricing* data details the price per impression of each ad broadcast on each of the two commercial channels during prime-time throughout the time period.

3.1 Viewing characteristics

Several important features of this industry arise from inspection of the data. Table 2 presents the persistence parameter, i.e. the coefficient on the one-period lag of the market share across several specifications. The results are aligned with the findings in previous research (e.g. Anand and Shachar (2004); Shachar and Emerson (2000)) and exhibit several stark features. First, the market share of television viewing exhibits a strong within day correlation - as shown in specifications (2) and (4) - while a substantially weaker correlation across days - specifications (1) and (3). Consequently, market share flows much more between channels than with alternative leisure activities. Inspection of the difference between specifications (2) and (4) provide insight into the extent to which viewers flow between broadcast channels and multi-channel viewing alternatives, i.e. cable

[13]The broadcasting and viewer data was acquired from Kantar Media, a private firm that collects data on television viewer shares in Israel. The ad price data was acquired from Ifat, a subsidiary of FIBEP media intelligence.

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

[15]The broadcast data was aggregated to the minute level. I.e. whenever more than one program was broadcast on a given minute, the program allocated the largest amount of time was associated with the specific minute of broadcast.

Table 2: Market share persistence

	TV		Broadcast TV			Channel	
	Between days	Within day	Between days	Within day	Total	Between episodes	Within episodes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Lagged MS	.175	.980	.171	.968	.942	.143	.952
	(.031)	(.001)	(.031)	(.001)	(.001)	(.014)	(.001)
Controls	+	+	+	+	+	+	+
Adj. R^2	.578	.997	.360	.983	.990	.772	.991
Obs.	1,095	130,300	1,095	130,424	384,929	6,550	359,191

Notes: The table presents the persistence in market shares across several scenarios. The unit of observation in specifications (1) and (3) is the mean daily market share across all minutes of prime-time; in specification (6) it's the mean market share within an episode; the remaining specifications focus on the minutes within prime-time. All specifications include data from 2003-2005. The controls are date & time (week, weekday, day, minute) dummies and broadcasting characteristics (program dummies and percentage of the program). Standard errors in parentheses.

and satellite (**MCTV**). Viewing is to a large extent constrained within the broadcast television alternatives. Channels' market share also exhibits a strong persistence, while this is mainly driven by the persistence in market share within episodes - specification (7) - and much less so between episodes - specification (6). Note that the non-negligent persistence in market share between episodes in specification (6) is indicative of the "lead-in" effect (Moshkin and Shachar (2002)).

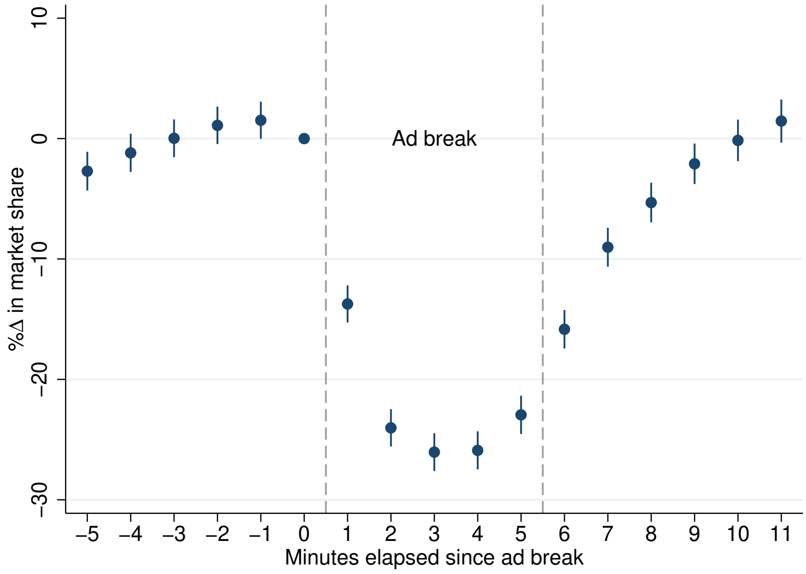
A second phenomenon of the viewing data is incomplete information on the side of the viewers. Figure 2 presents the percentage change in the market share of channels before and after the beginning of a commercial break. Figure 2a presents the percentage change of the market share of the channel broadcasting the ad while Figure 2b displays the effect on the market shares of the other two broadcast channels. Figure 2a shows that the market share of a channel is steady until the beginning of an ad break, upon which viewers depart en-masse. Their return occurs gradually, where some viewers return while the channel is still broadcasting commercials, whereas others return after the program has resumed. Moreover, the market share fully recuperates approximately five minutes after the end of the commercial break. The form of the change in market share implies that viewers imperfectly allocate their viewing due to uncertainty pertaining to the length of a commercial break.

Figure 2b provides additional insight into viewer behavior. While the data used in the analysis is aggregated to market shares and does not follow individual viewers, the trends in market shares show that approximately 62% of the market share flows within the broadcast television channels and not to other viewing alternatives (or non-viewing alternatives). This is aligned with the evidence in Table 2 whereby viewership is to a large extent confined within the broadcast television channels. The shape of the change in market shares closely follows that presented in Figure 2a providing stark evidence to viewers ad avoidance behavior.

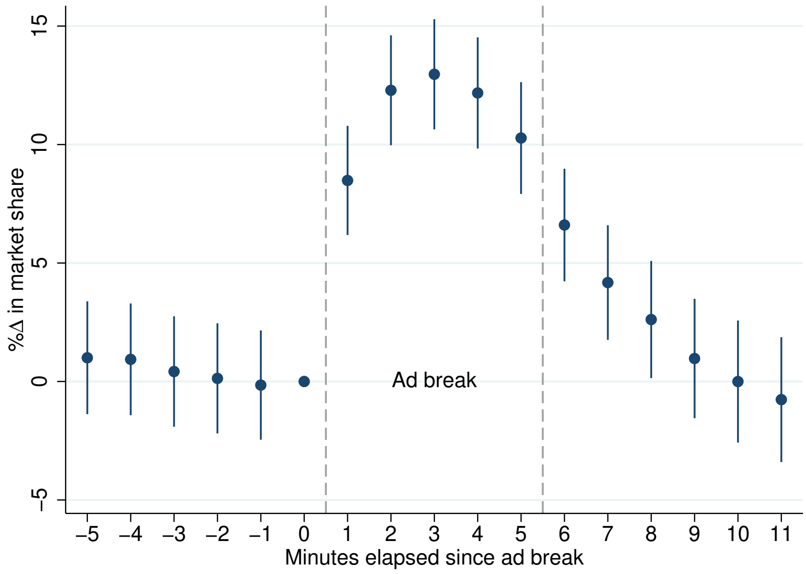
The viewership measures detailed in the viewing data allow me to construct mea-

Figure 2: Percent change in market shares

(a) Own effect



(b) Cross effect



Notes: The figure displays the average percentage change in market share for each minute leading up to and following a five minute ad break using various controls (percentage of the program, percentage of the program squared, as well as channel, minute, day, and program fixed effects). The horizontal axis presents the minutes elapsed from the beginning of an ad break. The vertical axis presents the mean percentage change in 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 $t = 0$. Only within-program ad breaks were included in the figure. The data used pertains to the commercial channels (2 & 10) throughout 2005. Figure 2a presents the percentage change in market share of a channel following an ad break broadcast on that channel. Figure 2b displays the percentage change in market share for the other two broadcast channels.

sures for both the total number of households' owning a TV, i.e. the *potential market size*, as well as the total number of households watching television in a certain day, i.e.

the *daily market size*.[16] Figure 3a displays the distribution of the share of households watching television and the share of households watching a broadcast channel on each day throughout 2005. These two series —the share of households watching television on a given day and the average market share of households watching a broadcast channel —will provide important variation to identify viewers' leisure decisions. Specifically, controlling for week and weekday effects, the two series display a correlation of 0.98 (with a standard error of 0.04).

Figure 3b displays the distribution of the viewer shares among the viewing alternatives, i.e. the average daily market share for each of the alternatives - channels 1, 2, 10, and the outside option - throughout 2005. The market shares were calculated with respect to the daily market size, i.e. the maximal number of viewers watching television during any minute of the day's prime-time. The market shares are consistent throughout the time period, with the outside option accounting for roughly 60%, channel 2 for 20% and channels 1 & 10 for 10% each.

3.2 Advertising characteristics

The commercial placement behavior of the channels exhibits heterogeneity across several dimensions: a. across days; b. across minutes within a day; c. across programs; and d. across episodes of a given program. Table 3 presents summary statistics pertaining to the time allotted to advertising across days and across minutes. The advertising constraint determines a limit for the amount of ads within prime-time - 24 minutes. Ads are broadcast in bulks - both at the individual ad lasting roughly 20 seconds on average as well as the commercial break lasting approximately 4 minutes on average. Figure 4 presents the distribution of the daily advertising on both commercial broadcast channels throughout 2005. The figure clearly shows that the advertising constraint is binding for most days. With that, the figure also displays that channels advertise a more ads than permitted on some days.

Advertising on the public channel differs substantially from that on the commercial

[16] The potential market size was calculated as:

$$M = (D \cdot T \cdot J)^{-1} \sum_{d=1}^{D} \sum_{t=1}^{T} \sum_{j=1}^{J} \frac{N_{jtd}}{s_{jtd}} \tag{3.1}$$

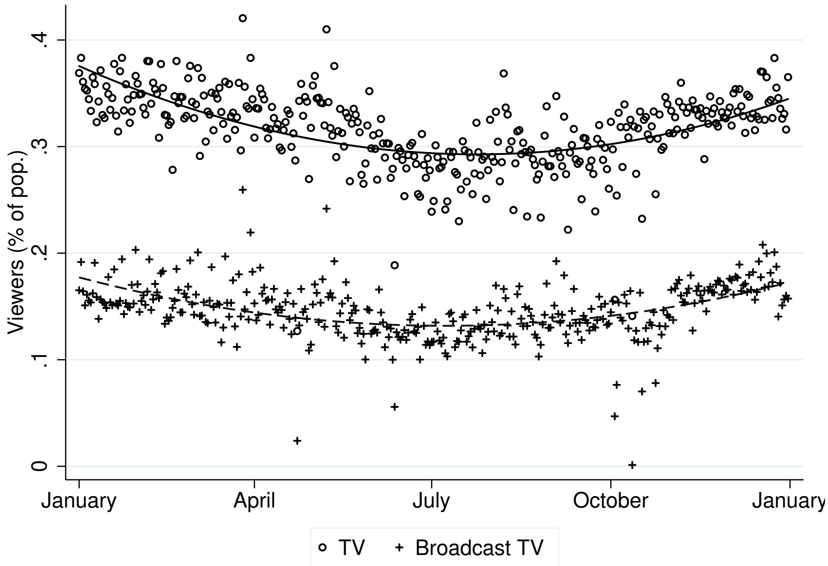
where M is the potential market size, N_{jtd} are the number of viewers watching channel j at minute t of day d and s_{jtd} is the market share with respect to the total number of households owning a television. The effective market size was calculated as:

$$M_d = \max_{t=1,\ldots,T} \left[J^{-1} \sum_{j=1}^{J} \frac{N_{jtd}}{v_{jtd}} \right] \tag{3.2}$$

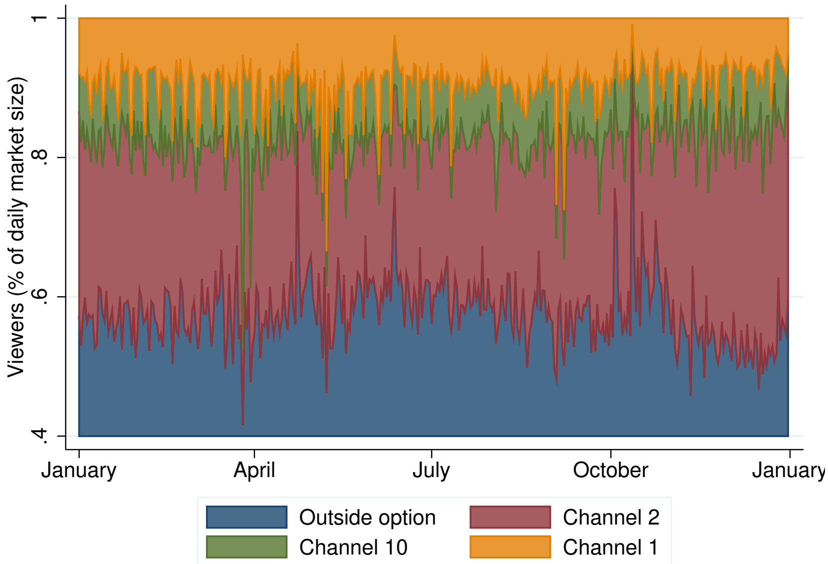
where M_d is the effective market size on day d and v_{jtd} is the market share with respect to the total number of households watching television.

Figure 3: Daily market shares

(a) Daily television viewership



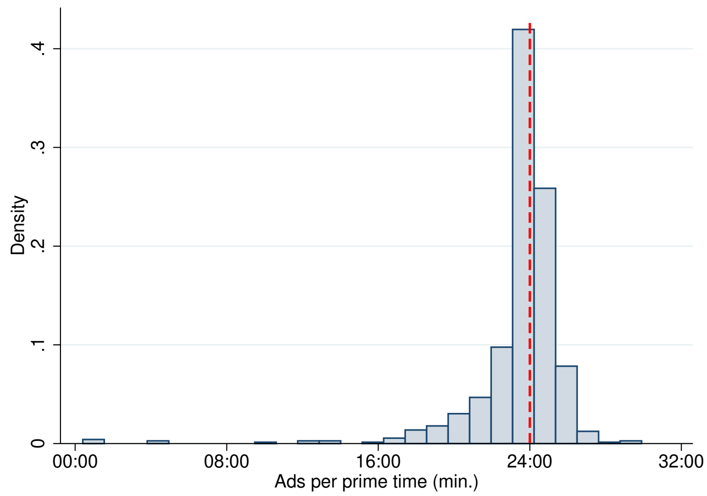
(b) Daily market shares



Notes: The figure displays the evolution of daily market shares across 2005. The vertical axis in panel 3a displays the share of households owning a TV while the horizontal axis displays days of the year. Each mark represents the share of households on a day of the year. The fit lines were derived from a fractional polynomial regression consistent with a quadratic specification. Panel 3b presents the average daily market share (based on the effective daily market size delineated in equation 3.2) throughout prime-time for each viewing alternative, i.e. the broadcast channels and the outside option. The outside option includes both multi-channel alternatives and non-television activities.

channels. This is evident both in the total time allotted to ads, whereby the public channel broadcasts 8% of total time, as opposed to 23% and 21% on channels 2 and 10 respectively. Furthermore, while the commercial channels allocate ads mainly within programs, with propensities of 73% and 87%, channel 1 mainly allocates non-program

Figure 4: Daily advertising on commercial broadcast channels



Notes: The figure displays distribution of daily advertising on channels 2 and 10 throughout 2005. The red line indicates the aggregate ad quantity constraint.

content between programs, with an in-program ad propensity of 16%.

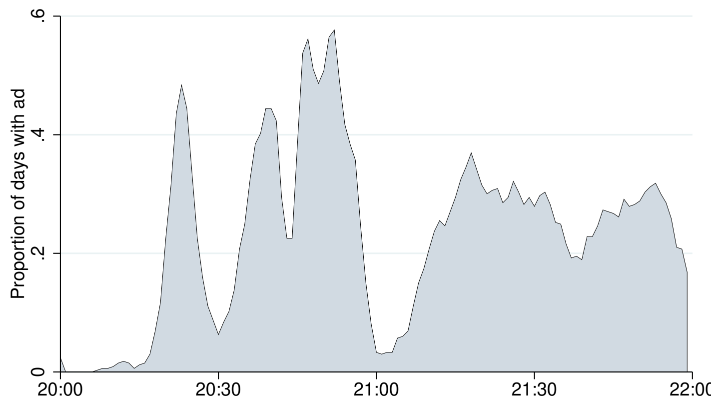
Advertising propensities differ across minutes of the day, as shown in Figure 5a. Specifically, the figure clearly shows that while some minutes are more prone to advertising than others, there is substantial variation in the ad propensities across days. Advertising heterogeneity across minutes is also evident in the break lengths. Figure 5b displays the variation in the length of ad breaks broadcast at each minute of prime-time across the year on channel 2. The shaded area depicts the interquartile range and shows that advertising in a given minute differs across days also in the length of the ad break. Together, these two advertising characteristics explain viewers' switching behavior during ads as documented in Figure 2. Namely, viewers face uncertainty regarding the advertising in each minute. This uncertainty is twofold, both regarding the probability an ad break will start at each minute, but also with respect to the length of the ad break.

3.3 Programming characteristics

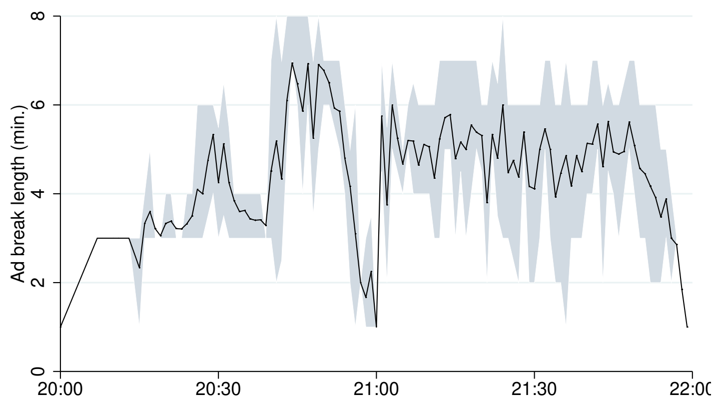
Programs also exhibit substantial variation in their advertising propensity. Figure 6 presents the distribution of the advertising propensities, i.e. the ads per minute across all episodes broadcast on the two commercial channels during prime-time throughout 2005. Figure 6a presents the total variation across episodes while Figure 6b presents the distribution of the variation between programs and within programs, i.e. between episodes of a given program. There is substantial variation across episodes in the sample, ranging from no ads and up to 50%. This variation is mainly driven by between program variation, accounting for 81% of the total variation in advertising propensities. Episodes of a program also exhibit variation, with up to 20% difference in the amount of ads relative to the program average.

Figure 5: Ad distributions across minutes

(a) Advertising propensities



(b) Ad break length



Notes: Figure 5a displays the share of days in 2005 in which channel 2 broadcast an ad in each minute of prime-time. Figure 5b displays the mean and IQR of the ad break lengths broadcast on channel 2 in each minute of prime-time throughout 2005. The solid line depicts the mean ad break length and the shaded area the IQR.

