A Household Level Model of Television Viewing with Implications for Advertising Targeting

by

Yiting Deng

Business Administration
Duke University

Date: ______________________
Approved:

____
Carl F. Mela, Co-Supervisor

____
Kenneth C. Wilbur, Co-Supervisor

____
Richard Staelin

____
Peng Sun

____
Yi (Daniel) Xu

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Business Administration
in the Graduate School of Duke University
2015

Abstract

A Household Level Model of Television Viewing with Implications for Advertising Targeting

by

Yiting Deng

Business Administration
Duke University

Date: ____________________

Approved:

____
Carl F. Mela, Co-Supervisor

____
Kenneth C. Wilbur, Co-Supervisor

____
Richard Staelin

____
Peng Sun

____
Yi (Daniel) Xu

An abstract of a dissertation submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in Business Administration
in the Graduate School of Duke University
2015

Abstract

Television (TV) is the predominant advertising medium, and recent technological advances such as digital video recorders (DVRs) and set-top boxes (STBs) have the potential to transform this industry by enabling household-specific advertising. Since exposure to TV represents a substantial share of consumer time and attention, this potential to micro-target communications represents an enormous opportunity for the TV advertising market.

This paper outlines an approach to facilitate the micro-targeting of TV advertising. We employ a unique dataset, integrating TV program and advertisement viewing at the household level with purchase data, to address the question of how advertisers can achieve better advertising targeting in the digital context. Based on this dataset, we first develop a model of household TV viewing behavior. The viewing model comprises three integrated components: TV show sampling and watching, TV show recording, and advertising viewing. All three components are motivated by the theoretical concept of flow utility, that is, the moment-by-moment enjoyment a household derives from different activities: watching a TV show, watching a TV advertisement, and other non-TV activities. This model has decent out-of-sample prediction power on show choices and time spent on each selected show. We then link household advertising exposure with purchase. Finally, the viewing model and identified advertising-sales relationship are utilized to conduct counterfactual policy experiments on advertising targeting. We consider several household-level targeting

scenarios by manipulating: 1) whether the advertising purchase is made in advance or in real time; and 2) whether the objective function is to minimize costs for a given set of exposures or to maximize incremental profit from advertising. Results indicate micro-targeting can lower advertising costs and raise incremental revenue.

The key contributions of this paper are as follows. Theoretically, we develop an integrated model on TV show viewing, TV advertising viewing, purchasing and advertising targeting. Methodologically, we propose a new modeling framework on media consumption by explicitly accounting for the role of uncertainty, and propose targeting strategies leveraging household-level data. Substantively, we offer policy recommendations to advertisers on micro-targeting which can be of great potential.

Keywords: TV advertising, targeting, DVR, sampling

JEL Classification Codes: M31, M37, L10, L82, C61

Contents

Abstract **iv**

List of Tables **i**

List of Figures **ii**

List of Abbreviations and Symbols **iv**

Acknowledgements **viii**

1 Introduction **1**

2 Relevant Literature **4**

3 Data **7**

3.1 Data Description 7

3.1.1 TiVo Log Files (Show and Advertising Viewing) 8

3.1.2 IRI DataSets (Purchase) 8

3.1.3 TNS Advertising Schedule Data (Advertising Exposures and Prices) 9

3.1.4 Nielsen Television Viewing Data 10

3.1.5 TMS Data (Program Characteristics) 10

3.2 Empirical Regularities in TV Viewing 10

3.2.1 Show Sampling and Watching 11

3.2.2 Show Recording 15

3.3 Empirical Regularities in Advertising Viewing 15

3.3.1 Advertising Viewing . . . 15
3.3.2 Advertising Costs . . . 18
3.4 The Advertising-Sales Relationship . . . 20
4 TV Viewing Model **22**
4.1 Flow Utilities . . . 22
4.1.1 Show Utility . . . 23
4.1.2 Advertisement Utility . . . 23
4.1.3 Outside Good Utility . . . 24
4.2 TV Show Viewing . . . 24
4.2.1 Sampling . . . 26
4.2.2 Watching . . . 28
4.3 TV Show Recording . . . 30
4.4 TV Advertising Viewing . . . 31
4.4.1 Zapping . . . 31
4.4.2 Zipping . . . 32
5 Estimation and Identification **33**
5.1 Estimation . . . 33
5.2 Identification . . . 34
6 Estimation Results **36**
7 Policy Experiments: Advertising Targeting **42**
7.1 Advertising Response . . . 42
7.1.1 Model Free Analysis . . . 43
7.1.1.1 Homogeneous Advertising Effects . . . 43
7.1.1.2 Heterogeneous Advertising Effects . . . 46
7.1.2 Mixture Logit Analysis . . . 48

7.2 Targeting Strategies 49
7.2.1 Cost-Minimizing Real-time Buy 51
7.2.2 Profit-Maximizing Real-time Buy 52
7.2.3 Cost-Minimizing Advance Buy 54
7.2.4 Profit-Maximizing Advance Buy 56
7.2.5 Targeting Summary 57

8 Conclusion **59**

A Details on Inferring Advertising Exposures **62**

A.1 Key Data Files 62
A.1.1 *Watchtv* (from TiVo log files) 62
A.1.2 *Trickplay* (from TiVo log files) 63
A.1.3 *Adschedule* (from TNS advertising schedule file) 63
A.2 Identify Idle Periods 64
A.3 Infer Viewing Time on TV Shows 65
A.4 Infer Advertising Exposure and Skipping 66

B Length of Viewing **68**

C Likelihood Function of the Viewing Model **70**

D A Latent Class Logit Model on Brand Choice (§7.1.2) **74**

D.1 Model 74
D.2 Estimation and Identification 76
D.2.1 Estimation 76
D.2.2 Identification 79
D.3 Results 80
D.4 Detailed Estimation Results of the Sales Response Model 82
D.4.1 Children's Yogurt Category 82

D.4.2 Children's Cereal Category 83

D.4.3 Regular Cola Category 85

D.4.4 Diet Cola Category 86

D.4.5 Sports Drink Category 87

D.4.6 Toothpaste Category 88

D.4.7 Bathroom Tissue Category 89

Bibliography **91**

Biography **95**

List of Tables

3.1 Analysis of Variance for Advertising Skipping (Recorded Shows Watched During 8 p.m.-Midnight) 18

3.2 Summary Statistics of Per-Exposure Advertising Price Based on Nielsen Ratings (July 2006) 19

6.1 Flow Utility Parameter Estimates for Shows 37

6.2 Flow Utility Parameter Estimates for Advertisements 38

6.3 Mean Estimates of Shock Arrival Rates 39

7.1 Purchase Data Summary 43

7.2 Advertising Effect by Product Category and Consumer Segment . . . 48

7.3 Percent of Significant Advertising Effects 48

7.4 Targeting Scenarios 50

7.5 Targeting Gains 57

C.1 Monte Carlo Simulation Results 73

D.1 Percent of Significant Advertising Effects 81

List of Figures

3.1 Fraction of Days With 8 p.m.-Midnight Viewing (by Household). The x-axis is the fraction of days with 8 p.m.-midnight viewing, and the y-axis is the percent of the particular fraction in all observations. . . 11

3.2 Frequency of Cumulative Daily TV Usage, 8 p.m.-Midnight (by Household-Day). The x-axis is the number of hours viewing, and the y-axis is the percent of the particular viewing time in all observations. 12

3.3 Show Completion Rate (by Household-Show). The x-axis is the show completion rate, and the y-axis is the percent of the particular completion rate in all observations. Panel (a) is based on all shows watched during 8 p.m.-midnight, and panel (b) excludes shows with a completion rate below 1% or above 99%. 13

3.4 Hazard Rate for One-hour Shows Watched from the Broadcast Start Time (across Household-Show) 14

3.5 Number of Shows Sampled Before Watching One. The x-axis is the number of samplings before watching a show, and the y-axis is the percent of the particular number in all observations. 14

3.6 Advertising Viewing and Skipping 16

3.7 Number of People Zipping (by Second) in One Episode of CSI 17

3.8 Distribution of Per-exposure Advertising Price Based on Nielsen Ratings (July 2006) 19

3.9 Advertising Exposure and Sales (Children's Yogurt) for Panelists. The dashed line depicts aggregated weekly sales volume (unit: pint). Solid lines represent aggregated weekly advertising exposures for selected campaigns. 21

4.1 Decision Making in One Viewing Session 26

6.1 Histogram of Estimated Zapping Cost 39

6.2 Empirical CDF (Across Households) of the Hit Rate on Show Sampling (a) and Watching (b) 40

6.3 Empirical CDF (Across Households) of MAE on Viewing Length . . 41

7.1 Histogram of Purchase Log Odds Ratio With and Without Advertising Exposure (by Brand) 46

7.2 Simulated Incremental Profit (for Sample Households, Real-time Buy) Under Different Unit Sales Margins 54

7.3 Simulated Incremental Profit (for Sample Households, Advance Buy) Under Different Unit Sales Margins 57

D.1 Distribution of the Asymptotic t-Statistics of Estimated Advertising Coefficients (by Advertising Campaign and Household Segment) . . . 80

D.2 Percentage Change in Own-Brand Market Share Following Elimination of an Advertising Campaign 82

List of Abbreviations and Symbols

Symbols

Symbol	Description	Role
Viewing Model		
i	Household	Index
t	Time	Index
n	TV network	Index
tn (show)	The show broadcasted on network n at time t	Index
td	The recorded show deleted from the DVR at time t	Index
tn (ad)	The advertisement aired on network n at time t	Index
y^S_{itn}	Indicator that show tn was sampled by household i	Data
y^W_{itn}	Indicator that show tn was watched by household i	Data
y^R_{itn}	Indicator that show tn was recorded by household i	Data
$y^A_{itn,L}$	Indicator that live advertisement tn was viewed by household i	Data
$y^A_{itn,R}$	Indicator that recorded advertisement tn was viewed by household i	Data
X^S_{itn}	A vector of characteristics associated with show tn for household i	Data

Symbol	Description	Role
X_{it}^{O}	A vector of characteristics associated with the outside good at time t for household i	Data
X_{itn}^{A}	A vector of characteristics associated with advertisement tn for household i	Data
u_{itn}^{S}	Flow utility of show tn for household i	Latent construct
u_{it}^{O}	Flow utility of the outside good at time t for household i	Latent construct
u_{itn}^{A}	Flow utility of advertisement tn for household i	Latent construct
$\bar{u}_{itn}^{S}$	Deterministic part of u_{itn}^{S}	Notational simplification
$\bar{u}_{it}^{O}$	Deterministic part of u_{it}^{O}	Notational simplification
$\bar{u}_{itn}^{A}$	Deterministic part of u_{itn}^{A}	Notational simplification
ν_{itn}^{S}	Idiosyncratic error term affecting u_{itn}^{S}, observed by the household but not by the researcher	Integrated out
ν_{it}^{O}	Idiosyncratic error term affecting u_{it}^{O}, observed by the household but not by the researcher	Integrated out
ϵ_{itn}^{A}	Idiosyncratic error term affecting u_{itn}^{A}, observed by the household (conditional on exposure) but not by the researcher	Integrated out
ϵ_{itn}^{S}	Uncertainty related to u_{itn}^{S}, unobservable to both the household and the researcher, but will become known to the household conditional on sampling	Integrated out
$\mathcal{N}_{it}$	The set of networks that have not been sampled by household i at time t	Notational simplification
$IV_{it\mathcal{N}_{it}}$	The inclusive value of the shows not yet sampled by household i at time t	Notational simplification
q_{itn}	Household i's probability of exiting show tn, conditional on receiving a shock	Notational simplification

Symbol	Description	Role
L_{itn}	Remaining length of show tn when sampled by household i	Data
l_{itn}^{W}	Amount of time that household i spends on show tn	Data
λ_{itn}	Arrival rate of external shocks for household i and show tn	Model primitive
ρ_i	Parameters related to λ_{int}	Model primitive
β_i^S	Parameters related to u_{itn}^S	Model primitive
β_i^O	Parameters related to u_{it}^O	Model primitive
β_i^A	Parameters related to u_{itn}^A	Model primitive
c_i	Zapping cost faced by household i	Model primitive
Targeting Counterfactual		
y_{ijm}^P	An indicator that household i purchases brand j in shopping trip m	Data
A_{ijm}	An indicator variable of whether household i is exposed to brand j's advertisement since the preceding trip (in the same store-week)	Data
x_{itn}	An indicator variable of selecting show tn for household i	Decision variable
c_{tn}	Per-exposure advertising price associated with show tn	Data
e_{itn}	An indicator variable of whether household i watches show tn	Data
N_{iw}	Number of shopping trips for household i in week w	Data
Y_i	The current expected total advertising exposures for household i	Notational simplification
m	Unit sales margin	Assumption

Symbol	Description	Role
Δ_{ij}	Effect of brand j's advertising on household i's purchase probability	Notational simplification
r_{itn}	The probability that household i watches the advertisement placed in show tn	Notational simplification

Abbreviations

BIC	Bayesian information criterion.
CPG	Consumer packaged goods.
CPM	Cost per thousand viewers.
DMA	Designated market area.
DVR	Digital video recorder.
IRI	Information Resources Incorporated.
STB	Set-top box.
TMS	Tribune Media Services.
TV	Television.

Acknowledgements

I am indebted to my esteemed committee: Carl Mela (Co-Chair), Ken Wilbur (Co-Chair), Rick Staelin, Peng Sun, and Yi (Daniel) Xu. Without their constant help and guidance, I would not have completed this work. Carl has contributed immense amount of time in guiding me towards my goals. His insights, work ethics and integrity have been my inspiration along the way. Ken spent countless days helping me stay focused and think creatively. Rick has been the cheerleader in the marathon of my PhD journey, and his invaluable advice helps me become a better researcher. Peng has been extremely encouraging and generous with his time and knowledge. Daniel has provided insightful comments and I have always been inspired by his rigorous research.

My deep appreciation also goes to the rest of the Marketing faculty at Fuqua for their support and training. They made my journal in the PhD program fulfilling and enjoyable. I especially thank Jim Bettman and Wilfred Amaldoss for their encouragement and feedback throughout the PhD Program, Bryan Bollinger for many helpful comments on my dissertation, and Bill Boulding for encouraging me to be a broader researcher.

I am also grateful to fellow PhD students for their comments on my work and for the friendship that will continue.

Finally, I thank my family for their boundless love and support.

1

Introduction

Television (TV) is the most prominent modality for the transmission and reception of video content. According to Nielsen's most recent cross-platform report (Nielsen (2014)), American adults spend about 40 hours watching traditional TV each week. In contrast, the weekly average time spent on the Internet (both on computers and mobile devices) is about 11 hours. TV also remains, by far, the largest advertising medium. eMarketer (2014) forecasts that $68.5 billion will be spent on TV advertising in U.S. in 2014, accounting for 38% of total U.S. advertising spending. In contrast, online and mobile advertising accounts for 25% of total U.S. advertising spending, though it is expected to overtake TV advertising in 2018.

Moreover, as a result of recent digital innovations in transmission (specifically, digital set-top boxes (STBs) and video recorders (DVRs)), this transmission modality may continue to remain the preeminent channel of choice for viewing video content. 46% of TV households own DVRs as of February 2013 and many others own STBs (Nielsen (2013)). These technological advances not only offer households more ways to access TV programs and enhance the TV-viewing experience, but also transform the media industry by enabling household-specific advertising. From the advertiser's point of view, growing STB and DVR penetration improves opportunities to target and customize TV advertising to households.

Though uncommon in TV, micro-targeting has been extensively applied in online advertising. Targeting tools have been developed by Google, Yahoo!,[1] Facebook,[2] YouTube,[3] Yelp[4] and a number of other Internet companies/websites. Micro-targeting on the Internet is highly successful,[5] and this targeting model can now also be applied to TV. Several companies have already started offering targeted TV advertising solutions to advertisers. For instance, TiVo Research and Analytics, Inc. (TRA) developed a software platform "Media TRAnalytics®", which combines household-level TV tuning and purchase data to help advertisers achieve higher ROI. Similarly, Nielsen Catalina Solutions launched "AdVantics on Demand (TM)" that helps advertisers achieve better targeting based on retail purchase data.

Given DVR proliferation, micro-targeting in TV advertising is of substantial economic import, but relatively underattended in the literature. Hence, it is our goal to learn about households' TV consumption behavior in these new digital contexts, and based upon the consumption behavior, propose ways for advertisers to improve targeting. This is achieved by employing a unique dataset that integrates household-level media consumption with purchase behavior (specifically, DVR data, purchase data, show data and advertising data). Based on this dataset, we first develop a model of household TV viewing behavior which allows for household-specific pa-

[1] `http://advertising.yahoo.com/targeting/`

[2] `http://www.facebook.com/help/131834970288134/`

[3] `http://www.youtube.com/yt/advertise/targeting.html`

[4] `http://www.yelp.com/advertise/agency/targeting`

[5] Ansari and Mela (2003) develop a statistical and optimization approach for customization of information on the Internet, and apply the model in the context of permission-based email marketing. They find that the content-targeting can potentially increase the expected number of click-throughs by 62%. Likewise, based on a natural field experiment from the large rectangular advertising unit on the Yahoo! front page, Farahat and Bailey (2012) find that the advertising effect for the targeted population doubles that of the untargeted population.

rameters. We then link household advertising exposure with purchase. Finally, the viewing model and identified advertising-sales relationship are utilized to conduct counterfactual policy experiments on advertising targeting. We consider several household-level targeting scenarios by manipulating: 1) whether the advertising purchase is made in advance or in real time; and 2) whether the objective function is to minimize costs for a given set of exposures or to maximize incremental profit from advertising. Results indicate micro-targeting can lower advertising costs and raise incremental revenue.

The key contributions of this paper are as follows. From a theoretical perspective, we develop an integrated model of TV show viewing, TV advertising viewing, purchasing and advertising targeting. From a substantive point of view, we offer policy recommendations to advertisers on household-level targeting. From a methodological view, we propose a new modeling framework in the TV viewing context that explicitly accounts for the role of uncertainty, and viewers' use of sampling to reduce consumption uncertainty. It provides a theoretically motivated, internally consistent framework on TV show viewing, recording, and advertising viewing decisions. In addition, we develop targeting strategies that leverage household-level data.

The remainder of the paper is organized as follows. First, §2 reviews the relevant literature to link our study to previous research. §3 then describes TV viewing behavior to motivate the viewing model presented in §4. §5 discusses estimation and identification, and §6 describes the estimation results. Based on these results and a purchase model, §7 conducts counterfactual policy experiments and evaluates potential gain to be realized from targeting. Finally, we conclude with a schedule of next steps in §8.

