Micro-costs: Inertia in television viewing*

Constança Esteves-Sorenson
Yale University

Fabrizio Perretti
Bocconi University

January 2012

Abstract

We document substantial default effects despite negligible switching costs in a novel setting: television program choice in Italy. Despite the low costs of clicking the remote and of searching across only six channels and despite viewers extensive experience with the decision, show choice depends strongly on whether viewers happened to watch the previous programme on the channel. Specifically, (i) male and female viewership of the news depends on whether the preceding programme appealed to men or women, and (ii) a show's audience increases by 2-4% with an increase of 10% in the demand for the preceding program. These results are robust to endogenous scheduling. This behaviour appears most consistent with procrastination in switching, which stations fully exploit in their scheduling.

*Corresponding author: Constança Esteves-Sorenson, Yale School of Management, 135 Prospect Street, New Haven, CT, 06520 (constanca.esteves-sorenson@yale.edu). We thank Stefano DellaVigna, Steven Tadelis and Catherine Wolfram for their valuable advice. We also thank Gregorio Caetano, Arthur Campbell, Urmila Chatterjee, Keith Chen, Judy Chevalier, Liran Einav, Pedro Gardete, Jeff Greenbaum, Rachita Gullapalli, Ahmed Khwaja, Botond Kőszegi, Kory Kroft, Rosario Macera, Alex Mas, Amy Nguyen-Chyung, Miguel Palacios, Gisela Rua, Rob Seamans, Olav Sorenson, Betsy Stevenson, Justin Wolfers and participants in the Berkeley Psychology & Economics, Berkeley Haas School of Business, Emory Goizueta Business School, Melbourne Business School, University of Pennsylvania Wharton School, University of Toronto Rotman School of Management, and Yale School of Management seminars for valuable suggestions at different stages of this project. We thank Mediaset S.p.A. for generously providing the data for this project.

Airing immediately after the hit show Seinfeld, Frasier's initial time slot was Thursdays at 9:30 P.M. ... as good a scheduling slot as existed in prime-time television ... Steve Sternberg, an advertising executive, quipped that "you could read the phone book after Seinfeld and get a 25% viewer share."[1]

1 Introduction

Default options substantially affect choices, and potentially welfare, due to their stickiness. Whatever the default in, for example, retirement plans, health plans, membership plans, organ donations, agents persistently choose it; that is, they display inertia. Moreover, they do so even when the default option appears suboptimal (e.g., Choi et al., 2011).[2]

The standard explanation for inertia is high switching costs: both the direct cost of enacting the change and the indirect cost of evaluating the alternatives before doing so.[3] When these costs exceed the marginal benefit of switching, agents should rationally choose the default. If curbing default stickiness is desirable, then the main lever under this model is to shrink switching costs by, for example, increasing agents' literacy or experience with the decision.

Another model, however, suggests inertia can persist even when switching costs are very low: low costs induce procrastination in switching due to dynamically inconsistent time preferences (Strotz, 1956; Phelps & Pollak, 1968; Akerlof, 1991; Laibson, 1997; O'Donoghue & Rabin, 1999). If procrastination accounts for inertia, other measures to curb default stickiness, such as abolishing defaults in favour of active decisions may be warranted (e.g., Carroll et al., 2009).

It is an open empirical question whether inertia would persist in an environment with both extremely low direct and indirect switching costs. And if yes, what is the most plausible mechanism for this phenomenon, as different mechanisms may warrant different policy prescriptions.

We address this question by first documenting substantial inertia in a novel environment with both extremely low direct and indirect switching costs: the choice of television programs in Italy. Italy is uniquely well-suited for this question. The direct cost of switching is nearly negligible: remote controls are ubiquitous with switching requiring just a press of a button. Further,

[1]Harvard Business School case,"Frasier" (A), 2001, p.2.

[2]See Samuelson & Zeckhauser (1988), Madrian & Shea (2001), Handel (2009) for research on inertia in retirement and health plans choices; Johnson & Goldstein (2003) and Abadie & Gay (2006) for inertia in organ donation choices; DellaVigna & Malmendier (2004, 2006), for inertia in cancelling automatic memberships after stopping using the service. C. Anderson (2003) surveys lab findings on inertia.

[3]These direct and indirect costs can be broadly thought of as, respectively, transactional and learning switching costs, as defined in Farrel & Klemperer (2007).

the indirect switching costs stemming from the number and complexity of alternatives or from inexperience with the decision are very low. Italian viewers choose among only a few channels – mainly six – which account for 90% of viewership. Moreover, viewers are very experienced: Italians, as Americans, watch, on average, more than four hours of television daily.[4]

We find that the choice of a show is substantially determined by whether consumers happened to watch the previous program on that channel. We use a novel dataset of demand for television shows in Italy, supplied by Italy's superior audience tracking system, containing: (i) minute-by-minute aggregate audiences for women and men between 6:00 P.M. and 12:00 A.M. for 2002-2003 for each of the six channels and television overall; (ii) aggregate audiences and shares for all shows on these channels between 6:00 P.M. and 12:00 A.M., from 1990 to 2003.

Our evidence stems from two complementary approaches, examining how the audience of a focal show is systematically affected by variations in the audience of the preceding program on the same channel, conditional on a set of controls. Though a positive correlation between the demand for these consecutive shows is consistent with inertia, this correlation could also arise from correlated omitted shocks, such as from endogenous scheduling by channels, or weather (e.g., it rains that day so more people watch television) or reverse causality, as consumers tune-in to the earlier program to ensure seeing the subsequent show from the beginning. Our tests overcome these challenges to identification and estimate the causal effect of the demand for a program on that of the subsequent show, unbiased by these confounds.

First, a minute-by-minute event-study exploits the variation in the differing appeal of programs to women and men before the news in 2002-2003. When a female show, such as a romantic series, precedes the news, more women watch the news than men. In contrast, when a male show, such as soccer, precedes the news, more men watch the news than women. Female and male viewership converges during the news, however, suggesting that inertia decays over time. These results are robust to unobservables affecting male and female viewership.

Second, ordinary least squares (OLS) and instrumental variables (IV) estimation of the demand for programs for the larger 1990-2003 sample, finds that the audience for a show increases by 2-4% with an increase of 10% in the demand for the preceding program. Although OLS yields an estimate of 3.8%, which is statistically significant despite numerous controls, correlated omitted shocks in the demand for adjacent programs – such as from weather, endogenous

[4]Prior studies have broached the topic of channel persistence on program choice in the U.S., ascribing inertia primarily to asymmetric information: advertising of the current show during the preceding program on the channel. We will describe them later.

scheduling or viewers tuning to the channel early – could bias this estimate.

We address these correlated shocks with two IV specifications. First, we analyse how the audience for a show varies with that of the movie preceding it, by instrumenting the audience of each movie with its theatrical audience in Italy. Further, we address endogenous scheduling by channels by focusing this IV analysis on the relationship between the daily news and their preceding movies since the daily news popularity is arguably not susceptible to manipulation by channels. The IV estimate of 3.9% is marginally significant despite the small sample size and is 44% lower than the OLS estimate of 7.0% on this sample, indicating that correlated unobserved shocks are a significant source of bias.

Second, we analyse how the audience for a show varies with that of the program preceding it, by instrumenting the audience of each preceding program by its average audience in the previous month. This instrument provides a larger sample on which to conduct robustness checks and test for explanations of inertia. Again, we address the endogenous scheduling by channels by focusing our IV analysis on the relationship between the daily news and their preceding programs. The IV estimate of 2.2% is significant and 37% lower than the OLS estimate of 3.5% on this sample, again indicating that correlated unobserved shocks, such as from weather or reverse causality, are a significant source of bias.

Having established that considerable inertia in program choice exists, we try to identify the mechanisms underpinning it: (i) asymmetric information – advertising of a show during the preceding program, (ii) unsynchronized start times – competing shows not starting at the same time, (iii) quasi-indifference – consumers' near-indifference between the focal show on the default channel and competing programs on other channels, and (iv) naïve quasi-hyperbolic preferences – consumers' procrastination in switching channels.

We rule out the first two mechanisms and investigate the plausibility of quasi-indifference or quasi-hyperbolic preferences via the predictions of a dynamic choice model. On the one hand, viewers may have time-consistent preferences, but be almost indifferent between the focal show on the default channel and those on competing channels. They persistently choose the focal show until a utility shock leads them to switch. On the other hand, they may have quasi-hyperbolic preferences, discounting the immediate future at a steeper rate than when discounting between two future adjacent time periods (Strotz, 1956; Phelps & Pollak, 1968; Akerlof, 1991; Laibson, 1997; O'Donoghue & Rabin, 1999). They plan on switching at some future minute m but at m the steeper discount between the present and immediate future leads them to persist in the default. They repeatedly delay switching, planning to switch in the future. These time inconsistent preferences, coupled with naïveté about one's own behaviour, have been used to rationalise inertia in empirical settings where the marginal benefit of switching appears to exceed the cost (e.g., DellaVigna & Malmendier (2004, 2006), Carroll et al. (2009)).

Two findings suggest that the results are most consistent with naïve quasi-hyperbolic preferences. First, women's inertia into the news after a female-oriented show is insensitive to whether competing channels offer female-oriented shows during the news. Quasi-indifference, by contrast, suggests that it should decline as more channels offer female-oriented shows. Second, inertia into a focal show is also insensitive to whether competing programs to the focal show are new and thus have higher variance in their benefits. Quasi-indifference, by contrast, suggests it should decline as more channels offer novel programs since the option value of switching increases.

Lastly, we document that channels fully exploit viewer inertia when scheduling programs, via a simple model and enumeration exercise. If, instead, they ignored viewer inertia in their scheduling, this could affect 20%-40% of their profits.

We build on and contribute to the literature on inertia in the demand for television shows (Horen, 1980; Rust & Alpert, 1984; Shachar & Emerson, 2000; Goettler & Shachar, 2001; Moshkin & Shachar, 2002). The two most recent and relevant studies are by Goettler & Shachar (2001) and Moshkin & Shachar (2002). Each uses a single week of individual viewing choices in the U.S. to estimate structurally how the demand for the focal show is affected by that of the preceding program. Goettler & Shachar (2001), using one week of data in 1992, find that the strongest predictor of watching a show is whether the viewer watched that channel in the previous period. A follow-up paper by Moshkin & Shachar (2002), using one week of data in 1995, concludes that this persistence stems from two sources. The first and predominant source is asymmetric information, as U.S. channels frequently advertise the upcoming show with clips during the preceding program, inducing viewers to stay on the channel. The second, but much smaller magnitude source, is the cost of evaluating programs on other channels before switching, as the typical U.S. consumer enjoyed dozens of channels in the 1990s.

Our contributions to this literature are fourfold. First, the Italian setting allows us to document that sizeable inertia can persist in the absence of asymmetric information as advertising of a show during the previous program, though common in the U.S., is very rare in Italy. We show that asymmetric information is not a mechanism underpinning inertia in our setting.

Second, the Italian setting allows us to document that sizeable inertia can persist even when the indirect switching costs of evaluating alternatives are very low. One of the advantages of our setting is that during our 1990 through 2003 data, Italians chose among only a few channels –

six – accounting for 90% of the viewership. In contrast, Americans chose from among almost 40 channels (1992) to 70 channels (2001), with viewership increasingly diffused from the four largest networks to cable channels.[5] Hence, Italians faced very small indirect switching costs compared to Americans. Thus, large indirect switching costs is also not a mechanism underpinning inertia in our setting.[6]

Our third contribution is to conceptualise the consumer choice problem in a more flexible way. We conceptualise it as a dynamic problem, where procrastination arising from time inconsistency can also explain our findings. The previous research modelled the consumer choice over the single week as static, precluding this dynamic explanation, and constraining inertia to be due to either asymmetric information or the indirect costs of evaluating programs on other channels.

Our fourth contribution to this literature is methodological. Our event-study and IV approaches address how other forward-looking behaviour may bias the estimates of inertia. For example, inertia may be confounded with viewers tuning-in early to a channel to ensure seeing a show from its start. The previous research does not address this reverse causality: the utility of the focal show depends on whether viewers watched the channel in the previous period, but not the reverse (the utility of the preceding program may depend on that of the focal show, so that viewers tune-in early to the preceding program).[7]

In addition to the television literature, our results also contribute to the literatures on default stickiness and its mechanisms and, in particular, on stickiness induced by quasi-hyperbolic preferences (e.g., Samuelson & Zeckhauser (1988), DellaVigna & Malmendier (2004, 2006), Madrian & Shea (2001), Carroll et al. (2009)). Given the low direct cost of pressing a button and the

[5]The networks and cable shares changed from 52.4% and 25.7% shares in 1990, respectively, to 37.4% and 49.7% shares in 2001 (Source: "Ratings Analysis" by Webster, Phalen and Lichty and FCC working paper 37, page 21).

[6]The introduction of Electronic Programs Guides in the U.S., which allow viewers to check other programs' schedules without leaving the current program was only patented in 1996 (Lennon, 2000). Before 1996, when the data for this prior research was collected (one week in 1992 and another week in 1995), there were two main ways to obtain TV program information: "(1) Channel surfing: The user switches from channel to channel to get an overview of the programs which are currently running. (2) Using Printed Program Guides: The user studies paper guides including descriptions and indices to the TV program in order to get (a priori) TV program information" (page 243 of Ehrmantraut et al. (1996)). There are no such electronic guides in Italy, given the small number of channels, though as described later, there is abundant print information about programming. Further, we document later in Section 2.2 that viewers, when browsing to obtain full information on the programs on other channels, learn about and evaluate each program in a few seconds. Thus, browsing through five main alternative channels to the default channel, as is the case with Italy, represents a very small cost compared to the cost of browsing through almost forty channels, as was the case in the U.S. in the early and mid-nineties.

[7]Fang et al. (2007) discuss the usefulness of IV approaches to test for latent mechanisms in structural models.

low indirect costs owing to only six channels and highly experienced consumers, it is hard to ascribe the persistent choice of a show in a channel where consumers just happened to watch the previous program to a standard model of time-consistent preferences with switching costs. We document how this behaviour appears more consistent with procrastination in the default channel induced by quasi-hyperbolic preferences, as these can trigger substantial stickiness even when switching costs are very low. Moreover, we contribute to the literature documenting inertia induced by these preferences: in this literature, the direct costs of switching have not been as small as a thumb press but rather a phone call or mailing a letter (e.g., DellaVigna & Malmendier (2006), Carroll et al. (2009)). Or, the indirect costs of switching stemming from the number and complexity of choices and consumer inexperience have been typically larger: consumers do not evaluate six alternatives in a setting in which are very experienced – as is the case with Italy television viewing – but rather they choose, occasionally, from among several complex options (e.g., 401k choice in Carroll et al. (2009)). Further, we show that the stickiness induced by these preferences is fully exploited by stations. In this context, our analysis also relates to the literature on how firms exploit non-standard features of consumer behaviour.[8]

Finally, our results contribute to the debate on optimal defaults (e.g., Cronqvist & Thaler (2004), Carroll et al. (2009), Ericson (2011)). Should policy makers set defaults? Or should they require active choices? Would reducing the costs of switching away from defaults reduce their stickiness? Our results suggest that extremely small costs still induce significant inertia. If curbing default stickiness is desirable as, for example, some defaults may be suboptimal (e.g., Choi et al. (2011)), abolishing defaults in favour of active choices – instead of merely shrinking the costs of switching away from these defaults – may be warranted.

2 Background, audience measurement and data

Italy is an excellent setting for studying inertia in program choice. Italy's unique television industry structure allows us to isolate the mechanisms underpinning inertia and its sophisticated measurement system offers highly reliable data.

[8] See for example, DellaVigna & Malmendier (2004), Heidhues & Kőszegi (2008), Gabaix & Laibson (2006). See Ellison (2006) for a review.

2.1. Institutional background

Italy's television industry consists of a duopoly. State-owned Rai competes mainly with publicly-listed Mediaset, partly owned and controlled by the recurring prime-minister Silvio Berlusconi.[9] Each owns three channels and together these six channels captured 90% of the audience between 1990-2003. Rai owns flagship Rai 1 with 25% share; and Rai 2 and Rai 3 with 11% share each. Mediaset owns flagship Canale 5 with 24% share; Italia 1 with 11%; and Rete 4 with 8%. The remaining audience was chiefly split among MTV, LA 7 (mainly older movies), and local channels.[10] Thus, the limited number of channels induce low search costs.

These six channels follow a generalist strategy. They air shows with broad appeal, not focusing on specific demographics, as MTV with teens and pre-teens, or topics, as the Discovery Channel with science. Still, they appeal to slightly different audiences. Rai 1 and Canale 5 target a broad audience whereas Rai 2 and Italia 1 target younger viewers, for example. The schedule for a typical day appears in Appendix Figure E.1; Appendix Table E.1 describes the program genres (e.g., news, sitcom).