Table 3 presents summary statistics regarding the channels' broadcasting detailing a large variety of programs, with channel 2 and channel 10 broadcasting 190 and 292 programs respectively. Furthermore, programs are observed multiple times, with a large variation in the number of episodes between programs. Finally, Table 4 presents the ranking and proportional prevalence of each genre on the three broadcast channels. *News & Current Affairs* programs account for approximately half of the prime-time on all channels. Differentiation is apparent in the remaining rankings. The public channel - channel 1 - allocates relatively similar time to the rest of the genres with *Sports* and *Culture, Leisure & Education* as the second and third. It is noteworthy that *Culture, Leisure & Education* programs rank lowest on the two commercial channels. After news

Table 3: Summary statistics

Category	Variable	Channel 1 Mean	Channel 1 SD	Channel 2 Mean	Channel 2 SD	Channel 10 Mean	Channel 10 SD
Programs	Episode length (min.)	51.46	26.85	54.72	23.74	53.84	31.36
	Programs per day	3.14	0.90	2.62	0.58	3.02	0.86
	Episodes per program	4.13	16.12	5.04	17.79	3.75	16.25
Ads	Total (min.)	9.29	3.70	27.82	3.78	25.31	2.89
	In-program (%)	.16	.29	.73	.17	.87	.12
	Between-program (%)	.84	.29	.27	.17	.13	.12
	Minute ad propensities (%)	.08	.12	.23	.15	.21	.12
	Break length (min.)	3.44	1.68	4.58	2.17	3.88	1.50
	Breaks per hour	1.35	0.86	3.04	0.68	3.26	0.83
	PPI (2005 USD)	-	-	94.60	36.78	163.33	133.23
Observations	Channel-minute	39,960		39,960		39,960	
	Programs	277		190		292	
	Ad prices	-		723		723	

Notes: The table presents summary statistics regarding the broadcasting behavior of the three broadcast channels throughout 2005.

Table 4: Genre prevalence

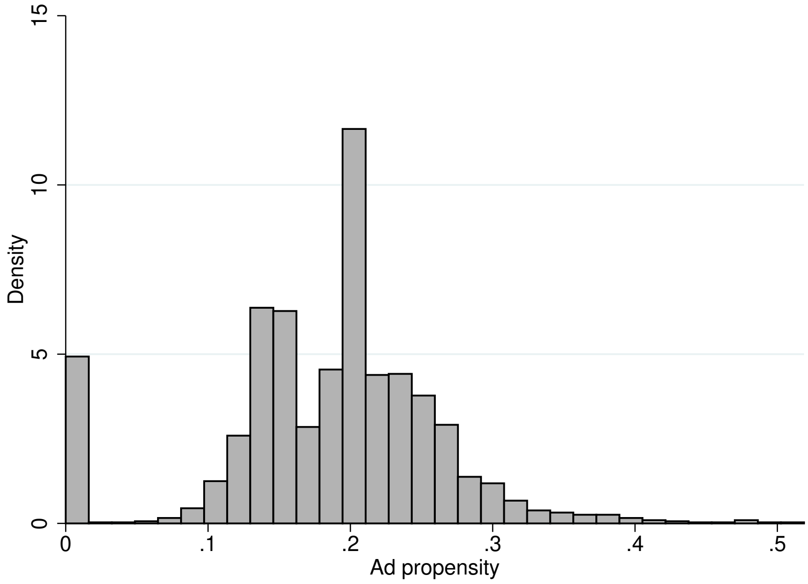
Rank	Channel 1	Channel 2	Channel 10
1	News & Current Affairs (.651)	News & Current Affairs (.481)	News & Current Affairs (.515)
2	Sports (.088)	Entertainment (.309)	Entertainment (.111)
3	Culture, Leisure & Education (.078)	Documentary (.143)	Television Drama (.108)
4	Entertainment (.064)	Television Drama (.038)	Documentary (.093)
5	Documentary (.058)	Cinema (.016)	Sports (.089)
6	Television Drama (.031)	Other (.009)	Cinema (.068)
7	Cinema (.017)	Culture, Leisure & Education (.003)	Other (.009)
8	Other (.013)	-	Culture, Leisure & Education (.007)

Notes: Proportion of prime-time minutes throughout 2005 allocated to each genre presented in brackets. The data used in calculating the genre prevalence was limited to the two hour prime-time window (20:00-22:00). Programs commencing prior to 20:00 or ending after 22:00 were censored. Between program commercial breaks were excluded from the analysis.

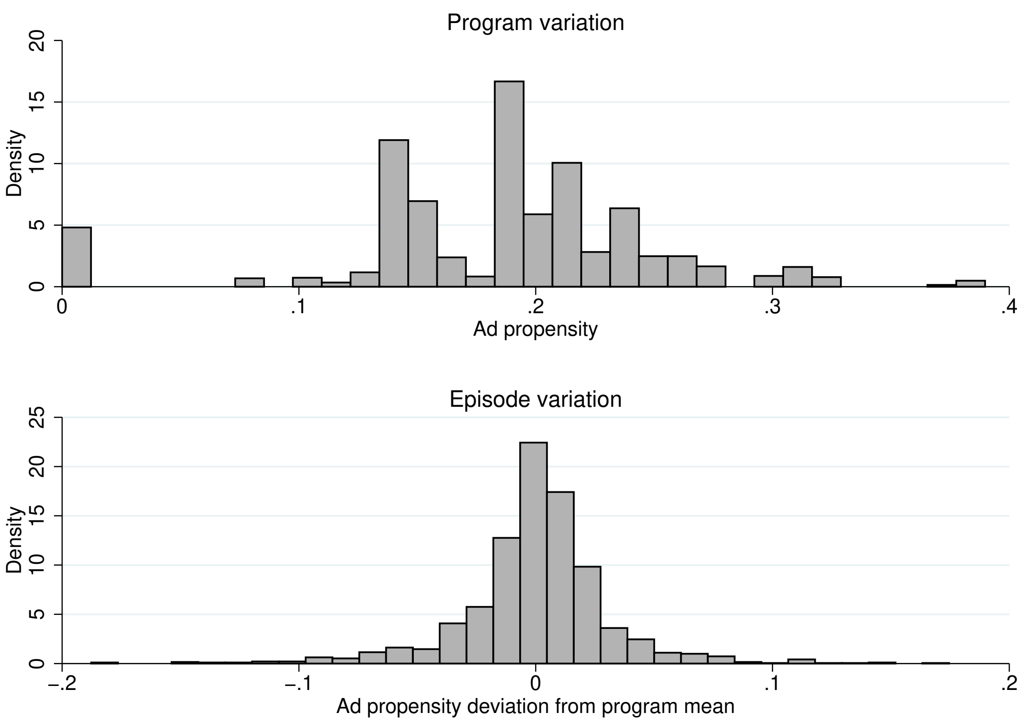
programming, channel 10 allocates the same amount of time to *Entertainment*, *Drama* and *Sports* as channel 2 allocates to *Entertainment* programming.

Figure 6: Ad variation across programs

(a) Total variation



(b) Variation breakdown



Notes: Figure 6a displays the distribution of ads per minute of program across all episodes broadcast on the two commercial channels during prime-time throughout 2005. The data in the figure does not include commercial breaks between episodes. The top panel of Figure 6b presents the distribution mean ad propensities across programs. The bottom panel of Figure 6b displays the ad propensity in an episode relative to its program average.

4 A model of the television market

The following section develops a model for a television market in which demand is with respect to both viewers and advertisers, i.e. a two-sided market. The viewer demand model has two stages, at the beginning of each day, households make their *leisure decision*, i.e. whether to watch television, according to their expectations regarding the utility from TV, determining the daily market size. At this stage households are uncertain regarding both the quality of the episodes as well as the amount (and timing) of advertising. At each minute within a day, viewers make their *viewing decision* among the viewing alternatives, determining the market share of each channel. Once watching television, viewers know the quality of the programming while uncertainty remains regarding the timing of advertising. Consequently, viewers base their viewing choices on their expectations that update with the broadcasting of the channels and their individual information sets. The gradual updating of viewers' ad assessments incorporates a learning process similar to that in Moshkin and Shachar (2002).

Viewers have several options from which they can choose: watch a commercial broadcast channel, $\mathcal{J}_m$; a public broadcast channel, $\mathcal{J}_o$; or opt for the outside option denoted by $j = 0$.[17] Within each minute, viewers are assumed to choose one of the $j \in \mathcal{J} = \{0, \mathcal{J}_m, \mathcal{J}_o\}$ alternatives. Channels broadcast a set of *programs* denoted by $p \in \mathcal{P}_j$ at specific minutes and days —denoted by (t, d) respectively. The set of all times in which a program p is broadcast is denoted by $(t, d) \in \mathcal{T}_p$. The quality of each program is a random variable G_p with $g_p = \mathbb{E}[G_p]$. Each day, channels broadcast *episodes* of a subset of their programs with corresponding realized qualities. Commercial channels choose the mean quality of each of the programs they broadcast as well as the expected amount of ads within each episode.[18] Channels' quality choice is made ex-ante, shifting the distribution of episode qualities; advertising choices are made after observing the daily quality realizations on all channels.

4.1 Viewing decision(s)

4.1.1 Utilities

Denote minutes within a day (discrete and finite) by $t = 1, ..., T$ and days (also discrete) by $d = 1, ...\infty$. Viewers' maximize their contemporaneous indirect utility. Denote the

[17] The outside option includes non-television activities as well as multi-channel alternatives, i.e. cable or DBS.

[18] Note that while channels cannot choose the total amount of ads freely, they choose how to allocate their permitted amount of ads across the episodes they broadcast.

indirect utility of viewer i from watching channel j in minute t of day d:

$$u_{ijtd} \equiv u(a_{jtd}, q_{jtd}, x_{jtd}, \varepsilon_{ijtd}; \alpha, \beta, \gamma) = \alpha a_{jtd} + x_{jtd}\beta + \gamma \underbrace{q_{jtd}}_{=g_p+\xi_{jtd}} + \varepsilon_{ijtd} \quad (4.1)$$

$a_{jtd} \in \{0, 1\}$ is an indicator of whether an ad is being broadcast on channel j at minute t of day d; x_{jtd} is a vector of observed characteristics; q_{jtd} is the quality of the broadcasting on channel j at minute t on day d where $q_{jtd} = g_p + \xi_{jtd}$, in which g_p is the quality of the program broadcast on channel j at minute t of day d, i.e. $g_{jtd} = g_p$ for all $(t, d) \in \mathcal{T}_p$ and ξ_{jtd} is the stochastic quality component; ε_{ijtd} is an idiosyncratic taste shock iid across individuals, alternatives and time. The parameter α is a scalar representing the mean viewer's utility from a commercial; β is a vector of mean taste coefficients for a channel's observed characteristics; and γ is viewers' mean quality preference. The common part of the indirect utility is normalized to zero for the outside option, yielding the viewer a reservation utility of $u_{i0t} = \varepsilon_{i0t}$.

Viewers make their viewing decision prior to observing the ad state. The ambiguity surrounding the commercial state at the time of making their viewing choice propels viewers to consider their expected utilities, conditional on their information set at time t denoted by Ω_{ijtd}:

$$\mathbb{E}(u_{ijtd}|\Omega_{ijtd}, \varepsilon_{ijtd}) = \alpha\mathbb{E}(a_{jtd}|\Omega_{ijtd}) + x_{jtd}\beta + \gamma q_{jtd} + \varepsilon_{ijtd}$$

4.1.2 Beliefs

The viewers perceive ads to be allocated as a non-stationary 1$^{\text{st}}$ order Markov process. Consequently, their knowledge of the ad state on previous minutes affect their perception of the ad probability. Viewers share a common prior, whereas a viewer's belief (posterior) diverges from the common prior according to their information set.

Viewers' common prior specifies the instantaneous probability of an ad at each minute. The common beliefs are rational in the sense that they are consistent with the programming of the channels. Viewers perceive the ad process as a non-stationary 1$^{\text{st}}$ order Markov process independent across channels and programs. Furthermore, the transition probabilities are perceived at the program level and therefore constant across episodes.

Assumption (Common beliefs)**.** *The viewers' common prior regarding the probability of an ad being broadcast on channel j at time t of day d depends only on 1. the previous ad state on the channel; 2. the deterministic program quality on the channel; and 3. the* *broadcast characteristics. Formally:*

$$\Pr[a_{jtd}|\underbrace{a_{11d},\ldots,a_{Jt-1d}}_{\mathbf{a}_{td}},\underbrace{q_{11d},\ldots,q_{JTd}}_{\mathbf{q}_{d}},\underbrace{x_{11d},\ldots,x_{JTd}}_{\mathbf{x}_{d}}] = \Pr[a_{jtd}|a_{jt-1d}, g_p, x_{jtd}]$$

Denote viewers' prior regarding channel j at time t of day d by $\lambda_{jtd}(a_{jt-1d}) = \Pr[a_{jtd}|a_{jt-1d}, g_p, x_{jtd}]$ and let Λ_{jtd} be the 2×2 matrix of transition probabilities. The rows of Λ_{jtd} define the outgoing state (in minute $t-1$) and the columns define the incoming state (at time t):

$$\Lambda_{jtd} = \begin{array}{c c} & \begin{array}{cc} a_{jtd}=0 & a_{jtd}=1 \end{array} \\ \begin{array}{c} a_{jt-1d}=0 \\ a_{jt-1d}=1 \end{array} & \begin{pmatrix} 1-\lambda_{jtd}(0) & \lambda_{jtd}(0) \\ 1-\lambda_{jtd}(1) & \lambda_{jtd}(1) \end{pmatrix} \end{array}$$

The probability of being in each of the two states in the beginning of the evening is represented by a 1×2 vector of the marginal ad probability denoted accordingly by Λ_{j1d}:

$$\Lambda_{j1d} = \begin{pmatrix} 1-\lambda_{j1d} & \lambda_{j1d} \end{pmatrix}$$

Viewers update their assessment of the ad probability, i.e. posterior, according to a Bayesian procedure, by which they take into account all possible paths leading to an ad in the upcoming minute. The amount of time since last viewed (τ_{ijt}) determines the individual specific transition probability matrix. Multiplication of the objective transition probability matrices creates probabilities to reach any state from any initial state via all possible routes. E.g. a viewer who has viewed a channel in the previous period has only 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 resulting from the several possible broadcasting routes. Viewer i's transition probability matrix regarding channel j after not viewing the channel for τ_{ijt} periods is:

$$\Lambda_{ijtd}(\tau_{ijtd}; \boldsymbol{\lambda}_{jd}) = \prod_{\ell=0}^{\tau_{ijtd}} \Lambda_{jt-\ell d} \tag{4.2}$$

Finally, viewers' posterior, denoted by μ_{ijt}, is defined by the ad state in their last viewed time, $a_{ijt} \equiv a_{jt-\tau_{ijt}}$. The last observed state determines the row within the matrix, where the ad probability is given by the second column of the relevant row:

$$\mu_{ijt}(a_{ijtd}, \tau_{ijtd}; \boldsymbol{\lambda}_{jd}) = \begin{pmatrix} 1-a_{ijtd} & a_{ijtd} \end{pmatrix} \cdot \Lambda_{ijtd} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.3}$$

Viewer i's *information set* at time t is:

$$\Omega_{itd} = \{q_{jtd}, x_{jtd}, \boldsymbol{\lambda}_{jd}, \tau_{ijtd}, a_{ijtd}\}_{j \in \mathcal{J}}$$

4.1.3 Choices

Using the ex-ante utility of viewer i of watching channel j together with the individual ad belief derived above, we can rewrite the viewers' ex-ante utility as:

$$U_{ijtd} \equiv \mathbb{E}[u_{ijtd}|\Omega_{itd}, \varepsilon_{ijtd}] = \underbrace{\delta_{jtd}(q_{jtd}, x_{jtd}; \beta, \gamma)}_{\text{common utility}} + \underbrace{\alpha \cdot \mu_{ijtd}(a_{ijtd}, \tau_{ijtd}; \boldsymbol{\lambda}_{jd}) + \varepsilon_{ijtd}}_{\text{idiosyncratic utility}} \quad (4.4)$$
$$\text{where} \quad \delta_{jtd}(q_{jtd}, x_{jtd}; \beta, \gamma) = x_{jtd}\beta + \gamma q_{jtd}$$

Accordingly, at each minute, viewers choose the alternative deriving them the highest contemporaneous utility:

$$y_{it} = \arg\max_{j \in \mathcal{J}} \{U_{i0td}(\Omega_{itd}), \ldots, U_{iJtd}(\Omega_{itd})\}$$

Aggregation of viewers' choices at each minute generate the choice probabilities at each minute unconditional on the viewing history:

$$s_{jtd} = \int \mathbb{1}\{U_{ijtd} \geq U_{iktd} \; \forall \, k \neq j\} \, dF_{\varepsilon} \quad (4.5)$$

where F_{ε} denotes the population distribution of the taste shocks.

4.2 Leisure decision

Households' decision to watch television excludes alternative leisure activities, e.g. going to the beach. While not watching television remains an option at each minute, the decision to forgo alternative leisure activities alters the value of the outside option. Households' decision to watch television relies on their assessment of the utility from doing so as well as characteristics of the outside option. This first stage endogenizes the market size whereby channels' choices affect the viewership throughout the day on all channels.[19]

Households make their leisure decision under two dimensions of uncertainty, pertaining to the ad timing —as detailed in Section 4.1 —and the daily realized content quality. Once viewers begun watching television, they know the realized quality while making their viewing decision. Before the day has begun, they are not aware of the specific quality values.[20] As such, they make their leisure decisions based on the expected

[19]The market size common to all minutes of the day distinguishes this model from the Nested Logit framework (Berry (1994)).