2

Relevant Literature

Since Lehmann (1971)'s seminal work, a rich body of literature has identified various factors affecting viewers' utility from watching TV programs. Such factors include viewer demographics, program genre, cast demographics, advertising time, viewer's previous program choices, and spouse's choice (Rust and Alpert (1984); Rust et al. (1992); Shachar and Emerson (2000); Goettler and Shachar (2001); Moshkin and Shachar (2002); Yang et al. (2006); Wilbur (2008); Anand and Shachar (2011); Esteves-Sorenson and Perretti (2012)). Collectively, this line of literature suggests that person, show and time factors explain substantial variation in show viewing. Our approach will integrate such factors and leverage them in predicting viewers' future TV viewing patterns. In addition, this paper differs from most of the viewing literature in that it explicitly captures uncertainty pertaining to show quality, whereas most existing literature has made a stronger assumption that viewers have complete information about show utility. Exceptions include Moshkin and Shachar (2002), Anand and Shachar (2011), Esteves-Sorenson and Perretti (2012) and Yao et al. (2015). Moshkin and Shachar (2002) assume that the consumer knows the attributes of a new show offered by the network which she was watching in the previous period, but is uncertain about the new shows offered by other networks. In addition, the consumer can obtain full information of the attributes of all shows through a costly search. Anand and Shachar (2011) use a Bayesian learning framework to model consumer learning about the quality of future TV shows through advertising (promos). Esteves-Sorenson and Perretti (2012) assume that viewers do not know *ex-ante* the benefit they will obtain from an alternative TV channel but have priors on it from previous experience. Yao et al. (2015) argues that there exists uncertainty in show utility and commercial breaks provide opportunities for consumers to resolve the uncertainty through a costly search. In this paper, we capture the uncertainty with a sampling process, in which a viewer watches a show for a short time to learn about its quality, and then decides if to continue watching or sample other shows. This process is similar to a search process in terms of sampling order, but as we will discuss later, there are features that differentiate this sampling process from a search process. This paper also differs from most existing studies in that the rich data allow us to fully accommodate household-level heterogeneity by performing estimation household-by-household, thereby separating state dependence and heterogeneity.

Conditional on watching TV shows, viewers inevitably encounter commercial breaks, which often trigger zapping (i.e., channel switching, leaving the room, etc.), and in the case of recorded shows, zipping (i.e., fast-forwarding). Hence, a second related stream of literature looks into viewers' advertising avoidance behavior, and has identified various viewer- and ad-specific factors that affect such behavior. Identified viewer-specific factors include household category purchase history and the media weight of a campaign (i.e., the number of times that a household had previously been exposed to a commercial) (Siddarth and Chattopadhyay (1998); Gustafson and Siddarth (2007)). Identified ad-specific factors include the frequency of the commercial, length and content of the commercial, program genre, commercial location, as well as the congruity between the commercial and the program (Norris and Colman

(1993); Siddarth and Chattopadhyay (1998); Furnham et al. (2002); Furnham and Price (2006); Moore et al. (2005); Gustafson and Siddarth (2007); Teixeira et al. (2010); Schweidel and Kent (2010)). In a more recent research based on a dataset that links a panel of consumers' TV advertising consumption with product choice behavior, Tuchman et al. (2015) allow for complementarities in consumption of advertising and products, and find the consumption of goods increases the utility of its advertisements. This line of literature has further enriched our insights about person and show factors that potentially account for variation in commercial viewing. Our research contributes to the advertising viewing literature by modeling both zipping and zapping decisions under a coherent framework, which is built upon the flow utilities a household derives from different TV and non-TV activities.

Finally, given our goal to explore the potential of micro-targeting in TV advertising, this work also relates to the growing literature in advertising targeting. A number of papers have examined why and how targeted TV advertising works from a theoretical point of view (e.g., Gal-Or et al. (2006); Anand and Shachar (2009); Ghosh and Stock (2010)). A few papers address the issue of geographically or demographically targeted TV advertising from an empirical point of view (e.g., Kitts et al. (2010); Anand and Shachar (2011); Lovett and Peress (2014)). Our focus is instead targeting at more granular levels, i.e., the individual household. Finer targeting affords better opportunities to incorporate individual households' past viewing and purchase data in targeting decisions. In this regard, the most closely related study is Tuchman et al. (2015), who explore implications of individual-level targeting to consumers who are less likely to skip the targeted advertisements and have positive marginal advertising effect on purchase. Our paper differs from theirs because we also make suggestions on how to reach targeted consumers (i.e., which advertising slots to purchase) based on their past viewing behavior.

3

Data

Micro-targeting in television is facilitated by historical viewing and purchase data at the household level. Viewing data are useful to forecast which advertisements are viewed, and purchase data enable one to link household advertising exposure to household purchases. These data are described in this chapter. Specifically, we first overview the sources of the household-level viewing and purchase data (§3.1), then describe the TV program viewing data and the advertising viewing data to generate model-free insights regarding household viewing behavior (§3.2 and §3.3) and next consider advertising response in §3.4.

3.1 Data Description

Several sets of data are integrated for this study, including DVR usage data, purchase data, advertising data, and programming data. This combination is often called single source data because it covers the entire TV viewing and purchase experience for a set of households. The DVR data (TiVo log files) track each household's complete usage of a TiVo DVR and therefore all viewing behavior. The purchase data are from Information Resources Incorporated (IRI), and contain each household's store visits and purchase history in 77 consumer packaged goods (CPG) categories, as well as store causal data. The advertising data are obtained from TNS Media Research, and include the timing and advertising costs for national TV advertisements airing during the duration of the data. The TNS data are supplemented with national viewing data from AC Nielsen in order to normalize shows' advertising rates to the exposure level. The programming data come from Tribune Media Services (TMS) and contain information on popular TV programs. The DVR data, advertising data and programming data will be used to estimate the viewing model (§4-§6). The purchase data and the Nielsen data will be used along with viewing model estimates in policy experiments on targeting (§7). We describe each dataset in turn.

3.1.1 TiVo Log Files (Show and Advertising Viewing)

The viewing data are from a field study conducted by IRI, TiVo, and a consortium of major CPG manufacturers.[1] The TiVo log files track each household's moment-by-moment usage of a TiVo DVR. They record every keystroke of the DVR as well as all TV content viewed and whether it was live or recorded. Among other things, the keystrokes are used to determine which content was fast-forwarded. We use data in the period of July 2005 - July 2006, keeping the 834 households that have both viewing and purchase information.

3.1.2 IRI DataSets (Purchase)

The panel data used to link with TV viewing and advertising exposure data are provided by IRI and include purchase data, trip data, and store data in the period of June 2005 - June 2006.[2]

[1] For a more detailed description of the field study and the data, please see Bronnenberg et al. (2010).

[2] Of note, the starting and ending date of the IRI datasets are both earlier than those of the three datasets related to TV viewing. All datasets intersect during the period of July 2005 - June 2006. We retain the IRI data in June 2005, one month before the start of the TV data, in order to

The first component, the IRI purchase panel data contain the purchase history for panelists in 77 CPG categories. Organized by panelist-category-item-transaction time, the data include store, item, item attributes, price, and promotional status (display or feature).

The second component, the IRI trip panel data record panelists' store visits. Organized by panelist-transaction time, the data include store visited and total amount spent. Combined with the purchase panel data, these store visit data enable us to infer non-purchases, defined as no purchase in a category on a given store visit.

The third component, the IRI store causal data report store sales for each item (in 77 CPG categories) sold in 57 stores. Organized by store-week-item, the data include weekly price, promotional status (display or feature), and units sold. By matching these data with transactions in the purchase panel data and store visits in the trip panel data, we can construct a choice set with associated causal variables for each purchase occasion.

3.1.3 TNS Advertising Schedule Data (Advertising Exposures and Prices)

The TNS advertising schedule data describe advertising schedules for 61 national broadcast and cable TV networks. For each advertisement, the data report the precise air time, network, length, a brief description of the advertisement, attributes of the advertised product (e.g., product category, company and brand), name and genre of the associated show, location of the commercial break within the show (i.e., pod) and the slot within the break (i.e., pod location or slot), and the estimated price of the advertisement. We infer advertising exposures by noting the time and channel of the advertisement, and assessing whether or not the channel was viewed at that time. Appendix A provides details on advertising exposure inference.

initialize behavioral measures such as last brand purchased. We retain the TV data in July 2006, one month after the end of the IRI data, for hold-out validation and policy experiments.

3.1.4 Nielsen Television Viewing Data

Because advertising rates furnished by TNS are at the show level, they do not yield a per-impression cost. As micro-targeting is at the impression level, we need to translate the show cost to an exposure cost, and do so by collecting information on advertising exposures. Specifically, we supplement the TNS advertising schedule data with Nielsen ratings to obtain per-exposure price for each advertisement. The Nielsen ratings are manually collected from the Broadcasting & Cable magazine and report the audience size of top TV programs on broadcasting networks. We collect these data in July 2006, the period during which the policy experiments are run. These data include 376 shows on 4 networks: ABC, CBS, NBC and FOX. TMS Data (Program Characteristics)

3.1.5 TMS Data (Program Characteristics)

TMS data contain descriptive information (e.g., program name, genre, cast, plot description) for 55,684 programs accounting for 90% of the TiVo viewing observations related to the top 27 TV networks (6 broadcast networks and 21 cable networks). Each view in the TiVo Log Files is tagged with a unique TMS identifier, which is used to match with the TMS data, resulting in a description of each show viewed.

3.2 Empirical Regularities in TV Viewing

This section reports a descriptive analysis of TV viewing both to illustrate the nature of the data and to motivate the ensuing model. Households' viewing behavior can be described by a series of conditional decisions, and we organize the discussion along this progression of decisions.

When starting to watch TV, a household first chooses which show to watch. Due to incomplete knowledge of program and episode quality, the household usually samples a show (either live or recorded) for a brief duration to decide whether to watch. After viewing the show for a short period of time, the household decides whether to continue watching, switch or exit. If deciding not to view the first show sampled, another show is sampled and the process repeats. To illustrate this process, §3.2.1 describes patterns observed from the data on how households sample and watch TV shows.

While sampling and watching shows, households construct and maintain an inventory of recorded shows. Patterns related to show recording will be examined in §3.2.2.

3.2.1 Show Sampling and Watching

Owing to the observation that most viewing (and advertising spending) takes place in prime time (defined as 8 p.m. - midnight in this paper), our ensuing discussion pertains to this daypart. Figure 3.1 shows most households watch TV most evenings: on average, households watch prime-time TV on 85% of the days in the sample.

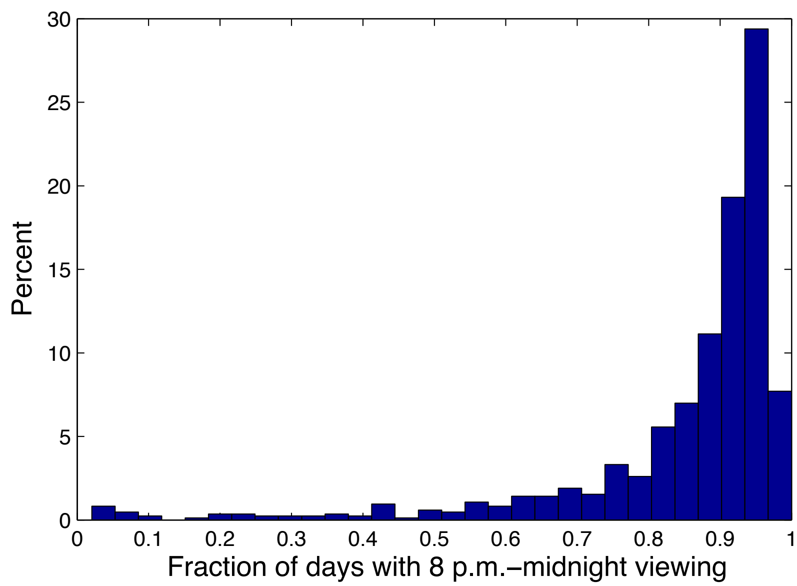


FIGURE 3.1: Fraction of Days With 8 p.m.-Midnight Viewing (by Household). The x-axis is the fraction of days with 8 p.m.-midnight viewing, and the y-axis is the percent of the particular fraction in all observations.

Figure 3.2 depicts the hours of prime-time TV viewing across household-days, conditional on watching TV. Over 75% of the household-day viewing time exceeds 3 hours, suggesting people watch a lot of prime-time TV.

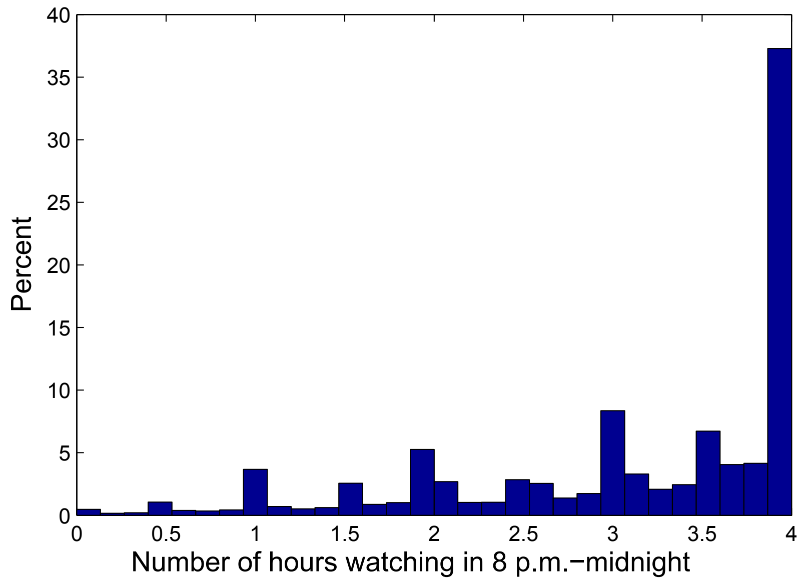
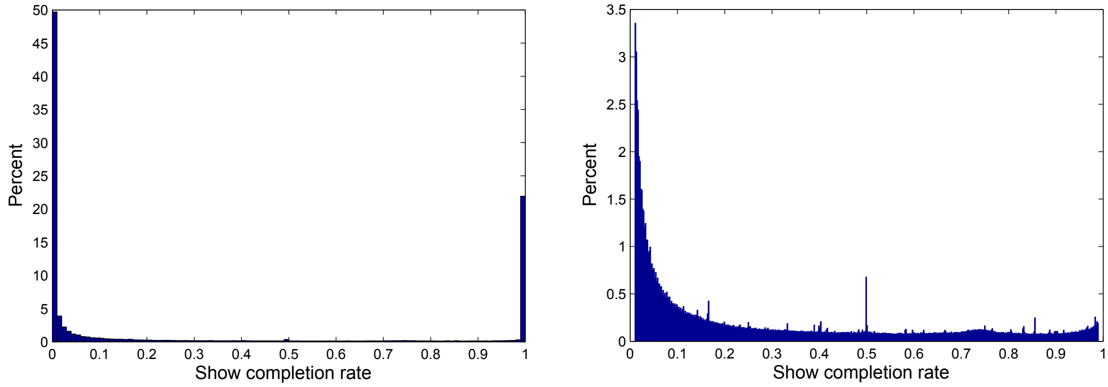


FIGURE 3.2: Frequency of Cumulative Daily TV Usage, 8 p.m.-Midnight (by Household-Day). The x-axis is the number of hours viewing, and the y-axis is the percent of the particular viewing time in all observations.

To characterize the process of show sampling, we define the "show completion rate" as the ratio of the time a show is actually viewed relative to its total broadcast length. Panel (a) of Figure 3.3 illustrates that completion rate is bimodal: it tends to be either very low or very high. This dichotomy is consistent with a process wherein people first sample a set of shows and then proceed to watch the one that is liked. Truncating the non-views and the completed views (the endpoints in panel (a)) enables us to zoom in on the sampling behavior (panel (b)). It suggests a power law with most sampling not exceeding a few minutes per show.[3]

[3] The small spike at 50% occurs because people watching a one-hour show might exit in the middle to watch another show that just started, as most shows start or end around the hour or the half hour.



(a) All 8 p.m.-midnight viewing

(b) 8 p.m.-midnight viewing with show completion rate within 1-99%

FIGURE 3.3: Show Completion Rate (by Household-Show). The x-axis is the show completion rate, and the y-axis is the percent of the particular completion rate in all observations. Panel (a) is based on all shows watched during 8 p.m.-midnight, and panel (b) excludes shows with a completion rate below 1% or above 99%.

Further illustration of the apparent sampling process requires a definition of sampling events. Because excessively short viewing durations (e.g., 30 seconds or less) likely reflect channel surfing (i.e., using the up or down button to shift channels), a sampling duration is defined to be 30 seconds or longer. To obtain the threshold that differentiates watching from sampling, we collect one-hour shows watched from the broadcast start time, and compute the hazard rate (i.e., the fraction of surviving viewers that leave) by time into show (Figure 3.4). The hazard rate decreases drastically within the first 3 minutes, and remains relatively stable afterwards. Similar patterns are observed for half-hour shows. Therefore, we categorize viewing durations between 30 seconds and 3 minutes as sampling events. Based on this categorization, Figure 3.5 indicates that, in 70% of the cases, a household decides to watch the first show sampled. Hence, sampling is informative of preferences.

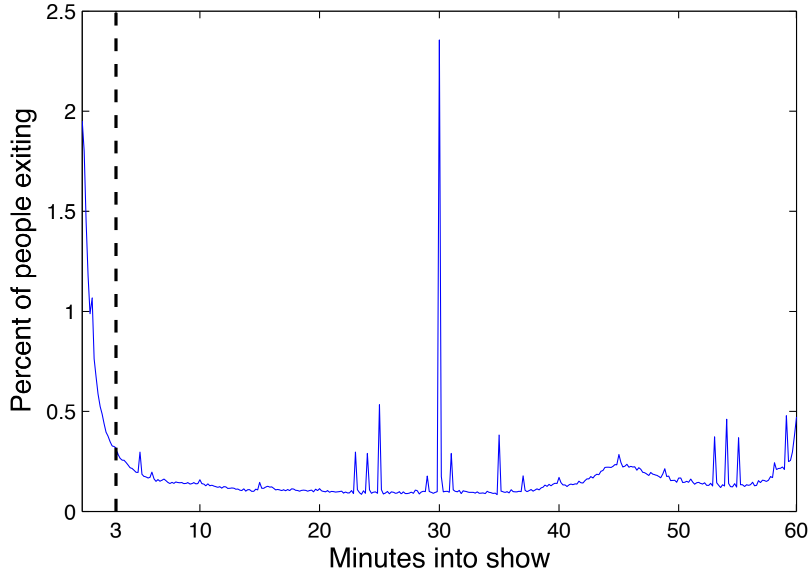


FIGURE 3.4: Hazard Rate for One-hour Shows Watched from the Broadcast Start Time (across Household-Show)

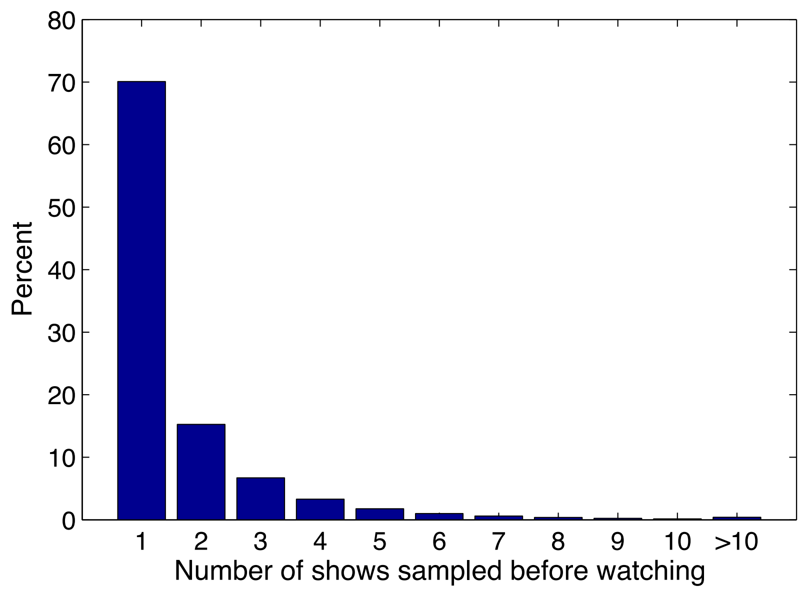


FIGURE 3.5: Number of Shows Sampled Before Watching One. The x-axis is the number of samplings before watching a show, and the y-axis is the percent of the particular number in all observations.