Advertising of a show during the preceding program is rare, occurring only in two instances. Flagships Rai 1 and Canale 5 advertise their 8:00 P.M. news with a clip during the preceding program. Canale 5 started this in 1995; Rai 1 followed suit. The anchors for both news programs also verbally announce the next program; Rai 1's anchor also announces the topic and guests of the news talk-show Porta-a-Porta, which usually airs after the 11:00 P.M. news.[11] Thus, inertia due to advertising is not a plausible explanation for most of our findings.

Advertising of programs is costly for stations – it crowds out regular paid advertising because of the regulatory cap on the amount of advertising allowed per hour – thus limiting its use (Dematté & Perretti, 2002). Nevertheless, viewers have substantial information about programs. Newspapers and TV guides advertise programming and publish stations' schedules.

Italians, as Americans are very experienced with the decision of which program to watch. Italians watched 4 hours and 43 minutes in 2006, up 21 minutes from 1997. Americans meanwhile watched 4 hours and 35 minutes during 2005-2006, up 3 minutes from the previous season.[12]

Italian audiences peak at prime-time (8:00-11:00 P.M.), as in the U.S. This is when adver-

[9] Prime-minister in 1994, 2001-2006, and 2008-present.

[10] Cable and satellite shares were negligible in this period. Satellite grew to 9% of market share in 2007. Shares of the six main channels were still high at 82% in 2007.

[11] Sources: Interview with Giovanni Modina, channel manager of Canale 5, the flagship channel for Mediaset (March 11, 2008) and interview with the Roberta Delpasso, Marketing Division of Mediaset (March 13, 2008).

[12] Italy: Auditel estimates, via Mediaset; U.S: Nielsen estimates, via Mediaweek, September 21, 2006.

tising prices also peak and competition among channels is the most fierce.

2.2. Audience measurement

Italy had a technologically superior audience measurement system during 1990-2003, supplied by Audits of Great Britain (AGB). Nielsen only upgraded to a system approximating the Italian one after AGB threatened to enter the U.S. in 1985.[13] The accuracy of the Italian data is also buttressed by the fact that Italy's measurement organisation (Auditel) is co-owned by stakeholders in the outcome of the measurement: stations, the national advertisers association and media buyers. Nielsen, in the U.S., is not and is the sole provider of measurements. Program ratings and prices for advertising in Italy are based on Auditel's audience measurements. This paper uses the same data.

Auditel has several procedures in place to ensure the accuracy of the data. Auditel's panel comprises a stratified representative sample of the Italian viewing population, with 5,101 households, 14,000 viewers and 8,000 meters (1 meter per television, with 1.6 televisions per household). Auditel rewards participation with household goods. They monitor viewing behaviour daily and check for abnormal patterns. They interview members twice per year and also call them at random and ask whether and what they are watching. Their answers are compared with the meters. Misbehaviour, though rare, leads to expulsion. Auditel refreshes the panel yearly with new members.

Viewers interact with the meter using a remote. Most interactions require two to three clicks. Once the television starts, the TV screen asks who is watching ("Registration prompt"). When viewers browse and settle on a channel for thirty seconds, the screen asks if there have been any changes in viewers ("Action prompt"). The thirty-second threshold arises from observed browsing behaviour: viewers generally evaluate programs in less than thirty seconds.[14] No viewers are counted as watching until "Yes" or "No" is punched. If there are no channel changes for fifteen minutes, the screen asks if there have been any changes in viewers ("No action prompt").[15] No

[13] The New York Times, October 8, 1990: Black Hole in Television; Nielsen's 'People Meter' Has Engendered A Revolution by Showing a Fall in Viewers.

[14] A study on Internet television by Cha et al. in 2008 supports this assertion. They observed the behaviour of 250,000 Internet television viewers choosing over 150 channels. They found that: (i) over 60% of users switch channels within ten seconds, (ii) the average time before switching is nine seconds, when switching within one minute. Earlier research in the design of television systems has assumed that consumers take ten seconds to evaluate programming (Ehrmantraut et al., 1996).

[15] With Nielsen, in the U.S., this prompt appears only after 70 minutes of no activity. Thus the Italian meter offers a more accurate measurement of audiences.

viewers are counted until "Yes" or "No" is punched. The prompt appears either as a translucent screen over current programming or as a bar at the bottom.[16] See Appendix A for more details.

2.3. Data

The data comprise two related datasets. The first contains the male and female aggregate audiences for every minute between 6:00 P.M. and 12:00 A.M. in 2002-2003 for each of the six channels and for television overall. The unit of analysis is audience by channel, gender, calendar-day and minute within the calendar-day and the data include about 2.5 million observations.

The second contains the audience, market share (percentage of total television audience), genre (if a sitcom, reality show, etc.) and the starting and ending times for each show between 6:00 P.M. and 12:00 A.M. from 1990 to 2003, for the six channels. The audience for each show is the average of the recorded audience at each minute.[17,18] The unit of analysis is *episode of show* and the data include almost 200,000 observations (we exclude weekends).[19] Most are news (22% of total), followed by variety shows (11%), talk-shows (10%), TV series (9%) and movies (7%). The average program lasts 45 minutes and the average show comprises 16 episodes (Appendix Table E.1).

3 Empirical analyses and identification strategies

We use our unique data first to show graphical evidence of inertia. A minute-by-minute event-study for 2002-2003 reveals that the relative popularity of the news among women and men flips according to the appeal of the preceding program to women and men.

Next, we estimate how the audience for a show increases with the audience for the preceding program using OLS and IV estimation on the 1990-2003 sample of shows. We corroborate the event study's magnitudes and calibrate the effect of viewer inertia on channel profitability. This larger sample also allows further testing for explanations of inertia.

[16]Before August 1997, about half the panel had the three prompts – Registration, Action and No Action; the remainder only had the Registration prompt. After 1997, the whole panel had the three prompts. This difference in measurement does not affect our estimates.

[17]Total show audience $= 1/M \sum_m$ Show Audience$_m$, m=1, ... M, $m \equiv$ minutes of duration of show

[18]A typical data point is "Show: 8:00 P.M. news Rai 1; Genre: news; Start of show: 8:00 P.M.; End of show: 8:29 P.M.; Audience: 4.5 million viewers; Share of total television viewers: 33%".

[19]Weekend line-ups and audiences differ dramatically from those on weekdays.

3.1. Event-study with minute-by-minute audience for women and men

Null hypothesis. We investigate whether more women than men watch a focal show when the preceding program appeals mainly to women, and, conversely, more men than women watch it when the preceding program appeals mainly to men. In the absence of inertia, the focal show's audience should be insensitive to variations in the relative appeal of the preceding program. It should only reflect the intrinsic appeal of the focal show to women or men.

Sampling scheme and identification strategy. The variation in the appeal to men and women of programs preceding the daily late news (starting around 11:00 P.M.) on Rai 1 in 2002-2003 offers the ideal setting. The late news are preceded by female shows (such as romantic series, where women always outnumber men for every episode) on 127 days, by male shows (primarily soccer games, where men always outnumber women) on 16 days, and by neutral shows (where women outnumber men on some episodes but not others) on 12 days.[20]

We document the minute-by-minute evolution of male and female audience from 60 minutes before the start of the late news (the event) to 60 minutes after, for the three types of days: days with female shows, days with soccer and days with neutral shows. We define an event window as (m_1, m_2): minute m_1 to minute m_2 relative to the start of the late news; therefore the event window is (-60,+60).

The late news lasts on average eight minutes. It is usually followed by the daily news talk-show Porta-a-Porta (covering political and current affairs), which lasts more than one hour. Appendix Table B.1 details the sample construction.

Unadjusted audience analysis. Female and male viewership of the late news flips according to the appeal to women and men of the program that precedes it (Figure I). During female shows (e.g., a romantic series), women outnumber men and this trend persists through the news into the subsequent news talk show (left panel). Conversely, during soccer games, always followed by the short sports news program (Rai Sport), men outnumber women and this trend continues through the news and into the news talk-show (centre panel). Male viewership, however, converges to female viewership over time. During neutral shows, women outnumber men slightly and this trend persists through the news into the subsequent news talk-show (right panel).

Adjusted audience analysis. Neutral shows establish that the baseline viewership for the

[20]Soccer is the only program that lasts longer than one hour where the audience of men consistently and significantly outnumbers that of women and that alternates with female shows in the same slot before the same program.

channel during the (-60,+60) event window is slightly higher for women than for men. Figure II therefore adjusts each left and centre panels of Figure I for (i) women and men watching Rai 1 on neutral show days (baseline) and (ii) calendar-day-by-minute-of-the-calendar-day-by-gender fixed effects to control for unobservables affecting women and men at each minute of the calendar day. These unobservables include the appeal of competing shows or of the outside option of not watching television. We add audience by calendar-day-by-minute-of-the-calendar-day-by-gender observations from the other five channels to facilitate the estimation of these effects.

Adjusted male and female audience for female show days. Men and women watching Rai 1 on female show days (left panel of Figure I) are adjusted by:

$$\begin{aligned} Audience_{\tau,channel,calday,min,gender} &= \alpha_{0,\tau} + \alpha_{1,\tau}\text{Women.Rai 1} + \alpha_{2,\tau}\text{Women.Rai 1.FemShow} \\ &+ \beta_{0,\tau}\text{Men} + \beta_{1,\tau}\text{Men.Rai 1} + \beta_{2,\tau}\text{Men.Rai 1.FemShow} \\ &+ \Gamma_{calday}\Gamma_{min}\Gamma_{gender} + \epsilon_{\tau,channel,calday,min,gender} \end{aligned}$$

where Women, Men, Rai 1 and FemShow are indicator variables for women, men, channel Rai 1 (versus any other channel) and whether the calendar day is one when a female show precedes the late news on Rai 1 (versus one when a neutral show precedes them), respectively; $\tau \equiv$ Time from start of the late news on Rai 1 $= -60, ..., +60$. The sample comprises 127 female show days and 12 neutral show days.[21]

The top left panel of Figure II depicts the adjusted difference between women and men on female show days by plotting the estimates of the key coefficients $\alpha_{2,\tau}$ and $\beta_{2,\tau}$; $\alpha_{2,\tau}$ measures the difference between women on female show days and women on neutral show days (baseline) at minute τ from the start of the news, adjusted for the fixed effects; $\beta_{2,\tau}$ measures this difference for men. We run a regression for each $\tau = -60, ..., -1, +1, ..., +60$, to obtain these estimates. For each τ, we pool women and men on each minute of the calendar day, on each of the 6 channels.[22] The standard errors are clustered by calendar day to account for serial correlation

[21]This specification combines two specifications. One adjusts women on Rai 1 on female show days: $Audience_{\tau,channel,calday,min,women} = \alpha_{0,\tau} + \alpha_{1,\tau}\text{Women.Rai 1} + \alpha_{2,\tau}\text{Women.Rai 1.FemShow} + \Gamma_{day}\Gamma_{min}\Gamma_{women} + \epsilon_{\tau,channel,calday,min,women}$ where $\alpha_{0,\tau}$=(adjusted) women on all channels except Rai 1 on both female and neutral show days; $\alpha_{0,\tau} + \alpha_{1,\tau}$=(adjusted) women on Rai 1 on neutral days and $\alpha_{0,\tau} + \alpha_{1,\tau} + \alpha_{2,\tau}$=(adjusted) women on Rai 1 on female show days. The coefficient of interest is $\alpha_{2,\tau}$, the adjusted difference women on Rai 1 on female show days versus women on Rai 1 on neutral show days. The other adjusts men on Rai 1 on female show days in a similar way: $Audience_{\tau,channel,calday,min,men} = \beta_{0,\tau} + \beta_{1,\tau}\text{Men.Rai 1} + \beta_{2,\tau}\text{Men.Rai 1.FemShow} + \Gamma_{day}\Gamma_{min}\Gamma_{men} + \epsilon_{\tau,channel,calday,min,men}$.

[22]For example, for τ=10, we pool minutes of the calendar day corresponding to: (i) 10 minutes after the start

in minute-by-minute audiences within the day (Bertrand et al., 2004).

The adjusted difference between women and men on female show days is significant during the hour before the start of the late news and persists but declines during the news and the news talk show. This difference is significant at the 5% level for most of the 29 minutes after the start of the news (bottom left panel), except for a brief period where it is significant at only the 10% level.

Adjusted male and female audience for soccer days. Adjusting men and women on soccer days with a similar specification – by replacing FemShow by Soccer, an indicator of whether the calendar day is a day when a soccer game precedes the late news on Rai 1 – but now on the sample of 16 soccer days and 12 neutral show days, reveals that the difference between men and women on soccer days is high during the hour before the start of the late news and persists but declines during the late news and the news talk show (Figure II, top right panel). The difference is significant at the 5% level for the first 26 minutes after the start of the late news (bottom right panel).

Magnitude of inertia. When female shows precede the news, the adjusted average female-male viewer gap on Rai 1 thirty minutes after the start of the news ($1/T \sum_{\tau=1}^{T} \alpha_{2,\tau} - \beta_{2,\tau}$, for $T = +30$) is 201,000 viewers, 18% of the adjusted average 1,120,000 female-male gap during female shows (Row 3-Row 6 of columns (1) and (2) in Table I). When soccer precedes the news, the average male-female viewer gap thirty minutes after the start of the news is 354,000 viewers, 20% of the average 1,787,000 male-female gap during soccer games (Row 6-Row 3 of columns (3) and (4) in Table I).[23]

Insensitivity of inertia to competition on gender-specific content. Women's inertia into the late news and the news talk show on Rai 1 appears impervious to the number of competing channels offering female shows during the news. We conduct two tests on the sample of female show days on Rai 1, since their large number (127) allows the creation of subsamples.[24]

First, women's persistence into the news and news talk-show after a female show is not statistically different on days when more than one competing channel offers female shows during the news (above-median competition) from days when at most one channel does so (below-median competition). We control for day-of-the-week unobservables affecting women on Rai 1

of the late news on Rai 1, (ii) for calendar days with female and neutral shows preceding the late news on Rai 1, (iii) and for each of the 6 channels.

[23]Thirty minutes marks the midpoint for the hour after the start of the news.

[24]We do not conduct the analysis on the days with soccer before the news on Rai 1 as the sample size is much smaller – sixteen days – preventing the creation of subsamples.

(Table II, column (3)).

Second, a more stringent test documents that women's persistence into the news and news talk-show after a female show is not statistically different on days when one or more competing channels start female shows during the commercial break before the start of the news to 5 minutes into the news on Rai 1 (high outset competition) from days when no channel does so (low outset competition). We control for day-of-the-week unobservables affecting women on Rai 1 (Table II, column (6)).

Substitution mostly between channels and not from the outside option of not watching television. The one million extra women watching female shows on Rai 1 before the news compared to the number of women watching soccer (Table III, column (3)) come from other channels and not from non-viewers, since a nearly identical 12.1 to 12.5 million women watch television on both types of days (Table III, column (6)). Otherwise, we would observe a higher total viewership of women in days with female shows airing on Rai 1, which is not the case. As for men, the 1.5 million extra men watching soccer on Rai 1 before the news compared to the number of men watching female shows on Rai 1 (Table III, column (3)) come mostly from other channels: the total number of men watching television when soccer games air on Rai 1 only exceeds the total number of men watching television when female shows air on Rai 1 by 0.6 million, indicating the number of new television male viewers due to soccer on Rai 1 is 0.6 million, whereas most of the remaining men (0.9 million) come from other channels (Table III, column (6)).

Robustness checks. Rai 1 could cause the gender flip in viewership of the late news and news talk-show if it tailored the news talk-show to women on female show days and to men on soccer days (endogenous scheduling). Inspection of a random sample of topics on female show days and of all topics on soccer and neutral show days reveals that this is not the case. Topics on female show days ranged from corruption and politics to an interview with the prime-minister. Those on soccer days included a discussion on Osama Bin Laden, Mad Cow disease and a review of the life of Pope John XXIII. Those on neutral show days spanned the U.S. attack on Iraq to coverage of the elections (Appendix Table E.2).

Rai 1 could also generate the gender flip in viewership by announcing the topic of the news talk-show at the end of the 8:00 P.M. news. If more women than men watch the 8:00 P.M. news on female show days, then more women than men might be persuaded to watch the news talk-show (and the late news beforehand). Conversely, if more men that women watch the 8:00 P.M. news on soccer days, then more men than women might be persuaded to watch the news talk-show. Rai 1's 8:00 P.M. news, however, has identical female and male viewership levels on female show days and soccer days (Appendix Figure E.2).