[20]In the model, the quality at each minute is determined by the realization of ξ.

quality:

$$\mathbb{E}_{\xi}\left[U_{ijtd}\right] = \alpha\mu_{ijtd}(\Omega_{itd}) + x_{jtd}\beta + \gamma g_p + \varepsilon_{ijtd} \tag{4.6}$$

At the time of making their leisure decision, households' are aware of the incomplete information they will face throughout their viewing regarding advertising. Furthermore, they know that they are not naive in their viewing, but learn throughout. The ex-ante expected maximal utility from television viewing in minute t, conditional on a viewing history that induces an information set Ω_t:

$$\upsilon_{td}(\Omega_{td}) = \mathbb{E}_{\varepsilon}\left\{\max_{j\in\mathcal{J}}\left[\alpha\mu_{jtd}(\Omega_{td}) + x_{jtd}\beta + \gamma g_p + \varepsilon_{jtd}\right]\right\} \tag{4.7}$$

The beliefs associated with a viewing path together with the known channel-time characteristics determine the choice probabilities for each alternative, $(\mathscr{s}_{0td}(\Omega_{td}), \ldots, \mathscr{s}_{Jtd}(\Omega_{td}))$. The expected value of watching television throughout a day is given by the sum of the values, weighted by their respective choice probabilities. Furthermore, the viewing paths are dependent on a realized ad sequence, which is unknown ex-ante, requiring integration over ad sequences consistent with the viewers' expectations:

$$\begin{aligned}\Upsilon_d(\mathbf{g}_d, \mathbf{x}_d, \boldsymbol{\lambda}_d) = \mathbb{E}_{\Lambda}\Bigg\{\upsilon_{1d} + \sum_{j_1\in\mathcal{J}} \mathscr{s}_{j_1 1d}\Bigg[\upsilon_{2d}(\Omega_{2d}) + \sum_{j_2\in\mathcal{J}} \mathscr{s}_{j_2 2d}(\Omega_{2d})\Bigg(\upsilon_{3d}(\Omega_{3d}) + \sum_{j_3\in\mathcal{J}} \mathscr{s}_{j_3 3d}(\Omega_{3d})\ldots \\ \ldots\upsilon_{T-1d}(\Omega_{T-1d}) + \sum_{j_{T-1}\in\mathcal{J}} \mathscr{s}_{j_{T-1}T-1d}(\Omega_{T-1d})\upsilon_{Td}(\Omega_{Td})\Bigg)\Bigg]\Bigg\}\end{aligned} \tag{4.8}$$

Calculating the summand in equation 4.8 for the entirety of a day directly will take more than a lifetime. In the empirical application I overcome this by averaging simulated expected utilities for many simulated ad sequences consistent with the viewers' beliefs. Households' expected value of watching television and value of participating in alternative activities, i.e. devoting all the individual's leisure time to non-television activities, are given by:

$$\begin{aligned} V_{TVd} &= \phi_V \Upsilon_d + \zeta_{TVd} \\ V_{0d} &= \phi_d + \zeta_{0d} \end{aligned} \tag{4.9}$$

The parameter ϕ_V captures the effect of the perceived value of watching television on the market size and ϕ_d is a day effect to be specified in the empirical section. The idiosyncratic taste shocks, ζ, are independent across alternatives, days and the within day taste shocks ε. The market size for a given day is determined by integration over households' choices:

$$\mathscr{s}_{TVd} = \int \mathbb{1}\{V_{TVd} \geq V_{0d}\} dF_{\zeta} \tag{4.10}$$

where F_{ζ} denotes the distribution of the taste shocks. The market size on day d is $M_d = \mathscr{s}_{TVd} \cdot M$, where M are the number of households owning a TV. Channel j's number of viewers at minute t on day d is $N_{jtd} = \mathscr{s}_{jtd} \cdot M_d$.

4.3 Advertiser demand

The demand for advertising is modeled similarly to that in Rysman (2004); Fan (2013). Differently from them, I assume impressions on all media outlets constitute perfect substitutes. There is a continuum of price-taking advertisers, whose demand for impressions on day d is:

$$D_d(r_d, \chi_d; \eta) = \chi_d r_d^{-\eta} \tag{4.11}$$

where r_d is the price per impression; χ_d is an advertiser demand shifter; and η is the advertiser demand elasticity. Assume an infinite elasticity of substitution between media outlets and between impressions on the different channels. The advertising equilibrium condition is:

$$\underbrace{D_d(r_d, \chi_d; \eta)}_{\text{demand for impressions}} = \underbrace{B_d + \mathbf{a}_d \cdot \mathbf{N}_d}_{\text{supply of impressions}}$$

where B_d are the impressions supplied by alternative media outlets and $\mathbf{a}_d \cdot \mathbf{N}_d = \sum_{j \in \mathcal{J}_m} \sum_{t=1}^{T} a_{jtd} \cdot N_{jtd}$ is the supply of impressions provided by the commercial channels. The resulting inverse advertiser demand is:

$$r_d = \left[\frac{\chi_d}{B_d + \mathbf{a}_d \cdot \mathbf{N}_d} \right]^{1/\eta} \tag{4.12}$$

4.4 Channels' supply

Commercial channels, $j \in \mathcal{J}_m$, choose the mean quality of each program and the amount of ads per episode with the goal of maximizing ad revenue.[21] Although content quality is hard to determine, channels can influence quality by choosing the investment in the content they provide whereas a higher investment will generate a stochastically higher quality program. For a given ad distribution, quality shifts viewership as well as the costs associated with content provision. Viewership shifts at the intensive margin —by reallocating market shares among alternatives —and the extensive margin —by changing the market size. The learning framework implies that increases in ad probabilities at any given minute will affect viewership in the minute itself but also in ensuing minutes.

[21]The public channels may have other goals not aligned with profit or viewership maximization. Multichannel alternatives' revenues are derived primarily from subscription fees as opposed to advertising.

4.4.1 Strategies

Channels' *quality strategy* consists of a vector of mean qualities for each of their programs, $\mathbf{g}_j = (g_p)_{\forall p \in \mathcal{P}_j} \in \mathbb{R}^{|\mathcal{P}_j|}$. Channels' *advertising strategy* is a vector consisting of the ad probabilities at each minute, $\boldsymbol{\sigma}_j = (\sigma_{jtd})_{\forall t,d} \in [0,1]^{T \times D}$. In total, each commercial channel's strategy vector consists of $T \times D$ advertising elements and $|\mathcal{P}_j|$ quality elements which they choose to maximize their expected payoff from advertising.

4.4.2 Payoffs

Channels' revenue is determined by the price per impression and the number of impressions. Content provision has an associated cost varying with the quality while there is no direct cost associated with advertising. Channel j's realized daily payoff on day d is:

$$\pi_{jd}(\mathbf{a}_d, \mathbf{q}_d; \boldsymbol{\lambda}_d, \chi_d) = \sum_{t=1}^{T} [\ \underbrace{a_{jtd} \cdot N_{jtd}(\mathbf{a}_d, \mathbf{q}_d; \boldsymbol{\lambda}_d)}_{\substack{\text{channel } j\text{'s supply} \\ \text{of impressions}}} \cdot \underbrace{r_d(\mathbf{a}_d, \mathbf{N}_d; \chi_d)}_{\substack{\text{price per} \\ \text{impression}}} - \underbrace{C\left(g_{jtd}\right)}_{\substack{\text{mean program} \\ \text{variable cost}}}\ ]$$

$$\text{where} \quad N_{jtd}(\mathbf{a}_d, \mathbf{q}_d; \boldsymbol{\lambda}_d) = s_{jtd}(\mathbf{a}_d, \mathbf{q}; \boldsymbol{\lambda}_d) \cdot M(\mathbf{g}_d; \boldsymbol{\lambda}_d)$$

The realized payoff is based on realizations of the channels' advertising strategies. The cost is a per-minute average cost of production, determined by the quality of the program being broadcast ($g_{jtd} = g_p\ \forall\ (t,d) \in \mathcal{T}_p$). The expected daily payoff is determined prior to ad realizations:

$$\Pi_{jd}(\boldsymbol{\sigma}_d, \mathbf{q}_d; \boldsymbol{\lambda}_d, \chi_d) = \mathbb{E}_{F_a(\boldsymbol{\sigma}_d)}\left[\pi_{jd}(\mathbf{a}_d, \mathbf{q}_d; \boldsymbol{\lambda}_d, \chi_d)\right] \tag{4.13}$$

Viewer learning framework implies that π_{jtd} relies on the entire advertising history up to minute t. Consequently, determination of expected payoff requires specification of joint advertising distribution (F_a). I assume that ads are allocated as a non-stationary 1st order Markov process in which the persistence parameter is exogenously determined and fixed across all minutes and days for a given channel.

$$\Pr\left(a_{jtd} = 1 | a_{jt-1d}, \sigma_{jtd}\right) = \sigma_{jtd} + \rho_j a_{jt-1d} \tag{4.14}$$

Therefore, channels' choice of the ad probability σ_{jtd} fully determines the joint advertising distribution.

Channels choose program qualities ex-ante, prior to observing quality realizations or the advertiser demand shifter. Therefore, they choose the quality to maximize their ex-ante payoffs. In doing so, they take into account the advertising policy function, i.e. they take into account that each day they will choose their advertising probabilities optimally.

The channels equilibrium advertising policy is denoted by σ^*:

$$\Pi_p(g_p, \sigma^*, \mathbf{g}_{-p}; \boldsymbol{\lambda}) = \mathbb{E}_{\xi,\chi}\left[\sum_{d\in\mathcal{T}_p} \Pi_{jd}(\sigma^*, \mathbf{g}_d; \boldsymbol{\lambda}_d, \boldsymbol{\xi}_d, \chi_d)\right] \tag{4.15}$$

4.5 Timing, information & equilibrium

4.5.1 Timing

The timing of the model is detailed below and illustrated in Figure 7:

1. Commercial channels choose mean program qualities, $\mathbf{g}_j$, determining quality distributions $\mathbf{G}_j$

For each $d = 1, \ldots, D$:

At $t = 0$:

2. Households make their leisure decision based on expected qualities and advertising, Υ_d, determining the daily market size, M_d
3. Advertiser demand shock, χ_d, and qualities, $\boldsymbol{\xi}_d$, are realized and observed by channels
4. Channels choose daily advertising, $\boldsymbol{\sigma}_d$, and ad prices are determined r_d

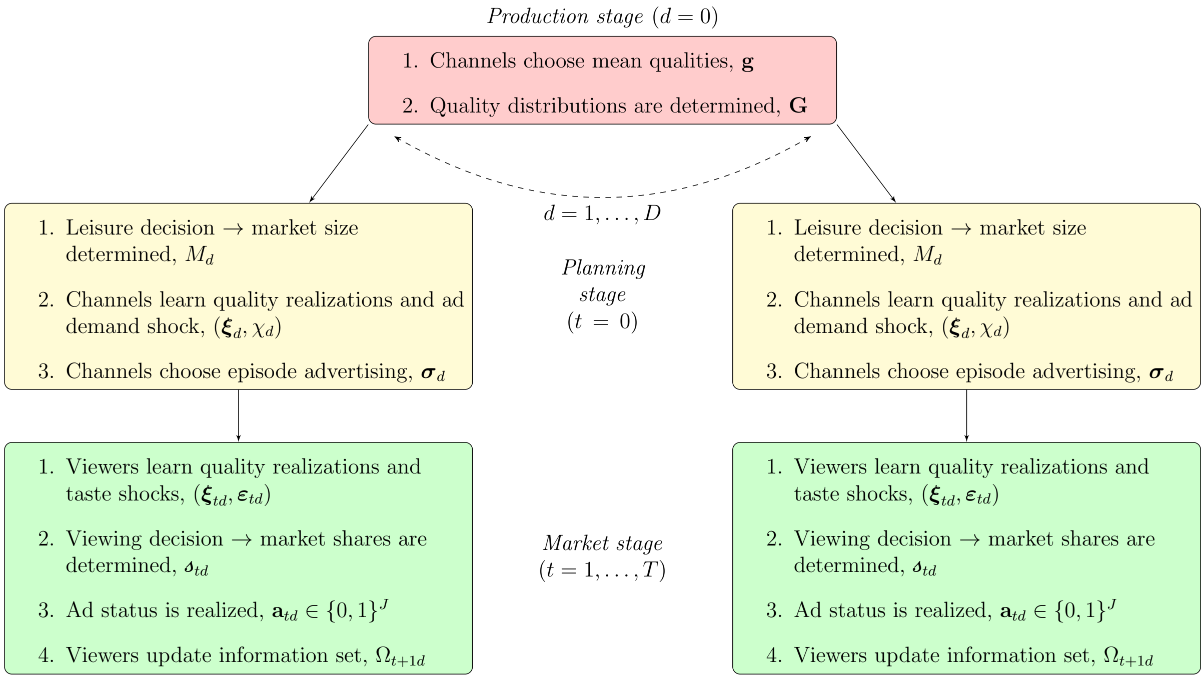
For each $t = 1, \ldots, T$:

5. Viewers learn qualities, $\boldsymbol{\xi}_{td}$, and taste shocks, $\boldsymbol{\varepsilon}_{td}$
6. Viewers make choices based on information set $y_{td}\left(\Omega_{td}, \boldsymbol{\varepsilon}_{td}\right)$ and market shares are realized $\boldsymbol{s}_{td}$
7. Ad state is realized $\mathbf{a}_{td}$ and viewers update information set, Ω_{t+1d}

4.5.2 Information

As apparent in the model timing, the knowledge in the model is such that all players - both channels and viewers - know the quality strategies and program quality distributions. At the beginning of each day, prior to making their daily advertising choices, channels learn the realized quality of all episodes on all channels while viewers learn the realized quality only after making their leisure decisions. Viewers' common prior regarding the ad probabilities is consistent with the program level advertising. I.e. viewers know the quality and associated expected amount of ads on each program. This is consistent with viewers' knowing channels' equilibrium advertising policy rule mapping quality to

Figure 7: Model timing



advertising. With that, viewers' perception of the ad process differs form the channels' advertising policy in two regards. First, viewers don't know the advertising demand. Furthermore, viewers don't update their assessment of the ad probability once learning the quality realizations. This can be rationalized by viewers not knowing the quality in later minutes and therefore how channels will choose to allocate ads across the different episodes broadcast within a day. Finally, channels know the advertising strategies of each other but not the specific realization of ads throughout a day.

4.5.3 Equilibrium

The commercial channels are subject to a binding advertising constraint, $\bar{a}$. Although the total amount of ads is exogenously determined per day, the channels account for episode qualities in their ad allocation across episodes. I will concentrate on a symmetric equilibrium. The channels' equilibrium advertising policy is a function mapping realized qualities, ad demand and an ad quantity constraint to minute level advertising probabilities:[22]

$$\sigma^*(\mathbf{q}, \chi; \bar{a}) = \arg\max_{\boldsymbol{\sigma}_{jd}} \Pi_{jd}(\boldsymbol{\sigma}_d; \mathbf{q}_d, \chi_d) \quad s.t. \quad \underbrace{\sum_{t=1}^{T} \sigma_{jtd} \leq \bar{a}}_{\text{ad quantity constraint}} \tag{4.16}$$

[22] A full characterization of the equilibrium advertising strategies is available upon request.

Optimal quality choice is determined by:

$$\underbrace{\mathbb{E}_{F_a(\sigma^*)}\left\{ |\mathcal{T}_p|^{-1} \sum_{d \in \mathcal{T}_p} \sum_{t=1}^{T} a_{jtd} \cdot \left[\overbrace{\frac{dN_{jtd}}{dg_p} \cdot r_d}^{\text{viewer demand}} + \overbrace{N_{jtd} \cdot \frac{dr_d}{dg_p}}^{\text{advertiser demand}} \right] \right\}}_{\text{expected average per minute marginal revenue from program } p} = \underbrace{c_p(g_p)}_{\substack{\text{per minute marginal} \\ \text{cost of program } p}}$$

where

$$\frac{dN_{jtd}}{dg_p} = \underbrace{\frac{\partial M_d}{\partial g_p} \cdot s_{jtd}}_{\text{market expansion effect}} + \underbrace{M_d \cdot \frac{\partial s_{jtd}}{\partial g_p}}_{\text{ad avoidance effect}} \tag{4.17}$$

Equation 4.17 provides insight into the equilibrium relationship between advertising and quality. Specifically, for any ad probabilities there exists a corresponding quality level that maximizes channels' payoff. Furthermore, equation 4.17 equates the marginal cost of content provision with the program's expected marginal revenue. The marginal cost varies with the observed program characteristics, w_p, program quality, and unobserved program characteristics, ω_p, according to:

$$\ln\left[c(g_p)\right] = w_p \kappa_w + \kappa_g g_p + \omega_p \tag{4.18}$$

5 Empirical analysis

Table 5 provides an overview of the model components that need to be estimated, the data source used to estimate each component and the estimation method. Identification of the model components is discussed in the respective sections.

Table 5: Overview of empirical analysis

Model component		Parameters	Data source(s)	Method
(1) Viewer demand	(1.1) Beliefs	Λ	Broadcasting data	Logit
	(1.2) Preferences	α, β, γ	Viewer shares & broadcasting data	Simulated moments (MSM)
	(1.3) Leisure	ϕ_V, ϕ_d	Viewer shares supplemental data	Logit
(2) Advertiser demand		η, χ	Ad prices & supplemental ad data	OLS
(3) Production variable costs		κ_g, κ_w	Model outputs	First-order condition (FOC)

5.1 Viewer demand

Estimation of the viewer demand parameters follows a two-step procedure. In the first step, viewers' common priors are estimated from the channels' broadcasting behavior. In the second step, the viewing utility parameters are estimated using a method of simulated moments (McFadden (1989); Pakes and Pollard (1989)).

5.1.1 Identification

Identification of viewers' response to advertising makes use of the dynamic nature of the regulation to construct a *shadow cost shifter* that varies the amount of advertising in an episode independently of its quality and viewer preferences. Specifically, the advertising quantity constraint links channels' advertising decisions among the episodes broadcast each day, beyond the mutual effects arising from common ad prices and the market expansion effect.[23] This shadow marginal cost identification is illustrated with a simple example. Consider a day in which a channel broadcasts episodes of equal length from two programs, $p = 1, 2$. The channel's maximization problem is:

$$\max_{\sigma_1,\sigma_2}(\sigma_1 \cdot N_1 + \sigma_2 \cdot N_2) \cdot r(\sigma_1 + \sigma_2) \quad s.t. \quad \sigma_1 + \sigma_2 \leq \bar{a}$$

Consider a binding advertising constraint, $\sigma_1 + \sigma_2 = \bar{a}$. Hence the ad price is unaffected by the channel's choice of ads, $r(\sigma_1+\sigma_2) = r(\bar{a})$ and the channel chooses how to distribute the ads across the two episodes. Advertising on each episode is determined by:

$$\underbrace{N_p + \sigma_p \cdot \frac{\partial N_p}{\partial \sigma_p}}_{\text{marginal revenue}} = \underbrace{\varpi}_{\substack{\text{shadow}\\ \text{marginal cost}}} \qquad \Rightarrow \qquad N_1 + \sigma_1 \cdot \frac{\partial N_1}{\partial \sigma_1} = N_2 + \sigma_2 \cdot \frac{\partial N_2}{\partial \sigma_2} \tag{5.1}$$

I.e. the marginal revenue from advertising is equal on the episodes from both programs. Examine the comparative static resulting from equation 5.1. Consider two days: d and d' with $q_{1d} = q_{1d'}$ and $q_{2d} > q_{2d'}$. On both days the ad quantity constraint is binding, $\sigma_{1d} + \sigma_{2d} = \sigma_{1d'} + \sigma_{2d'} = \bar{a}$. According to the condition in 5.1: $\sigma_{1d} < \sigma_{1d'}$ and $\sigma_{2d} > \sigma_{2d'}$.[24] Hence, episode 2 quality affects advertising on episode 1 through the ad quantity constraint without affecting viewer behavior, i.e. $\varpi \perp \xi$. As such, the quality of later episodes or dynamics in the usage of the ad quota can be used as shadow cost shifters

[23] Households' leisure decision generates television viewing persistence. In this form, channels' choices influence all minutes within a day.