In sum, the data suggest that the frequency of prime-time viewing is high, that households spend considerable time watching in the evening, and when viewing they tend to sample shows until finding one of interest. With this characterization of show viewing in mind, we turn to the recording decision.

3.2.2 Show Recording

Households can choose to record shows in any given day. Across households and days, 3% of the household-day observations are associated with recording only, 53% are associated with viewing only, and 44% are associated with both viewing and recording. Hence, recording is pervasive.

The TiVo DVRs' storage capacity was able to store 40 hours of programming. DVR program inventories are near capacity (>90%) on most (82%) household-day pairs in the data, implying households usually have to delete one show before recording another. These recording and deletion decisions are informative about show preferences.

Show choice ultimately determines the advertisements a household might see as these advertisements are embedded within the shows. As viewing, sampling and recording are common, people have considerable potential to be exposed to advertisements; yet in recorded contexts, there is ample opportunity to forward past them. Because this attenuates exposure, recording is a critical factor in advertising avoidance, which we explore next.

3.3 Empirical Regularities in Advertising Viewing

3.3.1 Advertising Viewing

Advertising viewership is predicated on the series of decisions shown in Figure 3.6. First, advertising exposure requires one be watching the show when the advertisement airs (*expose*). Second, a household must decide whether to watch a show live or recorded. As most viewing is live, 78% of all advertising exposures are live. Third, when confronted with an advertisement, a household can decide to avoid it. Complete skipping occurs when a fast-forward starts before the advertisement and ends after it. Partial skipping occurs when the household starts or stops fast-forwarding (i.e., zipping) during the advertisement and/or switches channels into or out of the advertisement (i.e., zapping).[4] As expected, recorded shows are more subject to skipping, as evidenced by the nearly 80% skipping rate (mostly zipping) for recorded shows and a lower than 15% skipping rate (mostly zapping) for live or near-live shows.

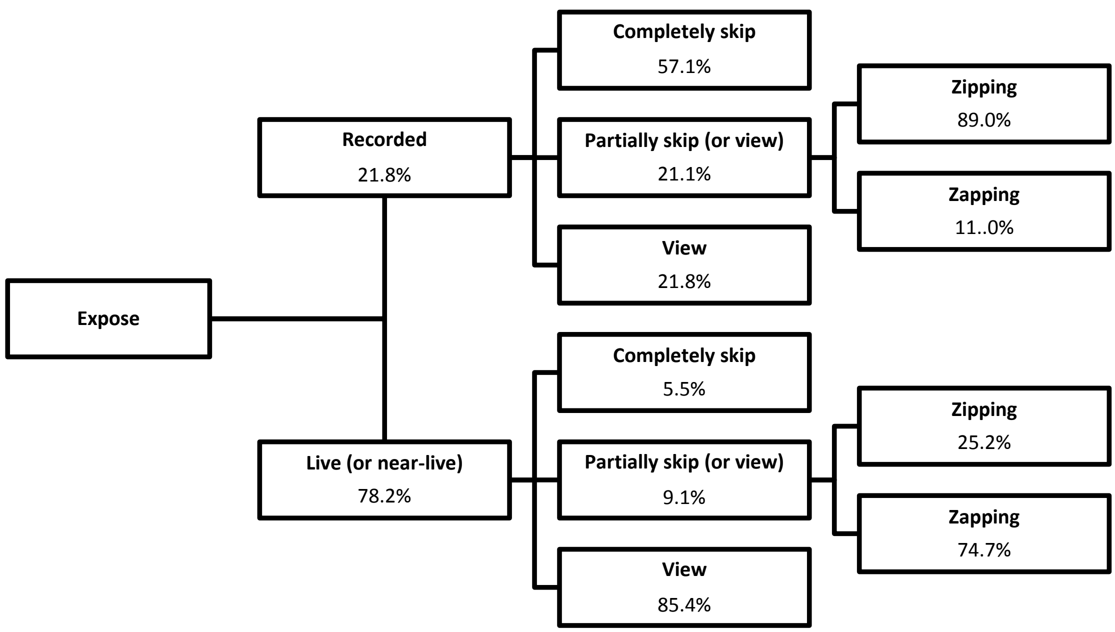


FIGURE 3.6: Advertising Viewing and Skipping

To further exemplify skipping patterns in recorded shows, we depict in Figure 3.7 the timing of zips and commercial breaks during a one-hour-long episode of "CSI: Crime Scene Investigation" on December 8, 2005. Figure 3.7 indicates zips coincide with commercial breaks.[5] Similar patterns exist in other shows.

[4] Zapping can also be done by other means to avoid paying attention, such as leaving the room. However, as in most previous studies, we are unable to observe such behavior, and hence focus only on channel switching.

[5] Zipping durations exceed the durations of the national advertising breaks because the TNS

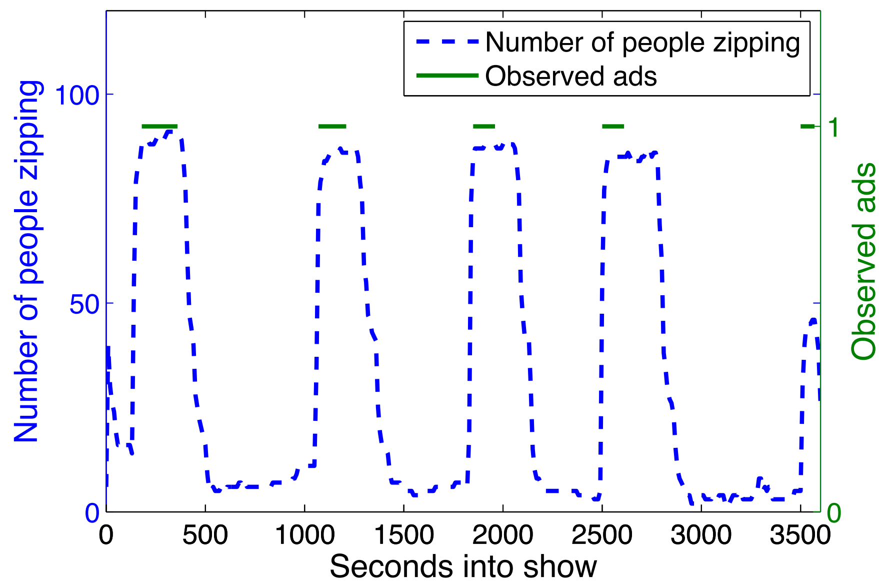


FIGURE 3.7: Number of People Zipping (by Second) in One Episode of CSI

To ascertain the factors that explain variation in prime-time advertising skipping, we conduct a variance decomposition using household, brand, show genre, network, product category, location of the commercial break within the show (i.e., pod) and the slot within the break (i.e., pod position or slot), day of week, hour, and past skipping (Table 3.1). If all the variation in skipping can be apportioned to shows or time, then standard aggregate methods of targeting based on purchasing slots in shows should be effective tools to address avoidance. If, in contrast, there remains substantial household-specific variation, then the efficacy of micro-targeting is amplified.

We find the latter to be the case. The factors incorporated in the model account for 34.5% of the overall variance in advertisement viewing and skipping. Most importantly, household fixed effects account for 59.1% of the explained variance. Moreover, demographic variables alone are not sufficient in explaining the variation. If household fixed effects are replaced by a set of demographic variables, the total

advertising schedule data do not include local cable advertisements or advertisements for upcoming cable network programs ("promos" or "tune-ins") and these are aired at the beginning or the end of the commercial break (Wilbur et al. (2013)).

explained variance drops from 34.5% to 27.8%, and observed demographic variables account for only 11.4% of the explained variance. All in all, these results suggest extant models of demographic based advertising buys are relatively ineffective for addressing advertising avoidance, and there exist systematic differences in household advertising avoidance that suggest substantial potential gains to micro-targeting.

Table 3.1: Analysis of Variance for Advertising Skipping (Recorded Shows Watched During 8 p.m.-Midnight)

Source	DF	Type I SS	Mean Square	F Value	Pr > F	% Variance
Household	777	113065.5	145.5	1160.6	<.0001	20.4%
Brand	1920	1586.4	0.8	6.6	<.0001	0.3%
Genre	14	4135.9	295.4	2356.2	<.0001	0.7%
Network	55	2129.1	38.7	308.7	<.0001	0.4%
Product category	573	168.6	0.3	2.4	<.0001	0.0%
Pod	27	3964.1	146.8	1171.0	<.0001	0.7%
Pod position (Slot)	33	415.7	12.6	100.5	<.0001	0.1%
Day of week	6	83.8	14.0	111.5	<.0001	0.0%
Hour	3	114.0	38.0	303.0	<.0001	0.0%
Previous advertisement skipped	1	65691.3	65691.3	523923.0	<.0001	11.9%

3.3.2 Advertising Costs

Table 3.2 and Figure 3.8 respectively provide summary statistics and distribution of per-exposure price for 15-second advertising slots for the shows discussed in §3.1.4. The median per-exposure price is about 1 cent for ABC, CBS and NBC, and is slightly above 1 cent for FOX. There also exists moderate price variation within each network.

Table 3.2: Summary Statistics of Per-Exposure Advertising Price Based on Nielsen Ratings (July 2006)

Network	Number of Observations	Mean	Median	S.D.
ABC	94	0.011	0.010	0.005
CBS	102	0.012	0.011	0.004
NBC	93	0.011	0.010	0.006
FOX	87	0.015	0.014	0.010

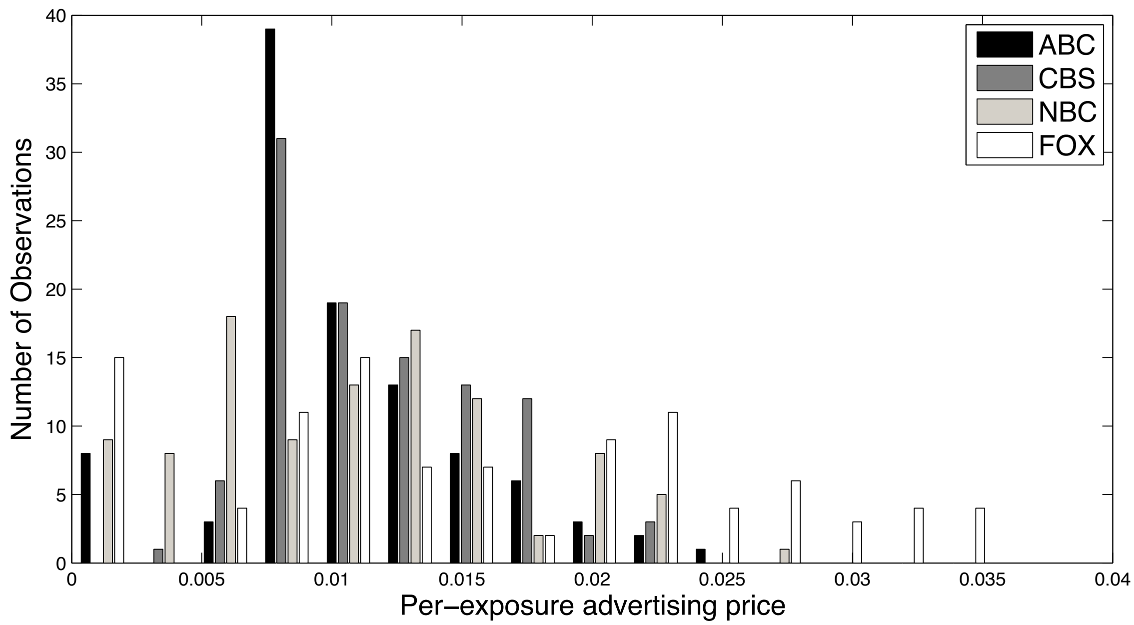
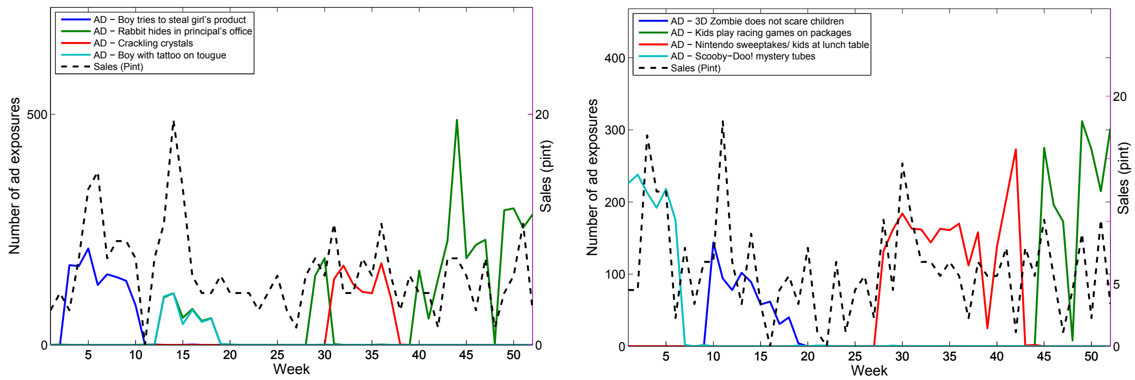


Figure 3.8: Distribution of Per-exposure Advertising Price Based on Nielsen Ratings (July 2006)

To ascertain whether advertising prices relate to viewership, we regress the advertisements' prices on ratings and network dummies. Results indicate a positive and significant relationship between price and rating, and the network coefficient is the highest for FOX, followed in turn by CBS, NBC and ABC. Presumably the differences across networks relate to the demographics of the shows viewed (e.g., Goettler (2012)). It is interesting to note that micro-targeting in the context of DVRs can facilitate the reach to these specific demographics.

3.4 The Advertising-Sales Relationship

For advertising targeting to be effective, it is necessary for advertising to correlate with sales. To provide an example on the advertising-sales relationship, Figure 3.9 plots panelists' weekly purchases and TV advertising exposures to selected campaigns for two popular children's yogurt brands, Yoplait Trix (panel (a)) and Yoplait Go-Gurt (panel (b)), between July 2005 and June 2006. The correlation between sales volume and brand advertising exposures (aggregated across campaigns) is 0.36 ($p=0.008$) for Yoplait Trix and 0.26 ($p=0.066$) for Yoplait Go-Gurt, suggesting a positive correlation between advertising and sales for these items. A similar analysis that disaggregates advertising exposures across campaigns yields similar insights, but suggests campaigns vary in their efficacy - and that it may be sensible to measure advertising response at the campaign level (Homer and Yoon (1992); Malaviya et al. (1996); Vakratsas and Ambler (1999)). It should be noted this correlation might reflect endogenous demographic based targeting; however, given observed viewing is not highly correlated with observed demographics, this effect is likely to be modest. Further analyses on short-term advertising effects will be presented in §7 to facilitate development of targeting strategies.



(a) Yoplait Trix (b) Yoplait Go-Gurt

FIGURE 3.9: Advertising Exposure and Sales (Children's Yogurt) for Panelists. The dashed line depicts aggregated weekly sales volume (unit: pint). Solid lines represent aggregated weekly advertising exposures for selected campaigns.

4

TV Viewing Model

This chapter describes the TV viewing model, which comprises three integrated components: TV show sampling and watching, TV show recording, and advertising viewing. All three components are motivated by the theoretical concept of flow utility, that is, the moment-by-moment consumption benefit a household derives from watching a TV show or advertisement, and non-TV activities (the outside good).

We first introduce the three flow utilities in §4.1. Based on these flow utilities, we then describe the household choice process (show viewing, show recording, and advertising viewing) in §4.2, §4.3 and §4.4 that follow.

4.1 Flow Utilities

We model flow utilities of TV shows, TV advertisements, and the outside good. All utilities are household specific.

4.1.1 Show Utility

Each show can be represented by a unique combination of t (time) and n (network). The flow utility that household i derives from show tn is defined as:

$$\begin{aligned} u_{itn}^{S} &= X_{itn}^{S}\beta_i^{S} + \nu_{itn}^{S} + \epsilon_{itn}^{S} \\ &\equiv \bar{u}_{itn}^{S} + \nu_{itn}^{S} + \epsilon_{itn}^{S}, \end{aligned} \tag{4.1}$$

where X_{itn}^{S} is a vector that captures show characteristics and household i's past viewing behavior, including genre, network, show length, the number of previous episodes of the program that the household has sampled in the preceding week, and the percent of show aired when sampling begins (i.e., viewing offset). ν_{itn}^{S} represents a household-show specific error term observed by the household but not by the researcher. ϵ_{itn}^{S} represents household uncertainty pertaining to program and episode quality prior to sampling that is revealed to the household only after sampling the show. We assume both ν_{itn}^{S} and ϵ_{itn}^{S} are i.i.d. standard Type I Extreme Value distributed.

4.1.2 Advertisement Utility

Each advertisement insertion can also be represented by a unique combination of t (time) and n (network). The flow utility that household i obtains from advertisement tn is a function of the characteristics of the advertisement, X_{itn}^{A}, and is given by:

$$\begin{aligned} u_{itn}^{A} &= X_{itn}^{A}\beta_i^{A} + \epsilon_{itn}^{A} \\ &\equiv \bar{u}_{itn}^{A} + \epsilon_{itn}^{A}, \end{aligned} \tag{4.2}$$

where X_{itn}^{A} includes: pod, pod position (slot), genre of the associated show, product category, and whether the preceding advertisement is skipped. ϵ_{itn}^{A} is an idiosyncratic error term affecting the inherent valuation of advertisement tn (conditional on exposure), and it is observed by the household but not by the researcher. We assume ϵ^A_{itn} to be i.i.d. standard Type I Extreme Value distributed.