3.2. OLS and IV on a panel of television shows

3.2.1. Main analyses and results

Null hypothesis. Absent inertia, the demand for an episode of a show should not vary systematically with the demand for the preceding program on the channel. It should only depend on the show's intrinsic characteristics – for example, the cast, genre, year, month and time slot – and those of competing shows on other channels.

OLS estimation. There is a high (.66) correlation between the audience of adjacent shows. For example, the audience of the 8:00 P.M. news on Canale 5 covaries with that of the preceding Wheel of Fortune; the same for Hitchcock Presents and the movies that precede it (Figure III).

To ascertain the causal link between the demand for a show and that of its preceding program, we exploit the (unbalanced) panel structure of the data: most shows have more than one episode per show. Thus, we can control for time-invariant unobserved factors influencing the demand for an episode of a show.

We postulate that logged *Demand for episode e of show i, on channel c, in calendar year y, month m and half-hour time slot s* should be a flexibly linear function of *Show i*'s (i) intrinsic attributes, such as cast and genre (Γ_i), (ii) channel (Γ_c), (iii) year and month (Γ_y and Γ_m, respectively), (iv) half-hour time slot (Γ_s), and (v) intensity of competition, by either popular shows on other channels (*Competition on popularity*) or shows of the same genre (*Genre overlap*):

$$\text{Demand}_{e\ of\ i,\ c,\ y,\ m,\ s} = \alpha_0 + \alpha_1 \text{Demand preceding program} + \alpha_2 \text{Comp. popularity} + \\ + \alpha_3 \text{Genre overlap} + \Gamma_i \Gamma_c \Gamma_y \Gamma_m \Gamma_s + \epsilon_{i,c,y,m,s}$$

After accounting for these factors, demand should not depend on the logged demand for the preceding program on the channel, i.e., the null is $\alpha_1 = 0$.

The controls for show (Γ_i), channel (Γ_c), year (Γ_y), month (Γ_m) and half-hour slot (Γ_s) enter the estimation as fully interacted time-invariant characteristics (a total of 16,965 fixed effects).[25]

[25] Audience does not vary significantly by day of the week, except on weekends, which are excluded from the

The interaction between channel and show accommodates the few shows playing across different channels in the same group.[26] The further interaction with year, month and half-hour fixed effects accounts for unobservable factors affecting demand for that show within the calendar month and half-hour slot. We, therefore, only estimate demand for shows that air at least twice within the same channel, calendar-month and half-hour slot.

Competition on popularity is the logged index of the average audience that competing shows garnered in the previous month. For example, an index of 1.5=(2.5+2.0+1.0+1.0+1.0)/(5 channels) for a half-hour show on Rai 1 would mean that Rai 2 aired a program that averaged 2.5 million viewers in the previous month, Rai 3 one that averaged 2.0 million, and Rete 4, Canale 5, and Italia 1 ones that averaged 1.0 million each. For competing new or single-episode shows we approximate their audiences with the audience of programs of the same genre, on the same channel and starting on the same half-hour slot during the previous month. This variable enters the specification as a proxy for the popularity of competing programs to the focal show.[27] See Table IV for summary statistics.

Genre overlap is an index of the fraction of time the focal show competes against shows of the same genre. For example, 0.2=(1+0+0+0+0)/(5 channels) would mean that during the news on Rai 1, Rai 2 aired news, but Rai 3, Rete 4, Canale 5 and Italia 1 did not. See Table IV for summary statistics.

The estimate of α_1, conditional on these controls, is 0.38 and significant at the 1% level: demand for a focal show increases by 3.8% with an increase of 10% in the demand for its preceding program. We obtain this result by adding controls sequentially, reducing the estimate of α_1 from 0.67 down to 0.38-0.40 (Table V). The standard errors are conservatively clustered by day to account for correlation among the demand for shows in the same time slot.

The OLS estimate may nonetheless be biased due to simultaneity and omitted variable bias. Simultaneity occurs because *Demand for episode e of show i* may influence *Demand for preceding program*: viewers, for instance, may tune to the channel earlier so that they do not miss the beginning of the episode. Omitted variables, such as weather or other unobserved shocks affecting the demand for adjacent shows may also bias the estimate of α_1. We address these biases with two instrumental variable specifications.

analysis. Nevertheless, specifications including day-of-the-week fixed effects yield equivalent results.

[26]For example, Walker Texas Ranger aired on Mediaset's Rete 4 in 1996 and on Mediaset's Italia 1 in 2003.

[27]Its inclusion, instead, as an instrument for the popularity of competing programs to the focal show does not change the estimate of the coefficient of interest, α_1.

IV estimation using the theatrical audience of movies screened in Italy. We analyse how the audience of a show's episodes varies with the audience of the movies preceding them, by instrumenting the audience of each movie with its theatrical audience in Italy. About 2000 movies released in Italian theatres between 1990–2000 later aired on television.

The theatrical audience of a movie is significantly correlated with its television audience on its first airing. It is also arguably uncorrelated with shocks in the demand for the show airing after the movie. For example, an omitted weather shock may influence the audience of an episode and that of its preceding movie the day they air, but it does not affect the past theatrical audience of the movie. Or, the popularity of the current episode may influence that of the preceding movie because viewers tune-in earlier to the channel to watch the episode, but this early tuning-in also does not affect the theatrical audience of a movie.

$$\begin{aligned}\text{First stage}: Demand_{\text{preceding program(movie)}} &= \theta_0 + \theta_1 \text{Theatrical Audience} + \theta_2 \text{Comp. popularity} + \\ &+ \theta_3 \text{Genre overlap} + \Gamma_i \Gamma_c \Gamma_y \Gamma_m \Gamma_s + \upsilon_{i,c,y,m,s}\end{aligned}$$

$$\begin{aligned}\text{Second stage}: Demand_{e\ of\ i,c,y,m,s} &= \beta_0 + \beta_1 Demand_{\text{preceding program(movie)}} + \beta_2 \text{Comp. popularity} + \\ &+ \beta_3 \text{Genre overlap} + \Gamma_i \Gamma_c \Gamma_y \Gamma_m \Gamma_s + \eta_{i,c,y,m,s}\end{aligned}$$

The demand for an episode of a show increases by 4.8% with an increase of 10% in the demand for the movies preceding them (Table VI, column (2)). This estimate is significant at the 1% level and smaller than its OLS counterpart of 5.7%, suggesting a fair amount of OLS bias. The OLS and IV estimates are not statistically different though due to the large standard errors (Table VI, column (1)).

In the first stage, the television audience of a movie on its first airing increases by 0.62% with an increase of 10% in its theatrical audience (Table VI, column (3)), an estimate significant at the 1% level, with a t-statistic of 5.63 (corresponding to an F-statistic of 31.7).[28] The instrument, therefore, appears strong (Stock et al., 2002).

Stations, however, may endogenously schedule more popular episodes of a show after more popular movies and, conversely, less popular episodes of a show after less popular movies. Thus, we restrict the sample of shows preceded by movies to the news since its daily popularity is

[28] Movies tend to air, on average, three times on television. The partial correlation between a movie's theatrical audience and television audience is only statistically different from zero on the first airing.

arguably not susceptible to manipulation by stations.

The demand for the news increases by 3.9% with an increase of 10% in the demand for movies preceding them (Table VI, column (5)). This estimate is significant only at the 10% level since the number of observations declines to 143. This estimate is also almost half of its OLS counterpart of 7%, suggesting again a significant amount of OLS bias. The two estimates are not statistically different, however, due to the larger standard errors caused by the smaller sample size.

Given the small number of observations available with this instrument we cannot use it to conduct robustness checks or to test for explanations of inertia. We therefore use another instrument whose case for exogeneity is slightly weaker, but which yields a larger sample on which to conduct tests. Its estimates are smaller but not statistically different from those obtained with the first instrument.

IV estimation using the average demand in the previous calendar month. We analyse how the audience of a focal show's episodes varies with that of their preceding programs, by instrumenting the audience of each preceding program with its average audience in the previous month.

The average demand for a program in the previous calendar month correlates highly with its current demand. It is also uncorrelated with weather shocks affecting the demand for the focal show. But its exogeneity may be violated if unobserved factors affecting both the focal show in the current month and its preceding program in the previous month vary within the current month. The estimates we obtain with this instrument, however, will be similar to those obtained previously, allaying these concerns.

We estimate how the demand for the main daily news – starting at 6:30, 7:00, 7:30, 8:00 and 8:30 P.M. – varies with that of the programs preceding them.[29] Restricting the outcome to demand for the main news has two advantages. It is hard for channels to manipulate the daily popularity of the news. Further, we can test whether inertia holds on different samples of news: the news preceded by movies (in the previous analysis) tend to air later, at about 11:00 P.M., and to be shorter; the main daily news air earlier and tend to be longer.

This instrument yields results that are lower than – but not statistically different from – those

[29] We differentiate between the main daily news and the short, late night, daily news. The main daily news have longer lengths – no less than 20 minutes, averaging 32 minutes – and start every day at the same time. The late night news average 11 minutes and usually air at the end of prime-time, but at no fixed time, such as the eight minute news around 11:00 P.M. on Rai 1 (discussed in the minute-by-minute estimation).

using the theatrical audience instrument, suggesting that the unobserved factors described above do not appear to bias the estimates.

First stage:

$$Demand_{Program\ preceding\ main\ news} = \theta_0^0 + \theta_1^0 \text{Average demand previous month}_{Program\ preceding\ the\ main\ news}$$
$$+ \theta_2^0 \text{Comp. popularity} + \theta_3^0 \text{Genre overlap} + \Gamma_i \Gamma_c \Gamma_y \Gamma_m \Gamma_s + \upsilon_{i,c,y,m,s}^0$$

Second stage:

$$Demand_{Main\ news,c,y,m,s} = \beta_0^0 + \beta_1^0 Demand_{Program\ preceding\ main\ news} +$$
$$+ \beta_2^0 \text{Comp. popularity} + \beta_3^0 \text{Genre overlap} + \Gamma_i \Gamma_c \Gamma_y \Gamma_m \Gamma_s + \eta_{i,c,y,m,s}^0$$

The demand for the main news increases by 2.2% with an increase of 10% in the demand for the program preceding it (Table VII, column (2)). These estimates are 37% smaller than the OLS estimates at 3.5%, suggesting again a fair amount of OLS bias, and statistically different given the large sample size. The first stage (Table VII, column (3)) is strong with at t-statistic of 10.9 on the instrument, corresponding to an F-statistic of 118.8 (Stock et al., 2002).

3.2.2. Other findings on OLS/IV analysis

Decay rate of inertia. We test for the decay in inertia by splitting the main news into news below- and above-median length. Their respective average lengths are 28 and 38 minutes. Inertia decreases in the duration of the news. The audience of the main news with below-median length increases by 3.0% with an increase of 10% in the audience of the preceding program. But this magnitude falls to 1.5% for above-median-length news (Table VII, columns (4)-(5)). The magnitude of inertia from movies into the late news, averaging eleven minutes, is 3.9%.

This corroborates the event-study finding that inertia has a decay rate. The magnitude of inertia in the event-study is 18-20% thirty minutes after the start of the news. It is smaller than the 30% effect on shows averaging 28 minutes in length but not statistically different from the 22% effect for the whole sample of main news, averaging 32 minutes in length (Table VII, column (2)).

As in the event-study, viewer inertia persists into the program following the news. The demand for the show that follows the main news increases by 0.9% with an increase of 10% in the demand for the program that precedes the main news (Table VII, column (6)).

Effect of uncertainty about competing shows on inertia. We test whether more uncertainty about competing programs to the focal show (the main news) decreases inertia into it. The higher the uncertainty about competing programs, the higher the potential reward of switching. If competing programs are worse than the focal show on the default, viewers can click back to it. Thus, the upside of switching can be high whereas the downside is truncated at the cost of clicking. The next section models this process explicitly.

We test whether inertia into the news changes when no novel shows start at the same time as the news (within one minute) versus when one or two do so.[30] We classify a show as novel if it is in its first quartile of episodes.[31]

Inertia appears insensitive to novelty in competing programs. The coefficient on the interaction between the instrumented *Demand for preceding program to the main news* and *No uncertainty about the competing programs*, measuring the difference in inertia between no competing channels showing novel programs versus one or two channels doing so, is not statistically different from zero (Table VIII, column (2)).

4 Mechanisms

The empirical analyses established (i) a causal link between the demand for consecutive programs on channel, (ii) a decay rate to inertia and (iii) that inertia appears insensitive both to competition on gender-specific content and to novel programs on other channels. We now discuss mechanisms consistent with these findings.

I. Asymmetric information. Advertising of a show during the preceding program could cause inertia by persuading viewers to remain on the channel. Canale 5 started to advertise its 8:00 P.M. news with a clip during the preceding program in 1995; Rai 1 followed suit.

We remove these two news shows from the sample focusing only on the main news on the other four channels, which do not advertise during the preceding program. Inertia in this subsample is 2.5%, and not statistically different from the 2.2% estimated for the full sample of main news shows (Table VIII, column (3)). Thus, inertia exists in the absence of asymmetric

[30] Cases when three or more channels start programs at the same time (within one minute) as the focal news are rare, fewer than 5% of cases.

[31] We also tested specifications where we classified a show as novel if it was in its first decile of episodes or if it was on its first half of episodes. Both yielded equivalent results to those obtained with the classification of novel shows in the text.

information on the programming of the default and competing channels.

II. Unsynchronized start times for programs. Differing starts of programs might also generate inertia. If some viewers experience disutility from not watching a show from its beginning, they may remain on the default channel until a program starts on another station. Thus, inertia may stem from no programs starting around the same time as the focal show.

Unsynchronized start times for programs do not affect inertia. We split the main news show sample into news that have no competing programs starting at the same time (within one minute) versus those with one or more programs starting at the same time. We test whether inertia is higher when no competing shows start at the same time as the focal show (the main news). The interaction between the instrumented *Demand for program preceding the main news* and *No shows starting in the 1 minute vicinity of the main news*, measuring the difference in inertia between the two conditions, is not statistically significant (Table VIII, column (4)).

III. Quasi-indifference versus naïve quasi-hyperbolic preferences. Two other possible mechanisms are time-consistent preferences with quasi-indifference between the focal show on the default channel and competing programs, and quasi-hyperbolic preferences. Viewers with the former preferences stay in the default because they are nearly indifferent between the focal show and programs on other channels. Viewers with the latter preferences, coupled with naïveté about their procrastinating tendencies, procrastinate in switching away from the default.

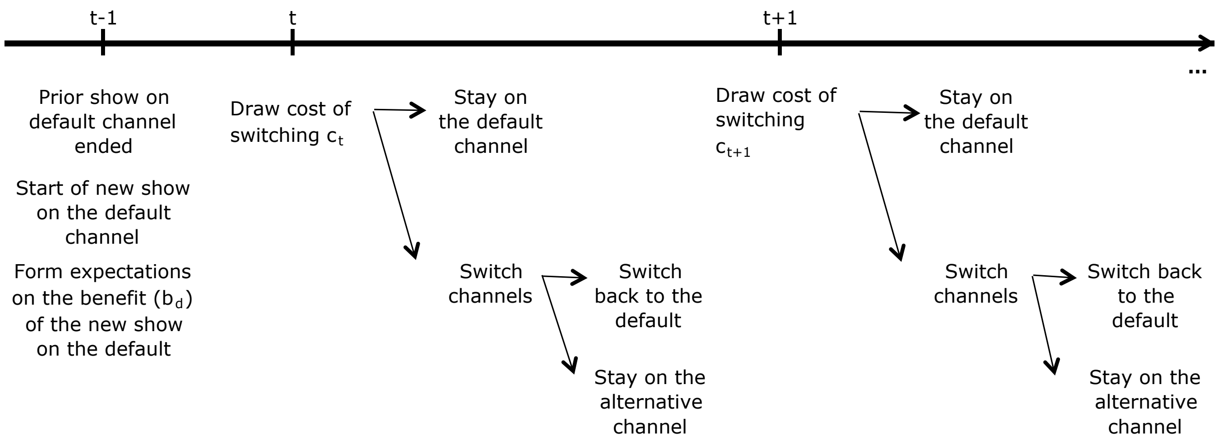
To fix ideas, the dynamic model below describes the relationship between these two types of preferences, switching costs, the option value to switching and delays in the default. We then relate the predictions of the model to the findings in the event-study and OLS and IV estimation.