[24] The formal conditions under which this holds are:

1. $\frac{\partial N_p}{\partial g_p} > \frac{\partial N_p}{\partial g_{-p}} \geq 0$, $\frac{\partial N_p}{\partial \sigma_p} < 0$ and $\frac{\partial^2 N_p}{\partial \sigma_p \partial g_p} \geq 0$; or
2. $\frac{\partial N_p}{\partial g_p} \geq \frac{\partial N_p}{\partial g_{-p}} > 0$, $\frac{\partial N_p}{\partial \sigma_p} < 0$ and $\frac{\partial^2 N_p}{\partial \sigma_p \partial g_p} > 0$.

Note that this could also hold under oppositely signed conditions.

to identify viewers' sensitivity to ads. This identification strategy is similar in nature to that used in Dubois et al. (2018).

Identification of viewers' response to quality is derived from the timing of the model (Ackerberg and Crawford (2009)). Channels choose the mean quality of each program generating a distribution of episode qualities. In each day, the episode quality is pre-determined and not affected by channels' choices, whilst viewers respond to the pre-determined episode qualities. Hence, the timing difference between channels' quality decisions and viewers' response to quality identifies the viewers' responses to quality.

5.1.2 Estimation

The utility parameters are estimated via the method of simulated moments where the basis of the moments is the structural term (ξ_{jt}) as proposed by Berry (1994). The individual expectations form a viewer specific component ($\alpha\mu_{ijt}$) and 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). To allow simulated viewers to develop differing information sets, I run the model over the whole two hours of prime-time while excluding the first 10 minutes, i.e. 20:00-20:10, from the estimation of the non-linear parameter (α) (Ching et al. (2013)). Program qualities are estimated by decomposing the structural error term into a program specific component and a residual component (Nevo (2000)).[25]

The individuals' expectations introduce a dynamic component in the viewers' choices, requiring the model be solved via forward simulation.[26] That is, for fixed $\hat{\boldsymbol{\lambda}}$, and a generic parameter vector, $\theta = (\alpha, \beta, \gamma)$, we begin by solving the utilities for the first period. Lacking viewing histories, the beliefs equate across all simulated individuals and all heterogeneity across individuals is driven by the taste shocks, $\boldsymbol{\varepsilon}_{j1}$. Upon generating the first period's choices, individuals begin to gather a viewing history, thereby creating differing information sets and heterogeneous beliefs. The expected utilities at $t = 1$ and $t > 1$:

$$\begin{aligned} U_{ij1} &= \alpha\hat{\lambda}_{j1} + \delta_{j1} + \varepsilon_{ij1} \\ U_{ijt} &= \alpha\hat{\mu}_{ijt} + \delta_{jt} + \varepsilon_{ijt} \quad \forall \quad t = 2, ..., T \end{aligned} \tag{5.2}$$

where $\hat{\lambda}_{j1}$ is the unconditional ad probability at the beginning of the evening and $\hat{\mu}_{ijt} =$

[25]In practice, this is done by estimating a program specific parameter for each program broadcast by the channels. As noted in Nevo (2000), this may induce identification issues. These issues are resolved in this application through some programs being broadcast on more than one channel, allowing me to identify both channel specific effects along with program effects.

[26]Initially I simulated $ns \times D \times T \times J$ taste shocks ε using Halton draws to reduce simulation error for $ns = 1000$ households after discarding the first 500 draws (Train (2009); Reynaert and Verboven (2014)).

$\mu_{ijt}(\Omega_{ijt}; \hat{\boldsymbol{\lambda}})$. The viewing choice of simulated viewer i in period t is determined by:

$$y_{it}(\theta; \hat{\boldsymbol{\lambda}}) = \arg\max_{j \in \mathcal{J}} \{U_{i0t}(\theta; \hat{\boldsymbol{\lambda}}), ..., U_{iJt}(\theta; \hat{\boldsymbol{\lambda}})\} \tag{5.3}$$

Assume the taste shocks follow a type 1 extreme value distribution, hence the choice probability of each simulated viewer follow a logistic distribution. The channels' simulated choice probabilities in each period are given by integrating over the choice probabilities of the ns simulated viewers:

$$\mathscr{s}_{jt}(\delta_t, \alpha) = \frac{1}{ns} \sum_{i=1}^{ns} \mathscr{s}_{ijt}(\delta_t, \hat{\mu}_{it}, \alpha) = \frac{1}{ns} \sum_{i=1}^{ns} \frac{e^{\delta_{jt} + \alpha \hat{\mu}_{ijt}}}{\sum_{k \in \mathcal{J}} e^{\delta_{kt} + \alpha \hat{\mu}_{ikt}}} \tag{5.4}$$

The contraction mapping isolates the structural error ξ_{jt} by solving the system of equations $\mathscr{s}(\delta) = S$, for any generic value of α where S are the observed market shares. The contraction mapping requires iterating over the series:[27]

$$\delta^{\ell+1} = \delta^{\ell} + \ln\left[S\right] - \ln\left[\mathscr{s}(\delta^{\ell})\right] \quad \forall \quad \ell = 0, ..., L$$

Until the difference between two successive iterations is smaller than a pre-defined tolerance level:

$$\sup\left|\delta^{L} - \delta^{L-1}\right| \leq \epsilon$$

Finally, the structural error term is defined as:

$$\xi_{jt} = \delta_{jt} - \gamma_p - x'_{jt}\beta \tag{5.5}$$

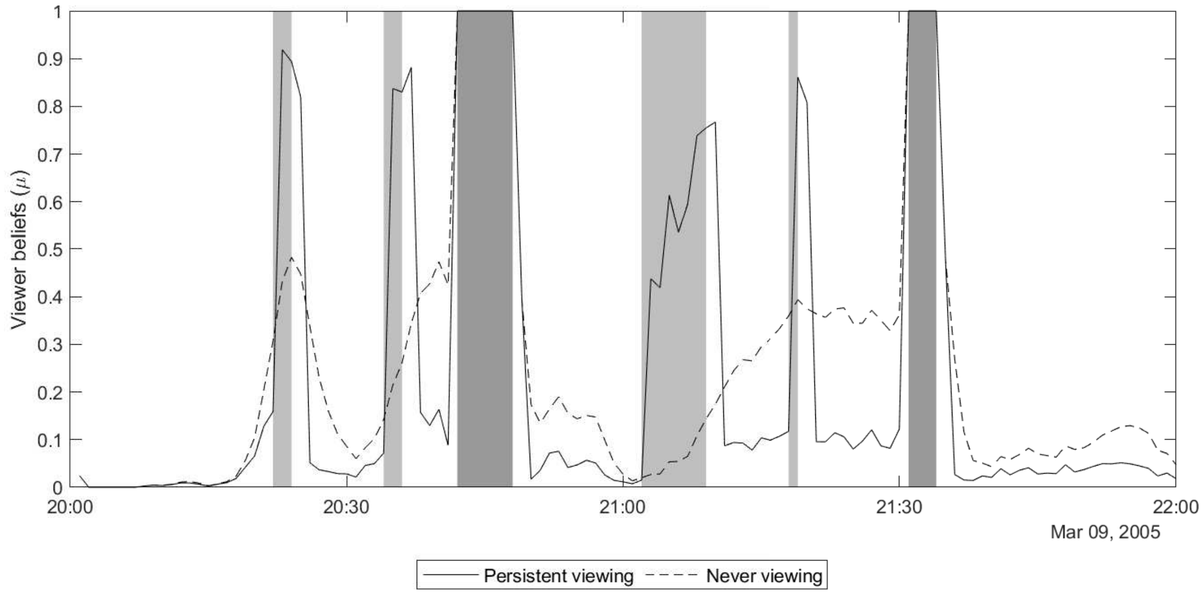
The structural shocks were stacked to a $D \cdot (T - 10) \cdot J \times 1$ vector $\xi(\theta)$. Define the set of moments corresponding to the orthogonality of the structural error term from a set of instruments, Z, as $m(Z, \xi(\theta)) = E[Z'\xi(\theta)]$. The identification assumption is that $m(Z, \xi(\theta^0)) = 0$ at the true parameter vector, θ^0. The instruments used are regulation based variables[28] and Berry et al. (1995) instruments.[29] The resulting MSM estimates

[27] I iterated over the exponentiated series which saved computation time. This affected the stopping criteria to $\sup\left|e^{\delta^{L} - \delta^{L-1}} - 1\right| \leq \epsilon$.

[28] Specifically, the variables which were used are: (a) the remaining amount of commercials left for each channel to broadcast within prime-time; (b) the remaining number of breaks left for each channel to broadcast within prime-time.

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

Figure 8: Estimated viewer beliefs



Notes: The figure portrays the possible range of viewers' beliefs along each minute of channel 2 broadcast. The horizontal axis displays the minutes within the day and the vertical axis the perceived probability of an ad at each minute. The shaded areas indicate commercial broadcast and the dark shading indicate between program commercial breaks.

are the solution to:[30]

$$\hat{\theta}_1 = \arg\min_{\theta} \bar{m}(Z, \xi(\theta))' W \bar{m}(Z, \xi(\theta))$$

where $\bar{m}(\cdot)$ is the sample counterpart of $m(\cdot)$. The initial weighting matrix was computed as $W^0 = \mathbb{E}[ZZ']$. The efficient weighting matrix was computed as $W^e = \mathbb{E}\left[\bar{m}(Z, \xi; \hat{\theta}_1)\bar{m}(Z, \xi; \hat{\theta}_1)'\right]$. Standard errors were computed using the methods in Newey and McFadden (1994).

5.1.3 Results

Figure 8 compares the ad beliefs between two extremes, an individual who views a channel continuously and an individual who never views the channel. All commercial probabilities will lie in between the two extremes.[31] The figure illustrates the informational advantage of continuous viewing. The common priors underlying the beliefs are visible in the co-movement of the two lines. The core informational advantage in viewing arises during ad breaks within a program, while the beliefs are similar across viewers throughout between-program commercial breaks. Continuous viewing displays an initial lagged spike in the ad probability during ad breaks and a lagged decline, similarly to Figure 2.

[30]Following Nevo (2000), the parameter search was conducted only over α. For a given parameter estimate α, the rest of the parameters are estimated with a weighted least squares procedure $\hat{\beta} = [X'ZWZ'X]^{-1}[X'ZWZ'\delta(\alpha)]$ where the weights are the same as those used in the moment value function, i.e. $W = \{W^0, W^e\}$.

[31]In estimating the common beliefs , the dependent variable is the ad state on channel j at minute t of day d (a_{jtd}), and the independent variables were the lagged ad state (a_{jt-1d}), program dummies, the percentage of the program and minute dummies. The link function is a Logit function.

The results from estimation of the viewer demand model are presented in Table 6. The first column, *Perfect foresight (OLS)*, refers to a model with perfect foresight in which the ad status of the channels is known to the viewers at the time of making their viewing decision. The second specification, *Perfect foresight (2SLS)*, extends the perfect foresight model by accounting for the endogeneity while maintaining perfect foresight of the viewers. The third column, labeled *Incomplete information* incorporates the estimated ad probabilities that constitute the common priors without a learning framework. In this specification, all viewers are uniformly aware of the previous ad state on all channels. The final column, *Learning*, is based on the learning model with incomplete information detailed in section 4. The same instruments are used in the latter three models. Finally, the number of days used in the empirical exercise is less than the whole year, since only days in which all prime-time minutes on the three channels are observed.

The magnitude of the estimate for the sensitivity to ads increases by an order of magnitude when controlling for endogeneity using the instruments, implying the existence of an attenuation bias in the estimate resulting from simultaneity.[32] By construction, the 2SLS procedure in column (2) transforms the binary ad state to a continuous measure, resulting in a slight decrease in the sensitivity to ads in the transition to column (3), the incomplete information specification. When accounting for the dynamic nature of viewing through learning, the estimate increases in magnitude substantially. While the parameter estimates of the other variables also exhibit changes between specifications, $\hat{\alpha}$ is the most substantial and clearly shows the economic importance of controlling for the heterogeneous information in estimating viewers' relation towards ads. This result quantifies the intuition from Figure 2 that portrayed the percentage change in the market share following the beginning of an ad break. Specifically, the figure illustrated that the change in market shares exhibits an initial sharp drop that recuperates gradually, also after the end of the ad break, consistent with asymmetric information due to individual information. The *% program* variable implies that viewers are more committed to a program the farther along it is, this is most notable in the final specification but is also apparent in the former specifications. As expected from the popularity of the channels measured by their respective market shares, the utility parameter for channel 2 is positive

[32]Demand estimation across a larger sample period ranging from the January 1$^{\text{st}}$, 2001 and throughout 2005 yield very similar estimates to those presented in the first column of Table 6 with an estimate for the ad sensitivity parameter of -.238 (and a standard error of .002). Throughout this time frame three regulatory changes came into effect. First, channel 10 entered, commencing broadcasrt on January 28$^{\text{th}}$, 2002. Furthermore, on May 22$^{\text{nd}}$, 2002 the ad regulation changed allowing commercial channels to allocate their ads across the two hours of prime-time at their discretion, i.e. 24 minutes across two hours as opposed to 12 minutes per hour. Finally, on August 18$^{\text{th}}$, 2003 a unified cable operator - HOT - began operating throughout Israel. Incorporating these instruments in addition to the instruments used in the second column of Table 6 shows the efficacy of the employed identification strategy. The additional instruments - all of which are policy based - provide no additional identification power, changing the parameter estimate of viewers' sensitivity to ads only slightly, yielding a parameter estimate of -.919 (with a corresponding standard error of .016).

Table 6: Viewer demand results

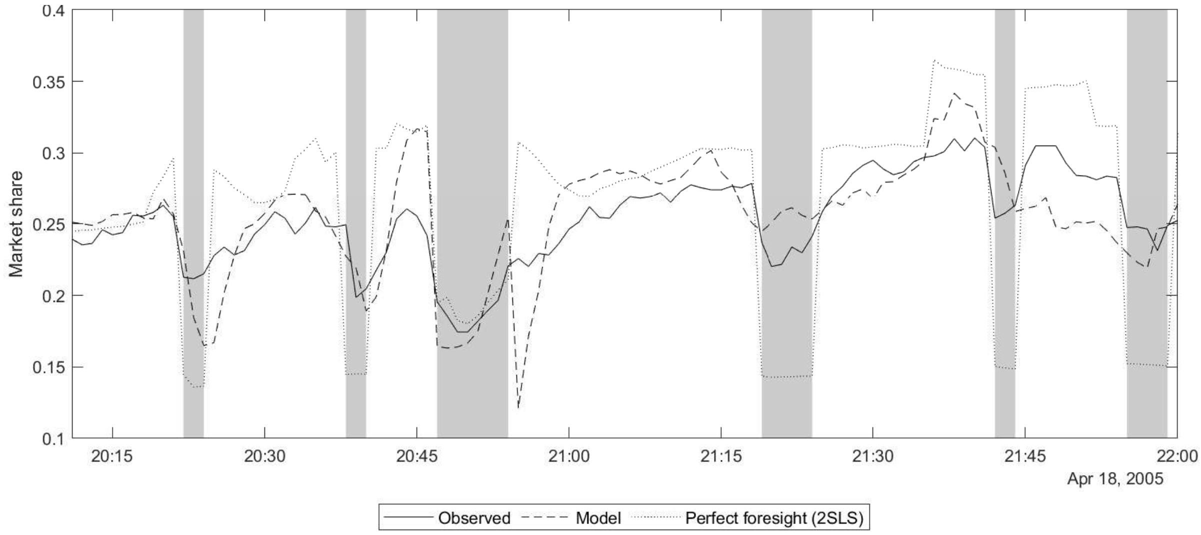
Variable	Perfect foresight OLS (1)	Perfect foresight 2SLS (2)	Incomplete information 2SLS (3)	Learning MSM (4)
Commercial (α)	-0.239	-0.957	-1.050	-2.753
	(0.002)	(0.021)	(0.017)	(0.010)
% program	0.087	0.180	0.184	0.275
	(0.004)	(0.006)	(0.005)	(0.010)
Channel 2	1.109	1.106	1.102	1.067
	(0.007)	(0.010)	(0.008)	(0.007)
Channel 10	-0.267	-0.255	-0.253	-0.234
	(0.009)	(0.012)	(0.011)	(0.007)
Episode duration				
Between 30 min. and 1 hour	0.017	0.021	0.023	0.039
	(0.008)	(0.010)	(0.009)	(0.010)
Between 1 and 1.5 hours	0.043	0.062	0.053	0.039
	(0.008)	(0.011)	(0.010)	(0.011)
Between 1.5 and 2 hours	0.033	0.056	0.053	0.061
	(0.013)	(0.017)	(0.015)	(0.014)
Longer than 2 hours	-0.022	0.131	0.038	-0.116
	(0.025)	(0.033)	(0.029)	(0.022)
Rerun	-0.134	-0.497	-0.264	0.003
	(0.008)	(0.015)	(0.010)	(0.015)
Constant	-2.399	-1.744	-1.676	-0.154
	(0.018)	(0.030)	(0.025)	(0.010)
Time FE	+	+	+	+
Date FE	+	+	+	+
Program values (γ) (IQR)	[-0.444, 0.311]	[-1.090, -0.299]	[-1.194, -0.408]	[-2.884, -1.920]
RMSE	0.027	0.038	0.040	0.038
RMSE commercial channels	0.029	0.044	0.046	0.040
Obs. (N)	118800	118800	118800	118800

Notes: Table 6 presents the estimated viewer preference parameters. The models include 632 *Program value* parameters. Specification (4) used 108,900 observations in estimating viewers' sensitivity to ads ($\hat{\alpha}$) and the whole dataset in estimating the remaining parameters. The GMM procedure in column (4) used 1,094 moment conditions: 1,089 own broadcast characteristics + 3 cross characteristics (Berry et al. (1995) instruments) + 2 regulation based instruments. Standard errors in parenthesis.

while that of channel 10 is negative.