4.1.3 Outside Good Utility

When allocating time, the household contrasts the utility from viewing TV to the best available alternative (i.e., the outside good). If the utility from viewing TV exceeds that of the outside good, the household will watch TV. As such, the flow utility of the outside good is tantamount to the "opportunity cost" of time, posited to vary by household (i) and time (t), and denoted as:

$$\begin{aligned} u^O_{it} &= X^O_{it}\beta^O_i + \nu^O_{it} \\ &\equiv \bar{u}^O_{it} + \nu^O_{it}, \end{aligned} \tag{4.3}$$

where X^O_{it} is a vector of observable day characteristics specific to household i and time t. These characteristics include weekday fixed effects and indicators for previous day TV viewing and previous weekday TV viewing. Month fixed effects are added to control for seasonality. ν^O_{it} is an idiosyncratic error term affecting the utility from the outside good at time t, and is observed by the household but not by the researcher. For instance, a rainy evening might be accompanied by a smaller value of ν^O_{it} as outdoor activities become less appealing. We assume ν^O_{it} to be i.i.d. standard Type I Extreme Value distributed.

4.2 TV Show Viewing

Based on flow utilities introduced above, this section describes household decisions associated with TV show viewing.

Owing to ϵ^S_{itn}, households have *ex-ante* uncertainty in the utility of viewing that can only be resolved by sampling a show (that is, a brief viewing of the show). While many possible heuristics exist for the order in which shows can be sampled, we assume that a household first samples the alternative with the highest *ex-ante* expected viewing utility. This ordering bears low cognitive cost and, in the absence of learning, is also consistent with the optimal search ordering implied by Weitzman (1979) and Kim et al. (2010).[1] After the short exposure to the show, the household observes its utility shock, and makes the decision of whether to continue watching by comparing the flow utility of this show with the expected highest flow utility to be obtained from other shows available at that time. If the flow utility of the current show is lower, the household samples other shows.

If the flow utility of the current show is higher than the expected best remaining alternative, the household selects the show and enters a "flow" state of watching. During this flow state, the flow utility of the show remains constant until an external "shock" arrives that changes the flow utility by perturbing ϵ^S_{itn}. Such a perturbation might reflect a change in a story on a news show, for example. When this occurs, a new error is drawn in the flow utility model in Equation (4.1). If the resulting flow utility is lower, the household compares this new utility to the expected highest flow utility that could be obtained from switching to another show or the outside good.

[1] However, the search models of Kim et al. (2010) and Weitzman (1979) and the sampling model in this paper differ in a number of respects. In the search models, consumers cease searching when the maximum utility among searched alternatives exceeds the reservation utility of the next best alternative, and the consumer chooses the option with the highest utility among all items that have been searched. In contrast, our sampling model presumes households choose between the current sampled show and the remaining shows that have yet to be sampled. Thus, the consumer always chooses the last show sampled in our model. This specification is predicated upon our observation of the data which reveals that consumers rarely, if ever, return to watch previously sampled shows (occuring in only 4% of the shows selected). Hence, the search models are not consistent with the show selection process we observe. While it is theoretically possible to add a cost of revisiting a previously sampled show to the utility model to explain this phenomenon, identification and computation become problematic. Hence we opt for a more parsimonious representation of the consumer problem.

If these alternatives yield higher utility, the household switches.

Figure 4.1 overviews this process and the sections of the paper that elaborates on the sampling and viewing decisions. We present these decisions in turn.

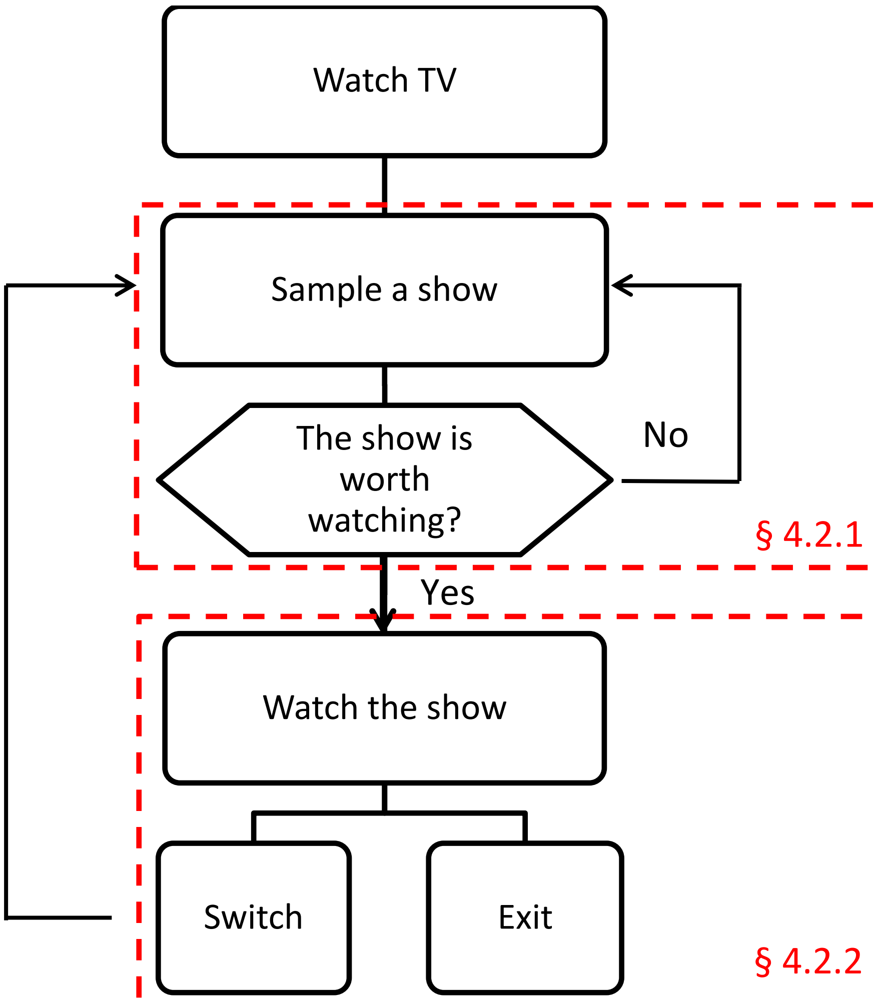


FIGURE 4.1: Decision Making in One Viewing Session

4.2.1 Sampling

As noted above, households sample shows to ascertain whether the flow utility exceeds that of other options.

At time t, the set of available shows to be sampled consists of all current live programs and the current menu of recorded programs. The household starts by sampling the show with the highest expected flow utility. Expected flow utility of show tn is $\bar{u}^S_{itn} + \nu^S_{itn} + E\left(\epsilon^S_{itn}\right)$. Since ϵ^S_{itn} $(\forall n)$ are i.i.d. distributed, $E\left(\epsilon^S_{itn}\right)$ is equal across shows and the decision is based on $\bar{u}^S_{itn} + \nu^S_{itn}$. Show tn is sampled if $\bar{u}^S_{itn} + \nu^S_{itn} > \bar{u}^S_{itn'} + \nu^S_{itn'}, \forall n'$ and $\bar{u}^S_{itn} + \nu^S_{itn} > \bar{u}^O_{it} + \nu^O_{it}$. Under the assumption that $\nu^S_{itn}\,(\forall n)$ and ν^O_{it} in Equations (4.1) and (4.3) are i.i.d. standard Type I Extreme Value distributed, this probability is:

$$Pr\left(y^S_{itn} = 1\right) = \frac{exp\left(\bar{u}^S_{itn}\right)}{exp\left(\bar{u}^O_{it}\right) + \sum_{n'} exp\left(\bar{u}^S_{itn'}\right)}, \tag{4.4}$$

where y^S_{itn} is an indicator that show tn is sampled by household i.

After sampling show tn,[2] household i observes ϵ^S_{itn} and therefore u^S_{itn}. The viewer then compares u^S_{itn} with the expected highest flow utility (i.e., the inclusive value) to be obtained from the remaining non-sampled shows available at that time and the outside good:

$$\begin{aligned} IV_{it\mathcal{N}_{it}} \quad | \left\{\nu^S_{itn}\right\}_n, \nu^O_{it} &= E_{\left\{\epsilon^S_{itn'}\right\}_{n' \in \mathcal{N}_{it}}} max_{n' \in \mathcal{N}_{it}} \left\{\bar{u}^S_{itn'} + \nu^S_{itn'} + \epsilon^S_{itn'},\ \bar{u}^O_{it} + \nu^O_{it}\right\} \\ &= max\left\{\bar{u}^O_{it} + \nu^O_{it},\ ln\left(\sum_{n' \in \mathcal{N}_{it}} exp\left(\bar{u}^S_{itn'} + \nu^S_{itn'}\right)\right)\right\}, \end{aligned}$$

where $\mathcal{N}_{it}$ denotes the set of networks at time t that have not been sampled.

If $u_{itn} \geq IV_{it\mathcal{N}_{it}}$, show tn is worth watching, and household i watches it. The associated probability is:

[2] For reasons indicated in §3.2.1, the duration of a sampling event is assumed to be between 30 seconds and 3 minutes. Within this interval, the choice set is assumed to be (and normally is) constant.

$$
\begin{aligned}
& Pr\left(y_{itn}^{W}=1 \mid y_{itn}^{S}=1,\left\{\nu_{itn}^{S}\right\}_{n},\nu_{it}^{O}\right) && (4.5)\\
= \; & Pr\left(u_{itn}^{S}\geq IV_{it\mathcal{N}_{it}} \mid \left\{\nu_{itn}^{S}\right\}_{n},\nu_{it}^{O}\right)\\
= \; & 1-F_{\epsilon_{itn}^{S}}\left(max\left\{\bar{u}_{it}^{O}+\nu_{it}^{O},\ ln\left(\sum_{n'\in\mathcal{N}_{it}}exp\left(\bar{u}_{itn'}^{S}+\nu_{itn'}^{S}\right)\right)\right\}-\bar{u}_{itn}^{S}-\nu_{itn}^{S}\right),
\end{aligned}
$$

where y_{itn}^{W} is an indicator that household i watches show tn, $F_{\epsilon_{itn}^{S}}(\cdot)$ denotes the cumulative distribution function (CDF) of ϵ_{itn}^{S}, which follows the standard Type I Extreme Value distribution.

Computation of $Pr\left(y_{itn}^{W}=1 \mid y_{itn}^{S}=1\right)$ involves integrating out $\left\{\nu_{itn}^{S}\right\}_{n}$ and ν_{it}^{O} in $Pr\left(y_{itn}^{W}=1 \mid y_{itn}^{S}=1,\left\{\nu_{itn}^{S}\right\}_{n},\nu_{it}^{O}\right)$. The sampling order implies $\nu_{itn'}^{S}$ and ν_{it}^{O} in Equation (4.5) are truncated respectively below $\bar{u}_{itn}^{S}+\nu_{itn}^{S}-\bar{u}_{itn'}^{S}$ and below $\bar{u}_{itn}^{S}+\nu_{itn}^{S}-\bar{u}_{it}^{O}$.

If $u_{itn}^{S}<IV_{it\mathcal{N}_{it}}$, household i samples another show if the inclusive value of remaining shows is higher than the value of the outside good. The choice set is now $\tilde{\mathcal{N}}_{it}=\mathcal{N}_{it}\setminus n$, and the probability of sampling show $t\tilde{n}$ ($\tilde{n}\in\tilde{\mathcal{N}}_{it}$) is:

$$
Pr\left(y_{it\tilde{n}}^{S}=1\right)=\frac{exp\left(\bar{u}_{it\tilde{n}}^{S}\right)}{exp\left(\bar{u}_{it}^{O}\right)+\sum_{n'\in\tilde{\mathcal{N}}_{it}}exp\left(\bar{u}_{itn'}^{S}\right)}. \tag{4.6}
$$

The sampling process repeats until the household identifies a show that is worth watching (in which case the household watches the show), or when the value of the outside good exceeds the inclusive value of remaining shows (in which case the household ends the viewing session).

4.2.2 Watching

Upon selecting show tn to view, household i obtains viewing flow utility u_{itn}^{S} until the show ends, or when deciding to stop watching, whichever comes first. The decision to stop watching is driven by arrival of external shocks that change the flow utility. If the external shock is sufficiently negative, the household terminates the show.

Specifically, at some time $t' > t$, household i encounters an external shock $\epsilon^S_{it'n}$ (e.g., change in plot, actor, or scene), which replaces ϵ^S_{itn} and changes the flow utility of show tn from $\bar{u}^S_{itn} + \nu^S_{itn} + \epsilon^S_{itn}$ to $\bar{u}^S_{itn} + \nu^S_{itn} + \epsilon^S_{it'n}$. If this new flow utility falls below the inclusive value of remaining alternatives, the household will exit the show.[3]

Assume external shocks arrive via a homogeneous Poisson process with rate λ_{itn} for household i and show tn.[4] Conditional on receiving a shock at time t' into show, the probability of exiting is:

$$
\begin{aligned}
& q_{itn,t'} \mid \left\{\nu^S_{itn}\right\}_n, \nu^O_{it} \qquad (4.7) \\
& = Pr\left(\bar{u}^S_{itn} + \nu^S_{itn} + \epsilon^S_{it'n} < max\left\{\bar{u}^O_{it} + \nu^O_{it},\ ln\left(\sum_{n' \in \mathcal{N}_{it'}} exp\left(\bar{u}^S_{it'n'} + \nu^S_{it'n'}\right)\right)\right\}\right) \\
& = F_{\epsilon^S_{it'n}}\left(max\left\{\bar{u}^O_{it} + \nu^O_{it},\ ln\left(\sum_{n' \in \mathcal{N}_{it'}} exp\left(\bar{u}^S_{it'n'} + \nu^S_{it'n'}\right)\right)\right\} - \bar{u}^S_{itn} - \nu^S_{itn}\right).
\end{aligned}
$$

Computation of $q_{itn,t'}$ involves integrating out $\left\{\nu^S_{itn}\right\}_n$ and ν^O_{it} in $q_{itn,t'} \mid \left\{\nu^S_{itn}\right\}_n, \nu^O_{it}$. $q_{itn,t'}$ is not necessarily fixed through the duration of show tn, and can alter when available shows on alternative networks ($\mathcal{N}_{it'}$) change. For instance, when a new show $t'n'$ starts on network n', $\bar{u}^S_{it'n'} + \nu^S_{it'n'}$ changes and $q_{itn,t'}$ would change accordingly. Hence, $q_{itn,t'}$ is piecewise constant and changes whenever shows on alternative networks ($\mathcal{N}_{it'}$) change. We denote segments of a show as intervals in which $q_{itn,t'}$ stays fixed.

Using an approach similar to Arcidiacono et al. (2013) and Nevskaya and Albuquerque (2013), Appendix B shows for $q_{itn,t'}$ that is piecewise constant with segment

[3] This characterization of the show exiting decision is in essence similar to the First-Hitting-Time (FHT) models, see Lee and Whitmore (2006) for a comprehensive review of the literature.

[4] The homogeneous Poisson process implies that within a show, external shocks arrive at a constant rate for the household.

$1, ..., M$, the CDF of the viewing length $l^{\star}_{itn}$ is:

$$F_{l^{\star}_{itn}}\left(\bar{t}\right) \equiv Pr\left\{l^{\star}_{itn} \leq \bar{t}\right\} = 1 - e^{-\lambda_{itn}\sum_{m=1}^{M} l^{m}_{itn} q^{m}_{itn}}, \tag{4.8}$$

where q^{m}_{itn} is the exiting probability in segment m, and l^{m}_{itn} is the length of segment m up to time $\bar{t}$, $\sum_{m=1}^{M} l^{m}_{itn} = \bar{t}$.

The probability density function of $l^{\star}_{itn}$ is therefore:

$$f_{l^{\star}_{itn}}\left(\bar{t}\right) = \lambda_{itn} q^{\bar{m}}_{itn} e^{-\lambda_{itn}\sum_{m=1}^{M} l^{m}_{itn} q^{m}_{itn}}, \tag{4.9}$$

where $\bar{m}$ is the segment that $\bar{t}$ falls into.

As it is not possible to watch the show past its end,

$$l^{W}_{itn} = min\left(l^{\star}_{itn}, L_{itn}\right), \tag{4.10}$$

where l^{W}_{itn} is the time household i spends watching show tn, and L_{itn} is the remaining length of show tn when sampled.

4.3 TV Show Recording

As discussed in §3.2.2, DVRs are typically filled to capacity. Therefore, the decision to record a new show is typically accompanied by the deletion of an older recorded show, revealing information about show preferences.

The TiVo model in the data only supports the ability to record one show at a time. Therefore, a newly recorded show tn is supposed to have i) higher expected flow utility than the show that is replaced ($u^{S}_{itn} > u^{S}_{itd}$); and ii) higher expected flow utility than all shows that air at time t but are not recorded ($u^{S}_{itn} > u^{S}_{itn'}, \forall n' \neq n$).[5]

[5] If the recording is set up by the same household member who is watching a live show when recording takes place, then the recorded show is supposed to have lower expected flow utility than the show being watched. However, if the recording is set up by a different household member, then

Based on the flow utility specified in Equation (4.1), conditions i) and ii) imply the probability that household i records show tn is given by:

$$Pr\left(y^R_{itn}=1\right)=\frac{exp\left(\bar{u}^S_{itn}\right)}{exp\left(\bar{u}^S_{itd}\right)+\sum_{n'}exp\left(\bar{u}^S_{itn'}\right)}, \tag{4.11}$$

where y^R_{itn} is an indicator that show tn is recorded by household i.

4.4 TV Advertising Viewing

4.4.1 Zapping

In the viewing model introduced above, the household can choose to avoid advertisements in a live show by channel switching (zapping). Similar to §4.2.2 where households can make channel switching decisions during program content upon receiving an external shock, households make zapping decisions during each advertisement. The zapping decision depends on the relative attractiveness of the advertisement as compared with other alternatives, and the cost of zapping. The probability of zapping advertisement tn (in a live show) can be written as:

$$\begin{aligned} &Pr\left(y^A_{itn,L}=0 \mid \left\{\nu^S_{itn}\right\}_n, \nu^O_{it}\right) \\ &= Pr\left(\bar{u}^A_{itn}+\epsilon^A_{itn} < max\left\{\bar{u}^O_{it}+\nu^O_{it},\ ln\left(\sum_{n'\in\mathcal{N}_{it}} exp\left(\bar{u}^S_{itn'}+\nu^S_{itn'}\right)\right)\right\}-c_i\right) \\ &= F_{\epsilon^A_{itn}}\left(max\left\{\bar{u}^O_{it}+\nu^O_{it},\ ln\left(\sum_{n'\in\mathcal{N}_{it}} exp\left(\bar{u}^S_{itn'}+\nu^S_{itn'}\right)\right)\right\}-c_i-\bar{u}^A_{itn}\right), \end{aligned} \tag{4.12}$$

the relationship between the recorded show and the live show may be reversed. In the data, over 96% of the user-specified recordings are of shows that air at a time when the household does not watch live TV, so we choose not to model the relationship between the recorded show and the live show watched when recording takes place.

where $y^A_{itn,L}$ is an indicator that advertisement tn is viewed (not zapped) by household i, c_i is the zapping cost faced by household i.[6] Computation of $Pr\left(y^A_{itn,L}=0\right)$ involves integrating out $\left\{\nu^S_{itn}\right\}_n$ and ν^O_{it} in $Pr\left(y^A_{itn,L}=0 \mid \left\{\nu^S_{itn}\right\}_n, \nu^O_{it}\right)$.