Model setup. Consumers are on the default channel at the end of a program. The decision problem is whether to switch to an alternative channel. The decision–making horizon is infinity. At minute $t-1$, a new show on the default channel starts and the consumer gathers information about it. The information gathered during minute $t-1$ allows her to form unbiased expectations on the benefit $\hat{b_d}$ she will derive every minute thereafter from this show. This setup is consistent with the finding that consumers evaluate programs in less than one minute, updating their priors on the current programming almost instantaneously (Cha et al., 2008). We assume, for simplicity, that $\hat{b_d} > 0$.

At the beginning of minute t the consumer draws a cost c_t of clicking to another channel. The cost of clicking at each minute (c_t) is stochastic, i.i.d, drawn from a continuous distribution with cumulative density $F(c)$ known to the consumer, with support $[\bar{c}, \underline{c}]$, where $0 \leq \underline{c} < \bar{c}$.

The consumer does not know ex-ante the benefit b_a she will obtain on the alternative channel.

Timing of game



She has priors on it from previous experience or other information, but she only observes b_a by sampling the program.

She compares the benefits versus the costs of switching at each minute, discounting future time periods by δ, where $\delta \in (0, 1)$. At minute t she can switch by incurring c_t, the cost of clicking at t. If the program on the alternative channel is at least as preferred as the show on the default channel, she stays on the alternative channel and gains b_a at minute t and for all the minutes thereafter, reaping $b_a + b_a \frac{\delta}{(1-\delta)}$. If the show on the alternative channel is less preferred than that on the default, she returns to the default channel, gaining b_a in minute t and $\hat{b_d}$ from $t+1$ onwards, reaping $b_a + \hat{b_d} \frac{\delta}{(1-\delta)}$. We assume that it is costless to switch back to the default channel, so she has an even greater incentive to switch. Therefore, the upside of switching could be high compared to the downside, which is truncated below at $-c$.[32]

The standard model. The payoffs, at time t, associated with the actions of switching channels and not switching channels, are, respectively,

$$V(c_t, \hat{b_d}, \delta) = \begin{cases} -c_t + E[b_a] + E[b_a | b_a \geq \hat{b_d}] P(b_a \geq \hat{b_d}) \frac{\delta}{1-\delta} + \hat{b_d} P(b_a < \hat{b_d}) \frac{\delta}{1-\delta} & \text{if switch} \\ \hat{b_d} + \delta E[V(c_{t+1}, \hat{b_d}, \delta)] & \text{if not switch} \end{cases}$$

[32] A model with, instead, non-stochastic costs c and stochastic benefits at each minute (b_t) would yield similar predictions.

Solving the model. Let $G \equiv E[b_a] + E[b_a | b_a \geq \hat{b_d}]P(b_a \geq \hat{b_d})\frac{\delta}{1-\delta} + \hat{b_d}P(b_a < \hat{b_d})\frac{\delta}{1-\delta}$, the gain associated with the option value of switching. The consumer switches if $-c_t + G \geq \hat{b_d} + \delta E[V(c_{t+1}, b_a, \hat{b_d}, \delta)]$. The solution to this problem is a cut-off c^* whereby the consumer is indifferent between switching and not switching channels: If the cost of switching at each period is less than or equal to c^* the consumer switches the channel, and stays on the default channel otherwise. See the appendix for proofs of the propositions below.

Proposition 1. The consumer's game has a unique stationary equilibrium:

i. If $G \leq \underline{c} + \frac{\hat{b_d}}{1-\delta}$, then $c^* = \underline{c}$ (the consumer never switches)

ii. If $G \geq \frac{\bar{c} - \delta E[c] + \hat{b_d}}{1-\delta}$, then $c^* = \bar{c}$ (the consumer switches, no matter the cost)

iii. otherwise, c^* is the unique solution to the equation:

$$c^* = \frac{1}{1-\delta+\delta P(c_{t+1} \leq c^*)}\{G(1-\delta) - \hat{b_d} + \delta E[c_{t+1} | c_{t+1} \leq c^*]P(c_{t+1} \leq c^*)\}$$

Proposition 2. In the region where $\underline{c} < c^* < \bar{c}$, c^* is strictly increasing in G and strictly decreasing in $\hat{b_d}$.

The cut-off c^* is increasing in G, the gain associated with the option of switching, and decreasing in the attractiveness of the show on the default channel $\hat{b_d}$.

We assume that $b_a - \hat{b_d} \sim \mathrm{U}[\Delta - \sigma, \Delta + \sigma]$, where Δ is the mean difference between the benefit of the show on the alternative channel and that on the default channel, with Δ being any real number and $\sigma \geq 0$, the spread around the mean.[33] For $\sigma = 0$, this difference is deterministic with $\Delta = b_a - \hat{b_d}$. For $\sigma > 0$, we have three cases. When $\Delta \leq -\sigma$ the consumer never switches since switching would leave her worse off. When $\Delta \geq \sigma$ the difference in benefits is always

[33] We assume television consumers have imperfect information about the programs, as each airing is unique, unless it is a rerun. Full information can only be attained by sampling the program in a few seconds (see the previous discussion in Section 2.2. on the evidence that consumers evaluate programs in a few seconds). In this context, Δ can be seen as the mean of the prior distribution of the difference in benefits between the already sampled show in the default channel, about which consumers are perfectly informed, and the show on the alternative channel about which they are imperfectly informed (but whose average benefit they can infer from having seen this show advertised in print, or from sampling it in previous periods, or from the type of program the competing channel tends to air). The variance of this prior distribution is $\sigma \geq 0$. New programs on the alternative channel offer a higher variance in the difference in benefits than established programs since the consumer has less information, and therefore more uncertainty, about the benefits of newer programs on the alternative channel.

positive, but she has to incur the cost of switching. When $-\sigma < \Delta < \sigma$ the alternative program may be less attractive on average (e.g., $\Delta < 0$) but its spread high enough to induce switching.

We focus on cases two and three in the region $\underline{c} < c^* < \bar{c}$, which are the most interesting. Assume, for simplicity, that $c \sim U[0, 1]$ and δ arbitrarily close to 1.

Quasi-indifference between channels. The cut-off in case two ($\Delta \geq \sigma$) is $c^* \simeq \sqrt{2\Delta}$. The consumer will delay switching if the expected difference in benefits $\Delta = b_a - \hat{b_d}$ is small, causing c^* to be small. This quasi-indifference between channels leads to long delays in the default, because the consumer has to wait longer to draw a cost lower than c^*. The cut-off in case three ($-\sigma < \Delta < \sigma$) is $c^* \simeq \frac{\Delta+\sigma}{\sqrt{2\sigma}}$. It is increasing in the mean Δ and spread σ in the difference in benefits. The consumer will delay switching if she believes the other channel is only slightly better (e.g. both Δ and σ are small).

Procrastination in the default channel. Another model predicts delays in switching even when the mean and spread in the difference in benefits is significant. It focuses on time-inconsistency of preferences, whereby the consumer plans to change the channel at some future minute by incurring the cost of switching in return for watching a better show. When this minute arrives, however, the immediate cost of switching looms larger than the more distant benefit. The consumer delays switching, planning to switch in the future. She will repeatedly do so until a random shock in utility leads her to switch.

Time-inconsistent preferences, coupled with naïveté about one's own behaviour explain delays in the status-quo even when the marginal benefit of switching appears higher than the cost.[34] Not making a phone call to enrol in an employer's 401k plan, foregoing the employer's matching contributions (Madrian & Shea, 2001), or not cancelling a gym membership when no longer using the gym (DellaVigna & Malmendier, 2006) are consistent with these preferences.

McClure et al. in 2007 document time-inconsistency even when the delay in rewards lasts only minutes. Thirsty lab subjects preferred immediate squirts of juice or water to waiting five minutes for them. However, when choosing between squirts of juice and water in 10 minutes versus 15 minutes, or 20 minutes versus 25 minutes, they showed no such preference for the earlier rewards, though the lag between them was still five minutes.

[34] Time-inconsistent consumers can be divided into two categories. Sophisticates, who know they have time-inconsistent preferences (Strotz, 1956; Phelps & Pollak, 1968; Laibson, 1997; O'Donoghue & Rabin, 1999), and naïve or partially-naïve consumers (Akerlof,1991; O'Donoghue & Rabin, 2001), who naively believe they are more time-consistent than they actually are. Both types of consumers will show longer delays in the status-quo than time-consistent agents. The procrastination for naïve or partially-naïve consumers, however, is longer than for sophisticates because the latter are aware they will procrastinate and therefore switch earlier.

The fully naïve consumer sets a lower threshold c^* than the optimal, believing she will procrastinate less than she actually does, underestimating her losses from procrastination. This lower threshold leads to longer delays in the default.

A consumer with this type of intertemporal preferences postpones one-time tasks with immediate costs and delayed benefits, captured in the discount function 1, $\beta\delta$, $\beta\delta^2$, $\beta\delta^3$ .... where $\beta \in (0, 1]$. She believes she is time-consistent: her belief about her β, defined as $\hat{\beta}$, is that $\hat{\beta} = 1 \geq \beta$. Thus, she optimises over future time periods as a time-consistent agent, not recognising that she will procrastinate when the future becomes the present. The lower the β, the higher her procrastination.

Proposition 3 (i) The cut-off for the fully naïve consumer is: $c^{*,naive} = \beta c^{*,exp} + (1 - \beta)\Delta$; (ii) $c^{*,naive} = c^{*,exp}$ if $\beta = 1$; (iii) $c^{*,naive} < c^{*,exp}$ if $\beta \in (0, 1)$, (iv) $c^{*,naive}$ is strictly increasing in β if $\beta \in (0, 1)$.

If $\beta = 1$, she does not procrastinate – the cut-off for the naïve consumer ($c^{*,naive}$) is the same as that for the time-consistent consumer ($c^{*,exp}$). If $\beta \in (0, 1)$ the naïve consumer has a lower cut-off cost that the time-consistent consumer, taking longer to draw a small enough cost to switch channels. And the smaller the β, the smaller $c^{*,naive}$, the more she procrastinates switching, delaying longer in the default.

Thus, even when the average difference in benefits is high and its spread significant we could observe long delays in the default for a naïve consumer with small enough β.

Discussion and match with findings in the event-study and the IV estimation. Time-consistent consumers persist in the default if they are quasi-indifferent between the show in the default channel and the program in the alternative channel. However, inertia should decline as the expected difference in benefits (Δ) or the spread in this difference (σ) increases since their cut-off $c^* = c^{*,exp}$ also increases. Naïve quasi-hyperbolic consumers' inertia varies less with the increase in the expected gain from switching and its spread, as their cut-off is always smaller than that of the time-consistent consumer and further declines as their procrastination tendencies increase (smaller β).

Our findings seem more consistent with naïve quasi-hyperbolic preferences than with quasi-indifference. As documented in section 3.1. women's inertia into the news after a female show is insensitive to whether competing channels offer female shows during the news. Quasi-indifference, however, suggests that it should decline as more channels offer female shows: if women switched to Rai 1 to watch a female show they should switch at an increasing rate out of Rai 1 during the subsequent news as more channels offer female shows (higher potential difference in benefits). Inertia also appears insensitive to whether competing programs to the focal show are new (more uncertain) and thus have higher variance in their benefits, as documented in section 3.2. Quasi-indifference, however, suggests that it should decline as more channels offer novel programs since the option value of switching rises.[35]

5 Calibration of value of consumer inertia for channels

The previous sections established that viewer inertia affects the audience of television shows: the demand for an episode of a show increases by 2%-4% with an increase of 10% in the demand for the preceding program. We established that the most plausible model for behaviour in this environment is one of naïve quasi-hyperbolic preferences. We now show, with a simple static model and schedule enumeration that channels fully exploit consumer inertia in program choice. Though the model is static and does not model strategic interactions, it captures the essential features of this environment.

5.1. A simple model

Model setup. A channel has three consecutive time slots of equal length: s_1, s_2 and s_3. It wants to allocate three programs (1, 2 and 3) to these three time slots. The programs vary in their intrinsic appeal: $a_1 < a_2 < a_3$, where $a_i \equiv$ intrinsic appeal of program i. The intrinsic

[35] An alternative mechanism for viewer inertia is the purported hypnotic effect of television. The popular press has claimed that television's frequent edits induce an hypnotic state of apathy (e.g., Scientific American, 2003). This could induce persistence in the default channel. Evidence from studies on the effect of television on cognition in adults, conducted mostly in the 1970s and 1980s, is mixed. Krugman (1971) compared electroencephalographic (EEG) activity of a subject exposed to both television and magazine ads. Television created mainly slow (e.g. alpha) waves, indicating low mental effort. Reading created beta-wave activity, indicating higher effort. The conclusion that television led to a trance-like state ensued, though the study comprised a single subject. Weinstein et al. (1980) found that magazine ads generated more mental processing (beta-waves) than television ads. However, research varying reading and television content (Radlick, 1980) found television required higher mental effort when its content was more complex than that of books. D. R. Anderson et al. (2000) functional magnetic resonance imaging (fMRI) data suggest that the interpretation of montages requires the coordinated effort of several brain areas. The evidence for children suggests cognitive engagement with television. Violent content leads to violent behaviour (e.g., Bushman & Huesmann, 2001) and preschoolers' exposure to television affects positively their later achievement (e.g., Gentzkow & Shapiro, 2008). A review by Anderson and Burns (1990) concludes that educational content "seems to educate in the manner intended" and that, in general, "television viewing is a cognitively active behaviour".

appeal of a program is its audience net of scheduling effects. The audience of program 1 is normalised to one ($a_1 = 1$). There are no strategic interactions with competing channels. The channel's problem is to maximise average audience across the three time slots (daypart), since advertising revenues increase monotonically in audience.[36]

Optimal scheduling in the absence of viewer inertia. In the absence of viewer inertia, any allocation of the three shows across time slots yields the same average total audience $S(a_i, a_j, a_k) = \frac{1}{3}[1 + a_2 + a_3]$, for $i, j, k = 1, 2, 3$ and i $\neq j \neq k$.

Optimal schedule with viewer inertia. With viewer inertia, the focal show inherits a fraction ρ of the audience of the preceding program. Therefore, it is optimal to order shows in decreasing order of intrinsic appeal: the best program, 3, in the first slot, program 2 in the second slot and the weakest program, 1, in third slot, yielding an average audience: $S(a_3, a_2, a_1)^{Optimal} = \frac{1}{3}[a_3 + (a_2 + \rho a_3) + (1 + \rho(a_2 + \rho a_3)] = \frac{1}{3}[1 + a_2 + a_3 + \rho(a_2 + a_3 + \rho a_3)]$. The worst schedule orders shows in reverse, from weakest to best, yielding $S(a_1, a_2, a_3)^{Worst} = \frac{1}{3}[1 + (a_2 + \rho) + (a_3 + \rho(a_2 + \rho)] = \frac{1}{3}[1 + a_2 + a_3 + \rho(1 + \rho + a_2)]$. This strategy is consistent with anecdotal evidence. The *lead-in strategy* – placing a weak or new show after a popular one to inherit its audience – is described in industry books as a common strategy to exploit inertia.[37]

The difference in audiences between the optimal and worst schedules corresponds to a difference in revenues, since ad rates increase monotonically in audience.

Optimal schedule when show lengths are unequal. Prime-time shows in Italy vary in length. When show lengths vary, the optimal schedule depends on the relative ratio of intrinsic show audiences a_1, a_2, a_3, their lengths l_1, l_2, l_3 and the magnitude of the inertia parameter ρ. Therefore, analysis of the optimal schedule requires computing all possible combinations of shows and ascertaining which yields the highest and lowest average audiences during prime-time, conditional on viewer inertia.

5.2. Optimal schedule during prime-time: audience effects

We focus on the flagships, state-owned Rai 1 and Mediaset's Canale 5, which captured 50% of the audience in prime-time for 2003. We use the estimated inertia parameter to derive the audience, net of inertia, of each type of program in prime-time. Then we enumerate all possible

[36]The slots between 8:00 and 11:00 P.M. comprise the prime-time daypart, for example.

[37]This media scheduling strategy is common knowledge in the television industry and has been discussed extensively in many books. A leading book, *Ratings Analysis*, by Webster et al. in 2006 writes: *"... a lead-in strategy is the most common [strategy]..."*.

schedules in prime-time (six schedules since there are only three program types), calculate inertia from one program to the next, and estimate the average audience for each combination. The optimal schedule is the combination of programs yielding the highest audience.

Using an estimated inertia parameter of $\rho = 0.3$ for a 28 minute program (Table VII, column 4) we find that the prime-time schedules for both Rai 1 and Canale 5 in 2003 are close to the optimal: the average audience falls short of that of the optimal schedule by only 0.1% and 0.9%, respectively. Even though this exercise is in a static setting and does not account for strategic interactions, it shows that channels exploit viewer inertia when scheduling programs. The small deviations from optimality in our exercise likely arise from the lack of dynamics in scheduling and/or strategic interactions.