The viewer demand specification from the learning model (column (4) in Table 6) exhibits a good fit to the observed market shares. The RMSE of the learning model fits the data better than the other two models that account for the endogenity, and provides a substantially better fit with regards to the commercial channels. Figure 9 compares the observed market shares, those predicted by the learning model, and those from a perfect foresight model for a specific day within the data. The observed market shares exhibit a lot of variation throughout an evening complicating the model fit. The model does a good job of capturing the fluctuations in the market share throughout the evening as well as the dips in market shares during ad breaks. This figure illustrates

Figure 9: Viewer demand model fit



Notes: The figure displays the observed and the predicted market shares of channel 2. The model's RMSE of channel 2 on the specific date is 0.0273. The RMSE of the perfect foresight specification on the day is 0.05. The shaded areas indicate commercial broadcast.

the different predictions that arise from the perfect foresight model and a learning model during ad breaks. The perfect foresight model assumes viewers perfectly allocate their viewing during ads implying an immediate and constant drop in market share during an ad break. Furthermore, the market share fully recuperates at the end of the ad break. Alternatively, the learning model is able to generate a smoother recuperation of market shares during and after ad breaks, making it better aligned with Figure 2.

The quality measure implied by the model is a distribution of qualities for each of the programs broadcast on the three channels. Assessing the validity of this measure is difficult, since quality is not salient in this industry. As a case study, I compare the qualities of films implied by the model to several external rankings: Google, IMDb, Rotten Tomatoes, Metacritic, and Seret.[33] Table 7 provides the correlation of the estimated quality measure with the film rankings. The comparison set is small, ranging from a maximum of 71 comparisons with the IMDb ranking and up to 34 with the Seret ranking. The correlations are all positive, very strong, and statistically significant. The relationships imply that a one percent increase in the IMDb ranking is associated with a .87% increase in the estimated quality measure. These relationships provide support for the validity of the estimated quality measure.

[33]These rankings are loosely correlated amongst themselves, therefore are likely to capture different preference aspects. E.g. the correlations of the Google ranking with the remaining rankings are: IMDb (.15); Rotten Tomatoes (.016); Metacritic (-.041); Seret (.189).

Table 7: Comparison of quality measure to film rankings

	Google	IMdB	Rotten	Metacritic	Seret
Film ranking	0.669***	0.867***	0.767***	0.957***	0.738***
	(0.048)	(0.068)	(0.105)	(0.105)	(0.064)
Film age	-0.004	-0.005	-0.001	-0.006	-0.008
	(0.004)	(0.004)	(0.006)	(0.005)	(0.006)
Local film	0.214	0.034	0.003	-0.086	-0.046
	(0.176)	(0.107)	(0.330)	(0.295)	(0.221)
Adjusted R^2	0.860	0.835	0.686	0.752	0.855
Observations N	67	71	66	65	34

Notes: The table presents the correlation of the estimated quality measure to film rankings. The dependent variable is the estimated quality ($\hat{\gamma}$) derived from the learning model (column (4) in Table 6). The title of each column details the ranking used in the *film ranking* variable. Both film rankings and estimated quality were normalized to $[0, 1]$. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.2 Leisure decision

Once the common priors and the viewing preference parameters have been estimated, we can form an estimate for the expected value from watching television at each minute. The type 1 extreme value distribution of the taste shocks results in a closed form expression for the maximal expected utility defined in equation 4.7, i.e. the inclusive value:

$$\upsilon_t(\Omega_t) = \mathbb{E}_\varepsilon \left\{ \max_{j \in \mathcal{J}} \left[\alpha \mu_{jt}(\Omega_t) + x_{jt}\beta + \gamma g_p + \varepsilon_{jt} \right] \right\} = \log \left[\sum_{j \in \mathcal{J}} e^{\alpha \mu_{jt}(\Omega_t) + x_{jt}\beta + \gamma g_p} \right] \tag{5.6}$$

Averaging over the inclusive value of all simulated viewers allows me to integrate out the viewer specific viewing paths, generating an estimate for the unconditional inclusive value. The expected value of watching television in equation 4.8 was constructed using a pure frequency estimator, i.e. by averaging over many simulated ad sequences consistent with viewers expectations:

$$\Upsilon(\lambda) = \mathbb{E}_{F_a(\lambda)} \left[\Upsilon(\mathbf{a}, \lambda) \right] = \mathbb{E}_{F_a(\lambda)} \left[\sum_{t=1}^{T} \underbrace{\frac{1}{ns} \sum_{i=1}^{ns} \upsilon_t(\Omega_t)}_{= \upsilon_t(\mathbf{a})} \right] \tag{5.7}$$

There are two aspects of uncertainty at the outset of each day: with respect to quality and to ads. The corresponding value of watching television under each of the informational specifications is detailed in section A. Incomplete quality information reduces households' value of watching television by up to 15% relative to the perfect foresight scenario (full quality information and perfect ad foresight) with a median loss of approximately 4%. Incomplete information regarding advertising drives a larger wedge in viewers expected value of watching television, with a median loss of 6% and very few cases in which the value is greater than the perfect foresight scenario. This implies that the main loss in value from watching television results from imprecise allocation of viewing time among alternatives and to a lesser extent from the incomplete quality information.

Using the simulated value of watching television, the leisure equation in 4.10 is then estimated with various controls affecting the value of alternative leisure activities. Table 8 presents the parameter estimates for households' leisure decision. Models (1) - (3) use the expected inclusive value, i.e. $\hat{\Upsilon}(\lambda, g)$, as the informational framework for calculating the *TV viewing value*. Model (4) uses full quality knowledge, i.e. $\hat{\Upsilon}(\lambda, q)$. Model (5) transitions to perfect ad foresight with incomplete quality information and finally model (6) incorporates both perfect ad foresight as well as full quality information.

The effect of the expected inclusive value of the market share of households viewing television each day, $\hat{\phi}_V$ is in the expected direction, i.e. positive and statistically significant. Furthermore, the value is consistent across specifications. This implies that channels providing higher quality programming or less advertising have not only a business stealing effect but also a market expansion effect. In all models the temperature has a negative effect on households' television viewing. Meaning that on warmer evenings, people prefer engaging in alternative activities. Meanwhile, precipitation decreases the value of alternative leisure activities, increasing television viewing. With that, these effects are statistically significant only in model (2).

5.3 Advertiser demand

The advertiser inverse demand in equation 4.12 can be rewritten as:

$$\ln(r_d) = \tilde{\chi}_d - \frac{1}{\eta} \ln\left(B_d + \mathbf{a}_d \cdot \mathbf{N}_d\right) + \nu_d \tag{5.8}$$

where $\tilde{\chi}$ captures the observed variation in advertiser demand, e.g. time of year, while ν is the unobserved variation in advertiser demand, e.g. ad campaign due to roll-out of new product. The outer supply (B) includes the supply of advertising in online, print, radio and outdoor (billboards). Assume that the outer supply is fixed, i.e. $B_d = B$. As such, all variation in ad prices across days is due to variation in demand shocks $(\tilde{\chi}, \nu)$ and channels' supply of impressions $(\mathbf{a} \cdot \mathbf{N})$.[34] The price per impression was constructed as the weighted average daily price per impression across all ads where the weights are the relative length of each ad. To construct a measure of the value of the outer advertising supply (B), I use the Israeli Marketing Association's yearly reports for 2004 & 2005.

[34] I use the supply of impressions (and ads) during prime-time while the total supply of impressions is determined throughout the whole day. Industry professionals have stressed that the lion share of ad revenues are generated during prime-time and they, as well as advertising agencies rely only on the prime-time advertising. This trend is also apparent in the ad prices which are greater by an order of magnitude during prime-time.

Table 8: Leisure choice results

Variable	(1)	$\Upsilon(\boldsymbol{\lambda}, \boldsymbol{\gamma})$ (2)	(3)	$\Upsilon(\boldsymbol{\lambda}, \boldsymbol{\delta})$ (4)	$\Upsilon(\mathbf{a}, \boldsymbol{\gamma})$ (5)	$\Upsilon(\mathbf{a}, \boldsymbol{\delta})$ (6)
TV viewing value (ϕ_V)	0.005	0.007	0.008	0.008	0.008	0.008
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Rain (0/1)	-	-0.001	0.007	0.005	0.009	0.005
		(0.016)	(0.018)	(0.018)	(0.018)	(0.018)
Rain (mm)	-	0.005	0.006	0.006	0.006	0.006
		(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Temperature (C°)	-	-0.015	-0.002	-0.003	-0.002	-0.003
		(0.001)	(0.004)	(0.004)	(0.004)	(0.004)
Constant	-0.695	-0.709	-0.916	-0.973	-0.938	-0.994
	(0.067)	(0.064)	(0.139)	(0.137)	(0.139)	(0.137)
Weekday FE	-	+	+	+	+	+
Week FE	-	-	+	+	+	+
Adjusted R^2	0.349	0.522	0.564	0.582	0.570	0.585
Obs. (N)	330	330	330	330	330	330

Notes: The table displays the parameter estimates for households' leisure decision delineated in equations 4.9 and 4.10. The dependent variable is the share of households watching television each day (based on equations 3.2 and 3.1). The value of the inclusive values for each day in specifications (1)-(3) (the expected value of watching television) were calculated according to equation 5.7. The exact formulas used to calculate the inclusive values in specification (4)-(6) are detailed in Section A. The inclusive value in all specifications was averaged over 100 ad realizations consistent with observed ad probabilities. Standard errors in parentheses.

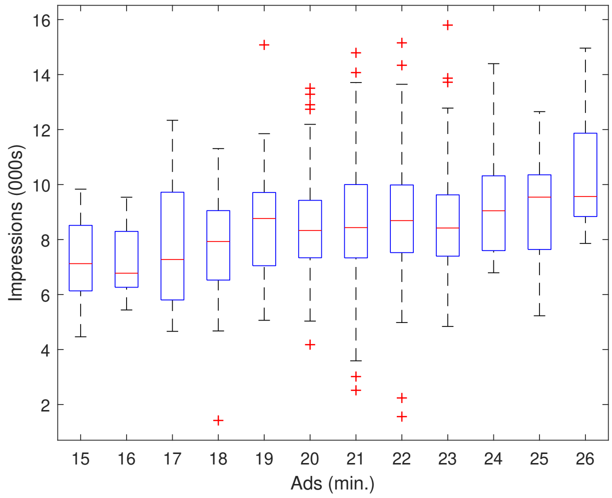
These reports detail the total advertising expenditure in Israel and the share of revenues derived from each medium.

The supply of impressions may be correlated with unobserved components making them endogenous. To overcome this, consider a day in which the advertising constraint is binding for both channels, $\mathbf{a}_2 = \mathbf{a}_{10} = \bar{a}$. Variation in supply of impressions by the channels is only determined by episode quality that is pre-determined and is therefore exogenous, allowing identification of the advertiser demand elasticity. Hence, I use a subsample of the days in which the ad constraint was binding for both channels to estimate the advertiser demand elasticity, $\hat{\eta}$.

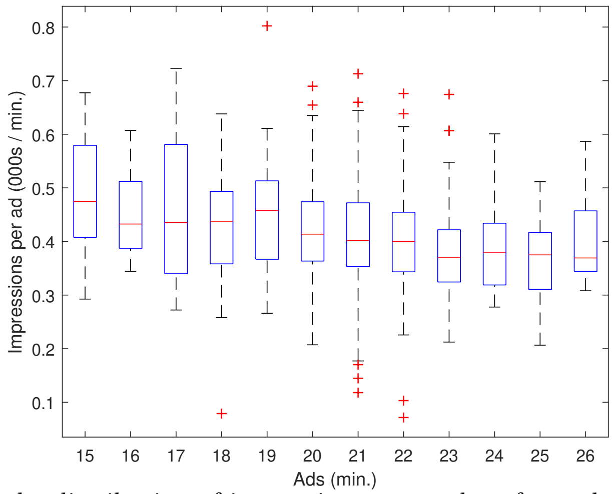
Figure 10 presents the distribution of impressions for each amount of advertising across days, Figure 10a presents the total amount of impressions throughout the prime-time while Figure 10b displays the average amount of impressions per ad. The most important characteristic arising from Figure 10a is that for any amount of ads, there is a large variation in the amount of impressions supplied by the channels across days. This variation, specifically in the upper tail of the ad distribution will be used to identify the advertisers' demand elasticity. Figure 10a also exhibits a positive correlation between ads and impressions, while this is in part mechanical - more ads implies more impressions

Figure 10: Distribution of impressions

(a) Total impressions



(b) Impressions per ad



Notes: Figure 10 display the distribution of impressions across days for each daily level of advertising, denominated in minutes. Figure 10a displays the total impressions in thousands and Figure 10b presents the average impressions, in thousands, per ad.

by construction - it is also indicative of the endogeneity by which channels advertise more in higher quality programs, which attract more viewers and thereby generating more impressions. Figure 10b complements by presenting the average impressions per ad for different advertising levels. Impressions per ad partly corrects for this simultaneity, exhibiting a negative correlation.

Table 9 presents the results from the estimation of advertisers' demand. In all cases the advertiser demand elasticity is significant. Models (1) and (2) do not impose any restrictions on the sample while models (3) and (4) impose a selection criteria according to the identification strategy. Specifically, these two models focus on the days in which the amount of ads is bounded by the advertising constraint. I use a constraint below the formal 24 minutes since ads are broadcast in blocks and therefore days broadcasting less than 24 minutes can still face a binding ad constraint. The results show that the identi-

Table 9: Advertiser demand results

Variable	(1)	(2)	(3)	(4)
Elasticity (η)	0.570	0.498	0.712	0.704
	(0.036)	(0.036)	(0.106)	(0.141)
Constant	24.204	26.876	20.538	20.712
	(1.143)	(1.513)	(2.160)	(2.950)
Time FE	-	+	-	-
Selection criterion	-	-	22+	22.5+
Obs. (N)	723	723	114	70

Notes: The data used in the estimation is daily *price per impressions* (PPI) and the number of impressions supplied in the market across the years 2004 and 2005. The PPI was calculated as the weighted price across all ad prices throughout prime-time. The impressions supplied by alternative media outlets was constructed using yearly reports by the Israeli Marketing Association for 2004 and 2005. The selection criteria for a level C (e.g. $C = 22$ minutes) are defined as $\sum_{j=2,10} \mathbf{a}_{jd} \geq C$. The elasticity standard errors were calculated using the delta method.

fication strategy eases the attenuation bias of the simultaneity, increasing the magnitude of the elasticity estimate by approximately 40%. The small number of observations in the restricted samples do not allow me to incorporate time controls.

5.4 Quality marginal costs

Equation 4.17 provided an equilibrium condition for quality choice based on demand fundamentals and the marginal cost. This equation can be used for two purposes, the first is to estimate the marginal costs associated with content provision. The equation provides a mapping from the expected marginal revenue associated with content provision to a marginal cost of content provision. Equation 4.18 provides a specification of the marginal costs according to program quality, observed program characteristics and unobserved program characteristics. Equation 4.17 can also be used to simulate quality under alternative advertising and informational scenarios, as will be discussed and implemented in section 7. Rewriting equations 4.17 and 4.18 incorporating the distributional assumptions made in the empirical analysis:

$$\underbrace{\mathbb{E}_{\mathbf{a}\in\boldsymbol{\sigma}}\left\{|\mathcal{T}_p|^{-1}\sum_{d\in\mathcal{T}_p}\sum_{t=1}^{T} a_{jtd}\cdot\left[\frac{dN_{jtd}}{dg_p}\cdot r_d + N_{jtd}\cdot\frac{dr_d}{dg_p}\right]\right\}}_{\text{Expected average per minute marginal revenue of program } p} = \underbrace{e^{w_p\kappa_w+\kappa_g g_p+\omega_p}}_{\substack{\text{Per minute marginal}\\ \text{cost of program } p}} \tag{5.9}$$

Derivations of the marginal revenue components are provided in section B. Table 10 provides summary statistics for the marginal revenue across programs for several model specifications. Both the *expected marginal revenue* and the *realized marginal revenue* are derived from the model specified in section 4 and presented in column (4) of Table 6; the

Table 10: Marginal revenue summary statistics

Specification	Mean	SD	Interquartile range
Expected marginal revenue	\$8.02k	\$7.25k	[\$2.68k, \$12.07k]
Realized marginal revenue (%Δ)	111.6%	90.5%	[96.9%, 117.1%]
Perfect foresight (%Δ)	81.7%	13.7%	[70.9%, 91.7%]
Narrow view (%Δ)	86.1%	46.5%	[99.9%, 113.8%]

Notes: Marginal revenues refer to the average, per minute marginal revenue of a program. The *expected marginal revenue* is denoted in thousands of 2005 USD; the remaining specifications are measured relative to the *expected marginal revenue*. Specifications (1), (3) and (4) were averaged over 100 simulated advertising realizations consistent with the observed advertising probabilities.

perfect foresight model refers to the model presented in column (2) of Table 6; the *narrow view* focuses on the effect a program's quality has only on the revenue from that specific program, without accounting for the effect on other episodes via the market size and ad price effects.[35]

Several points come across from Table 10, first follows from a comparison of the expected marginal revenue and the realized marginal revenue. The latter is larger by approximately 12% on average than the expected marginal revenue of programs. This result implies that there may be unobserved components in the channels' ad timing decisions affecting their payoffs.[36] Comparison of the expected marginal revenue with that derived from a *narrow view* exhibits that not accounting for market expansion effects attenuates the marginal revenue by 14% on average. The attenuation is most prevalent in the *perfect foresight* model since it does not properly account for viewers' switching behavior during ads, with a mean marginal revenue smaller by 18% relative to the expected marginal revenue.

Table 11 presents the parameter estimates of the components of the marginal cost associated with content provision. The four specifications coincide with the specifications in Table 10. Under the narrow view roughly 20% of the programs have a marginal revenue of zero, resulting in a smaller sample size than the other specifications. The model fit is also higher for the learning model than the perfect foresight specification (specification (3)), lending further support for the credibility of the learning model vis-á-vis a perfect foresight specification. Furthermore, the parameter estimate for the effect of quality on marginal costs ($\hat{\kappa}_g$) is much lower than in the other specifications. This is due to the

[35]The formal specification of the marginal revenue under the narrow view is:

$$|\mathcal{T}_p|^{-1} \sum_{\tau \in \mathcal{T}_p} \sigma_{j\tau} \cdot \left[\frac{dN_{j\tau}}{dg_p} \cdot r_\tau + N_{j\tau} \cdot \frac{dr_\tau}{dg_p} \right]$$

[36]E.g. Sweeting (2009); De Paula and Tang (2012) show that coordination incentives are prevalent in radio stations' ad timing decisions. Since this is not the focus of this paper, I will abstract form timing incentives beyond those implied by the advertising strategies explained in Section 4.