4.4.2 Zipping

Advertising avoidance is easier for recorded shows, where advertisements can be skipped by fast-forwarding (zipping). Because zipping reduces viewing time, the zipping decision is reached by comparing the flow utility of the advertisement with that of the outside good. If the advertisement provides higher flow utility than the outside good, the household watches it. The probability that household i zips advertisement tn (in a recorded show) is:

$$Pr\left(y^A_{itn,R}=0\right)=\frac{exp\left(\bar{u}^O_{it}\right)}{exp\left(\bar{u}^O_{it}\right)+exp\left(\bar{u}^A_{itn}\right)}, \tag{4.13}$$

where $y^A_{itn,R}$ is an indicator that advertisement tn is viewed (not zipped) by household i.[7]

The observed show sampling and watching decisions, show recording decisions and advertising viewing decisions enable us to recover flow utilities of TV shows, TV advertisements, and the outside good. We discuss estimation and identification in the next chapter.

---

[6] The zapping cost can reflect, for instance, the cost of missing the return of the current show, the cost of achieving consensus among multiple household members in determining which show to switch to, etc. Theoretically, there can also be a cost associated with channel switching during program content. We choose not to model this cost because it cannot be separately identified from the arrival rate of external shocks. A household observed to switch channel less often during program content can either have a low shock arrival rate or have a large switching cost.

[7] Theoretically, there can be a cost associated with zipping. However, it cannot be separately identified from the utility of the outside good, so we normalize the zipping cost to zero. Thus the zapping cost in essence measures the relative cost of zapping versus zipping, and is identified from the difference in zipping and zapping probabilities.

Note that one parameter in the advertisement flow utility measures “whether the preceding advertisement is skipped”, this parameter can be interpreted as the cost of starting to zip.

5

Estimation and Identification

This chapter discusses estimation and identification of the TV viewing model. All parameters are household-specific as indicated by subscript i. We perform estimation household-by-household, facilitated by the availability of panel data of relatively long cross-section and duration for each household.

5.1 Estimation

The viewing model is estimated by simulated maximum likelihood approach. The likelihood is derived in Appendix C. The sampling order implies $\nu^S_{itn'}$ and ν^O_{it} in Equations (4.5), (4.7), and (4.12) are truncated respectively below $\bar{u}^S_{itn} + \nu^S_{itn} - \bar{u}^S_{itn'}$ and below $\bar{u}^S_{itn} + \nu^S_{itn} - \bar{u}^O_{it}$. In the estimation, we simulate 100 sets of $\left(\left\{\nu^S_{itn}\right\}_n, \nu^O_{it}\right)$ to compute $Pr\left(y^W_{itn} = 1 \mid y^S_{itn} = 1\right)$, $q_{itn,t'}$ and $Pr\left(y^A_{itn,L} = 0\right)$.

It is well known that many households watch only a handful of available networks. We therefore construct household-specific consideration sets based on viewing history. For each household, the consideration set of networks consists of the smallest number of networks that collectively account for at least 90% of prime-time viewing time. On each viewing occasion, the choice set comprises the following two types of shows: i) live shows that are available on networks within the consideration set; and ii) shows stored on DVR that are recorded either manually or through a season pass.

To limit the size of β_i^S and β_i^A, we only estimate flow utility parameters associated with the 6 most popular genres (drama, comedy, reality TV, talk shows, news, and sports, together accounting for 68% of viewing time) and the 6 most popular networks (ABC, CBS, NBC, FOX, USA, and Comedy Central, together accounting for 60% of viewing time). There are 574 product categories in the advertisement sample. In an initial effort to capture the effects of product category on advertising preference, we classify the 574 product categories into four general categories: consumer packaged goods (CPG), service, drug, and other goods.

The arrival rate of external shocks, λ_{itn}, is parameterized as a function of genre:

$$\lambda_{itn} = exp\left(g_{tn}\rho_i\right), \tag{5.1}$$

where g_{tn} is a row vector on genre, the jth element being an indicator variable of whether show tn is of the jth genre.

5.2 Identification

Parameters that govern the flow utility of the show (β_i^S) and the outside good (β_i^O) are jointly identified by observed sampling, watching, and recording decisions, all revealing show preferences and time preferences. Similarly, parameters related to the flow utility from advertisements (β_i^A) are identified by variation in zipping and zapping decisions on advertisements.

Because the flow utility of the outside good varies by day, and the available programs usually also vary by day, one difficulty is to separate the value of the outside good from show quality in a day. If household i does not watch TV on a day, then there are two possibilities. For the first, all shows available are associated with low flow utility (i.e., small u^S_{itn} for all t, n). For the second, the opportunity cost of time is too high (i.e., large u^O_{it}). The recording behavior helps us disentangle these two factors. If no show is watched on a day but one or more shows are recorded, then it follows that the main reason for no viewing is high opportunity cost of time. If, on the other hand, none of the shows is watched or recorded, then the main reason for no viewing is low expected show quality in the day.

The parameters that determine the arrival rates of external shocks (ρ_i), and in turn, λ_{itn}, are identified by time spent on different types of shows. For instance, if a household switches more often in news than in dramas, then the household has a higher λ_{itn} if show tn is a news show than if it is a drama show, all else equal.

One concern is that people may switch less in shows that are more preferred, which leads to the question of whether the flow utility (β^S_i) and the shock arrival rate (λ_{itn}) can be separately identified. While sampling (show choice) decisions depend only on flow utility, viewing length depends on both shock arrival rate and flow utility. For instance, two shows with equal sampling probabilities can be watched for different lengths. All else equal, the show with a longer viewing length is associated with a lower shock arrival rate.

Finally, the zapping cost (c_i) is identified from the difference in zipping and zapping probabilities.

To assess whether the proposed estimation approach can recover known parameters, we develop a simulated dataset and implement the proposed estimation approach on it (Appendix C). The results show that the estimation approach works well in recovering known parameters.

6

Estimation Results

This chapter reports the estimation results of the viewing model. We use data from July 2005 to June 2006 for estimation, and reserve July 2006 for hold-out validation and policy experiments.

Table 6.1 reports the mean (across households) of parameter estimates in flow utilities of shows, as well as the percentages of households with significantly positive and negative estimates (5% level) respectively. It indicates the average viewer prefers shows that are short, familiar, live, and have just started (i.e., smaller viewing offset).

Table 6.1: Flow Utility Parameter Estimates for Shows

Variable	Mean Est	% Positive (5% level)	% Negative (5% level)
Length	-0.24	1.3%	75.8%
Number of episodes sampled in previous week	0.34	83.9%	0.6%
Live	8.88	86.5%	0.0%
Viewing offset	-5.25	0.1%	88.4%
Genre: Drama	-0.32	14.5%	43.2%
Genre: Comedy	-0.60	9.2%	57.2%
Genre: Reality TV	-0.91	8.4%	55.0%
Genre: Talk shows	0.12	34.9%	25.3%
Genre: News	0.01	29.0%	23.9%
Genre: Sports	-0.65	11.8%	39.5%
Network: ABC	0.96	48.6%	13.7%
Network: CBS	0.58	42.6%	20.0%
Network: NBC	0.92	52.0%	17.5%
Network: FOX	0.23	27.2%	21.8%
Network: USA	-0.35	8.5%	14.3%
Network: Comedy Central	-0.18	9.0%	11.2%

Similarly, Table 6.2 reports the mean (across households) of parameter estimates in flow utilities of advertisements, as well as the percentages of households with significantly positive and negative estimates (5% level) respectively. It indicates viewers forward blocks of advertisements successively, as demonstrated in the positive coefficient regarding whether the preceding commercial is viewed.[1] The first advertisement in a commercial break is less likely to be forwarded, presumably because it takes viewers some time after the commencement of a commercial break to initiate a forwarding action.

[1] As noted in §4.4.2, this positive coefficient also implies a positive cost associated with starting to zip.

Table 6.2: Flow Utility Parameter Estimates for Advertisements

Variable	Mean Est	% Positive (5% level)	% Negative (5% level)
Watch the preceding ad	4.89	74.0%	2.3%
First ad break	-0.25	10.2%	20.7%
Last ad break	-0.16	17.9%	15.4%
First slot in a break	1.63	63.9%	2.5%
Last slot in a break	-0.16	17.9%	15.4%
Genre: Drama	-1.12	14.6%	31.5%
Genre: Comedy	-0.32	15.4%	21.1%
Genre: Reality TV	-1.11	13.7%	30.5%
Genre: Talk shows	-0.37	13.3%	18.6%
Genre: News	-0.90	11.8%	20.9%
Genre: Sports	-0.56	11.5%	21.8%
Network: ABC	-0.69	17.5%	22.3%
Network: CBS	-0.57	15.8%	21.0%
Network: NBC	-1.03	15.2%	24.7%
Network: FOX	-0.75	13.6%	19.3%
Network: USA	-0.01	8.6%	8.5%
Network: Comedy Central	0.47	6.1%	4.1%
Category: CPG	-0.06	8.9%	8.2%
Category: Service	0.00	9.4%	5.6%
Category: Drug	-0.30	4.2%	7.2%

There also exists extensive heterogeneity in zapping cost across households, as reflected in the histogram of estimated zapping cost across households (Figure 6.1).

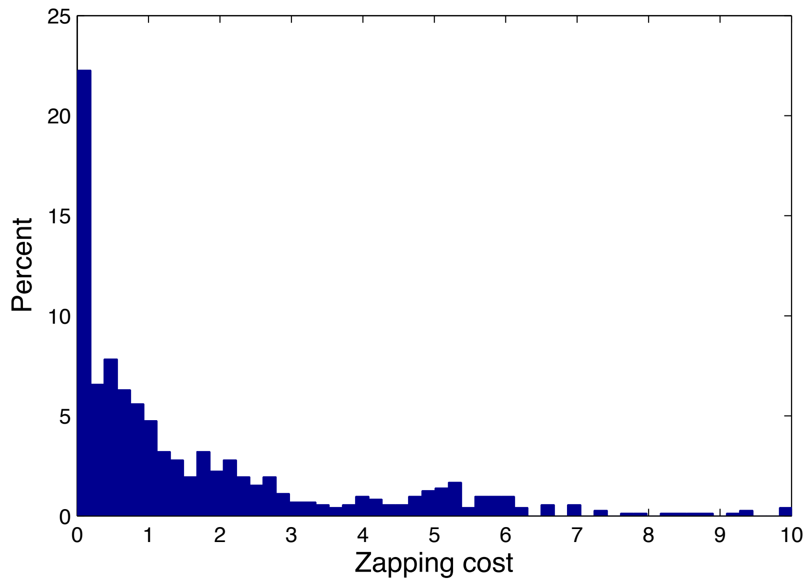


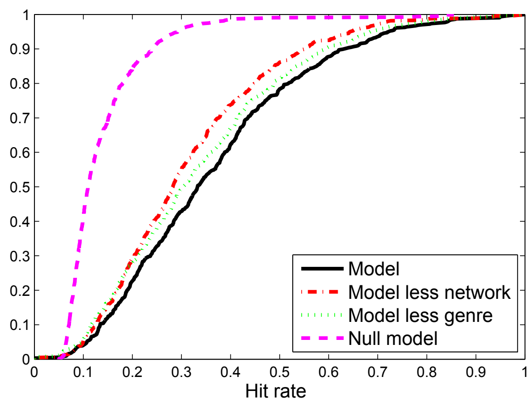
FIGURE 6.1: Histogram of Estimated Zapping Cost

Table 6.3 reports the mean estimates of shock arrival rates by genre. The average viewer is more likely to switch channel during news and less likely during drama shows, perhaps due to differences in program continuity. This finding is consistent with Shachar and Emerson (2000)'s finding that viewing persistence is higher for dramas and lower for news and sports.

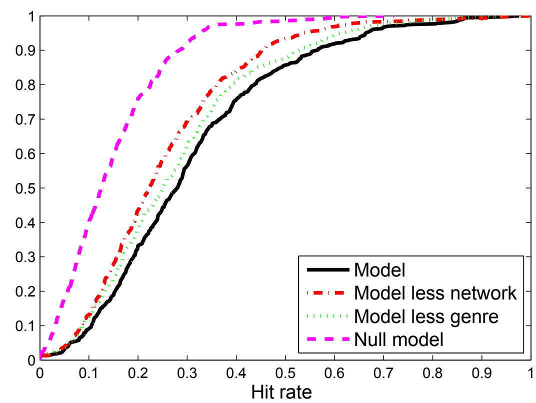
Table 6.3: Mean Estimates of Shock Arrival Rates

	Drama	Comedy	Reality	Talk	News	Sports
Number of Shocks per Hour	0.51	0.65	1.42	0.63	1.62	1.34

Next, we perform several model validity checks using the hold-out sample of July 2006. The first examines the hit rates (by household) in show sampling and watching (conditional on sampling) predictions, which can be obtained using Equations (4.4) and (4.5). We consider hit rates under several models on flow utility of show: including all covariates (the proposed model), excluding network covariates (model less network), excluding genre covariates (model less genre), and a null model with equal flow utilities across shows. Figure 6.2 indicates for both sampling and watching predictions, the hit rate under the proposed model first-order stochastically dominates the hit rate under the null model. Therefore, the proposed model performs better than the null model. In addition, networks are slightly more important than genres in predicting sampling and watching choices.



(a) Sampling Choices



(b) Watching Choices

FIGURE 6.2: Empirical CDF (Across Households) of the Hit Rate on Show Sampling (a) and Watching (b)

The second validity check concerns the viewing length (conditional on watching). Figure 6.3 compares the mean absolute error (MAE) of viewing length predictions under the proposed model and under a null model where exiting rate is constant throughout the show. The proposed model also outperforms the null model as reflected in the first-order stochastic dominance relationship.

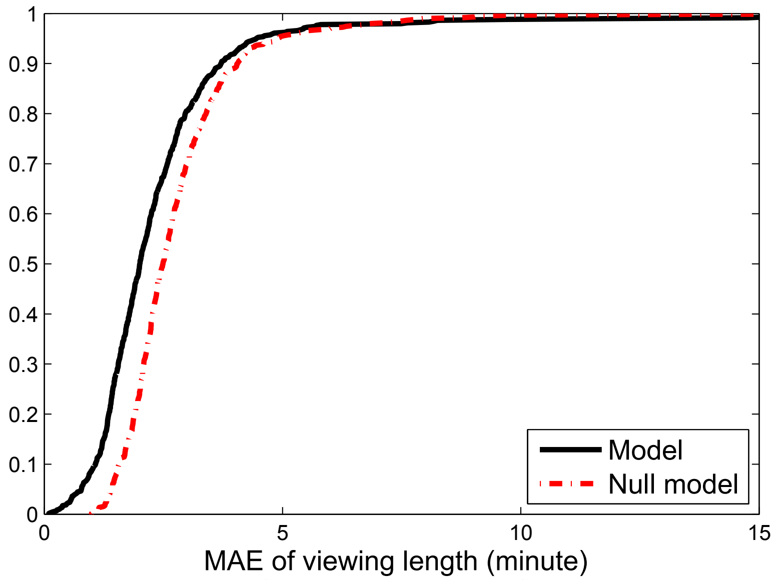


FIGURE 6.3: Empirical CDF (Across Households) of MAE on Viewing Length

Overall, there exists considerable heterogeneity across households in viewing preferences for genre, network, and advertising. Heterogeneity in viewing preferences, together with heterogeneity in advertising response, suggests the potential gains available from advertising targeting.

7

Policy Experiments: Advertising Targeting

This chapter conducts advertising targeting policy experiments. We first link advertising viewing with sales response in §7.1. Predicated on viewing behavior and advertising response, we discuss various targeting strategies in §7.2.

7.1 Advertising Response

To measure advertising response, we select seven product categories that evidence high variation in advertising and sales: children's yogurt, children's cereal, regular cola, diet cola, sports drink, toothpaste, and bathroom tissue. These categories are regularly purchased and frequently advertised, with sufficient cross-sectional and temporal variation in both purchase and advertising. Within these categories, we consider 22 leading brands that have non-negligible market share and advertised during the sample period. Table 7.1 presents descriptive information on these categories.

Table 7.1: Purchase Data Summary

Product Category	Number of Households	Number of Shopping Trips	Number of Purchases	Brands in Analysis
Children's yogurt	120	9,827	402	Yoplait Trix, Dannon Danimals, Yoplait Go Gurt
Children's cereal	525	47,022	2,829	Quaker Cap'n Crunch, Kellogg's Froot Loops, Kellogg's Frosted Flakes, General Mills Lucky Charms, General Mills Cinnamon Toast Crunch
Regular cola	446	40,679	3,189	Coke, Pepsi
Diet cola	416	38,409	4,071	Coke, Pepsi
Sports drink	298	26,874	1,377	Gatorade, Powerade
Toothpaste	386	39,268	1,017	Crest, Colgate, Aquafresh
Bathroom tissue	599	53,738	2,949	Charmin, Angel soft, Cottenelle, Quilted Northern, Scott

Product Category	Number of Purchases Per Household		Fraction of Purchases Made at Regular Price		Total Number of Weekly Advertisements		Average Number of Weekly Advertising Exposures Per Household	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Children's yogurt	3.4	3.9	0.84	0.37	104.3	41.9	1.9	1.5
Children's cereal	5.4	5.2	0.50	0.50	152.8	78.1	2.1	1.9
Regular cola	7.2	9.1	0.21	0.41	155.7	93.2	1.4	0.8
Diet cola	9.8	12.8	0.24	0.43	150.8	142.9	1.5	0.9
Sports drink	4.6	6.5	0.37	0.48	167.1	102.1	1.7	0.9
Toothpaste	2.6	2.0	0.70	0.46	870.9	293.0	1.6	1.1
Bathroom tissue	4.9	5.3	0.58	0.49	508.4	176.1	1.7	1.2

Two approaches are applied to measure advertising effects, a "model-free" approach, and a latent class approach. We discuss each in turn.

7.1.1 Model Free Analysis

7.1.1.1 Homogeneous Advertising Effects

A key issue with the measurement of advertising response is that advertising effects tend to be small relative to those of pricing and other potentially confounding factors (e.g., Tellis (1988); Mela et al. (1997)). Moreover, there are problems of advertising attribution, wherein the attribution of a sale to a particular advertisement is challenging and often involves various assumptions about decay (e.g., Clark et al. (2009)). Because of these problems, researchers typically measure advertising effects by using multivariate models to control for other covariates and past advertising. The downside of these approaches is that they involve numerous assumptions regarding functional form and error distributions. In contrast, our rich individual level panel data enable a different tact.