In contrast, the worst schedules fall short of the optimal average audience by 2.0% and 2.7%, respectively (see Table IX and Appendix Table D.1). If channels did not incorporate viewer inertia in their scheduling, this could result in a loss of at least 2% in audience.[38] We quantify below that this loss would be worth 20-40% of profits for for-profit channels.[39,40,41]

5.3. Relationship between audience and advertising rates

We document that changes in audience affect revenue significantly. We find that an increase of 1% in the forecasted audience for prime-time increases advertising revenues by 1.2%. We use the advertising rates charged by Mediaset's Canale 5 for 2002 and 2003 for a thirty-second ad in prime-time to calibrate this magnitude. The increasing return to audience conforms with the relationship between audience and the price for a thirty-second commercial in the U.S. in 2003, for all major networks, where an increase in audience by 1% increases advertising revenues by 1.4%, on average (Wilbur, 2008, Table 1, page 362). Appendix D details of these calculations.

[38] The average difference in audience across the six different schedules in prime-time is dampened by the large weight of a two-hour program in prime-time, which is less sensitive the audience of the preceding program since inertia decays over the duration of the two hour program.

[39] The results for the other four channels (not shown) are similar. Their schedules are at or close to the optimum and the percentage differences between the optimal and worst schedules are 2-4%.

[40] Using the estimated inertia parameter $\rho = 0.22$, for 32 minute program (Table VII, column 2), we find that the actual schedules for Rai 1 and Canale 5 come closer to the optimum: 0.1% and 0.5% difference, respectively. And the difference between the optimal and worse schedule shrinks to 1.0% and 1.6%, respectively. These results are not not shown, but available upon request.

[41] We do not calculate standard errors as this is an one-time enumeration exercise of three program types over three possible slots – and thus six possible schedules – using the average audience for the year for each program type. This is a good approximation as line-ups in Italy tend to be very stable, with the same program types airing daily, as is the case with the 8:00 p.m. news and the subsequent half-hour variety show.

5.4. Optimal schedule: impact on channel profitability

A 2% loss in audience would decrease for-profit Mediaset's channels' profits by 20% to 40%. First, a loss in 2% in audience represents a loss in advertising revenues of 2.4% estimated in the previous section. The average revenues for for-profit Mediaset in 2002 and 2003 were 2,280 and 3,029 million Euros, respectively, consisting mostly of advertising revenues (2,112 and 2,848 million Euros, respectively). Profits amounted to 258 and 175 million Euros, respectively (11.4% and 5.8% of revenue, respectively). Assuming that (i) a change in 2% in audience due to scheduling could be achieved for the whole day, not just for prime-time, and (ii) the estimated relationship between advertising rates and audience holds across all channels, not just Canale 5, a 2% decrease in audience resulting from ignoring inertia in scheduling, would decrease advertising revenue by 2.4%. This would have caused a 20% to 40% decline in profits in 2002 and 2003, respectively, given that programming costs are sunk for the year.[42]

Any decrease in audience for the state-owned Rai group would lead to negative profitability since its profit margin is close to zero – 0.2% and 1.0% in 2002 and 2003, respectively.

6 Conclusion

How sticky are defaults? Do they become inconsequential when *both* the direct and indirect costs of switching away from the default are extremely low? This paper suggests that not, by exploiting the unique features of the Italian media environment. Despite the low direct cost of switching – a click of a button – and the low indirect costs of switching owing to the limited number of channels to search – mainly six – and the vast experience of consumers with television, show choice in Italy depends significantly on whether viewers happened to view the previous program on that channel.

How can this phenomenon occur in a setting where " ... the switching costs ... are literally one thumb press?" (Thaler & Sunstein, 2008).[43] We test and discuss possible mechanisms underpinning inertia in such an environment, such as advertising of the subsequent show during the current program (asymmetric information), waiting until another program starts

[42]For $\rho = 0.22$ the difference between the optimal and worst schedule for Canale 5 would be 1%. Under the same set of assumptions, a decrease in audience of 1% would decrease advertising revenues by 1.2% and therefore Mediaset's 2003 profitability by 20%.

[43]Thaler and Sunstein in "Nudge" ask the question of why viewer inertia exists in television when the actual cost of switching away from the status quo channel is a thumb press of a button in the remote control.

on another channel (unsynchronized start times), caring little about which program to watch (quasi-indifference) or procrastination on a channel (naïve quasi-hyperbolic preferences). Procrastination in the default channel is the most consistent with the findings in our setting.

Further, we document that channels fully exploit this behaviour. Failing to have done so, however, would have cost them 20-40% of their profits. In fact, television stations design the choice environment – "choice architecture" (Thaler & Sunstein, 2008) – to fully exploit this behaviour. The *lead-in strategy* – placing a weak or new show after a popular one to inherit its audience – is described in industry books as a common strategy to exploit inertia. In the United Kingdom, where the BBC is statutorily mandated to air educational content, executives use popular lowbrow programs to nudge viewers into watching the subsequent educational shows: *"[...] it is a revival of the old idea of hammocking difficult programs between entertainment[...]"*.[44]

In sum, this paper documents that extremely low switching costs can still induce significant inertia. If curtailing default stickiness is desirable, abolishing defaults in favour of active choices – instead of merely shrinking the costs of switching away from these defaults – may be warranted.

Constança Esteves-Sorenson, Yale University
Fabrizio Perretti, Bocconi University

Submission date: April 2011
Acceptance date: July 2011

[44] Jana Bennet, BBC's director of Television, The Guardian, Media Section, February 2003; "hammocking" refers to scheduling a weak or new program between to popular ones, so that it inherits the audience of the preceding program and captures viewers tuning-in early to watch the succeeding program.

References

Abadie, A., & Gay, S. (2006). The impact of presumed consent legislation on cadaveric organ donation: A cross-country study. *Journal of Health Economics*, *25*(4), 599–620.

Akerlof, G. A. (1991). Procrastination and obedience. *American Economic Review*, *81*(2), 1–19.

Anderson, C. (2003). The psychology of doing nothing: Forms of decision avoidance results from reason and emotion. *Psychological Bulletin*, *129*(1), 139-167.

Anderson, D. R., Fite, C. V., Petrovich, N., & Hirsch, J. (2000). *Cortical activation during comprehension of visual sequences: An fMRI study".* (Unpublished manuscript; University of Massachusetts, Amherst, Massachusetts)

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, *119*(1), 249–275.

Burns, J. J., & Anderson, D. R. (1990). *Cognition and watching television* (Vol. The Neuropsychology of everyday life: Issues in development and rehabilitation; D. E. Tupper & K. D. Cicerone, Eds.). Springer.

Bushman, B. J., & Huesmann, L. R. (2001). *Effects of televised violence on agression* (Vol. Handbook of children and media (pp.223-268); D. G. Singer & J. Singer, Eds.). Sage.

Carroll, G. D., Choi, J. J., Laibson, D., Madrian, B. C., & Metrick, A. (2009, Nov). Optimal defaults and active decisions. *Quarterly Journal of Economics*, *124*(4), 1639–1674.

Cha, M., Gummadi, K., & Rodriguez, P. (2008, August 17-22). Channel selection problem in live IPTV systems. *SIGCOMM Conference, Seattle, Washington.*

Choi, J. J., Laibson, D., & Madrian, B. C. (2011). $100 bills in the sidewalk: suboptimal investment in 401(k) plans. *Forthcoming, Review of Economics and Statistics.*

Cronqvist, H., & Thaler, R. H. (2004). Design choices in privatized social-security systems: Learning from the Swedish experience. *American Economic Review*, *94*(2), 424–428.

DellaVigna, S., & Malmendier, U. (2004). Contract design and self-control: Theory and evidence. *Quarterly Journal of Economics*, *119*(2), 353–402.

DellaVigna, S., & Malmendier, U. (2006). Paying not to go to the gym. *American Economic Review*, *96*(3), 694–719.

Dematté, C., & Perretti, F. (2002). *L'impresa televisiva - principi economici and variabili strategiche* (Second ed.). Etas.

Ehrmantraut, M., Harder, T., Wittig, H., & Steinmetz, R. (1996). The personal electronic guide - towards the pre-selection of individual tv programs. Association of Computer Machinery.

Ellison, G. (2006). *Bounded rationality in industrial organization* (Vols. Advances in economics and econometrics, Theory and Applications, Ninth World Congress; W. K. N. Richard Blundell & T. Persson, Eds.) (No. Volume II). Cambridge University Press.

Ericson, K. M. (2011). *Market design when firms interact with inertial consumers: evidence form medicare part d.* (Working paper, Harvard University)

Fang, H. M., Keane, M., Khwaja, A., Salm, M., & Silverman, D. (2007). Testing the mechanisms of structural models: The case of the Mickey Mantle effect. *American Economic Review*, *97*(2), 53–59.

Farrel, J., & Klemperer, P. (2007). Handbook of industrial organization. In (Vol. 3, chap. Coordination and lock-in: Competition with switching costs and network effects). Elsevier.

Gabaix, X., & Laibson, D. (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *Quarterly Journal of Economics*, *121*(2), 505–540.

Gentzkow, M., & Shapiro, J. M. (2008, FEB). Preschool television viewing and adolescent test scores: Historical evidence from the Coleman study. *Quarterly Journal of Economics*, *123*(1), 279-323.

Goettler, R., & Shachar, R. (2001, Winter). Spatial competition in the network television industry. *Rand Journal of economics*, *32*(4), 624-656.

Handel, B. (2009). *Adverse selection and switching costs in health insurance markets: When nudging hurts.* (Working paper, Northwestern University)

Heidhues, P., & Kőszegi, B. (2008). Competition and price variation when consumers are loss averse. *American Economic Review*, *98*(4), 1245-1268.

Horen, J. (1980). Scheduling of network television programs. *Management Science*, *26*(4), 354-370.

Johnson, E., & Goldstein, D. (2003, NOV 21). Do defaults save lives? *Science*, *302*(5649), 1338-1339.

Krugman, H. (1971). Brain wave measures of media involvement. *Journal of Advertising Research*, *11*(3-9).

Kubey, R., & Csikszentmihalyi, M. (2003). Television addiction. *Scientific American Mind - Special Edition*, *14*(1).

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, *112*(2), 443–477.

Lennon, M. J. (2000). *Technology transfer guide.* Aspen Publishers.

Madrian, B., & Shea, D. (2001, November). The power of suggestion: Inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics*, *116*(4).

McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2007). Time discounting for primary rewards. *Journal of Neuroscience*, *27*(21), 5796–5804.

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

O'Donoghue, T., & Rabin, M. (1999, Mar). Doing it now or later. *American Economic Review*, *89*(1), 103–124.

O'Donoghue, T., & Rabin, M. (2001). Choice and procrastination. *Quarterly Journal of Economics*, *116*(1), 121–160.

Phelps, E. S., & Pollak, R. A. (1968). Second-best national saving and game-equilibrium growth. *Review of Economic Studies*, *35*(2), 185–199.

Radlick, M. (1980). *Processing demands of television: Neurophysiological correlates of television viewing.* Unpublished doctoral dissertation, Rensselaer Polytechnic Institute, Troy, NY.

Rust, R., & Alpert, M. (1984). An audience flow of television viewing choice. *Marketing Science*, *3*(2), 113-124.

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, *1*, 7-59.

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

Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics*, *20*(4), 518–529.

Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, *23*(62), 165–180.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth and happiness*. Yale University Press.

Webster, J., Phalen, P., & Lichty, L. (2006). *Ratings analysis: The theory and practice of audience research* (Third ed.). Routledge.

Weinstein, S., Appel, V., & Weinstein, C. (1980). Brain-activity responses to magazine and television advertising. *Journal of Advertising Research*, *20*(3), 57-63.

Wilbur, K. C. (2008). A two-sided, empirical model of television advertising and viewing markets. *Marketing Science*, *27*(3), 356–378.

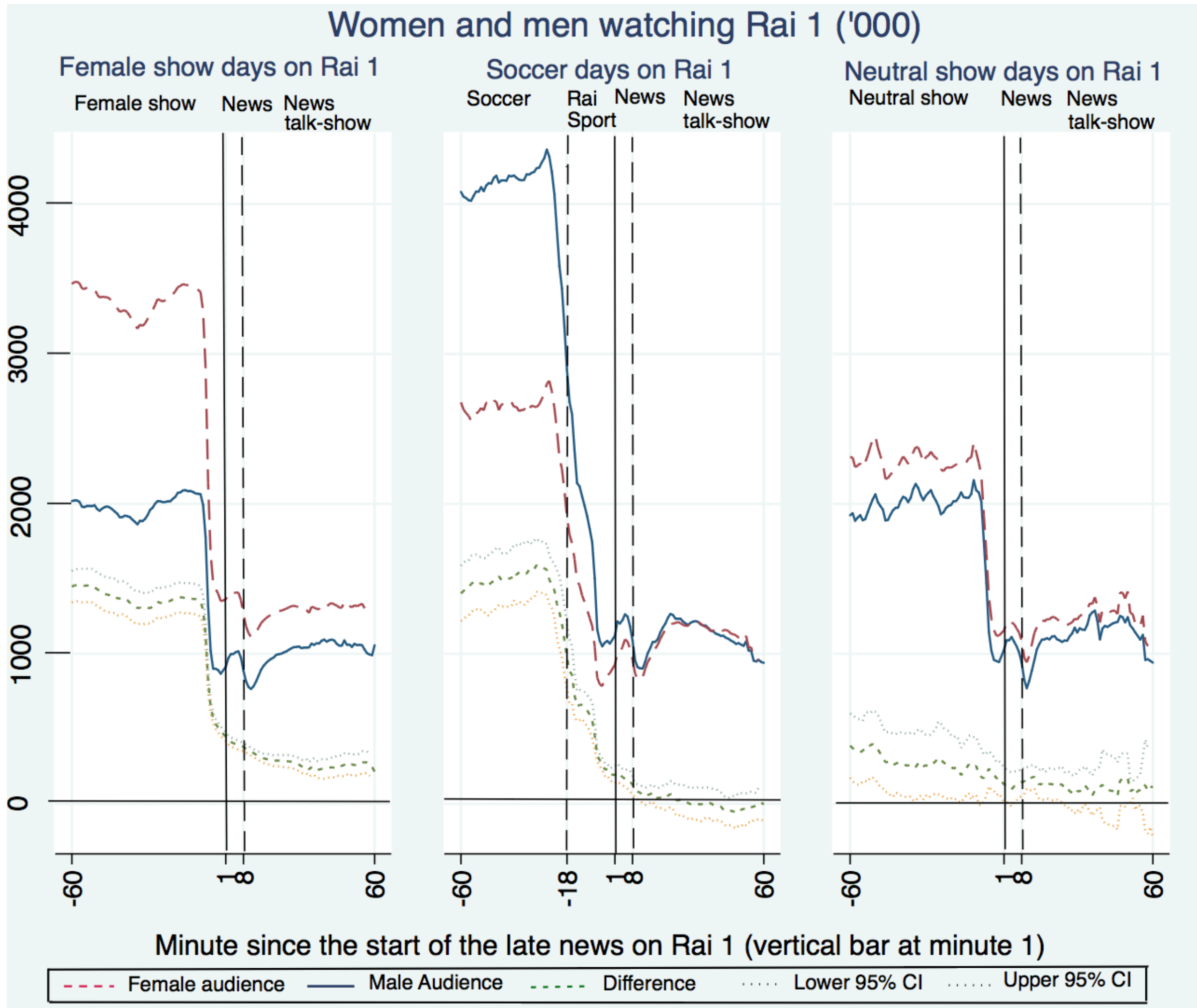


FIGURE I

Women and men watching Rai 1 one hour before and one hour after the start of the late news, unadjusted

All panels depict women and men watching Rai 1 from 60 minutes before to 60 minutes after the start of the late news on Rai 1: the left panel for the 127 days when female shows precede the news; the center panel for the 16 days when soccer games, always followed by the short sports news Rai Sport, precede the news; and the right panel for the 12 days when neutral shows precede the news. Standard errors are clustered by calendar day.