Table 11: Marginal cost results

Variable	Expected (1)	Realized (2)	Perfect foresight (3)	Narrow view (4)
Quality (κ_g)	1.052	0.955	0.611	0.916
	(0.104)	(0.094)	(0.122)	(0.103)
Number of epsiodes	0.003	0.002	0.004	0.007
	(0.003)	(0.003)	(0.004)	(0.003)
Mean episode length (min.)	0.013	0.014	0.014	-0.001
	(0.002)	(0.002)	(0.003)	(0.003)
SD episode length (min.)	0.027	0.026	0.033	0.010
	(0.011)	(0.010)	(0.012)	(0.010)
Channel 10	-1.125	-1.061	-1.314	-1.158
	(0.121)	(0.109)	(0.130)	(0.122)
Constant	10.662	10.200	8.606	11.463
	(0.353)	(0.318)	(0.289)	(0.363)
Genre FE	+	+	+	+
Adjusted R^2	0.491	0.541	0.441	0.449
Obs. (N)	408	406	408	326

Notes: Table 11 presents the parameter estimates for the cost components of the mean per minute program marginal cost of content provision. The dependent variable is the mean per minute marginal revenue delineated in the left hand side of equation 5.9 and the supplementary equations in **??**. Each column corresponds to a model specification in Table 10. The marginal revenue in specifications (1), (3), and (4) were averaged over 100 simulated ad realizations consistent with the channels' adverting probabilities. The marginal revenue in specification (2) is derived from the observed advertising. The value of quality in specifications (1), (2), and (4) were derived from the learning model while the quality values in specification (3) were derived from the perfect foresight 2SLS model. The number of observations differs across specifications due to non-positive marginal revenue for some programs. Only programs broadcast by the commercial channels were included in the analyses. Standard errors in parentheses.

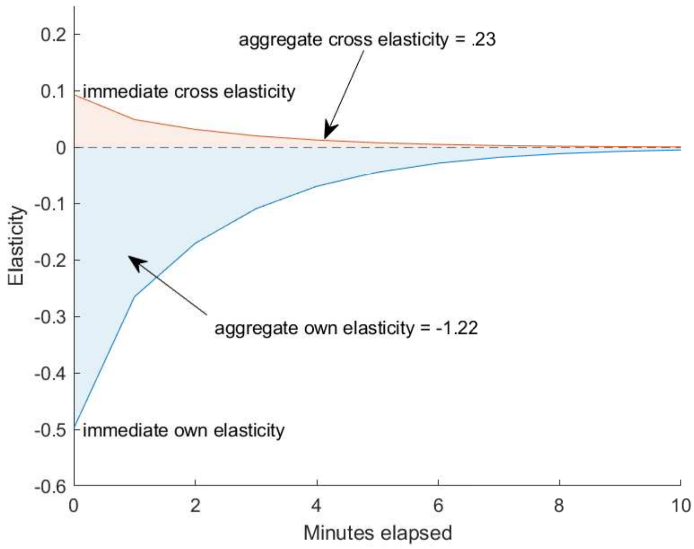
"transmitted bias" (Ackerberg and Crawford (2009)) in the perfect foresight specification. Namely, not accounting for the incomplete information in consumer choice introduces a bias in the estimate for sensitivity to advertising (α), which also propagates to other variables correlated with it, in this paper - content quality (γ).

6 Model outputs

6.1 Elasticities

There are two classes of elasticities that are interesting to examine: (a) ad probability elasticities; and (b) content quality elasticities. In simulating the ad probability elasticities and the program quality elasticities I made use of the special structure arising form the distributional assumptions to calculate the elasticities with an analytic formula. The details relating to calculation of the different elasticities are presented in Section C. All

Figure 11: Ad probability elasticity dynamics



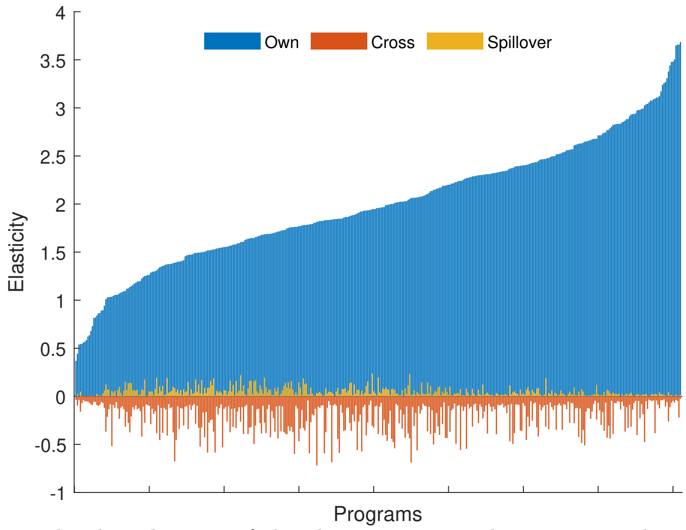
Notes: Figure 11 displays the elasticity dynamics over time. The horizontal axis is the time (in minutes) from the increase in the ad probability and the vertical axis is the minute specific elasticity. The lines refer to the mean effect.

elasticities were averaged over 100 simulated ad realizations consistent with the channels' observed ad probabilities.

Figure 11 depicts the mean ad probability elasticities. As expected from the learning framework in the model, the ad probability elasticities exhibit a strong dynamic effect, i.e. the increase of a commercial probability at a specific time persists roughly 10 minutes after the period in which the probability increased. The contemporaneous own elasticities are rather small with a mean of -0.5 (presented as the intercepts in Figure 11), while the aggregate effect is -1.2 (presented as the area bounded between the mean effects and the zero line in Figure 11). The cross elasticities exhibit a similar trend, whereby the aggregate effect is roughly three times the immediate effect.

The content quality elasticities —the distribution of which is presented in Figure 12 —show that there is a strong direct effect for the pivotal channel, i.e. the viewership during the program exhibits a mean elasticity of 2. Furthermore, the figure clearly displays the large heterogeneity in the elasticity across the different programs, ranging from 0.5 and up to 4. The spillover on other programs, while positive is very limited. These effects provide initial insight into the economic incentives arising in quality competition in broadcast media markets. Specifically, although this elasticity isn't sufficient to understand the strategic relationship of quality provision between channels, it implies that quality provision by one channel has dual effects on its competitors, the cross effect that is negative and the spillover effect that is positive.

Figure 12: Program elasticities



Notes: Figure 12 displays the distribution of the three program elasticities —the own elasticity, the cross elasticity (i.e. the elasticity on the programs broadcast at the same time on other channels), and the spillover elasticities (i.e. the effect on programs broadcast at other times of the day).

6.2 Quality competition

The strategic relationship in quality competition is unclear in the broadcast television industry. This ambiguity arises from the mixed strategic effects driving channels' quality decisions. The advertiser demand implies strategic substitutability in quality provision between the channels.[37] The viewer demand, on the other hand is ambiguous. Two main characteristics of the viewer demand affect channels' quality choice. First is the market expansion effect driven by the endogenous market size. In addition, viewers' ad avoidance decreases the return to quality provision. These two effects interact in the quality competition: enhanced quality provision by one channel increases the market size, in turn increasing the returns to quality provision. With that, higher quality programming induces more switching, thereby decreasing the quality provision incentive. I allow the model together with the data to provide insight into the strategic relationship in quality competition. Section F details the method by which the quality was simulated for hypothetical scenarios. The middle panel of Table 12 decomposes these three effects and their respective implications for the equilibrium quality provided by the channels. The top panel provides a baseline, that is the quality provision under the observed equilibrium. Each of the latter rows in the middle panel investigates the equilibrium quality while shutting down each of the respective demand features. E.g. the first row (of the middle panel) displays how equilibrium quality differs from the observed (in percent change) if

[37]This relationship follows from the quantity choice aspect of this side of the market, in which more quality increases the number of impressions, in turn driving down the price per impression across all impressions in the market, similarly to Cournot competition.

Table 12: Competition effects on equilibrium quality

Component	Scenario	Program expenditure (IQR)	Total expenditure
	Baseline	[$1.6 k, $13.5 k]	$503.2 mil.
Demand components	Market expansion	[-30.6%, +1.6%]	-10.2%
	Ad avoidance	[-37.1%, +17.2%]	-23.0%
	Ad price effect	[+5.8%, +158.0%]	+112.9%
Market structure	Symmetric channels	[-53.2%, +239.8%]	+11.1%
	Ownership consolidation	[-72.6%, -8.2%]	-32.5%

Notes: The table presents summary statistics for the simulated competitive effects of the demand factors: 1. The market expansion effect; 2. The ad avoidance effect; and 3. The ad price effect. As well as the market structure through: 1. Symmetric channels; and 2. Ownership consolidation. The first rows, title *baseline* displays the commercial channels' observed investment. The proceeding rows display the percentage change in the measure relative to the baseline investment. The *total expenditure* column aggregates over programs according tot heir prevalence in the data. In simulating the commercial channels' investment under symmetric channels, viewers' channel preference was equated between both channels (and equal to the mean of the two).

the market size were fixed and exogenously determined.[38] In this setting, the total quality expenditure of the commercial channels would decrease by 10.2%.

As expected, the *ad price effect* uniformly decreases the equilibrium quality provided in the market. Interestingly, in comparison to the demand side effects, the ad price effects has a substantially stronger effect, constituting the main competitive force. The viewer demand effects have ambiguous effects on equilibrium quality. This is more pronounced for the *ad avoidance* than the *market expansion*. Specifically, the market expansion effect, while increasing quality for some programs, the effect is mainly negative, both for the lion share of programs and also for the total program expenditure. The conclusion regarding the market expansion effect is that it has a clearly positive effect on the extent of quality provision in the market. The *ad avoidance* effect is more ambiguous than the other two effects: the interquartile range shows that some programs clearly decrease in quality while others' quality clearly increases. The total effect is strongly negative, i.e. the total effect of the ad avoidance is similar to the market expansion effect, albeit the magnitude of the effect is decidedly more pronounced.

The bottom panel of Table 12 illustrations the effects of two market structure experiments on equilibrium quality. The first presents the equilibrium quality effects of enhanced competition while the bottom presents the effects of diminished competition. The first analyzes the channels quality provision under symmetry. As shown in Table 6, viewers have an inherent preference for channel 2 over channel 10. This case equates viewers' preferences across both commercial channels[39], holding the value of the public

[38] An explanation of the method used to find the equilibrium quality under alternative settings is detailed in section E.

[39] This is done by setting viewers' preference for each of the channels to the mean of the preference across both channels.

channel and the outside option fixed. The effect on program expenditure is ambiguous, some programs expereience an increase in investment while others a decrease. In aggregate, quality investment increases by 11%. The second experiment relating to the market structure is one of a monopolist firm operating both commercial channels. This scenario clearly portrays that under ownership consolidation, quality competition decreases substantially, leading both channels to uniformly decrease their quality investment in all broadcast programs. In total, the results of the exercise in simulating the market structure's impact on quality investment implies that quality competition is beneficial for viewers, increasing quality investment.

7 Policy experiments

The framework laid out above allows us to conduct policy experiments relating to households' viewing behavior resulting from changes in the informational environment, advertising levels and content quality. Furthermore, the framework provides a structure to examine channels' equilibrium quality responses to differing environments. The main policy experiment of interest is constructing the *policy frontier*, i.e. the equilibrium content quality associated with differing regulatory environments. I will consider a social planner —to whom I will also refer as a regulator —that has two policy levers: the first is determining the permitted amount of ads per day and the second is the degree of information provision. I regard information provision as a continuous measure with the goal of capturing several information provision levels —no information as is the case in the data; partial information, e.g. a ticker on the screen displays the number of ads remaining in a break; and full information in which a ticker displays the amount of time remaining in an ad break.[40] Using the exogenous regulatory environment and the ensuing equilibrium quality, I will examine channels' profits, viewer surplus and social welfare generated from each regulatory environment.[41]

Prior to examining equilibrium responses to varying regulatory policies, I will analyze the effects of two counterfactual experiments. The first experiment focuses on the welfare implications of information provision. The results of this experiment highlight the importance of accounting for equilibrium quality in examining the welfare outcomes of a specific policy. Consider a *full information* experiment, I will analyze two possible scenarios within this experiment. Initially I will vary the informational environment with no quality response. This counterfactual experiment does not directly affect the

[40] Full information implies viewers are able to perfectly allocate their viewing choices according to the value of each channel at each minute. This is also consistent with recording technologies, e.g. TiVo in the US.

[41] In the machinery of the model, the policy variables determine an expected marginal revenue from each program. The variable of choice for the channels is the quality of each of the programs, while the remaining program attributes are considered as exogenous and fixed.

viewers' utility, only indirectly through their choices. As such, this experiment measures the extent to which households' welfare is decreased from making uninformed choices. Afterwards, I will consider the case with equilibrium quality responses, providing insight into equilibrium effects of such a policy as opposed to a demand side response alone.

The second experiment focuses on the welfare effects of viewers' lack of commitment. Specifically, this experiment examines the welfare implications of viewers committing to viewing the ads associated with a program. This is similar to the structure prevalent in many online content providers, examples include YouTube and Hulu. From a societal welfare perspective, this scenario is beneficial, it allows channels to internalize their quality investment more fully and in doing so, they can provide viewers more quality. With that, this scenario also alters the channels quality investment incentives. As was shown in Section 6.2, viewers' switching behavior is an important factor for quality provision. Consequently, the distributional effects of such a policy are ex-ante ambiguous.

Section E details the form by which the welfare measures were calculated. Section F details the form by which the equilibrium quality was simulated for alternative regulatory environments.

7.1 Biased regulation

Table 13 presents the welfare effects of the two counterfactual experiments relating to the same policy —full information. The details regarding the form by which viewers' choices are affected by the informational environment are detailed in section D. The upper panel of the table details summary statistics of the demand side counterfactual, i.e. full information without supply side reactions, and the bottom panel details the scenario in which program quality is updated accordingly. Extinguishing the uncertainty associated with the commercial broadcast channels ad timing results in significant welfare gains for viewers. The counterfatual experiment increases the mean share of households watching TV by 5% on average across days. While commercial channels enjoy a substantial increase in viewing across all minutes as shown by a mean increase in viewers of 24%, these gains are concentrated in non-ad minutes, as measured by the 63% decrease in impressions. The decrease in impressions leads to a substantial increase in ad prices at the scale of 52%. The welfare benefits to viewers are significant with a mean of increase of 7%. Hence, at face value, it seems that for a social planner aiming at maximizing households' surplus, such a policy may provide a large upside.

The bottom panel of Table 13 presents summary statistics for a similar policy intervention that accounts for the commercial channels' quality response. The main differences in the mechanism are captured in the two bottom rows detailing the change in content investment. As shown in the upper panel, the commercial channels lose a substantial part of the impressions they sell to advertisers. While ad prices increase, the increase in

Table 13: Counterfactual results

Scenario	Measure	Mean	SD
Full information & fixed quality	Market size	+5.0%	1.0%
	Viewers	+23.3%	20.2%
	Impressions	-62.7%	9.4%
	Ad prices	+52.1%	12.0%
	Channel profits	-78.5%	-
	Viewer surplus	+6.7%	-
Full information & equilibrium quality	Market size	-12.2%	5.2%
	Viewers	-31.0%	134.9%
	Impressions	-74.9%	9.0%
	Ad prices	+69.2%	17.8%
	Content expenditure	-48.7%	-
	Channel profits	-69.0%	-
	Viewer surplus	-14.9%	-

Notes: The table presents summary statistics of the effects across all days observed in 2005 except for *marginal cost* which refers to the change across all programs. The numbers represent the percentage change in each measure in relation to the baseline scenario observed in the data. The top panel refers to a full information scenario without a supply side response while the bottom panel includes commercial channels' equilibrium response to full information.

ad prices is far from sufficient to annul the decrease in impressions, resulting in drastic decreases in content quality investment, with a decrease in total investment of 79%. The decrease in program quality leads to a 10% decrease in market size. While the upper panel displayed clearly viewers optimal viewing allocations through the increase in viewership during non-ad minutes and decrease during ad minutes. In the bottom panel, the viewership across both types of minutes decreases. Viewership across all minutes decreases by 59% and impressions by 73% on average. This further decrease in impressions propagates to yet even higher ad prices that increase by 64%. Finally and most importantly, viewer surplus decreases, as opposed to the viewer surplus statistics while holding quality fixed. The expected decrease in households' surplus is 13% on average.

These two counterfactual experiments regarding the same policy illustrate the importance of accounting for the content quality implications of any policy interventions. Under fixed supply, the policy shows welfare gains to household, while allowing for equilibrium quality illustrate that changes to the channels' revenue structure will have substantial implications for quality provision incentives. As a result, not only would the channels suffer from such a policy, but households as well.

7.2 Ad avoidance

Section 7.1 has shown that information provision induces an equilibrium outcome that is detrimental to both viewers and channels. Whereas viewers are willing to trade-off more ads in order to receive higher quality programming, individually, they lack the commitment to watch the ads once they are broadcast. In this experiment I analyze the welfare implications of a mechanism requiring viewers to view the ad in order to also view the programs, similarly to the mechanism used in YouTube.[42]

The commitment mechanism has substantial effects on the channels' quality provision incentives, leading to a 35.6% decrease in content quality expenditure. The equilibrium outcome is detrimental to viewers, whose surplus decreases by $317 mil. (-7.1%). Channels, on the other hand, experience a $607.1 mil. (+72.2%) increase in profits. Aggregately, the social welfare implications of viewers' commitment is staggeringly positive, leading to a $290.1 mil. (+34.5%) increase in the social welfare. With that, the distributive implications are uneven, whereby technological capabilities that induce viewers to watch ads increase channels' profits, at the expense of viewers' welfare, in equilibrium.

7.3 Policy frontier

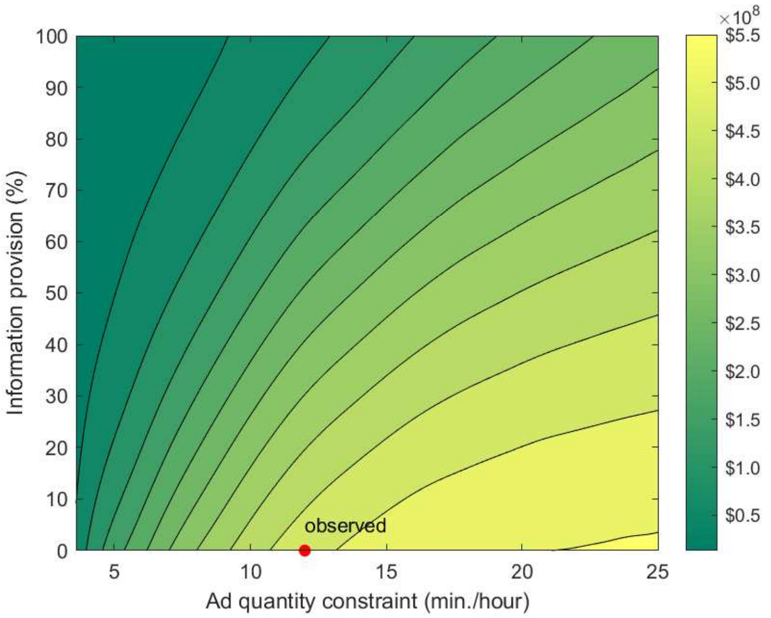
To construct the policy frontier, I first determine the equilibrium quality for any given regulation, i.e. pair of informational structure and ad quantity constraint. Figure 13 displays the mean equilibrium content quality in terms of marginal cost investment (2005 USD) for any pair of information provision and ad restriction.[43] It should be noted that in constructing the policy frontier, I didn't restrict the ads if the permitted amount was above the equilibrium amount, as such, in reality for any level above the equilibrium level, both the advertising and quality are constant. Figure 13 shows the increase in quality investment with relaxation of both constraints. It is interesting to note two things. First is that quality is sensitive to information. Information provision decreases revenues much more than ad constraints, leading to greater decreases in quality investment. Second, equilibrium quality's response the advertising constraint exhibits convexity for low advertising levels and concavity as the the advertising constraint relaxes.