Specifically, we can hold the in-store causal environment constant by considering two same-store visits within the same week. In the cases where a household did not purchase on the first trip, but did on the second,[1] the only major factors that can vary between trips are household inventory and advertising. As store causals are held constant, this removes a major source of variation in purchase behavior. Also, as the contrast is within household, comparing second and first trips removes unobserved individual effects, including advertising prior to the first visit. Two key factors that remain are inventory and advertising between visits. With regard to inventory, draw down is limited in one week meaning that inventory levels should be similar on the second visit. Hence, our strategy is to compare the difference in second-visit demand when an advertisement appears between weekly visits and when advertising does not. As long as advertising is not highly correlated with inventory draw down between visits, this should yield a model free, non-parametric estimate of advertising effects.

Using this approach, we compute each brand's second-visit purchase log odds ratio with and without advertising exposures received between the two visits (Figure 7.1). Under the assumption of no advertising effect, the log odds ratio should be

[1] A category purchase in the first store visit makes a category purchase in the second store visit very unlikely. If three or more same-store visits are observed within the same week, each visit is paired with its previous visit.

centered around zero.[2] As the distribution is instead skewed to the right of zero, it suggests a positive advertising effect for most brands. A one-sample Kolmogorov-Smirnov test indicates the brand-level log odds ratio is not normally distributed with mean zero (p<0.016).[34]

[2] Let N_a indicate the number of people exposed to a brand's advertising between the first and second visit, of which n_a purchased in the second visit. Similarly, denote N_n as the number of people who were not exposed to advertising between the first and second visit, of which n_n purchased in the second visit. Then the log odds ratio of fraction who purchased in each respecive case is $ln\left(\frac{n_a/N_a}{n_n/N_n}\right)$. If there is no advertising effect, then in expectation, $\frac{n_a}{N_a} = \frac{n_n}{N_n}$ and $ln\left(\frac{n_a/N_a}{n_n/N_n}\right) = 0$.

[3] As a robustness check, we consider endogeneity bias that might arise from targeting households that are more responsive to advertising (e.g., Narayanan and Manchanda (2009)). Exploiting the panel nature of our data, we estimate a linear probability model for each brand:

$$y^P_{ijm} = \alpha_{ij} + \gamma_j A_{ijm} + \epsilon_{ijm},$$

where i denotes household, j denotes brand, and m denotes shopping trip. y^P_{ijm} is an indicator variable of whether household i purchases brand j in shopping trip m. A_{ijm} is an indicator variable of whether household i is exposed to brand j's advertisement since the preceding trip (in the same store and week). α_{ij} is a household fixed effect that captures unobserved heterogeneity, and ϵ_{ijm} is an idiosyncratic error term. In this model, the parameter γ_j measures the advertising lift. A two-sample Kolmogorov-Smirnov test indicates the distribution of the estimated advertising lift across brands $(1, ..., J)$ $\hat{\gamma}_j$, is not significantly different from the distribution of $\hat{\gamma}_j$ obtained under an alternative model without household fixed effects. This suggests results are invariant to the inclusion of household fixed effects that control for unobserved household-specific factors (e.g., demographic based targeting).

[4] Our second robustness check addresses the concern that the gap between the two visits might drive both advertising exposure and purchase likelihood in the same direction. When the second visit is further apart from the first visit, the household is more likely to be exposed to advertising due to more TV viewing, and might also be more likely to purchase in the second visit due to inventory reduction. To see if this is the case, we compute purchase log odds ratio by brand and gap between visits (1 or 2 days, 3 or 4 days, and 5-6 days). If the identified advertising effect were simply correlation induced by the gap between two visits, then within each gap group, we would expect the distribution of log odds ratio across brands to be centered around zero. Instead, we find the distribution within each gap group to be skewed to the right of zero. Therefore, the causal effect of advertising is valid.

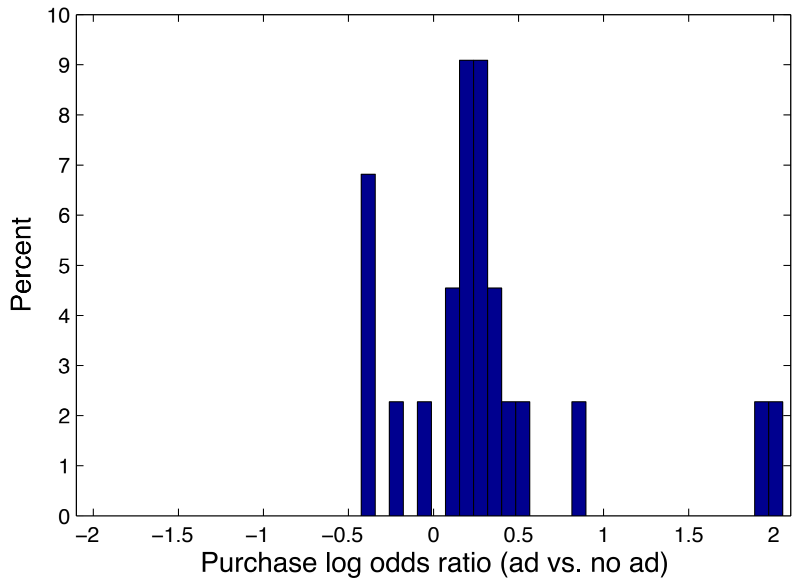


FIGURE 7.1: Histogram of Purchase Log Odds Ratio With and Without Advertising Exposure (by Brand)

7.1.1.2 Heterogeneous Advertising Effects

In the preceding subsection, we presumed that the increase in a brand's purchase probability conditioned on an advertising exposure is identical across households. We extend that model to consider a latent class representation of heterogeneity in advertising response. For household i who belongs to segment k ($k = 1, ..., K$), the second-visit probability of purchasing brand j is specified as:

$$Pr\left(y_{ijm}^{P} = 1 \mid i \in k\right) \quad = \quad \frac{exp\left(\theta_k A_{ijm} + \alpha_{jk}\right)}{1 + exp\left(\theta_k A_{ijm} + \alpha_{jk}\right)}, \tag{7.1}$$

where m denotes shopping trip, y_{ijm}^{P} is an indicator variable of whether household i purchases brand j in shopping trip m. A_{ijm} is an indicator variable of whether household i is exposed to brand j's advertisement since the preceding trip (in the same store and week). α_{jk} is a brand intercept for brand j and segment k.

The prior probability that household i is in ad-response segment k is specified as:

$$Pr\left(i \in k\right) = \frac{exp\left(\eta_k\right)}{\sum_{k'=1}^{K} exp\left(\eta_{k'}\right)}, \tag{7.2}$$

where η_1 is normalized to zero for identification. Household i's posterior probability of belonging to segment k can therefore be obtained by Bayes' rule:

$$Pr\left(i \in k \mid \left\{y_{ijm}^P\right\}_{j,m}\right) = \frac{Pr\left(i \in k\right) \prod_m \prod_j \left(Pr\left(y_{ijm}^P = 1 \mid i \in k\right)\right)^{y_{ijm}^P}}{\sum_{k'} Pr\left(i \in k'\right) \prod_m \prod_j \left(Pr\left(y_{ijm}^P = 1 \mid i \in k'\right)\right)^{y_{ijm}^P}}. \tag{7.3}$$

The likelihood function associated with household i's purchase decisions is:

$$L_i = \sum_k Pr\left(i \in k\right) \prod_m \prod_j \left(Pr\left(y_{ijm}^P = 1 \mid i \in k\right)\right)^{y_{ijm}^P}, \tag{7.4}$$

and the overall log likelihood is:

$$lnL \quad = \quad \sum_i lnL_i. \tag{7.5}$$

Because different product categories are associated with different purchase frequencies (Table 7.1), we estimate this model separately for each product category, and the number of segments is determined based on BIC.

For three product categories (children's yogurt, children's cereal, and toothpaste), a single segment is identified. For two product categories (bathroom tissue and sports drink), two segments are identified. For the rest two product categories (regular cola and diet cola), three segments are identified. Table 7.2 reports the estimated advertising effect (θ_k) by product category and consumer segment.

Table 7.2: Advertising Effect by Product Category and Consumer Segment

Product Category	Segment 1	Segment 2	Segment 3
Children's yogurt	-0.03 (0.52)	NA	NA
Children's cereal	0.23 (0.53)	NA	NA
Regular cola	0.51 (0.38)	0.20 (0.44)	-0.86 (0.58)
Diet cola	0.02 (0.28)	-0.68 (0.55)	-0.71 (0.69)
Sports drink	0.89 (0.33)	-0.12 (0.73)	NA
Toothpaste	-0.18 (0.31)	NA	NA
Bathroom tissue	0.27 (0.27)	-0.03 (0.32)	NA

Note: Standard errors in parentheses.

The estimates will be used in implementing household-level targeting strategies.

7.1.2 Mixture Logit Analysis

In a second approach to measuring advertising effects, we estimate a latent class brand choice model that includes campaign- and segment- specific advertising effectiveness parameters (see Appendix D). Compared to the model free approach outlined in §7.1.1, this analysis incorporates all shopping trips and controls for marketing mix variables, but at a cost of adding functional assumptions. As in the model free case, the results suggest a small, but generally positive short-term effect of advertising.

To assess the external validity of the findings, we calculate the percentages of estimated advertising effects that are significant at different p-values, and compare them with the percentages reported in existing review papers on advertising effects. Table 7.3 indicates our results are generally in line with previously reported results.

Table 7.3: Percent of Significant Advertising Effects

	P=0.05	P=0.2	P=0.4
Eastlack and Rao (1989)	24%		
Hu et al. (2007)		39% before 1995; 45% after 1995	
Lodish et al. (1995)			55% for new products; 36% for established brands
This study	15%	21%	36%

Overall, the short-term effects of advertising on sales appear to be small, and there is modest heterogeneity in advertising response. Of course, there can also exist long-term effects of advertising, suggesting that advertising targeting can focus on either advertising exposures or short-term profit. We discuss these possibilities below along with means to achieve them.

7.2 Targeting Strategies

Currently, national TV networks sell advertising inventory by show and in advance. In the upfront market, the advertiser contacts the TV network to purchase advertising space in bulk for the entire season. A typical request consists of a budget for the entire year and a negotiated cost per thousand viewers (CPM), with performance targets for specific periods and programs. In response to the advertiser's sales request, the TV network provides a list of commercials to be aired, by show and air date, although the exact location of the advertisement's placement within the show is decided at a later stage. Typically, advertising price varies with the number of viewers in a show as well as the demographic mix.

By enabling household-level targeting, digital TV advertising delivery offers the potential to relax several aspects of the advertising selling model along the follow two dimensions:

- Advance/Real-Time: The existing *advance purchase* approach can be improved via a viewing model such as developed in §4, as advertisers can better forecast advertising exposure among their target audience. Moreover, by tracking viewing in *real time*, advertisements can be targeted even more accurately. As viewing is known when buying real-time, a viewing model as developed above is not needed. Both real-time and advance buy are common in Internet settings, and we conjecture that TV will evolve similarly.

- Costs/Revenues: By tracking exposures and sales at the household level, advertisers can more easily track *costs* for a given set of exposures or *revenue and profits* arising from household-level sales response. When the objective is to minimize costs, the advertiser purchases advertising slots to maintain a set of exposures at the lowest cost. When the objective is to maximize profit, the advertiser takes into account both the imputed short-term profit gain from advertising, and the cost of advertising. While profits are more salient than costs to most firms, no advertising response model is needed for cost minimization.

Thus, there are four possible targeting scenarios (Advance/Real-Time $\times$ Costs/Revenue) arising from the use of digital TV, as shown in Table 7.4. Real-time cost minimization is the simplest to implement in the sense that no viewing model or sales data are required. To account for advertising skipping, we adjust observed/predicted advertising exposures with each household's mean zapping/zipping tendency.

Table 7.4: Targeting Scenarios

	Cost Minimization (No Ad Response Model)	Profit Maximization (Ad Response Model)
Real-time Buy (No Show Viewing Model)	§7.2.1	§7.2.2
Advance Buy (Show Viewing Model)	§7.2.3	§7.2.4

We illustrate these approaches using bathroom tissue brand Charmin in the holdout period of July 2006.[5]

[5] We select Charmin because: 1) the majority of households have made purchase in the bathroom tissue category (Table 7.1), and this category exhibits less fluctuation in demand than other

Several caveats of our simulations are worth noting. Our analysis abstracts away from competitive response (from the retailers, the competitors or the networks), and assumes a fixed (per-exposure) advertising price to a show. As such, our analysis can be best interpreted as a marginal improvement in advertising holding all else fixed. To the extent that usage of targeting tools becomes widespread, it is reasonable to conjecture rates will change in response. Yet even in that case, our approach solves the problem of an advertiser conditioned on rates, a first step in a general equilibrium approach.

We start with the two real-time purchase scenarios (§7.2.1 and §7.2.2), then illustrate the two advance purchase scenarios (§7.2.3 and §7.2.4).

7.2.1 Cost-Minimizing Real-time Buy

In this targeting scenario, the advertiser minimizes the total cost of maintaining the current number of exposures for each household, upon observing their show choices.

In the hold-out period of July 2006, sample households were exposed to 780 Charmin's advertisements at a total cost of $8.77. An alternative schedule that yields identical exposures within the same month can lower costs to $4.24, a 51% reduction in expenses. If the advertiser seeks to maintain these exposures on the same day, cost can still be decreased to $7.24, or by 18%. Scaled to the 1.1 million TV households in the three DMAs where the data were collected, the cost reduction will be $5,695 and $2,015 respectively.[6]

---

categories; 2) there exists moderate heterogeneity in household advertising response in the bathroom tissue category (Table 7.2), providing foundation for targeting; and 3) within this category, Charmin has the largest number of advertising exposures (share=46.8%) and the second largest sales (share=24.1%).

[6] Because we assume advertisers know *ex-ante* the show viewership, such cost reduction represents the "upper bound" of gains that can be realized under cost-minimizing buy. When implementing this strategy, the advertiser can set a threshold on advertising price, and when observing the household watch a show with a price below the threshold, purchase the slot.

7.2.2 Profit-Maximizing Real-time Buy

We use results from the model-free analysis with heterogeneous advertising effects (§7.1.1.2) to simulate incremental profit from advertising. Following Equation (7.1), the effect of advertising for household i on the purchase probability for brand j can be measured as:

$$\begin{aligned} \Delta_{ij} \mid i \in k &= Pr\left(y_{ij}^{P} = 1 \mid i \in k, A_{ij} = 1\right) - Pr\left(y_{ij}^{P} = 1 \mid i \in k, A_{ij} = 0\right) \\ &= \frac{exp\left(\theta_k + \alpha_{jk}\right)}{1 + exp\left(\theta_k + \alpha_{jk}\right)} - \frac{exp\left(\alpha_{jk}\right)}{1 + exp\left(\alpha_{jk}\right)}. \end{aligned} \tag{7.6}$$

Further, we use observed purchase data to impute each household's shopping frequency. For each household (i), we first calculate the mean and standard deviation of the number of weekly shopping trips. We then simulate the number of shopping trips in each week (w) of the hold-out period, N_{iw}. If there is an advertising exposure in week w, then the purchase probability of household i for brand j will change by $\Delta_{ij} N_{iw}$. Given a unit sales margin m, the advertising effect on weekly revenue becomes $\Delta_{ij} N_{iw} m$.

Summing over households and weeks, the optimization problem can be written as:

$$Max_{\{x_{itn}\}_{i,t,n}} \quad \sum_{i} \sum_{w} E\left\{\mathbf{1}\left(\sum_{t \in w} \sum_{n} x_{itn} e_{itn} \geq 1\right) \Delta_{ij} N_{iw} m - \sum_{t \in w} \sum_{n} x_{itn} c_{tn}\right\} \tag{7.7}$$

$$s.t. \quad x_{itn} \in \{0,1\}, \forall i, t, n, \tag{7.8}$$

where x_{itn} denotes household-show selection. $x_{itn} = 1$ if show tn is selected for household i, $x_{itn} = 0$ otherwise. e_{itn} is an indicator variable of whether household i watches the advertisement in show tn. Thus $\mathbf{1}\left(\sum_{t \in w} \sum_{n} x_{itn} e_{itn} \geq 1\right)$ indicates whether household i has at least one advertising exposure in week w, and

$\mathbf{1}\left(\sum_{t\in w}\sum_n x_{itn}e_{itn} \geq 1\right)\Delta_{ij}N_{iw}m$ measures the incremental weekly revenue from advertising. c_{tn} denotes the per-exposure advertising price associated with show tn. Hence $\sum_{t\in w}\sum_n x_{itn}c_{tn}$ is the weekly advertising cost associated with household i.

The expectation in Equation (7.7) is taken over the distribution of Δ_{ij} and N_{iw}. Because $E\left\{\mathbf{1}\left(\sum_{t\in w}\sum_n x_{itn}e_{itn} \geq 1\right)\Delta_{ij}N_{iw}m\right\} = m\mathbf{1}\left(\sum_{t\in w}\sum_n x_{itn}e_{itn} \geq 1\right)E\left(\Delta_{ij}N_{iw}\right)$, we first numerically compute $E\left(\Delta_{ij}N_{iw}\right)$, then solve the optimization problem[7] under different unit sales margins m.[8] Figure 7.2 plots the simulated incremental profit in current and optimal advertising schedules under these different margins. The simulated short-term advertising profit is strictly negative under the current schedule (the average loss ranges from $6.3 to $8.5). In contrast, optimal schedules produce small, but positive simulated incremental profit (the gain ranges from $0 to $1.2). Scaled to the 1.1 million TV households in the three DMAs where the data were collected, profit gain will range from $0 to $1,583. The ROI (defined as dollar of profit per dollar invested) under the current schedule ranges from -97% to -71%, whereas the ROI under the optimal schedules ranges from 0% to 46%.

[7] Specifically, we draw 50 sets of Δ_{ij} and N_{iw} for each household. $E\left(\Delta_{ij}N_{iw}\right)$ is taken as the mean of the 50 sets of $\Delta_{ij}N_{iw}$.

[8] Toilet paper has gross profit margin of roughly 13%. In the data, the mean and median prices of Charmin are respectively $6.8 and $6.3, leading to a unit sales margin of about $0.8-0.9.