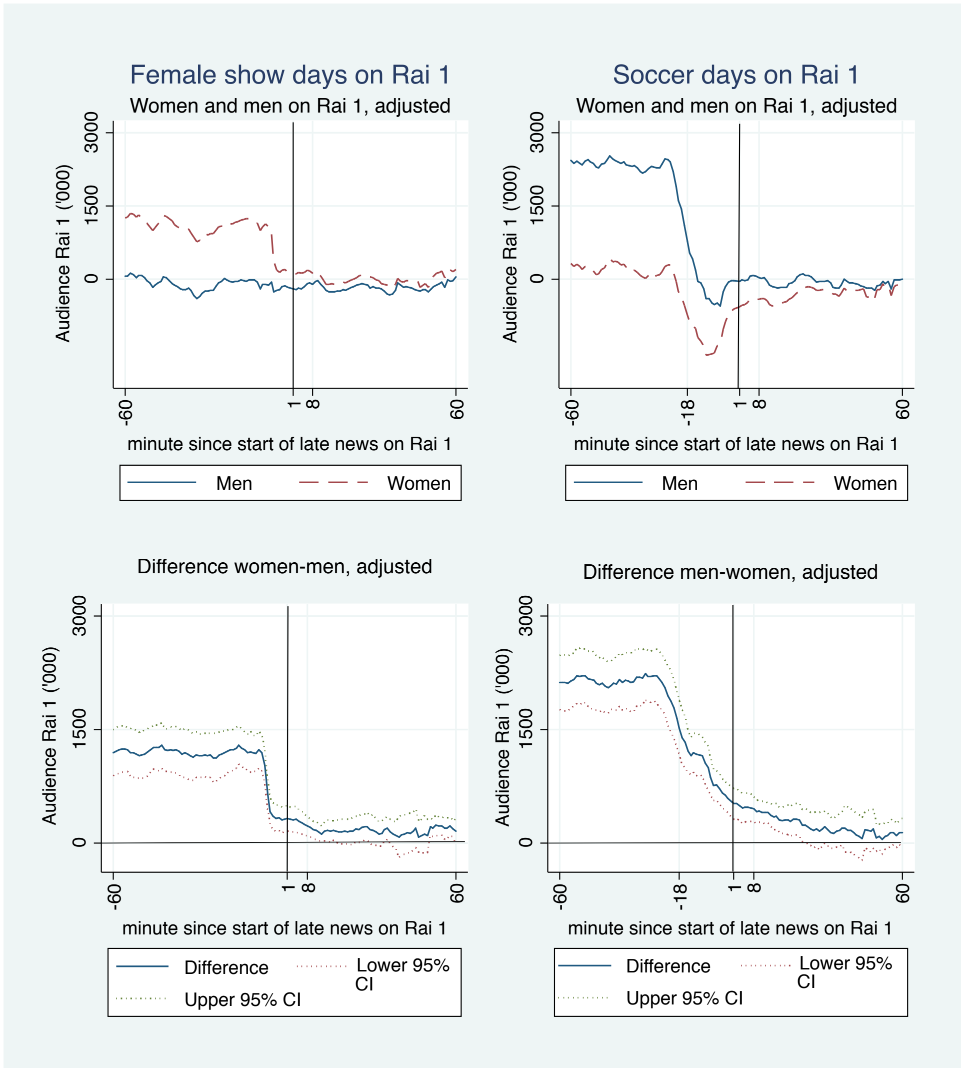


FIGURE II

Adjusted women and men on Rai 1 before and after the late news

The top left (right) panel depicts the adjusted audience for women and men from 60 minutes before to 60 minutes after the start of the late news on Rai 1 on female show (soccer) days. Men and women on female show (soccer) days were adjusted by (i) the baseline women and men on Rai 1 (women and men on neutral show days) and (ii) calendar-day-by-minute-of-the-calendar-day-by-gender fixed effects. The bottom left (right) panel shows the difference in the adjusted audience for women and men and its 95% confidence interval. Standard errors are clustered by calendar day. Details of the estimation are in the text.

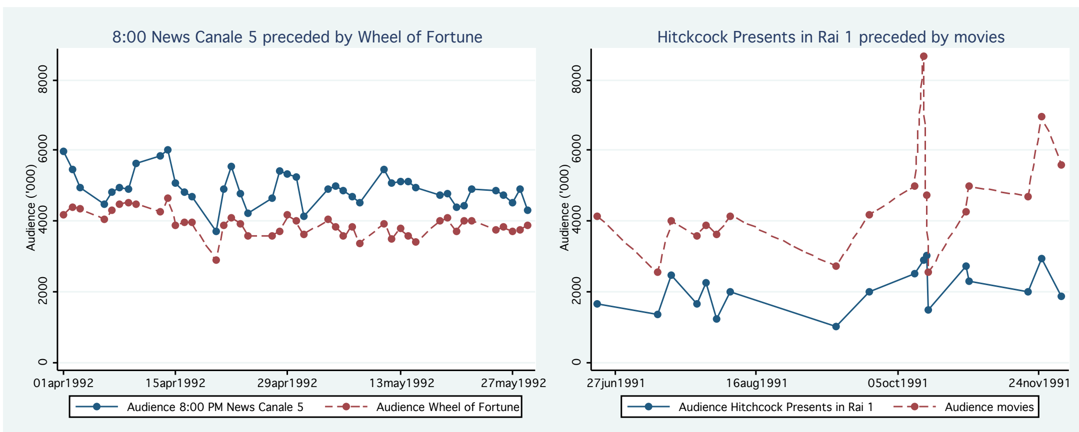


FIGURE III

Example of audience correlation between consecutive shows on a channel

The left panel depicts the audiences of the 8:00 P.M. news in Canale 5 and that of its preceding program, the game-show Wheel of Fortune. The right panel depicts the audiences of Hitchcock Presents on Rai 1 and that of its preceding program, movies.

Table I
Magnitude of decay rate of inertia on Rai 1 (2002-2003)

	Adjusted average Rai 1 viewers before and after the start of the late news - ('000)				
	On female show days			On soccer days	
	Event window			Event window	
	-60 to -1 minutes	+1 to +30 minutes		-60 to -1 minutes	+1 to +30 minutes
	(1)	(2)		(3)	(4)
Constant	1,441	901		1,470	945
	(12)***	(5)***		(22)***	(7)***
Women x Rai 1	583	262		583	262
	(129)***	(64)***		(131)***	(66)***
Women x Rai 1x x Female Show	1,015	30	Women x Rai 1x x Soccer	-226	-387
	(157)***	75		(222)	(106)***
Men	-	-		-	-
Men x Rai 1	679	289		679	289
	(167)***	(56)***		(170)	(57)***
Men x Rai 1x x Female Show	-104	-171	Men x Rai 1x x Soccer	1,560	-33
	182	(63)***		(294)**	(78)
Row 3 - Row 6	1,120	201	Row 6 - Row 3	1,787	354
	(128)***	(70)***		(145)***	(85)***
Cal day x minute x gender fixed effects	Yes	Yes		Yes	Yes
Calendar days	139	139		28	28
Channels	6	6		6	6
N	100,080	49,272		20,160	10,080

Notes: Row 3 - Row 6 in Column (1) shows the average adjusted gap in women-men during the hour preceding the start of the news on Rai 1 (-60 to -1 minutes event window) on female show days. Column (2) does so for the thirty minutes after the start of the news (+1 to +30 minutes event window). Row 6-Row 3 in Column (3) shows the average adjusted gap in men-women during the hour preceding the start of the news on Rai 1 (event window -60 to -1 minutes) on soccer days. Column (4) does so for the thirty minutes after the start of the news (event window +1 to +30 minutes). The sample for the adjusted audience estimation in columns (1) and (2) comprises 127 female show days and 12 neutral shows days and the six channels to facilitate the estimation of calendar day, minute of the calendar day and gender fixed effects for the adjusted estimation. The sample for the adjusted audience estimation in columns (3) and (4) comprises 16 soccer days and 12 neutral show days, as well at the fixed effects mentioned previously. Standard errors in parentheses, clustered by calendar day; ***significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

Table II

Effect of competition on women's inertia into the news and news talk-show on Rai 1 on female show days

	Average Rai 1 viewers in the half-hour after the start of the late news - ('000) +1 to +30 minutes event window					
	Below-median competition	Above-median competition	Difference (2)-(1)	Low outset competition	High outset competition	Difference (5)-(4)
	(1)	(2)	(3)	(4)	(5)	(6)
Women	1,214	1,234	20	1,200	1,300	100
	(47)***	(52)***	(52)	(43)***	(80)***	(73)
Men	886	906	20	873	982	109
	(43)***	(44)***	(39)	(36)***	(70)***	(62)
Calendar days	76	51	127	109	18	127
Day of the week dummies		Yes			Yes	
N (cal. dayXminuteXgender)		6,692			6,692	
Number of clusters (cal. days)		127			127	

Notes. Below-median competition indicates days when one or no competing channels are airing female shows during the news on Rai 1 (+1 to +30 minutes event window). Above-median competition indicates days when more than one channel does so. Low outset competition indicates days when no channels start airing a female show in the period from the commercial break before the start of the late news on Rai 1 to five minutes into the news. High outset competition indicates days when one or more channels do so. The 6,692 number of observations instead of the potential 127x30x2=7,620 is due to the audience data ending at midnight and the late news on Rai 1 starting after 11:00 P.M. on some days. Specifications control for time-invariant unobservable day-of-the-week factors affecting viewership. Standard errors in parentheses, clustered by calendar day; ***significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

Table III

Source of Rai 1 viewers – Rai 1 and total television audience before the news

	Average viewers in the hour before the start of the late news on Rai 1 - ('000) -60 to -1 minute event window					
	Rai 1			Total television		
	Female show days	Soccer days	Difference (1)-(2)	Female show days	Soccer days	Difference (4)-(5)
	(1)	(2)	(3)	(4)	(5)	(6)
Women	3,176	2,211	965	12,100	12,510	-411
			(170)***			(220)*
Men	1,884	3,407	-1,523	9,699	10,300	-601
			(200)***			(180)***

Notes. Estimation based on $N = 143$ days: 127 female show days plus 16 soccer days. The total number of women watching Rai 1 the hour before the start of the late news on female show days outnumber women watching soccer games in Rai 1 by almost 1 million. However, the total number of women watching television in the hour before the start of the late news on both types of days is only marginally different. The total number of men watching Rai 1 the hour before the start of the late news on soccer days outnumber men watching female shows by 1.5 million. However, the total number of men watching television when soccer games air on Rai 1 only outnumber the total number of men watching television when female shows air on Rai 1 by 0.6 million. Standard errors in parentheses, clustered by calendar day; ***significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

Table IV

Summary statistics for main variables by type of specification

	Mean	St. Dev.	Min	Max
Panel A. Sample: All shows, N=133,248				
Audience of focal show ('000)	2,374	1,812	55	23,543
Audience of preceding program ('000)	2,434	1,881	44	23,543
Index of competition on popularity for focal show ('000)	2,560	1,078	284	8,058
Genre overlap for focal show	0.09	0.12	0.00	1.00
Panel B. Sample: Show preceded by movies, N=305				
Audience of focal show ('000)	5,693	2,155	1,225	16,080
Audience of preceding program ('000)	1,890	717	519	6,284
Index of competition on popularity for focal show ('000)	1,995	863	547	4,411
Genre overlap for focal show	0.15	0.16	0.00	0.59
Panel C. Sample: News preceded by movies, N=143				
Audience of focal show ('000)	5,991	2,352	1,540	16,080
Audience of preceding program ('000)	2,051	777	662	6,284
Index of competition on popularity for focal show ('000)	2,373	803	991	4,411
Genre overlap for focal show (news)	0.08	0.10	0.00	0.38
Panel D. Sample: Main news preceded by any program, N=16,695				
Audience of focal show ('000)	2,267	1,486	122	7,392
Audience of preceding program ('000)	3,875	2,459	219	13,989
Index of competition on popularity for focal show ('000)	2,720	797	461	5,774
Genre overlap for focal show (main news)	0.18	0.10	0.00	0.63

Notes. Summary statistics for main variables by type of specification. Panel A focuses on the relationship between the demand for any two consecutive shows on a channel. Panel B focuses on the relationship between the demand for a show and that for its preceding movie. Panel C focuses on the relationship between the demand for the news (late night news and main news which play at 6:30, 7:00, 7.30, 8:00 and 8:30 P.M) at that for its preceding movie. Panel D focuses on the relationship between the demand for the main news and that for their preceding programs. Does not include weekends.

Table V
Partial correlation between the audience of consecutive shows on a channel

Sample[1]:	All shows 1990-2003						
Dependent variable:	Ln audience of focal show						
OLS specifications:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Demand for preceding program (in ln audience)	0.674	0.595	0.583	0.400	0.501	0.406	0.380
	(0.002)***	(0.002)***	(0.002)***	(0.003)***	(0.003)***	(0.004)***	(0.004)***
Controls:							
Competition on popularity[2]	-	0.365	0.377	0.565	0.157	-0.018	-0.112
(ln index audience of competing shows in prior month)	-	(0.004)***	(0.004)***	(0.004)***	(0.004)***	(0.006)***	(0.007)***
Genre overlap	-	-	0.775	0.591	-0.001	0.010	-0.015
(% of time genre overlaps with other channels')	-	-	(0.013)***	(0.012)***	(0.010)	(0.013)	(0.011)
Channel FE	-	-	-	Yes	-	-	-
Channel X Show FE	-	-	-	-	Yes	-	-
Channel X Show X Year X Month FE	-	-	-	-	-	Yes	-
Channel X Show X Year X Month X 1/2 hour slot FE	-	-	-	-	-	-	Yes
R-squared	0.43	0.48	0.50	0.58	0.89	0.95	0.96
N (number of distinct show episodes)	133,258	133,258	133,258	133,258	133,258	133,258	133,258
Number of days (clusters)	3,630	3,630	3,630	3,630	3,630	3,630	3,630

Notes.[1] Does not include weekends and titles that play sequentially (e.g., Friends followed by Friends); the focal show does not include titles that play only once in a given channel, year, month and half-hour slot, such as a special features or some movies. [2] Using index of ln audience of competing shows in the prior month as an instrument instead of as a proxy for the popularity of simultaneous competition to the focal show does not change the estimated coefficient on the variable of interest: Demand for the preceding program. Standard errors in parentheses, clustered by calendar day; ***significant at the 1% level; **significant at the 5% level.

Table VI

How the demand for a show or the news is affected by the preceding movie (1990-2003)

Sample:[1]	Any show preceded by movies			News preceded by movies		
Dependent variable:	Ln audience show preceded by movies		Ln audience movie	Ln audience news preceded by movies		Ln audience movie
Specification:	OLS	2SLS	1st stage: OLS	OLS	2SLS	1st stage: OLS
	(1)	(2)	(3)	(4)	(5)	(6)
Demand for preceding program (movie) - ln audience	0.566	0.483	-	0.695	0.388	-
(Instrumented with Ln Italian theatrical audience for movie)	(0.084)***	(0.107)***	-	(0.121)***	(0.205)*	-
1st stage: movie followed by news						
Ln Theatrical audience movies			0.062			0.060
			(0.011)***			(0.015)***
t-stat 1st stage			5.636			4.050
Controls:						
Competition on popularity [2]	-0.075	-0.077	0.043	-0.008	-0.01	0.062
(ln index audience of competing shows in prior month	(0.076)	(0.058)	(0.098)	(0.117)	(0.136)	(0.159)
Genre overlap	0.132	0.138	0.093	0.266	0.354	0.268
(% of time genre overlaps with other channels')	(0.127)	(0.097)	(0.171)	(0.210)	(0.226)	(0.298)
Channel X Show X Year X Month X 1/2 hour slot FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.83	-	0.82	0.79	-	0.82
N (Number of movie-show pairs)	305	305	305	143	143	143
Number of days (clusters)	268	268	268	129	129	129
Average length of show or news preceded by movie (minutes)	44			11		

Notes.[1] Does not include weekends;[2] Using index of ln audience of competing shows in the prior month as an instrument instead of as a proxy for the popularity of simultaneous competition to the focal show does not change the estimated coefficient on the variable of interest: Demand for the preceding program (movie). Standard errors in parentheses, clustered by day; ***significant at the 1% level; **significant at the 5% level.