Figure 14 displays the implied welfare effects along the possible policy alternatives across three measures: the top panel presents household surplus as measured by the compensating variation; the middle panel presents channels' profits; and the bottom panel presents the societal welfare. Information provision has a small effect on household sur-

[42]For the purposes of this experiment, it is unconsequential if viewers really watch the ads or leave the television on while doing something else.

[43]Note that I implemented the ad restriction in expectation. I.e. I updated the channels strategies and viewer beliefs according to the updated constraint and simulated 100 realizations from these ad strategies. In order not to change inherent ad variation across minutes that could be based on external factors - e.g. programming considerations - or unaccounted for strategic incentives, I updated the ad probability in each minute pro-rata.

Figure 13: Equilibrium content quality across regulatory environments



Notes: The figure displays the equilibrium quality response - measured in USD per minute of program broadcast - to different regulatory environments. The quality is measured as the total quality expenditure across all programs broadcast by the commercial channels weighted by their respective lengths.

plus. Viewers' increased ability to allocate their viewing across alternatives is balanced with the decrease in quality resulting from the enhanced ad avoidance. The viewers optimal advertising lies at 10 minutes of ads per hour. Viewer surplus would increase by 0.7% corresponding to a value increase of $50 million according to the compensating variation welfare measure. Under viewers' optimal regulation, content quality investment would decrease by 4.2% on average and channel profits would decrease by 8% ($88 million). The convexity of quality provision at low advertising levels illustrates viewers' willingness to trade-off relaxation in advertising regulation to increase the content quality they are provided. The concavity of the content provision at higher advertising levels leads viewer surplus to decrease as the advertising constraint is relaxed further. I.e. the increase in quality is insufficient to compensate viewers for the increased advertising.

The mid panel, displaying the channels' profits across alternative regulatory environments is consistent with the results in Table 13 that channels' profits are sensitive to information provision. Channels' optimal advertising is 22 minutes of advertising per hour. Channels' profits would increase by 13.5% ($148 million) from a transition to a non-regulated environment. Under the unregulated scenario, content quality investment would increase by 5.9% on average and viewer surplus would decrease by 7.3% corresponding to $487 million. Finally, the socially optimal advertising lies at 13 minutes of ads per hour. Under the socially optimal regulation, content quality investment would increase by 2.1% on average. The socially optimal regulation has the potential to increase the value derived from the industry by 1% corresponding to $34.9 million. Of

Table 14: Optimal regulation: welfare changes

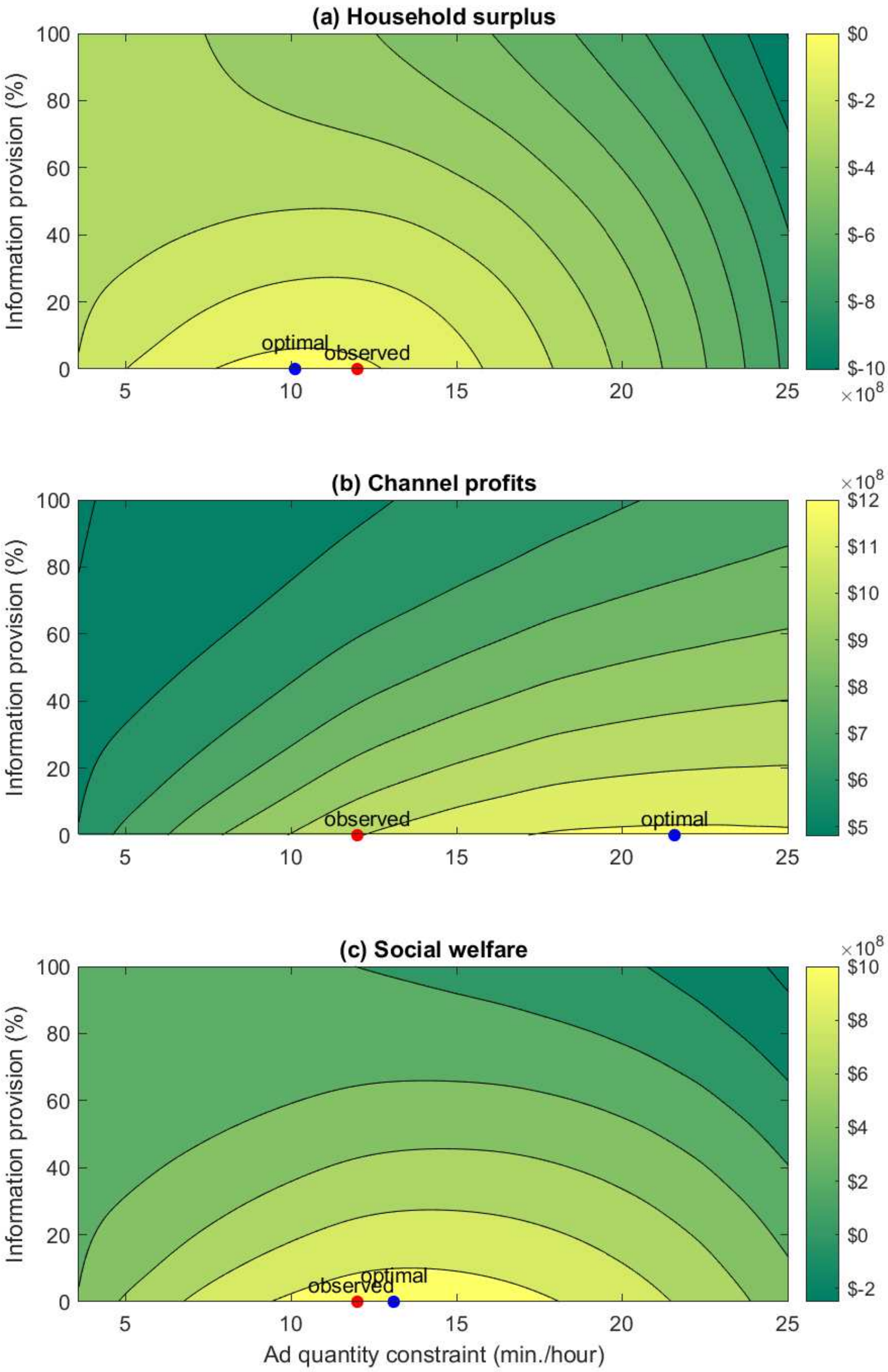
Scenario	Ads	Quality expenditure	Viewer surplus	Channel profits	Social welfare
Viewers optimal	10.1	-10.2% ($-48.8 mil.)	+0.1% ($+27.5 mil.)	-7.6% ($-82.9 mil.)	-6.2% ($-55.4 mil.)
Unregulated	21.6	+14.6% ($+69.9 mil.)	-6.7% ($-426.4 mil.)	+11.5% ($+125.5 mil.)	-28.4% ($-300.9 mil.)
Socially optimal	13.1	+3.9% ($+18.9 mil.)	-0.5% ($-7.9 mil.)	+2.9% ($+31.2 mil.)	+0.9% ($+23.3 mil.)
2^{nd} best socially optimal	12.7	+2.8% ($+13.3 mil.)	+0.0% ($+0.1 mil.)	+2.0% ($+22.2 mil.)	+0.8% ($+22.3 mil.)

Notes: The table presents the changes in welfare arising from different potential regulations. The numbers represent the percentage change in each measure in relation to the baseline scenario. The baseline scenario is that observed in the data corresponding to 24 minutes of ads broadcast over the two hour prime-time window (12 minutes per hour). *Ads* refers to the amount of ads of broadcast time. The *2^{nd} best socially optimal* refers to amount of ads in which viewer surplus is non-negative (between 8 and 13 minutes of ad time per hour).

the value derived, viewer surplus would decrease by .8% corresponding to $9.4 million, while channel profits would increase by 4% corresponding to an increase in profits of $44 million.

The measure of viewer surplus used in the social welfare measure allows for channels to compensate viewers for a change in regulation and vice versa. While this may not be realistic, we can examine the second best allocation in which we restrict to the policies for which viewer surplus increases above the baseline level. These are the advertising levels between 9.2 and 13 minutes of advertising per hour. The socially optimal advertising constraint under the second best solution is the upper bound, i.e. 13 minutes of ads. Under this policy, consumers are indifferent to the change, since the increase in content quality is sufficient to compensate for the increase in advertising. Content quality investment increases by 1.6% on average and channel profits increase by 3.1% corresponding to $34.1 million. An important aspect that is visible in the lower panel of Figure 14 is the asymmetric slopes between the two sides of the optimal point. Specifically, the social welfare decreases slower for higher advertising levels than for lower ones. Insofar as regulatory uncertainty, this implies that a laxer advertising constraint is socially preferable to a more restrictive one.

Figure 14: Welfare across regulatory environments



Notes: The horizontal axis displays the expected amount of ads a channel is permitted to advertise; the vertical axis displays the extent of information provision; and the coloring displays the welfare measure of interest. All measures were simulated for 100 ad realization consistent with a given advertising restriction.

8 Conclusion

In this paper I develop a model of households' television viewing demand, firms' advertising demand and channels' content quality supply, with the goal of assessing the effect of alternative regulatory environments on channels' equilibrium content quality choices and the ensuing welfare effects. The viewing demand model is characterized by incomplete information regarding the timing of ads and supplemented with a learning framework, generating a dynamic component in viewing decisions. The dynamic nature of viewers' choices generates channel viewing persistence, matching empirical regularities documented in the literature, as well as in the context of this paper. Households make their leisure decisions under an additional dimension of uncertainty, pertaining to the quality of the daily programming. Consequently, television viewing is also characterized by persistence, a phenomenon also documented in the literature. Advertisers are modeled as a continuum of price taking monopolist firms. Channels choose the allocation of ads across the day and the content quality of their programs with the goal of maximizing profit derived from advertising revenue and content quality cost.

The model is estimated using high-frequency data from Israel. Estimation is based on a simulated moments approach. Identification of viewers' response to both advertising and quality, as well as advertisers' response to impressions are discussed. The model estimates provide interesting insight into demand and supply fundamentals. Information provision on each channel in isolation decreases their respective market shares during ads but increases the share of households' watching television. Similarly with program quality, while the main beneficiary is the focal program, the market expansion effect benefits programs at other times in the day and of competing channels. Ad price effects are shown to be a major competitive force in this industry, curbing quality investment by 113%. Viewer demand characteristics have ambiguous effects on quality investment. While both the market expansion effect and the ad avoidance effects increase quality in aggregate - by 10% and 23% respectively - they exhibit heterogeneous effects among programs.

I simulate the welfare effects of alternative regulation, emphasizing the importance of accounting for equilibrium quality responses. In evaluating the effects of full information (i.e. a ticker on screen during ads), holding quality fixed results in a 7% increase in household surplus. Alternatively, incorporating equilibrium quality induces a 13% decrease in household surplus due to the decrease in quality. Furthermore, I simulate the welfare implications of viewers' commitment to watch ads, motivated by technological advances in online content provision. The effect on channels' quality provision incentives induces quality degradation. Consequently, while social surplus increases by 34.5%, the distributive implications are unbalanced —whereby channels experience a 72.2% increase in profits and viewers a 7.1% decrease in surplus.

Finally, I simulate the equilibrium quality for varying regulatory environments, characterized by an advertising quantity restriction. The welfare implications show the contrast between viewers' and channels' preferences - 10 vs. 22 minutes of ads per hour of broadcast time. The socially optimal regulation lies in the range of 13-14 minutes of ads per hour. This level advertising represents a 10% increase over the current regulation and is valued at $23 million.

This paper provided an empirical analysis illustrating the point made by Bork (1978) by which regulation aims easily measurable characteristics (ad quantities), at times disregarding the effects on less measurable features (content quality and welfare).

References

Ackerberg, D. A. and G. S. Crawford (2009). Estimating price elasticities in differentiated product demand models with endogenous characteristics.

Anand, B. N. and R. Shachar (2004). Brands as beacons: A new source of loyalty to multiproduct firms. *Journal of marketing Research 41*(2), 135–150.

Anderson, J. E. and M. Kraus (1981). Quality of service and the demand for air travel. *The Review of Economics and Statistics*, 533–540.

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

Armstrong, M. and D. E. Sappington (2007). Recent developments in the theory of regulation. *Handbook of industrial organization 3*, 1557–1700.

Ater, I. and O. Rigbi (2019). Price transparency, media and informative advertising. *Media and Informative Advertising (September 15, 2019)*.

Baker, M. and L. M. George (2019). The news hour: Welfare estimation in the market for local tv news. *Available at SSRN 2694687*.

Basso, L. J. (2008). Airport deregulation: Effects on pricing and capacity. *International Journal of Industrial Organization 26*(4), 1015–1031.

Belleflamme, P. and M. Peitz (2014). Asymmetric information and overinvestment in quality. *European Economic Review 66*, 127–143.

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

Bork, R. H. (1978). *The antitrust paradox: A policy at war with itself*, Volume 79. Basic books New York.

Brown, Z. Y. (2019). Equilibrium effects of health care price information. *Review of Economics and Statistics 101*(4), 699–712.

Chandra, A. and A. Collard-Wexler (2009). Mergers in two-sided markets: An application to the canadian newspaper industry. *Journal of Economics & Management Strategy 18*(4), 1045–1070.

Ching, A. T., T. Erdem, and M. P. Keane (2013). Learning models: An assessment of progress, challenges, and new developments. *Marketing Science 32*(6), 913–938.

Crawford, G. S., O. Shcherbakov, and M. Shum (2019). Quality overprovision in cable television markets. *American Economic Review 109*(3), 956–95.

Crawford, G. S. and M. Shum (2007). Monopoly quality degradation and regulation in cable television. *The Journal of Law and Economics 50*(1), 181–219.

Cuesta, J. I. and A. Sepúlveda (2019). Price regulation in credit markets: A trade-off between consumer protection and credit access. *Available at SSRN 3282910*.

De Paula, A. and X. Tang (2012). Inference of signs of interaction effects in simultaneous games with incomplete information. *Econometrica 80*(1), 143–172.

De Pinho, P. S. (2000). The impact of deregulation on price and non-price competition in the portuguese deposits market. *Journal of banking & finance 24*(9), 1515–1533.

Draganska, M., M. Mazzeo, and K. Seim (2009). Beyond plain vanilla: Modeling joint product assortment and pricing decisions. *QME 7*(2), 105–146.

Dubé, J.-P., G. J. Hitsch, and P. E. Rossi (2010). State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics 41*(3), 417–445.

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

Durante, R., M. Djourelova, and G. J. Martin (2020). Online competition, newspaper quality, and political accountability: Evidence from the introduction of craigslist.

Eizenberg, A. (2014). Upstream innovation and product variety in the us home pc market. *Review of Economic Studies 81*(3), 1003–1045.

Esteves-Sorenson, C. and F. Perretti (2012). Micro-costs: Inertia in television viewing. *The Economic Journal 122*(563), 867–902.

Fan, Y. (2013). Ownership consolidation and product characteristics: A study of the us daily newspaper market. *American Economic Review 103*(5), 1598–1628.

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

Galenianos, M. and A. Gavazza (2020). Regulatory interventions in consumer financial markets: The case of credit cards. *Available at SSRN 3405897*.

Gandhi, A., L. Froeb, S. Tschantz, and G. J. Werden (2008). Post-merger product repositioning. *The Journal of Industrial Economics 56*(1), 49–67.

Gaynor, M., K. Ho, and R. J. Town (2015). The industrial organization of health-care markets. *Journal of Economic Literature 53*(2), 235–84.

Gaynor, M. and R. J. Town (2011). Competition in health care markets. In *Handbook of health economics*, Volume 2, pp. 499–637. Elsevier.

George, L. (2007). What's fit to print: The effect of ownership concentration on product variety in daily newspaper markets. *Information Economics and Policy 19*(3-4), 285–303.

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

Grennan, M. and A. Swanson (2020). Transparency and negotiated prices: The value of information in hospital-supplier bargaining. *Journal of Political Economy 128*(4), 1234–1268.

Joskow, P. L. (2005). Regulation and deregulation after 25 years: Lessons learned for research in industrial organization. *Review of Industrial Organization 26*(2), 169–193.

Joskow, P. L. and N. L. Rose (1989). The effects of economic regulation. *Handbook of industrial organization 2*, 1449–1506.

Kinjo, K. and T. Ebina (2015). State-dependent choice model for tv programs with externality: analysis of viewing behavior. *Journal of Media Economics 28*(1), 20–40.

Levin, J. (2001). Information and the market for lemons. *RAND Journal of Economics*, 657–666.

Luco, F. (2019). Who benefits from information disclosure? the case of retail gasoline. *American Economic Journal: Microeconomics 11*(2), 277–305.

Matsa, D. A. (2011). Competition and product quality in the supermarket industry. *The Quarterly Journal of Economics 126*(3), 1539–1591.

Mazzeo, M. J. (2002). Product choice and oligopoly market structure. *RAND Journal of Economics*, 221–242.

McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica: Journal of the Econometric Society*, 995–1026.

Montag, F. and C. Winter (2019). Price transparency against market power. *Available at SSRN 3256476*.

Moshkin, N. V. and R. Shachar (2002). The asymmetric information model of state dependence. *Marketing Science 21*(4), 435–454.

Nevo, A. (2000). A practitioner's guide to estimation of random-coefficients logit models of demand. *Journal of economics & management strategy 9*(4), 513–548.

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

Pakes, A. and D. Pollard (1989). Simulation and the asymptotics of optimization estimators. *Econometrica: Journal of the Econometric Society*, 1027–1057.