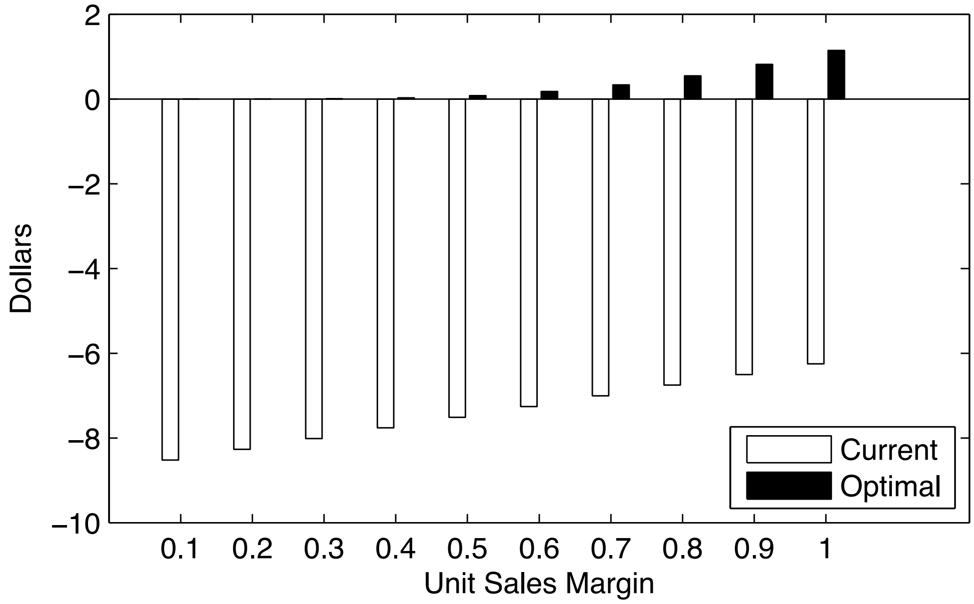


FIGURE 7.2: Simulated Incremental Profit (for Sample Households, Real-time Buy) Under Different Unit Sales Margins

7.2.3 Cost-Minimizing Advance Buy

Similar to §7.2.1, the advertiser minimizes the total cost of maintaining a set of exposures for each targeted household. Different from §7.2.1, the advertiser needs to predict show viewership when making decisions on which shows to buy for each household.[9]

For each household i, the advertiser selects shows that can (in expectation) maintain the current total advertising exposures at the lowest cost. Given the current expected number of advertising exposures to household i, Y_i, the optimization problem can be written as:

[9] We assume no network guarantee on minimal exposures to advertisers as is often common in practice.

$$Min_{\{x_{itn}\}_{t,n}} \quad \sum_t \sum_n c_{tn} x_{itn} \tag{7.9}$$

$$s.t. \qquad x_{itn} \in \{0,1\}, \forall t, n, \tag{7.10}$$

$$E\left(\sum_t \sum_n x_{itn} r_{itn}\right) \geq Y_i, \tag{7.11}$$

where c_{tn} and x_{itn} are defined similarly as in §7.2.2. r_{itn} is the probability that household i watches the advertisement placed in show tn. The constraint in Equation (7.11) ensures that the firm is minimizing costs for the same viewers as it currently targets with its show-level advance buying. Without this constraint, it is possible to buy much cheaper exposures, but to viewers who are not in the set currently targeted by advertisers. For example, Charmin could make cheaper purchases of exposures in children's shows, but these are a radically different target.

The expectation in Equation (7.11) is taken over the distribution of r_{itn}, which is associated with the uncertainty in viewing model estimates. r_{itn} can be obtained from the viewing model output:

$$\begin{aligned} r_{itn} &= Pr\left(y^S_{itn} = 1\right) Pr\left(y^W_{itn} = 1 \mid y^S_{itn} = 1\right) \\ &\quad \times \left(\int_{t'} Pr\left(l^W_{itn} \geq t'\right) f(t')\, dt'\right) \\ &\quad \times Pr\left(y^A_{itn} = 1\right), \end{aligned} \tag{7.12}$$

where t' denotes a possible advertising location (time into show) and $f(t')$ represents the probability density function of a uniform distribution, whose support is timing of commercial breaks in show tn. Assuming zero viewing offset, conditional on watching show tn, household i will be exposed to the advertisement placed at t' if the viewing length l_{itn}^{W} exceeds t'. $Pr\left(y_{itn}^{A}=1\right)$ denotes the probability that the advertisement will not be zapped (if show tn is live) or zipped (if show tn is recorded).

Because $E\left(\sum_t \sum_n x_{tn} r_{itn}\right) = \sum_t \sum_n x_{tn} E\left(r_{itn}\right)$, we first compute $E\left(r_{itn}\right)$ by simulation, and then use it to solve the optimization problem for each household.[10] The overall cost is reduced from \$8.77 to \$6.57, a 30% reduction in expenses. Scaled to the 1.1 million TV households in the three DMAs where the data were collected, the cost reduction will be \$2,902.

7.2.4 Profit-Maximizing Advance Buy

Finally, we consider the profit-maximizing advance buy scenario, which requires both the viewing model and the advertising response model. The optimization problem is similar to that outlined in §7.2.2, except that observed viewership e_{itn} is replaced by predicted viewership r_{itn} that is described in Equation (7.12):

$$Max_{\{x_{itn}\}_{i,t,n}} \quad \sum_i \sum_w E\left\{\sum_{t\in w}\sum_n x_{itn} r_{itn} \Delta_{ij} N_{iw} m - \sum_{t\in w}\sum_n x_{itn} c_{tn}\right\} \tag{7.13}$$

$$s.t. \quad x_{itn} \in \{0,1\}, \forall i,t,n. \tag{7.14}$$

The expectation in Equation (7.13) is taken over the distribution of r_{itn}, Δ_{ij} and N_{iw}. Because $E\left\{\sum_{t\in w}\sum_n x_{itn} r_{itn}\Delta_{ij}N_{iw}m\right\} = m\sum_{t\in w}\sum_n x_{itn} E\left(r_{itn}\Delta_{ij}N_{iw}\right)$, we first numerically compute $E\left(r_{itn}\Delta_{ij}N_{iw}\right)$, then solve the optimization problem under different margins m. Figure 7.3 plots the simulated incremental profit in optimal advertising schedules under these different margins. The simulated incremental profit ranges from \$0 to \$0.17. Scaled to the 1.1 million TV households in the three DMAs where the data were collected, the profit gain will range from \$0 to \$224. The ROI (defined as dollar of profit per dollar invested) ranges from 0 to 40%. The gain in

[10] Specifically, we draw 50 sets of household-specific parameter estimates, each leading to a prediction on households' show choices and viewing length conditional on choice. $E\left(r_{itn}\right)$ is taken as the average of this predicted viewership across the 50 sets of parameter draws.

ROI for advance buy is lower than the gain in ROI obtained under real-time buy because the advance buy optimization is based on predicted viewership instead of observed viewership.

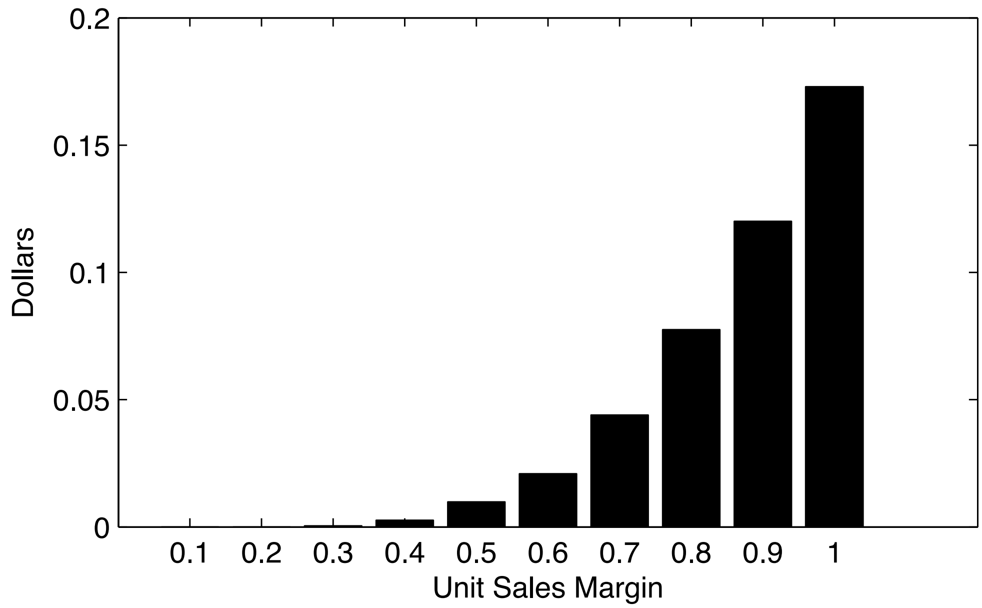


FIGURE 7.3: Simulated Incremental Profit (for Sample Households, Advance Buy) Under Different Unit Sales Margins

7.2.5 Targeting Summary

We summarize the targeting results in Table 7.5.

Table 7.5: Targeting Gains

	Cost Minimization	Profit Maximization
Real-time Buy	51% cost reduction (maintain each household's total exposures); 18% cost reduction (maintain each household's daily exposures)	(Depending on sales margin) ROI: 0%-46%
Advance Buy	30% cost reduction (maintain each household's total expected exposures)	(Depending on sales margin) ROI: 0%-40%

Results indicate by leveraging household viewing and purchase information, advertisers can achieve reduction in advertising costs and/or lift in incremental revenue with appropriate targeting strategies. The gain is especially considerable if advertisers can buy impressions in real time. This implies potential surplus in the TV advertising market and Pareto improvement to be achieved. The strategies we discussed allow the surplus to be fully captured by one advertiser. The allocation of surplus among multiple advertisers and TV networks in the long run depends on competition on both the product market and the advertising market, as well as the advertising pricing and allocation mechanisms.

8

Conclusion

TV remains, by far, the predominant modality for the transmission and reception of video content, and the largest advertising medium. More importantly, its pre-eminence stands to benefit from recent digital innovations such as DVRs. However, the DVR innovation is both a blessing and a curse. On the one hand, DVRs have greatly enhanced households' TV-viewing experiences, leading to increased viewing consumption. On the other hand, these boxes enable households to forward past advertisements. It is our goal to redress this limit by taking advantage of micro-level viewing data and micro-targeting capabilities inherent in DVRs to improve the efficacy of advertising. If targeting proves effective, it opens new paths to TV advertising pricing by allowing TV networks and cable companies to sell boxes as well as shows.

In this paper, we outline an approach to facilitate the targeting of TV advertising in the new digital contexts based upon a unique dataset that integrates several data sources (DVR data, purchase data, show data and advertising data). We estimate models of households' TV program sampling/consumption and advertising response, and then conduct counterfactual policy experiments to evaluate potential gains from targeting. We consider several household-level targeting scenarios by manipulating: 1) whether the advertising purchase is made in advance or in real time; and 2) whether the objective function is to minimize costs for a given set of exposures or to maximize incremental profit from advertising. Results indicate micro-targeting can lower advertising costs and raise incremental revenue.

This study makes several contributions. Theoretically, we develop an integrated model on TV show viewing, TV advertising viewing, purchasing and advertising targeting. Methodologically, we propose a new modeling framework on media consumption by explicitly accounting for the role of uncertainty, and propose targeting strategies leveraging household-level data. Substantively, we offer policy recommendations to advertisers on targeting based on micro-level data, which can be of great potential.

Several extensions are possible. First, the sales model does not consider context effects in advertising. As a result, the proposed targeting strategies do not account for context effects. Consumer behavior research suggests advertisements placed within a program of dissimilar content are recalled significantly better than if placed within a program of similar content (e.g., Furnham and Price (2006)). One extension is to incorporate context effects in measuring advertising effectiveness, and apply it in designing the targeting strategies.

Second, the targeting strategies we proposed do not take into account potential complementarities between the consumption of goods and advertisements (Becker and Murphy (1993)). According to a recent study by Tuchman et al. (2015), the quantity of the advertised product purchased recently can explain the advertising skip rates. They suggest that advertising efficacy would depend on the compliance of the targeted households, and the firm can target the subset of households whose purchases and welfare will likely change in response to the advertising campaign. Therefore, a fruitful extension to household-level targeting involves selection of households based on joint consumption of goods and advertisements. Another extension involves correlation between households' preferences for show attributes and for products.

Third, the targeting strategies are designed based on an integrated channel structure, under which advertisers and TV networks share the same information and same objectives. Therefore, we are unable to model the pricing mechanism and the advertising allocation problem from the TV networks' perspective. We do this because advertising allocation can be designed to align networks' and advertisers' incentives (e.g., Wilbur et al. (2013)). Still, another extension is to consider strategic interactions between advertisers and TV networks and how they affect gains available from micro-targeting.

Finally, the targeting strategies we proposed are designed in a partial-equilibrium framework. Future research can extend micro-targeting into a general equilibrium framework, taking into account competitive response and allowing for re-optimization of advertising and product prices.

Appendix A

Details on Inferring Advertising Exposures

A.1 Key Data Files

Inference on advertising exposure is based on three key data files: *Watchtv, Trickplay,* and *Adschedule*. We discuss each in turn.

A.1.1 Watchtv (from TiVo log files)

Watchtv contains all "watch TV" events. Watch TV events are logged whenever the TiVo recognizes that a new program is being watched. There are 2 types of Watch TV events, live and recorded.

- Watch TV live: Watch TV live events are logged in the following cases: 1) The user leaves the UI and enters the live TV context; 2) the user changes channels to a new program while in the live TV context; and 3) a new program begins while the user is sitting on a channel.
- Watch TV Recorded: Watch TV recorded events are logged whenever a user initiates playback of a recording from their Now Playing list.

The file includes the following key information on each event:

- Time of the event.
- The Tribune Media Services (TMS) identifier of the program.
- *Callsign*: The network that the program is on.
- *Dropin offset*: Number of seconds after the beginning of the program when viewing begins.
- *Duration*: Program duration.
- Reason for the recording based on disk (for recorded programs).

A.1.2 Trickplay (from TiVo log files)

Trickplay contains timing of all Trickplay events. Trickplay events are logged whenever the user uses the TiVo Trickplay buttons while watching video. Key events observed are:

- FORWARD (*Speed* = 3x, 18x, or 60x)
- PLAY
- PAUSE

A.1.3 Adschedule (from TNS advertising schedule file)

Adschedule contains the following information on all national TV advertisements:

- *Adstart*: Air time of the advertisement.
- *Adlength*: Length of the advertisement.
- Company name.
- Brand name.

- Product name.
- Product category.
- A brief (40 characters or shorter) description pertaining to the content/script of the advertisement.
- Name of the program that the advertisement airs in.
- Genre of the program that the advertisement airs in.
- *Showtime*: Air time of the program that the advertisement airs in.
- The estimated price of the advertisement.

A.2 Identify Idle Periods

When inferring advertising exposures, it is important to focus on exposures occurred when households actively watch TV. We do this based on usage of the remote control (keystrokes). If two observed keystrokes that occur at t_1 and t_2 are at least 60 minutes apart, we define the period between $t_1 + 60$ and t_2 as an idle period. We then delete from the log files all observations that occur during idle periods.[1] This way we obtain a relatively conservative measure on advertising exposure.

More specifically, we follow the steps below to identify idle periods:

- Step 1: In *Trickplay* data, compute the time between each keystroke (t_2) and the previous keystroke (t_1). Retain those observations with an over-one-hour gap. The period between $t_1 + 60$ and t_2 is considered idle.
- Step 2: Obtain the time of the last keystroke (T) observation for each household, the period between $T + 60$ and the end of the sample period is considered idle.

[1] We have tested other windows (15, 30, and 120 minutes), and our key results are robust to different windows.

- Step 3: Combine the two files obtained in Step 1 and 2. The new file contains the start time and the end time of each idle period.

- Step 4: Identify events in *Watchtv* data that occur during the idle periods, and delete these events.

A.3 Infer Viewing Time on TV Shows

Because *Watchtv* data only report the start time of each event, we have to infer the actual time a household spends on each show in order to infer the show content (and in turn commercials) the household is exposed to. This is not trivial since we need to incorporate pauses and forwards.

We follow the steps below to infer viewing time.

- Step 1: In *Trickplay* data, obtain the pause time (*Pausetime*) for each PAUSE as the time between the PAUSE and the next keystroke.

- Step 2: In *Trickplay* data, define the time associated with each FORWARD as the start time of a FORWARD, *Skipstart*. Define the end time of the FORWARD, *Skipend*, as the time of the next keystroke or the time of the next program change, whichever is earlier. The amount of program content skipped (*Skiptime*) depends on the forwarding speed: $Skiptime = Speed \times (Skipend - Skipstart)$.

- Step 3: In *Watchtv* data, compute the time between each Watch TV event and the next Watch TV event, name it *Gap*.

- Step 4: Compute *Timewatch* (length of the show that the household is exposed to, either at regular speed or at high speed)

  - When there is no pause or forward between two Watch TV events

    * *Timewatch* = Min (*Duration* - *Dropin offset*, *Gap*)

  - When there are pauses and/or forwards between two Watch TV events

    * *Timewatch* = Min (*Duration* - *Dropin offset*, *Gap* - *Pausetime*+ *Skiptime*)

A.4 Infer Advertising Exposure and Skipping

We use *Watchtv, Trickplay* and *Adschedule* to infer different types of skipping behavior, following the steps below:

- Step 1: Map *Skipstart* and *Skipend* in *Trickplay* data into program time in the corresponding show, taking into account pauses and forwards that occur in the show and prior to *Skipstart*.

- Step 2: Map *Adstart* and *Adend* (*Adend* = *Adstart* + *Adlength*) in *Adschedule* data into program time in the corresponding show (by *Showtime*).

- Step 3: Generate different types of zipping behavior using information generated in Step 1 and Step 2:

  - Start zipping =1 if: $Adstart < Skipstart < Adend$

  - Full zipping = 1 if $Skipstart \leq Adstart < Adend \leq Skipend$

  - Stop zipping = 1 if: $Skipstart < Adstart < Skipend < Adend$

- Step 4: Generate different types of zapping behavior using *Watchtv* data and information generated in Step 2:

  - Start zapping = 1 if: $Dropin\ offset \leq Adstart < Dropin\ offset + Timewatch < Adend$ or $Adstart \leq Dropin\ offset \leq Dropin\ offset + Timewatch \leq Adend$

- Stop zapping =1 if: $Adstart \leq Dropin\ offset < Adend < Dropin\ offset + Timewatch$

Appendix B

Length of Viewing

Assume external shocks arrive via a homogeneous Poisson process (rate λ), and assume the exiting probability is piecewise constant with M segments.