Table VII

Effect of the program preceding the main news on its demand and the decay rate to inertia (1990-2003)

Sample:[1]	Main news,1990-2003[3]					
Dependent variable:	Ln audience of news		Ln audience of program preceding the news	Ln audience news below median length	Ln audience news above median length	Ln audience show after the main news
Specification:	Pooled OLS	2SLS	1st stage: OLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Demand for program preceding the news - in ln audience	0.353	0.220	-	0.298	0.147	0.091
(Instrumented with ln average audience in the previous month for 2SLS)	(0.012)***	(0.028)***	-	(0.044)***	(0.034)***	(0.046)**
1st stage: show followed by news						
Ln of average audience in prior month			0.218			
			(0.020)***			
t-stat 1st stage			10.90			
Controls:						
Competition on popularity[2]	-0.089	-0.089	0.001	-0.090	-0.094	-0.095
(ln index of audience of competing shows in prior month)	(0.025)***	(0.026)***	(0.024)	(0.027)***	(0.035)***	(0.017)***
Genre overlap	0.182	0.162	-0.147	0.091	0.236	-0.062
(% of time genre overlaps with other channels')	(0.061)***	(0.062)***	(0.068)**	(0.062)	(0.118)**	(0.024)**
Channel X Show X Year X Month X 1/2 hour slot FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.98	-	0.95	-	-	-
N (Number of show-news pairs)	16,695	16,695	16,695	9,708	6,987	15,050
Number of days (clusters)	3,589	3,589	3,589	3,528	3,020	3,583
Average length of main news show (minutes)	32			28	38	24

Notes.[1] Does not include weekends; does not include news shows that play sequentially and news shows which are "Extraordinary Editions", like a news special on 9/11/2001;[2] Using index of ln audience of competing shows in the prior month as an instrument instead of as a proxy for the popularity of simultaneous competition to the focal show does not change the estimated coefficient on the variable of interest: Demand for the preceding program;[3] The main news are the standard daily news scheduled at 6:30 PM, 7:00 PM, 7:30 PM, 8:00 PM and 8:30 PM which last no less than 20 minutes. Standard errors in parentheses, clustered by day; ***significant at the 1% level; **significant at the 5% level; *significant at 10% level.

TABLE VIII

Other findings and mechanisms

Sample:[1]	Main news	Main news	Main news (except Rai 1 and Canale 5 at 8:00 P.M.)	Main news
Dependent variable:	Ln audience main news			
Specification:	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)
Demand for preceding program - in ln audience (Instrumented with ln average audience in the previous month)	0.220 (0.028)***	0.220 (.046)***	0.250 (0.036)***	0.223 (0.028)***
Demand of preceding program X no shows starting in same minute vicinity (In ln audience: Instrumented ln of average audience in the previous month)				-0.007 (0.004)
Demand of preceding program X No uncertainty about competing shows (In ln audience: Instrumented ln of average audience in the previous month)		-0.001 (0.007)		
Controls:				
Channels starting in same minute vicinity (=1 if no shows starting in the same minute vicinity)				0.046 (0.030)
No uncertainty about competing shows to current show (=1 if no competing programs are novel)		0.004 (0.056)		
Competition on popularity [2] (ln index aud of competing shows in prior month)	-0.089 (0.026)***	-0.090 (0.034)***	-0.115 (0.029)***	-0.089 (0.026)***
Genre overlap (% of time genre overlaps with other channels')	0.162 (0.062)***	0.098 0.076	0.164 (0.071)**	0.164 (0.062)***
Channel X Show X Year X Month X 1/2 hour slot FE	Yes	Yes	Yes	Yes
N (Number of show-news pairs)	16,695	7,424	10,759	16,695
Number of days (clusters)	3,589	3,178	3,575	3,589

Notes.[1]Does not include weekends;[2]Using index of ln audience of competing shows in the prior month as an instrument instead of as a proxy for the popularity of simultaneous competition to the focal show does not change the estimated coefficient on the variable of interest: demand for the preceding program. Column (1) re-displays the effect of inertia on the demand for main news found previously. Column (2) analyses the effect of uncertainty about competing shows to the main news on inertia into the news. Novel (with higher uncertainty) shows are classified as such if they are in their first quartile of episodes and not novel otherwise. Column (3) analyses whether inertia into the main news exists in the absence of asymmetric information by removing from the main news sample the main news for Canale 5 and Rai 1 which are advertised with a clip during the preceding program. Column (4) analyses whether having no competing shows starting at the same time (1 minute vicinity) as the main news versus having more than one affects inertia into the main news. Standard errors in parentheses, clustered by day; ***significant at the 1% level; **significant at the 5% level; *significant at 10% level.

Table IX

Current, optimal and worst schedule for flagship channels Canale 5 and Rai 1 (2003)

Panel A: Privately-controlled flagship Canale 5 - 2003 prime-time schedule							
	Show number	Time (P.M.)	Average Audience (millions)	Optimal Schedule	Average Audience (millions)	Worst Schedule	Average Audience (millions)
Pre-prime time							
Quiz show	s_0	6:40-8:00	3.753				
Prime-time							
8:00 PM News	s_1	8:00-8:30	6.422	Half-hour variety show	5.981	Two-hour programs[1]	5.376
Half-hour variety show	s_2	8:30-9:00	7.133	8:00 PM News	7.502	Half-hour variety show	6.754
Two-hour programs[1]	s_3	9:00-11:00	5.750	Two-hour programs[1]	5.780	8:00 PM News	7.632
Weighted average audience 8:00-11:00 PM			6.092		6.101		5.982
% Difference versus optimal schedule			0.1%				2.0%

Panel B: State-owned flagship Rai 1 - 2003 prime-time schedule							
	Show number	Time (P.M.)	Average Audience (millions)	Optimal Schedule	Average Audience (millions)	Worst Schedule	Average Audience (millions)
Pre-prime time							
Quiz show	s_0	6:40-8:00	4.330				
Prime-time							
8:00 PM News	s_1	8:00-8:30	6.924	Half-hour variety show	4.567	Two-hour programs[1]	5.248
Half-hour variety show	s_2	8:30-9:00	5.319	8:00 PM News	7.149	Half-hour variety show	4.890
Two-hour programs[1]	s_3	9:00-11:00	5.359	Two-hour programs[1]	5.530	8:00 PM News	7.155
Weighted average audience 8:00-11:00 PM			5.613		5.662		5.506
% Difference versus optimal schedule			0.9%				2.7%

Notes.[1] For example, Movies, Series, Miniseries. Backup calculations are in Appendix Table D.1.

A Appendix A – Audience measurement details

The most common remotes (Type I) have a button for each member of the household. Pushing once the member's button plus the OK button confirms that the person is watching; pushing twice plus OK indicates she is not. A less prevalent type (Type II) has one button for all members of the household plus an upward and downward arrow to interact with the meter.

Figure A.1
Version of Type I remote that interacts with the television meter

B Appendix B - Event-study sample construction

Table B.1
Sample for minute-by-minute event-study analysis in 2002-2003

January 1st 2002-December 31st 2003	Total	Average start (P.M.)	Standard deviation (min)
Days with late (11:00 PM news in Rai 1)[(1)]	498	11:08	16.8
Total days with late news followed by Porta-a-Porta in Rai 1[(2)]	253	11:12	13.9
Observations for analysis			
Total female show days[(3)]+News+Porta-a-Porta in Rai 1	127	11:17	13.5
Total male show (soccer) days[(4)] + News +Porta-a-Porta in Rai 1	16	11:03	6.7
Total neutral show days[(5)]+News+Porta-a-Porta in Rai 1	12	11:16	11.4
Total number of days for analysis[(6)]	155		

Notes: (1) Does not include weekends; (2) Porta-a-Porta is off the air in the summer and occasionally does play after the 11:00 P.M. news in Rai 1; (3) Female show is a show where women outnumber men on every episode (e.g., Incantesimo, a series on the romantic lives of doctors and nurses at a Roman hospital; I Racommandati, a singing talent show) and that lasts one hours or more; (4) Male show is a show where men always outnumber women, such as soccer, and that lasts one hour or more; (5) neutral show is a show where women outnumber men on some episodes but not on others (e.g. science show Superquark or Porta-a-Porta which occasionally airs before the late news in Rai 1) and that lasts one hour or more; (6) 57 days were removed from the analysis because the programs before the late news had only aired once and therefore we had no criteria to classify them; an additional 41 days were removed because the duration of the show before the start of the late news was less than one hour.

C Appendix C - Model proofs

The proofs for propositions 1 and 2 follow approximately the same logic as Carroll et al. (2009).

Proposition 1. The consumer's game has a unique stationary equilibrium:

i. If $G \leq \underline{c} + \frac{\hat{b_d}}{1-\delta}$, then $c^* = \underline{c}$ (the consumer never switches)

ii. If $G \geq \frac{\bar{c}-\delta E[c]+\hat{b_d}}{1-\delta}$, then $c^* = \bar{c}$ (the consumer switches, no matter the cost)

iii. otherwise, c^* is the unique solution to the equation:

$$c^* = \frac{1}{1-\delta+\delta P(c_{t+1} \leq c^*)}\{G(1-\delta) - \hat{b_d} + \delta E[c_{t+1}|c_{t+1} \leq c^*]P(c_{t+1} \leq c^*)\}$$

Proof. The solution to the problem $-c_t+G \geq \hat{b_d}+\delta E[V(c_{t+1}, b_a, \hat{b_d})]$ is a cut-off c^* whereby the consumer is indifferent between switching and not switching channels: If the cost of switching at each period is less than or equal to c^* the consumer switches the channel, and stays on the default channel otherwise.

The continuation value of the game if the consumer does not switch the channel is:

$$\begin{aligned}\phi \equiv E[V(c_{t+1}, \hat{b_d}, \delta, c^*)] &= \frac{1}{1-\delta+\delta P(c_{t+1} \leq c^*)}\{E[-c_{t+1}|c_{t+1} \leq c^*]P(c_{t+1} \leq c^*) \\ &+ GP(c_{t+1} \leq c^*) + \hat{b_d}(1-P(c \leq c^*)\}\end{aligned}$$

where G as defined in the main text is $G \equiv E[b_a]+E[b_a|b_a \geq \hat{b_d}]P(b_a \geq \hat{b_d})\frac{\delta}{1-\delta}+\hat{b_d}P(b_a < \hat{b_d})\frac{\delta}{1-\delta}$. We obtain the continuation value ϕ by noting that:

$$V(c_{t+1}, \hat{b_d}, \delta) = \begin{cases} -c_{t+1}+G & \text{if } c_{t+1} \leq c^* \\ \hat{b_d}+\delta E[V(c_{t+2}, \hat{b_d}, \delta)] & \text{if } c_{t+1} > c^* \end{cases}$$

Then $E[V(c_{t+1}, \hat{b_d}, \delta, c^*)] = E[-c_{t+1}+G|c_{t+1} \leq c^*]P(c_{t+1} \leq c^*)+(\hat{b_d}+\delta E[V(c_{t+2}, c^*)]P(c_{t+1} > c^*)$. Since c_t, c_{t+1}, c_{t+2} is i.i.d. then $E[V(c_{t+1}, \hat{b_d}, \delta, c^*)] = E[V(c_{t+2}, \hat{b_d}, \delta, c^*)]$. We then solve for $E[V(c_{t+1}, \hat{b_d}, \delta, c^*)]$ since c^* is the solution across all time periods.

Since c^* is constrained to belong to $[\underline{c}, \bar{c}]$,

$$c^* = \begin{cases} \underline{c} & \text{if } G-\hat{b_d}-\delta\phi \leq \underline{c} \\ G-\hat{b_d}-\delta\phi & \text{if } \underline{c} < G-\hat{b_d}-\delta\phi < \bar{c} \\ \bar{c} & \text{if } G-\hat{b_d}-\delta\phi \geq \bar{c} \end{cases}$$

Now we prove (i): $c^* = \underline{c}$ is an equilibrium to the game if and only if $G-\hat{b_d}-\delta\phi \leq \underline{c}$. For $c^* = \underline{c}$, $\phi = \frac{\hat{b_d}}{1-\delta}$ and $G \leq \underline{c}+\frac{\hat{b_d}}{1-\delta}$.

For (ii), $c^* = \bar{c}$ is an equilibrium to the game if and only if $G-\hat{b_d}-\delta\phi \geq \bar{c}$. For $c^* = \bar{c}$, $\phi = -E[c]+G$ and $G \geq \frac{\bar{c}-\delta E[c]+\hat{b_d}}{1-\delta}$.

For (iii), $c^* \in (\underline{c}, \bar{c})$ is an equilibrium of the game if and only if $G-\hat{b_d}-\delta\phi = c^*$. We prove that c^* exists and is unique. First, we obtain

$$c^* = \frac{1}{1-\delta+\delta P(c_{t+1} \leq c^*)}\{G(1-\delta)-\hat{b_d}+\delta E[c_{t+1}|c_{t+1} \leq c^*]P(c_{t+1} \leq c^*)\}$$

by noting that:

$$\begin{aligned}
-c^* &= -G + \hat{b_d} + \delta\phi \\
&= -G + \hat{b_d} + \frac{\delta}{1-\delta+\delta P(c_{t+1} \leq c^*)}\{E[-c_{t+1}|c_{t+1} \leq c^*]P(c_{t+1} \leq c^*) + GP(c_{t+1} \leq c^*) + \\
&+ \hat{b_d}(1 - P(c_{t+1} \leq c^*)\}
\end{aligned}$$

and further simplifying by eliminating common terms with G and $\hat{b_d}$.

Then we define $H(c^0) = g(c^0) - c^0$, where $g(c^0) = \frac{1}{1-\delta+\delta P(c \leq c^0)}\{G(1-\delta) - \hat{b_d} + \delta E[c|c \leq c^0]P(c \leq c^0)\}$, where c replaces c_{t+1} to simplify notation. The function g is differentiable with

$$g'(c^0) = (1-\delta+\delta F(c^0))^{-1}\delta c^0 f(c^0) - (1-\delta+\delta F(c^0))^{-2}\delta f(c^0)(G(1-\delta) - \hat{b_d} + \delta E[c|c \leq c^0]F(c^0)).$$

Since $H(\underline{c}) = g(\underline{c}) - \underline{c} > 0$ when $G > \frac{\hat{b_d}}{1-\delta} + \underline{c}$ as $g(\underline{c}) = G - \frac{\hat{b_d}}{1-\delta}$ and $H(\bar{c}) = g(\bar{c}) - \bar{c} < 0$ when $G < \frac{\bar{c}-\delta E[c]+\hat{b_d}}{(1-\delta)}$ as $g(\bar{c}) = G(1-\delta) - \hat{b_d} + \delta E[c]$, c^* exists. Further, when $c^0 = c^*$, $g'(c^*) = 0$. Therefore, $H'(c^*) = -1 < 0$. Thus $H(c^0) = g(c^0) - c^0$ has unique root c^* in $(\underline{c}, \bar{c})$.

Proposition 2. In the region where $\underline{c} < c^* < \bar{c}$, c^* is strictly increasing in G and strictly decreasing in $\hat{b_d}$.

Proof. Let $H(c^*) \equiv g(c^*) - c^* = 0$. Holding c^* constant, if we increase G, $g(c^*)$ increases. Therefore, c^* needs to increase by a positive amount to c^{**}, so that $H(c^{**}) \equiv g(c^{**}) - c^{**} = 0$. Since H is decreasing when cost $\in (\underline{c}, \bar{c})$, the new cut-off $c^{**} > c^*$. Conversely, if we increase $\hat{b_d}$, $g(c^*)$ decreases and the new cut-off $c^{**} < c^*$.

Proposition 3 (i) The cut-off for the fully naïve consumer is: $c^{*,naive} = \beta c^{*,exp} + (1-\beta)\Delta$; (ii) $c^{*,naive} = c^{*,exp}$ if $\beta = 1$; (iii) $c^{*,naive} < c^{*,exp}$ if $\beta \in (0,1)$, (iv) $c^{*,naive}$ is strictly increasing in β if $\beta \in (0,1)$.