Reynaert, M. and F. Verboven (2014). Improving the performance of random coefficients demand models: the role of optimal instruments. *Journal of Econometrics 179*(1), 83–98.

Romero, L. (2020). The effects of usury ceilings on consumers welfare: Evidence from the microcredit market in colombia.

Rossi, F. and P. K. Chintagunta (2016). Price transparency and retail prices: Evidence from fuel price signs in the italian highway system. *Journal of Marketing Research 53*(3), 407–423.

Rust, R. T. and M. I. Alpert (1984). An audience flow model of television viewing choice. *Marketing Science 3*(2), 113–124.

Rysman, M. (2004). Competition between networks: A study of the market for yellow pages. *The Review of Economic Studies 71*(2), 483–512.

Shachar, R. and J. W. Emerson (2000). Cast demographics, unobserved segments, and heterogeneous switching costs in a television viewing choice model. *Journal of Marketing Research 37*(2), 173–186.

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.

Sweeting, A. (2013). Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry. *Econometrica 81*(5), 1763–1803.

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

Whaley, C. M. (2019). Provider responses to online price transparency. *Journal of health economics 66*, 241–259.

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.

A Inclusive values

There are four possible knowledge specifications determining the value of watching television guided by viewers' knowledge of the ad timing —i.e. perfect foresight model or the learning model —and viewers' knowledge of the content quality. The table below presents the possible informational scenarios:

Quality / Ad timing	Perfect foresight	Learning
Known ex-ante	$\hat{\Upsilon}(\mathbf{a}, \delta)$	$\hat{\Upsilon}(\Lambda, \delta)$
Unknown ex-ante	$\hat{\Upsilon}(\mathbf{a}, \gamma)$	$\hat{\Upsilon}(\Lambda, \gamma)$

The main specification, including incomplete quality and advertising knowledge is delineated in equation 5.7. The value of watching television under perfect foresight with full quality information is given by:

$$\hat{\Upsilon}(\mathbf{a}, \delta) = \sum_{t=1}^{T} \log \left[\sum_{j \in \mathcal{J}} e^{\hat{\alpha} a_{jt} + \hat{\delta}_{jt}} \right] \tag{A.1}$$

where $\left(\hat{\alpha}, \hat{\delta}\right)$ are derived from the 2SLS perfect foresight model.

The scenario pertaining to perfect ad timing foresight and unknown content quality differs from the previous by exchanging $\hat{\delta}$ with its expected counterpart:

$$\mathbb{E}\left[\hat{\delta}_{jt}\right] = \hat{\gamma}_{jt} + x_{jt}\hat{\beta}$$

where $\left(\hat{\beta}, \hat{\gamma}\right)$ are derived from the 2SLS perfect foresight model.

Finally, the case of known content quality and unknown ad timing was calculated as:

$$\hat{\Upsilon}(\Lambda, \delta) = \sum_{t=1}^{T} \log \left[\sum_{j \in \mathcal{J}} \frac{1}{ns} \sum_{i=1}^{ns} e^{\hat{\alpha} \mu_{ijt} + \hat{\delta}_{jt}} \right] \tag{A.2}$$

where $\left(\hat{\alpha}, \hat{\delta}\right)$ are derived from the learning model.

B Marginal revenue derivations

The expected marginal revenue fo program p is:

$$\mathbb{E}_{\mathbf{a}\in\boldsymbol{\sigma}}\left\{|\mathcal{T}_p|^{-1}\sum_{d\in\mathcal{T}_p}\sum_{t=1}^{T}a_{jtd}\cdot\left[\frac{dN_{jtd}}{dg_p}\cdot r_d+N_{jtd}\cdot\frac{dr_d}{dg_p}\right]\right\}$$

where the total derivatives are:

$$\frac{dr_d}{dg_p}=-\frac{r_d}{\eta}\cdot\frac{\sum_{j\in\mathcal{J}_m}\sum_{t=1}^{T}a_{jtd}\cdot\frac{dN_{jtd}}{dg_p}}{B+\sum_{j\in\mathcal{J}_m}\sum_{t=1}^{T}a_{jtd}\cdot N_{jtd}} \tag{B.1}$$

$$\frac{dN_{jtd}}{dg_p}=\left[\frac{\partial s_{TVd}}{\partial g_p}\cdot s_{jtd|TV}+s_{TVd}\cdot\frac{\partial s_{jtd|TV}}{\partial g_p}\right]\cdot M \tag{B.2}$$

The partial derivatives in equation B.2 are:

$$\frac{\partial s_{TVd}}{\partial g_p}=\frac{\partial s_{TVd}}{\partial\Upsilon_d}\cdot\frac{\partial\Upsilon_d}{\partial g_p}=\underbrace{\phi_V\cdot(1-s_{TVd})\cdot s_{TVd}}_{=\frac{\partial s_{TVd}}{\partial\Upsilon_d}}\cdot\underbrace{\sum_{\tau\in\mathcal{T}_{pd}}s_{j\tau d|TV}}_{=\frac{\partial\Upsilon_d}{\partial g_p}} \tag{B.3}$$

$$\begin{aligned}\frac{\partial s_{jtd|TV}}{\partial g_p}&=\frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial s_{ijtd|TV}}{\partial g_p}\\&=\frac{1}{ns}\sum_{i=1}^{ns}\frac{\partial\left[\frac{e^{\alpha\mu_{ijt}+x_{jt}\beta+g_p}}{\sum_{k\in\mathcal{J}}e^{\alpha\mu_{ikt}+x_{kt}\beta+g_{kt}}}\right]}{\partial g_p}\\&=\frac{1}{ns}\sum_{i=1}^{ns}s_{ijtd|TV}\cdot\big(1-s_{ijtd|TV}\big)\end{aligned} \tag{B.4}$$

C Elasticity derivations

C.1 Program quality elasticities

The viewer demand elasticity of channel j at time t with respect to the quality of program p is:

$$e_p^{jt} = \frac{dN_{jt}}{dg_p} \cdot \frac{g_p}{N_{jt}} \tag{C.1}$$

where $\frac{dN_{jt}}{dg_p}$ is derived in equations B.2, B.3, and B.4.

C.2 Ad probability elasticities

The viewer demand elasticity of channel j at time t on day d with respect to the ad probability of channel k at minute s on day f is:

$$\boldsymbol{e}_{ksf}^{jtd} = \frac{dN_{jtd}}{d\lambda_{ksf}} \cdot \frac{\lambda_{ksf}}{N_{jtd}} \tag{C.2}$$

It should be observed that $\frac{dN_{jtd}}{d\lambda_{ksf}} = 0 \quad \forall \quad f \neq d \quad or \ f = d, \ s < t$. Hence, I will drop the day subscript. For the cases in which $f = d$ and $s \geq t$:

$$\frac{dN_{jt}}{d\lambda_{ks}} = \left[\frac{\partial \mathscr{s}_{TV}}{\partial \lambda_{ks}} \cdot \mathscr{s}_{jt|TV} + \mathscr{s}_{TV} \cdot \frac{\partial \mathscr{s}_{jt|TV}}{\partial \lambda_{ks}} \right] \cdot M \tag{C.3}$$

The partial derivatives in equation C.3 are:

$$\frac{\partial \mathscr{s}_{TV}}{\partial \lambda_{ks}} = \frac{\partial \mathscr{s}_{TV}}{\partial \Upsilon} \cdot \frac{\partial \Upsilon}{\partial \lambda_{ks}} = \phi_V \cdot (1 - \mathscr{s}_{TV}) \cdot \mathscr{s}_{TV} \cdot \alpha \sum_{\tau=1}^{T} \frac{1}{ns} \sum_{i=1}^{ns} \mathscr{s}_{ik\tau|TV} \cdot \frac{\partial \mu_{ikt}}{\partial \lambda_{ks}} \tag{C.4}$$

The effect of an increase in the (predicted) ad probability on channel k at minute s on the viewer share of channel j at minute t is:

$$\frac{\partial \mathscr{s}_{jt|TV}}{\partial \lambda_{ks}} = \frac{1}{ns} \sum_{i=1}^{ns} \frac{\partial \mathscr{s}_{ijt|TV}}{\partial \lambda_{ks}} = \frac{1}{ns} \sum_{i=1}^{ns} \alpha \cdot \mathscr{s}_{ijt|TV} \cdot \frac{\partial \mu_{ikt}}{\partial \lambda_{ks}} \tag{C.5}$$

Finally, the partial derivative of a viewers' posterior at minute t with respect to the prior at minute s (for $s \geq t$) is:

$$\frac{\partial \mu_{ikt}}{\partial \lambda_{ks}} = \begin{pmatrix} 1 - a_{ijt} & a_{ijt} \end{pmatrix} \cdot \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} \cdot \prod_{\ell=t}^{s} \Lambda_{k\ell} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{C.6}$$

D Viewers' informational structure

Increasing viewers' predictive ability on channel j, at minute t by Δ results in the common prior Λ'_{jt}:

$$\Lambda'_{jt} = \begin{array}{r} \\ a_{jt-1}=0 \\ a_{jt-1}=1 \end{array} \begin{array}{c} \begin{array}{cc} \qquad a_{jt}=0 \qquad\qquad & \qquad\qquad a_{jt}=1 \end{array} \\ \begin{pmatrix} 1-\{\lambda_{jt}(0)+\Delta\left[a_{jt}-\lambda_{jt}(0)\right]\} & \lambda_{jt}(0)+\Delta\left[a_{jt}-\lambda_{jt}(0)\right] \\ 1-\{\lambda_{jt}(1)+\Delta\left[a_{jt}-\lambda_{jt}(1)\right]\} & \lambda_{jt}(1)+\Delta\left[a_{jt}-\lambda_{jt}(1)\right] \end{pmatrix} \end{array} \tag{D.1}$$

At the limit, as $\Delta \to 1$, the structure approaches a full information specification in which the viewer demand is determined by a logistic formula:

$$\lim_{\Delta\to 1} \Lambda'_{jt} = \begin{pmatrix} 1-a_{jt} & a_{jt} \\ 1-a_{jt} & a_{jt} \end{pmatrix}$$

E Viewer surplus & social welfare

Denote by U_{jt} the utility derived in the existing framework (from viewing channel j at minute t) and by $\check{U}_{jt}$ the utility derived from an alternative regulatory framework. The values of watching television are denoted accordingly by Υ_d and $\check{\Upsilon}_d$ and are calculated using the formula in equation 5.7. The market size - denoted by $\mathscr{s}_{TVd}$ and $\check{\mathscr{s}}_{TVd}$ - is determined by equation 4.10. Finally denote the change to each of the variables by Δx, e.g. the change in the share of households watching television on day d is $\Delta \mathscr{s}_d = \check{\mathscr{s}}_{TVd} - \mathscr{s}_{TVd}$. Households' leisure decision rule is:

$$\{y_d \in \mathcal{J}\} \iff \{V_{TVd} \geq V_{0d}\} = \{\phi_V \Upsilon_d - \phi_d \geq \zeta_{0d} - \zeta_{TVd}\} \tag{E.1}$$

where y_d is the representative household's leisure choice and (V_{0d}, V_{TVd}) are the respective utilities delineated in equation 4.9. A policy change will change the set of households choosing to watch television as well as the welfare of the existing viewers. The set of viewers who choose to watch television in the existing regime is given by equation E.1, their welfare under the new regime is determined by integrating over the change in the viewers' welfare of the same set, formally:

$$\Delta CS_d^{old} = \left[\check{\Upsilon}_d(\mathbf{a}) - \Upsilon_d(\mathbf{a})\right] \cdot \mathscr{s}_{TVd} = \Delta \Upsilon_d(\mathbf{a}) \cdot \mathscr{s}_{TVd} \tag{E.2}$$

where as in equation 5.7, $\Upsilon_d(\mathbf{a})$ is the inclusive value for the realized advertising on day d as opposed to the expected inclusive value - denoted by Υ_d - that integrates over many possible ad realizations consistent with households' beliefs. The set of viewers who choose not to watch television throughout day d in the existing regime and choose to watch television under the counterfactual regime is given by:

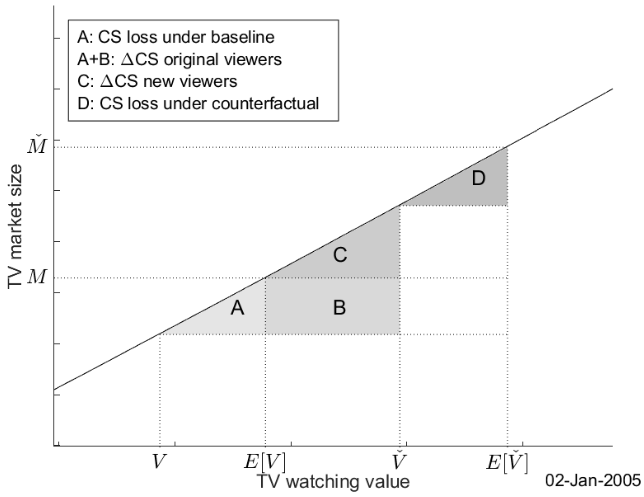
$$\phi_V \Upsilon_d - \phi_d \leq \tilde{\zeta}_d \leq \phi_V \check{\Upsilon}_d - \phi_d \tag{E.3}$$

where $\tilde{\zeta}_d = \zeta_{0d} - \zeta_{TVd}$ follows a logistic distribution and therefore the change in consumer surplus for the new viewers is given by integrating over the logistic distribution between the two bounds in equation E.3:

$$\begin{aligned}
\Delta CS_d^{new} &= \int_{\phi_V \Upsilon_d - \phi_d}^{\phi_V \check{\Upsilon}_d - \phi_d} \tilde{\zeta} dF_{\tilde{\zeta}} + \left[\check{\Upsilon}_d(\mathbf{a}) - \check{\Upsilon}_d\right] \cdot \Delta \mathscr{s}_{TVd} - \Delta \Upsilon_d \cdot \mathscr{s}_{TVd} \\
&= \int_{\Upsilon_d}^{\check{\Upsilon}_d} \frac{\tilde{\zeta} + \phi_d}{\phi_V} dF_{\tilde{\zeta}} + \left[\check{\Upsilon}_d(\mathbf{a}) - \check{\Upsilon}_d\right] \cdot \check{\mathscr{s}}_{TVd} - \left[\Upsilon_d(\mathbf{a}) - \Upsilon_d\right] \cdot \mathscr{s}_{TVd}
\end{aligned} \tag{E.4}$$

These two concepts, along with the consumer surplus loss resulting from incomplete information, are clearly depicted in Figure 15. The smooth line depicts households' valuation of viewing television vis-á-vis alternative leisure activities. A regulatory change

Figure 15: Change to consumer surplus



Notes: The figure displays change to the viewers' welfare following a transition to full information. The horizontal axis depicts households' valuation for TV watching. The vertical axis depicts the share of households' watching TV. The smooth line is the CDF of the viewers' valuations derived from the assumed TIEV distribution.

affects the inclusive value from watching, in this case a change from V to $\check{V}$ with an associated change to share of households watching television from M to $\check{M}$. The area generated by the increase to the value of TV is broken down into the welfare generated to the existing viewers, $\Delta CS^{old} = A + B$, the welfare generated to the new viewers, $\Delta CS^{new} = C$ and the welfare loss resulting from the incomplete information under the baseline and counterfactual, A and D.

To provide insight into optimal policies, the two welfare measures - channel profits and viewer surplus - must be normalized in order to be made comparable. While viewer surplus are measured in utile, channels' profits are in USD. Conversion of consumer surplus to USD is based on the compensating variation approach, upon which I measure *how much would it cost the channels to compensate viewers for the change in regulation through content provision.* Specifically, define the quality required to keep viewers indifferent between two potential policies, P and $\check{P}$ as:

$$\tilde{g} = \left\{ g : CS(g; \check{P}) = CS(g_0; P) \right\} \tag{E.5}$$

where g_0 is the equilibrium quality under policy P (i.e. the observed regulation). The compensating variation is therefore:

$$CV = C(g^*) - C(\tilde{g}) = \sum_{j \in \mathcal{J}_m} \sum_{d=1}^{D} \sum_{t=1}^{T} \frac{1}{\kappa_g} \left[e^{\kappa_g g^*_{jtd}} - e^{\kappa_g \tilde{g}_{jtd}} \right] \cdot e^{w_{jtd}\kappa_w + \omega_{jtd}} \tag{E.6}$$

where g^* denotes the equilibrium quality under the new regime. The measure of societal welfare is the sum of the change to channels' profits along with the compensating variation:[44]

$$\begin{aligned} SW(\check{P}; P) &= \Delta\Pi + CV = \Pi(g^*; \check{P}) - \Pi(g_0; P) + C(g^*) - C(\tilde{g}) \\ &= R(g^*; \check{P}) - [\Pi(g_0; P) - C(\tilde{g})] \end{aligned} \tag{E.7}$$

where $R(g; P)$ is the revenue derived from quality g under the regulatory environment P.

[44]This measure of social welfare abstracts away from the advertiser side of the demand. Assuming that advertising constitutes a transfer between monopolistic producers and the advertising platforms as in Anderson and Coate (2005).

F Equilibrium quality

The counterfactual experiments presented in Table 12 and in section 7 require finding the equilibrium level of program quality under counterfactual scenarios. In these scenarios, the channels' revenue structure or incentives change, with the goal of examining the effects on quality investment. The quality equilibrium condition in equations 4.17 and 5.9 provide a mapping from marginal revenue to marginal costs. Policy changes will influence the marginal revenue for given quality inputs and ad probabilities. As before, I consider an equilibrium concept in which viewer beliefs are consistent with the channels' advertising probabilities (i.e. strategies).

$$\log\left[MR(g;\sigma^*)\right] = g \cdot \kappa_g + w'\kappa_w + \omega \tag{F.1}$$

where $MR(g;\sigma^*)$ is defined in Section B and g is a vector of the program qualities. Equation F.1 constitutes a $|\mathcal{P}| \times 1$ vector of equilibrium conditions for each program observed in the data. A change to the revenue structure results in that the quality equilibrium condition in equation F.1 doesn't hold (for the original qualities). The lack of an analytic formula for the marginal revenue requires numerical solution. To find the equilibrium qualities associated with a counterfactual scenario I numerically solve the system:

$$g^* = \arg\min_{\mathbf{g}} \left(g\hat{\kappa}_g + \hat{K} - \log\left[MR(g;\sigma^*)\right]\right)'\left(g\hat{\kappa}_g + \hat{K} - \log\left[MR(g;\sigma^*)\right]\right) \tag{F.2}$$

where $\hat{K} = w'\hat{\kappa}_w + \hat{\omega}$.

This procedure produces the equilibrium quality for each program under differing regulatory conditions - i.e. ad quantities and informational framework - or differing incentives - e.g. ownership structure.