Proof. For (i) the fully naïve consumer discounts the immediate future by $\beta\delta$. Therefore, she solves:

$$V(c_t, b_a, \hat{b_d}, \beta, \delta) = \begin{cases} -c_t + E[b_a] + E[b_a|b_a \geq \hat{b_d}]P(b_a \geq \hat{b_d})\frac{\beta\delta}{1-\delta} + \hat{b_d}P(b_a < \hat{b_d})\frac{\beta\delta}{1-\delta} & \text{if switch} \\ \hat{b_d} + \beta\delta E[V(c_{t+1}, b_a, \hat{b_d}, \beta, \delta)] & \text{if not switch} \end{cases}$$

The cut-off for the fully naïve consumer is therefore,

$$-c^{*,naive} + E[b_a] - \hat{b_d} = \beta\{-G + E[b_a] + \delta\phi\} \tag{1}$$

whereas the cut-off c^* for the time-consistent consumer, obtain in the previous section (equation 1) was

$$-c^{*,exp} + E[b_a] - \hat{b_d} = -G + E[b_a] + \delta\phi \qquad \text{where} \qquad -c^{*,exp} = -c^* \tag{2}$$

Plugging equation (3) into (2), we find that $c^{*,naive} = \beta c^{*,exp} + (1-\beta)(E[b_a] - \hat{b_d})$ where $E[b_a] - \hat{b_d} = \Delta$ given the distributional assumptions of the difference in benefits. Therefore

$$c^{*,naive} = \beta c^{*,exp} + (1-\beta)\Delta$$

We obtain (ii) by plugging $\beta = 1$ in the expression in (i). For (iii) we show that $c^{*,exp} > \Delta$. Recall that $\sigma > 0$ and $\underline{c} < c^{*,exp} < \bar{c}$, that is, $0 < c^{*,exp} < 1$. In case two ($\Delta \geq \sigma$), $c^{*,exp} \simeq \sqrt{2\Delta} > \Delta$ for $\Delta \in (0,2)$. Since $0 < c^{*,exp} < 1$, then $\Delta \in (0, 1/2)$ and $c^{*,exp} > \Delta$. In case three ($-\sigma < \Delta < \sigma$), $c^{*,exp} \simeq \frac{\Delta+\sigma}{\sqrt{2\sigma}}$, therefore $\Delta \in (0, \frac{1}{2})$. We prove by contradiction that $c^{*,exp} > \Delta$. Suppose not: suppose $c^{*,exp} \simeq \frac{\Delta+\sigma}{\sqrt{2\sigma}} \leq \Delta$. Then $\Delta^2(1-2\Delta) + 2\sigma\Delta + \sigma^2 \leq 0$. For $\sigma = \frac{1}{2}$, $\Delta \leq -\frac{1}{4}$, contradicting $\Delta \in (0, \frac{1}{2})$. Otherwise factor into $(\Delta + \frac{\sigma}{1+\sqrt{2\sigma}})(\Delta + \frac{\sigma}{1-\sqrt{2\sigma}}) \leq 0$ where $\Delta \leq \frac{-\sigma}{1+\sqrt{2\sigma}} < 0$ and $\Delta \leq \frac{-\sigma}{1-\sqrt{2\sigma}} < 0$ for $\sigma < \frac{1}{2}$ violate $\Delta \in (0, \frac{1}{2})$. For (iv) the derivative of $c^{*,naive}$ with respect to β is positive.

D Appendix D - Enumeration of potential schedules and relationship between expected audiences and advertising rates

Table D.1

Enumeration of the six potential schedules for prime-time for Canale 5 and Rai 1 (2003)

Panel A: Enumeration of the six potential schedules for the privately-controlled flagship Canale 5 - 2003 prime-time schedule

	Current (s_0, s_1, s_2, s_3) (1)		(s_0, s_2, s_1, s_3) (2)		(s_0, s_3, s_2, s_1) (4)		(s_0, s_1, s_3, s_2) (5)		(s_0, s_2, s_1, s_3) (6)	
	Implicit ln baseline audience	Audience with inertia (millions)	Implicit ln baseline audience	Audience with inertia (millions)	Implicit ln baseline audience	Audience with inertia (millions)	Implicit ln baseline audience	Audience with inertia (millions)	Implicit ln baseline audience	Audience with inertia (millions)
s_0	1.323	1.000	1.323	1.000	1.323	3.753	1.323	3.753	1.323	3.753
s_i	1.463	6.422	1.407	5.981	1.540	5.376	1.407	5.981	1.463	6.422
s_j	1.407	7.133	1.463	7.502	1.407	6.754	1.540	5.669	1.540	5.712
s_k	1.540	5.750	1.540	5.780	1.463	7.632	1.463	7.242	1.407	6.762
Weighted average 8:00-11:00 PM		6.092		6.101		5.982		5.983		6.005

Panel B: Enumeration of the six potential schedules for the state-owned flagship Rai 1 - 2003 prime-time schedule

	Current (s_0, s_1, s_2, s_3) (1)		(s_0, s_2, s_1, s_3) (2)		(s_0, s_3, s_2, s_1) (4)		(s_0, s_1, s_3, s_2) (5)		(s_0, s_2, s_1, s_3) (6)	
	Implicit ln baseline audience	Audience with inertia (millions)	Implicit ln baseline audience	Audience with inertia (millions)	Implicit ln baseline audience	Audience with inertia (millions)	Implicit ln baseline audience	Audience with inertia (millions)	Implicit ln baseline audience	Audience with inertia (millions)
s_0	1.465	4.330	1.465	4.330	1.465	4.330	1.465	4.330	1.465	4.330
s_i	1.495	6.924	1.091	4.567	1.501	5.248	1.091	4.567	1.495	6.924
s_j	1.091	5.319	1.495	7.149	1.091	4.890	1.501	5.297	1.501	5.536
s_k	1.501	5.359	1.501	5.530	1.495	7.155	1.495	7.329	1.091	4.904
Weighted average 8:00-11:00 PM		5.613		5.639		5.506		5.514		5.662

Notes. s_0 denotes the show playing before prime-time in the original schedule; s_1 denotes the first show in prime-time; s_2 the second show; and s_3 the third and last show in prime-time. Column (1) shows the current schedule; columns (2) through (5) display the five other potential schedules, resulting from the five alternative sequences of programs. Implicit baseline audience=Ln average audience of current show - ρLn of Average audience of preceding program. Audience with inertia=exp[(Implicit log baseline audience)+ρ(log audience of preceding program show)]; $\rho = 0.3$ for a show averaging 28 minutes and $\rho = (0.3 + 0.3^2 + 0.3^3 + 0.3^4)/4 = 0.106$ for a two-hour show (which conforms with inertia into programs lasting 100 or more minutes). Program audiences are also adjusted for the average, for 2003, of total television audience during the time slot in which they play; variation in average total television audience across prime-time slots is no higher than 16%.

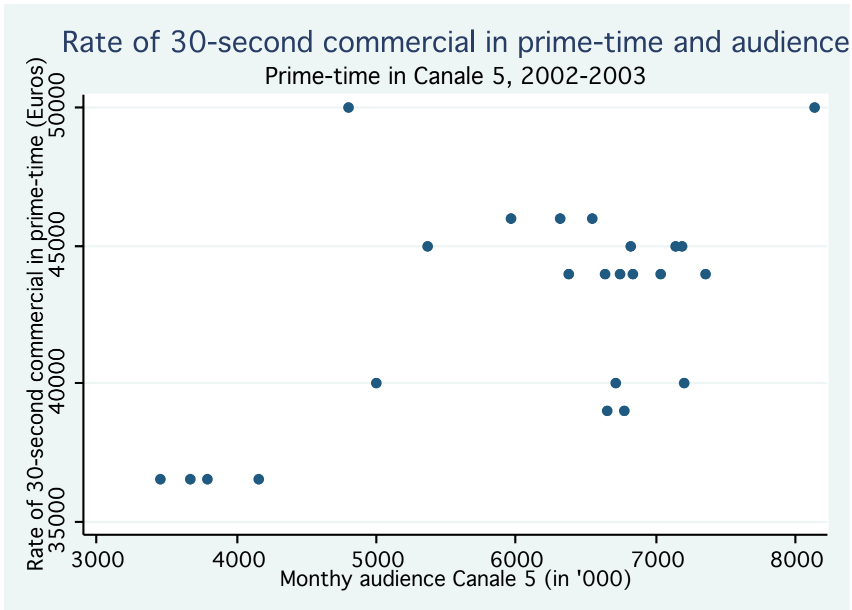
Estimation of the relationship between expected audience and advertising rates. Advertising rates are a function of the expected audiences for a daypart, in this case, prime-time. We only observe, however, the realised ex-post audiences.

To estimate the relationship between the expected audience and advertising rates, we assume that the relationship between the rate for a thirty-second commercial and expected audience is *ln rate*= β_1*ln expected audience* $+ v$, where *cov(v, ln expected audience)*=0 and *ln realised audience* = *ln expected audience*$-\epsilon$, where ϵ is the deviation from the logged expected audience. Therefore, *ln rate* = β_1 *ln realised audience*$+\beta_1\epsilon+v$. If *cov*(ϵ, *ln expected audience*)=0, then *cov*(ϵ, *ln realised audience*) $\neq 0$. The estimate of β_1 will be biased towards zero. This is the attenuation bias in the classical errors-in-variables. The OLS estimate of β_1 will be a lower bound on the effect of an increase in 1% in audience on the percent increase in the price of a thirty-second commercial.

Figure D.1 plots the relationship between advertising rates for a thirty-second commercial in prime-time and its audience for Canale 5 for each month in 2002-2003. The slope of the relationship between *Log advertising rate for a 30-second commercial* and *Log audience* is statistically significant at 1.2, suggesting that an increase in 1% in audience increases advertising rates by at least 1.2%.

This value is in the neighbourhood of that obtained in the U.S. Wilbur (2008), Table 1, page 362, shows the average advertising rates for thirty-second commercial and average audience for the six major networks in the U.S. during the sweeps in April 24-May 21 in 2003, from 8:00-10:00 P.M.. The audiences for UPN, WB, ABC, CBS, NBC and FOX in thousands of households were 2,793, 3,584, 5,693, 7,716, 7,361, and 8,058, respectively. The respective advertising rates for a thirty-second commercial in thousand of dollars, were 55, 71, 125,179, 212 and 241. A back of the envelope calculation, taking UPN as the baseline, shows that an increase in 1% in household audience versus UPN increases advertising rates by 1.4%, on average.

Figure D.1
Rate for a thirty-second ad and the monthly audience in prime-time for flagship Canale 5 (2002-2003)



E Appendix E - Additional Figures and Tables

FIGURE E.1

Typical Monday night line-up; Italy's 6 main channels are generalist with a broad range of genres

Date: Monday, October 7th, 2002

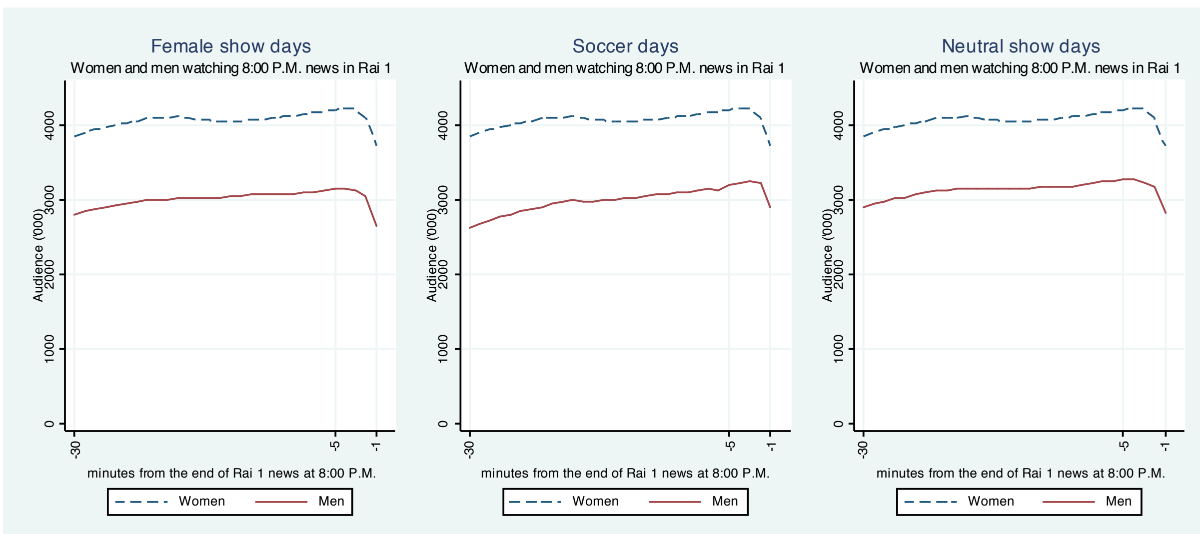
Starting at (PM): 6:00 | 6:30 | 7:00 | 7:30 | 8.00 | 8:30 | 9:00 | 9:30 | 10:00 | 10:30 | 11:00 | 11:30 | 12:00

Prime time: 8:00 PM to 11:00 PM

Channel									
Rai 1	Variety La vita in diretta	Game show/Quiz L'eredita	News Rai 1	Max & Tuck	Variety Super-variety	Film Public Enemy	News Rai 1	News talk show Porta-a-Porta	
Rai 2	Sport show Sport sera	Varie-ty	Soap Opera Cuori Rubati	TV series Cobra Special Squad	Cartoons Popeye	News Rai 2	Show Big Monday night	Variety Women stories	
Rai 3	Cultural Program Geo & Geo	News Rai 3	Regional News Rai 3	Sport Rai Sport	Varie-ty Blob di Tutto Piu	Soap Opera Posto al Sole	Variety show Chi l'ha visto	News Rai 3	News magazine Primo-Piano
Canale 5	Variety Verissimo	Game show Passaparola	News Canale 5	Variety Striscia	Miniserie Francesco	Talk-Show Maurizio Costanzo			
Italia 1	Sitcom Fresh Prince of Bel-Air	News Studio Aperto	Reality Show Operation Triumph	Sitcom Dharma & Greg	Game show/Quiz Sarabanda	TV series Carabinieri	Made for TV movie Da Cosa nasce Cosa		
Rete 4	Film	News Rete 4	Variety Sipario	Soap Opera Terra Nostra	Film El Dorado	Cultural program Traveling in the time machine			

Figure E.2

Do women and men watching the 8:00 P.M. news on Rai 1 on female show, soccer and neutral show days differ?



Notes. The 8:00 P.M. news on Rai 1 last on average almost thirty minutes.

Table E.1
Overview of composition of television shows, 1990-2003, in Italy's 6 main channels

Show genre	Description	Freq.	%	Average length minutes (minutes)	Average number of episodes per show
NEWS	Shows summarizing daily local and international news, such as the 6:00 P.M. news in the U.S.	43,602	22	22	206
VARIETY	Entertainment shows based on current events, such as mock news and missing persons mysteries	22,462	11	36	28
SHOW	Mostly talk-shows	19,422	10	74	25
TV SERIES	Mainly TV drama series such as CSI, the X-files, ER or Xena Warrior Princess	17,547	9	55	45
FILM	All movies except made for TV movies	14,152	7	108	3
GAME SHOW	Games shows	14,064	7	46	177
SPORTS SHOW	Mainly shows about current, past or future sports events e.g., past Olympic games	13,608	7	19	55
NEWS MAGAZINE	Mainly feature on current news events, such as 20/20 or 60 minutes in the U.S.	10,695	5	39	20
CARTOON	Mainly short animated features	7,056	4	16	38
SITCOM	Situational comedies, as in the U.S.; includes shows such as Friends	6,606	3	29	67
CULTURAL PROGRAM	Programs designed to educate viewers, such as documentaries on science, history or the arts	6,426	3	53	11
SOAP OPERA	Daily drama shows, similar to soap operas in the U.S.	5,956	3	43	109
SPORTS EVENT	Mainly the broadcast of sports events, such as soccer, basketball, tennis and volleyball	3,880	2	69	5
MADE FOR TV MOVIE	Movies made for television	2,828	1	102	2
MUSIC	Includes concerts, music festivals, and performances by well-know singers	2,320	1	65	4
PROMOTIONAL PROGRAM	Mainly short shows designed to sell a product, service	2,253	1	6	50
MOVIE COMMENTARY	Show commenting on movie	2,027	1	9	946
MINISERIE	TV series with usually fewer than thirteen episodes	1,230	1	101	5
REALITY TV	Non-scripted TV show based on real-life situations	1,212	1	42	34
Total		197,346	100	45	16

Note: Does not include weekends

Table E.2

Topics of Porta-a-Porta talk-show on female show days, soccer days and neutral show days on Rai 1 (2002-2003)

Female show days (N=16, from random sample of 127 days)	Soccer days (N=16, all days)	Neutral show days (N=12, all days)
Program on famous Italian film producer and politician	The hunt for Osama Bin Laden and Mullah Omar	Interview with television personality
Corruption and politics	The new custody law for children of separated parents	Love, loneliness and other issues for the elderly
The effects of power and who rules the home, men or women?	No description	Domestic and workplace violence
Euthanasia	Crisis in Argentina, what it means for its Italian population	Discussion of the elections
Interview with the current prime-minister, Silvio Berlusconi	Mad Cow disease	Not available
Love and war: reporters that die while covering war zones	The unsolved mystery of who killed little Samuel	The U.S. attack on Iraq
Interview with the president of Democrats for the Left	Anti-obesity drugs and dieting	Not available
The crisis of Fiat and the economic policies of the Government	Pope John XXIII, his life and works	Reform of pensions and labor markets
Interview with the Minister for the Economy	Not available	A review of pope John Paul II
The radical group that tried to derail the G8 summit	Prostitution and its links to celebrities	How do single people live in Italy
Discussion of the prostitution law	Interview with high-level political personality	The arts of seduction
Whether television should be regulated to protect children	Discussion on new law to increase labor market flexibility	India and its progress
The UN security council reaction to George Bush's Iraq ultimatum	Not available	
The return of the exiled Savoy royal family to Italy	Law that forbids smoke in offices	
Not available	Not available	
The soul and life after death	Broadcasting reform law	

Source: Porta-a-Porta archive, Rai 1.