Further denote:

- $N_m(t)$: the total cumulative number of external shocks occurred during segment m up to time t
- $N(t)$: the total cumulative number of external shocks up to time t, $N(t) = \sum_{m=1}^{M} N_m$
- $N_{exit}(t)$: the total cumulative number of external shocks that lead to exiting up to time t
- p_m: probability of keeping watching conditional on receiving an external shock during segment m
- $q_m(t)$: probability that a given shock is during segment m, which equals the share of segment m's time up to time t, $\frac{t_m}{t}$

Then we have:

$$
\begin{aligned}
Pr\{l \geq t\} &= Pr\{N_{exit}(t) = 0\} \\
&= \sum_{k=0}^{\infty} Pr\{N(t) = k\} \sum_{i=1}^{k} \Bigg[Pr\Bigg\{ N_1(t) = n_1,\ N_2(t) = n_2, \ldots, \\
&\quad N_M(t) = n_M \mid \sum_{m=1}^{M} N_m(t) = k \Bigg\} \times \prod_{m=1}^{M} (p_m)^{n_m} \Bigg] \\
&= \sum_{k=0}^{\infty} \frac{e^{-\lambda t} (\lambda t)^k}{k!} \sum_{i=0}^{k} \frac{k!}{n_1! \ldots n_M!} q_1^{n_1}(t) \cdots q_M^{n_M}(t) \prod_{m=1}^{M} (p_m)^{n_m} \\
&= \sum_{k=0}^{\infty} \frac{e^{-\lambda t} (\lambda t)^k}{k!} \sum_{i=0}^{k} \frac{k!}{n_1! \ldots n_M!} \prod_{m=1}^{M} [q_m(t)\, p_m]^{n_m} \\
&= \sum_{k=0}^{\infty} \frac{e^{-\lambda t} (\lambda t)^k}{k!} \left[\sum_{m=1}^{M} q_m(t)\, p_m \right]^k \\
&= e^{-\lambda t} \sum_{k=0}^{\infty} \frac{\left\{ \lambda t \left[\sum_{m=1}^{M} q_m(t)\, p_m \right] \right\}^k}{k!} \\
&= e^{-\lambda t} e^{\lambda t \left[\sum_{m=1}^{M} q_m(t) p_m \right]} \\
&= e^{-\left[1 - \sum_{m=1}^{M} \frac{t_m}{t} p_m \right] \lambda t} \\
&= e^{-\lambda \sum_{m=1}^{M} t_m (1 - p_m)}.
\end{aligned}
$$

Appendix C

Likelihood Function of the Viewing Model

We estimate the viewing model by simulated maximum likelihood approach. The sampling order implies $\nu^S_{itn'}$ and ν^O_{it} in Equations (4.5), (4.7), and (4.12) are truncated respectively below $\bar{u}^S_{itn} + \nu^S_{itn} - \bar{u}^S_{itn'}$ and below $\bar{u}^S_{itn} + \nu^S_{itn} - \bar{u}^O_{it}$. In the estimation, we first simulate $K = 100$ sets of $\left(\left\{\nu^S_{itn}\right\}_n, \nu^O_{it}\right)$, denoted as $\left\{\nu^{S1}_{itn}\right\}_n, \ldots, \left\{\nu^{SK}_{itn}\right\}_n, \nu^{O1}_{it}, \ldots, \nu^{OK}_{it}$. We then derive the simulated likelihood associated with sampling, watching, recording and advertising viewing decisions, and aggregate them to obtain the overall likelihood function.

Sampling

Each sampled show contributes to the likelihood with:

$$Pr\left(y^S_{itn} = 1\right) = \frac{exp\left(X^S_{itn}\beta^S_i\right)}{exp\left(X^O_{it}\beta^O_i\right) + \sum_{n' \in \mathcal{N}_{it}} exp\left(X^S_{itn'}\beta^S_i\right)}.$$

Watching

Each sampled show that is watched contributes to the simulated likelihood with:

$$
\begin{aligned}
&Pr\left(y^W_{itn}=1 \mid y^S_{itn}=1\right)\\
&= Pr\left(u^S_{itn} \geq IV_{it\mathcal{N}_{it}}\right)\\
&= \frac{1}{K}\sum_{k=1}^{K}\left[1-F_{\epsilon^S_{itn}}\left(max\left\{\bar{u}^O_{it}+\nu^{Ok}_{it},\ ln\left(\sum_{n'\in\mathcal{N}_{it}} exp\left(\bar{u}^S_{itn'}+\nu^{Sk}_{itn'}\right)\right)\right\}-\bar{u}^S_{itn}-\nu^{Sk}_{itn}\right)\right],
\end{aligned}
$$

and each sampled show that is not watched contributes to the likelihood with:

$$
Pr\left(y^W_{itn}=0 \mid y^S_{itn}=1\right) \quad = \quad 1-Pr\left(y^W_{itn}=1 \mid y^S_{itn}=1\right)
$$

Switching

Each show that is watched until the end ($l^W_{itn}=L_{itn}$) contributes to the likelihood with:

$$
\begin{aligned}
Pr\left(l^W_{itn} \mid l^W_{itn}=L_{itn}\right) \quad &= \quad Pr\left\{l^\star_{itn}>L_{itn}\right\}\\
&= \quad \frac{1}{K}\sum_{k=1}^{K} e^{-\lambda_{itn}\sum_{m=1}^{M} l^m_{itn}\left(q^m_{itn}\mid\left\{\nu^{Sk}_{itn'}\right\}_{n'},\nu^{Ok}_{it}\right)},
\end{aligned}
$$

where:

$$
\begin{aligned}
&q_{itn} \mid \left\{\nu^{Sk}_{itn'}\right\}_{n'}, \nu^{Ok}_{it}\\
&= F_{\epsilon^S_{it'n}}\left(max\left\{\bar{u}^O_{it}+\nu^{Ok}_{it},\ ln\left(\sum_{n'\in\mathcal{N}_{it'}} exp\left(\bar{u}^S_{it'n'}+\nu^{Sk}_{it'n'}\right)\right)\right\}-\bar{u}^S_{itn}-\nu^{Sk}_{itn}\right).
\end{aligned}
$$

Each show that is not watched until the end ($l^W_{itn}<L_{itn}$) contributes to the likelihood with:

$$
\begin{aligned}
Pr\left(l_{itn}^{W} \mid l_{itn}^{W} < L_{itn}\right) &= f\left(l_{itn}^{W}\right) \\
&= \frac{1}{K}\sum_{k=1}^{K}\lambda_{itn}q_{itn}^{\bar{m}}e^{-\lambda_{itn}\sum_{m=1}^{M}l_{itn}^{m}\left(q_{itn}^{m}|\left\{\nu_{itn'}^{Sk}\right\}_{n'},\nu_{it}^{Ok}\right)},
\end{aligned}
$$

where $\bar{m}$ is the segment that l_{itn}^{W} falls into.

Recording

Each recorded show contributes to the likelihood with:

$$
Pr\left(y_{itn}^{R} = 1\right) = \frac{exp\left(X_{itn}^{S}\beta_{i}^{S}\right)}{exp\left(X_{itd}^{S}\beta_{i}^{S}\right) + \sum_{n'} exp\left(X_{itn'}^{S}\beta_{i}^{S}\right)}.
$$

Advertising zapping (live shows)

Each non-zapped advertisement contributes to the likelihood with:

$$
\begin{aligned}
Pr\left(y_{itn,L}^{A} = 1\right) &= 1 - \frac{1}{K}\sum_{k=1}^{K}\left[F_{\epsilon_{itn}^{A}}\left(max\left\{\bar{u}_{it}^{O} + \nu_{it}^{Ok},\ ln\left(\sum_{n'\in\mathcal{N}_{it}} exp\left(\bar{u}_{itn'}^{S} + \nu_{itn'}^{Sk}\right)\right)\right\}\right.\right. \\
&\quad \left.\left. -c_{i} - \bar{u}_{itn}^{A}\right)\right],
\end{aligned}
$$

and each zapped advertisement contributes to the likelihood with:

$$
Pr\left(y_{itn,L}^{A} = 0\right) = 1 - Pr\left(y_{itn,L}^{A} = 1\right).
$$

Advertising zipping (recorded shows)

Each non-zipped advertisement contributes to the likelihood with:

$$Pr\left(y^A_{itn,R}=1\right)=\frac{exp\left(X^A_{itn}\beta^A_i\right)}{exp\left(X^A_{itn}\beta^A_i\right)+exp\left(X^O_{it}\beta^O_i\right)},$$

and each zipped advertisement contributes to the likelihood with:

$$Pr\left(y^A_{itn,R}=0\right)=1-Pr\left(y^A_{itn,R}=1\right).$$

Due to the high computational cost associated with simulation of $\left\{\nu^S_{itn}\right\}_n$ and ν^O_{it}, we take two steps in the estimation. In the first step, we estimate β^S_i, β^O_i and β^A_i using likelihoods associated with sampling, recording and advertising zipping decisions. These likelihoods do not rely on $\left\{\nu^S_{itn}\right\}_n$ and ν^O_{it}. In the second step, we estimate ρ_i (λ_i) and c_i, taken estimates $\hat{\beta}^S_i$, $\hat{\beta}^O_i$ and $\hat{\beta}^A_i$ from the first step as given.

Further, to assess whether the proposed estimation approach can recover known parameters, we simulate a synthetic dataset and implement the proposed estimation approach on the simulated data. Table C.1 presents the results. The results show that the estimation approach works well in recovering known parameters.

Table C.1: Monte Carlo Simulation Results

Parameter	True Value	Estimate	Standard Error
β^S_1	15.0	14.58	0.70
β^S_2	-1.5	-1.44	0.13
β^C	6.0	5.86	0.33
β^A_1	9.0	8.76	0.50
λ	0.5	0.48	0.09
c	20.0	19.76	0.80

Appendix D

A Latent Class Logit Model on Brand Choice (§7.1.2)

We apply a standard latent class logit model to estimate advertising effects on brand choice. Each observation is a shopping trip. The dependent variable is brand choice, and the choice set includes all brands in the category and an outside option of no purchase in the category.

D.1 Model

The utility that household i obtains from purchasing brand j ($j = 1, \ldots, J$) in shopping trip m is given by:

$$U_{ijm}^{P} = Z_{ijm}\theta_i + h\left(A_{ijm}\right) + \epsilon_{ijm}^{P}, \tag{D.1}$$

where Z_{ijm} is a vector that includes a brand fixed effect, price of brand j at shopping trip m, an indicator of whether brand j is on promotion (either display or feature) at shopping trip m, and an indicator of whether household i purchased brand j in the previous category purchase (loyalty). $h\left(A_{ijm}\right)$ captures the advertising effect.

Finally, ϵ^P_{ijm} is an idiosyncratic error term affecting the inherent valuation of brand j at shopping trip m, it is observed by the household but not by the researcher.

The function $h\left(A_{ijm}\right)$ is assumed to be linear in advertising exposures:

$$h\left(A_{ijm}\right) = \sum_{a=1}^{A} N^A_{ijam}\gamma_{ija} + \left(N^A_{ijm} - \sum_{a=1}^{A} N^A_{ijam}\right)\gamma_{ij0}, \tag{D.2}$$

where $a \in \{1, \ldots, A\}$ indexes advertising campaigns, and N^A_{ijam} is household i's number of exposures to advertising campaign a since the previous category purchase or in the past 7 days, whichever is smaller. We aggregate campaigns with smaller number of exposures. In particular, if N^A_{ijm} is household i's total number of advertising exposures across campaigns during this window, then $N^A_{ijm} - \sum_{a=1}^{A} N^A_{ijam}$ represents the total number of exposures to all other campaigns. The formulation above allows us to measure the differential effect of large advertising campaigns (e.g., Homer and Yoon (1992); Malaviya et al. (1996); Vakratsas and Ambler (1999); Tellis et al. (2005)).

The utility associated with the outside good (i.e., no purchase) is given by:

$$U^P_{i0m} = \epsilon^P_{i0m}. \tag{D.3}$$

Assuming the idiosyncratic error terms to be i.i.d. standard Type I Extreme Value distributed, the probability that household i chooses brand j in shopping trip m is:

$$Pr\left(y^P_{ijm} = 1\right) = \frac{exp\left(Z_{ijm}\theta_i + h\left(A_{ijm}\right)\right)}{1 + \sum_{j'=1}^{J} exp\left(Z_{ij'm}\theta_i + h\left(A_{ij'm}\right)\right)}. \tag{D.4}$$

D.2 Estimation and Identification

D.2.1 Estimation

We use a latent class model to capture household heterogeneity. Conditional on being in segment k ($k = 1, \ldots, K$), the probability that household i chooses brand j in shopping trip m is:

$$Pr\left(y_{ijm}^P = 1 \mid i \in k\right) = \frac{exp\left(Z_{ijm}\theta_k + h_k\left(A_{ijm}\right)\right)}{1 + \sum_{j'=1}^{J} exp\left(Z_{ij'm}\theta_k + h_k\left(A_{ij'm}\right)\right)}, \tag{D.5}$$

where $h_k\left(A_{ijm}\right) \equiv \sum_{a=1}^{A} N_{ijam}^A \gamma_{kja} + \left(N_{ijm}^A - \sum_{a=1}^{A} N_{ijam}^A\right)\gamma_{kj0}$.

The prior probability that household i is in segment k is modeled as:

$$Pr\left(i \in k\right) = \frac{exp\left(\eta_k\right)}{\sum_{k'=1}^{K} exp\left(\eta_{k'}\right)}, \tag{D.6}$$

where η_1 is normalized to zero for identification. Household i's posterior probability of belonging to segment k can therefore be obtained by Bayes' rule:

$$Pr\left(i \in k \mid \left\{y_{ijm}^P\right\}_{j,m}\right) = \frac{Pr\left(i \in k\right)\prod_m \prod_j \left(Pr\left(y_{ijm}^P = 1 \mid i \in k\right)\right)^{y_{ijm}^P}}{\sum_{k'} Pr\left(i \in k'\right)\prod_m \prod_j \left(Pr\left(y_{ijm}^P = 1 \mid i \in k'\right)\right)^{y_{ijm}^P}}. \tag{D.7}$$

The likelihood function associated with household i's purchase decisions is:

$$L_i^P = \sum_k Pr\left(i \in k\right)\prod_m \prod_j \left(Pr\left(y_{ijm}^P = 1 \mid i \in k\right)\right)^{y_{ijm}^P}, \tag{D.8}$$

And the total log likelihood is:

$$lnL^P = \sum_i lnL_i^P. \tag{D.9}$$

We estimate the sales model by MLE, maximizing lnL^P. The estimator is implemented using trust-region algorithm. Derivation of the gradients is provided below.

Derivation of the gradients

The likelihood function in the sales model is:

$$
\begin{aligned}
lnL^P &= \sum_i ln\left(\sum_k Pr\left(i \in k\right)\prod_m \prod_j \left(Pr\left(y_{ijm}^P = 1 \mid i \in k\right)\right)^{y_{ijm}^P}\right) \\
&\equiv \sum_i ln\left(\sum_k p_k \prod_m \prod_j \left(q_{ijm}^k\right)^{y_{ijm}^P}\right),
\end{aligned} \tag{D.10}
$$

where:

$$
q_{ijm}^k \equiv Pr\left(y_{ijm}^P = 1 \mid i \in k\right) = \frac{exp\left(Z_{ijm}\theta_k + h_k\left(A_{ijm}\right)\right)}{1 + \sum_{j'=1}^J exp\left(Z_{ij'm}\theta_k + h_k\left(A_{ijm}\right)\right)},
$$

$$
p_k \equiv Pr\left(i \in k\right) = \frac{exp\left(\eta_k\right)}{\sum_{k'=1}^K exp\left(\eta_{k'}\right)}.
$$

Further denote $l_i^k \equiv \prod_m \prod_j \left(q_{ijm}^k\right)^{y_{ijm}^P}$ as the likelihood of household i's choices conditional on being in segment k.

To compute $\frac{\partial lnL^P}{\partial \eta_k}$, we rewrite $\sum_k p_k l_i^k = \frac{\sum_{k'}\left\{exp(\eta_{k'})l_i^{k'}\right\}}{\sum_{k'} exp(\eta_{k'})}$, then:

$$
\begin{aligned}
\frac{\partial lnL^P}{\partial \eta_k} &= \sum_i \frac{\frac{exp(\eta_k)l_i^k \sum_{k'} exp(\eta_{k'}) - exp(\eta_k)\sum_{k'}\left\{exp(\eta_{k'})l_i^{k'}\right\}}{\left[\sum_{k'} exp(\eta_{k'})\right]^2}}{\sum_{k'} p_{k'} l_i^{k'}} \\
&= \sum_i \frac{p_k l_i^k - p_k \sum_{k'}\left(p_{k'} l_i^{k'}\right)}{\sum_{k'} p_{k'} l_i^{k'}}.
\end{aligned} \tag{D.11}
$$

Suppose $\theta_k = (\theta_{k1}, \dots, \theta_{kp})^T$, and $Z_{ijm} = (Z_{ijm1}, \dots, Z_{ijmp})$. To compute $\frac{\partial lnL^P}{\partial \theta_{kh}}$ $(h = 1, \dots, p)$, we note:

$$\begin{aligned}
&\frac{\partial q_{ijm}^k}{\partial \theta_{kh}} \\
&= \frac{Z_{ijmh} exp\left(Z_{ijm}\theta_k\right) \sum_{j'=1}^J exp\left(Z_{ij'm}\theta_k\right) - exp\left(Z_{ijm}\theta_k\right) \sum_{j'=1}^J \left(Z_{ij'mh} exp\left(Z_{ij'm}\theta_k\right)\right)}{\left[\sum_{j'=1}^J exp\left(Z_{ij'm}\theta_k\right)\right]^2} \\
&= Z_{ijmh} q_{ijm}^k - q_{ijm}^k \sum_{j'=1}^J Z_{ij'mh} q_{ij'm}^k.
\end{aligned}$$

This gives us:

$$\begin{aligned}
\frac{\partial \prod_m \prod_j \left(q_{ijm}^k\right)^{y_{ijm}^P}}{\partial \theta_{kh}} &= \prod_m \prod_j \left(Z_{ijmh} q_{ijm}^k - q_{ijm}^k \sum_{j'=1}^J Z_{ij'mh} q_{ij'm}^k\right)^{y_{ijm}^P} \prod_{m' \neq m} \prod_j \left(q_{ijm'}^k\right)^{y_{ijm}^P} \\
&= \prod_m \prod_j \left(q_{ijm}^k\right)^{y_{ijm}^P} \prod_{m' \neq m} \prod_j \left(q_{ijm'}^k\right)^{y_{ijm}^P} \\
&\quad \times \prod_j \left(Z_{ijmh} - \sum_{j'=1}^J Z_{ij'mh} q_{ij'm}^k\right)^{y_{ijm}^P} \\
&= \prod_m \prod_j \left(q_{ijm}^k\right)^{y_{ijm}^P} \left[\sum_m \prod_j \left(Z_{ijmh}\right)^{y_{ijm}^P} - \sum_m \sum_j \left(Z_{ijmh} q_{ijm}^k\right)\right].
\end{aligned}$$

Therefore:

$$\frac{\partial lnL^P}{\partial \theta_{kh}} = \sum_i \frac{p_k l_i^k}{\sum_{k'} p_{k'} l_i^{k'}} \left[\sum_m \prod_j \left(Z_{ijmh}\right)^{y_{ijm}^P} - \sum_m \sum_j \left(Z_{ijmh} q_{ijm}^k\right)\right]. \qquad \text{(D.12)}$$

D.2.2 Identification

The sales model is a latent class logit model, and the identification of the model is based on variations observed in marketing mix variables, prices, and household choices. More specifically, the brand fixed effects are identified by brand market share for different households. For instance, households who always choose Coke over Pepsi have a large Coke fixed effect. The price coefficient is identified by variation in prices within- and across- brands and associated variation in choice decisions. The coefficients related to promotion and loyalty are identified similarly.

The advertising effects are identified by variation in the number of campaign exposures and the associated variation in choice decisions. One concern with this analysis is endogeneity in advertising exposure, because firms may target households that are more responsive to advertising. To address this concern, we compute for each household and each brand: i) the brand's share in the household's category purchase, and ii) the brand's share in the household's category advertisement exposure. We do not find a correlation between the two variables, implying the advertising variables are unlikely to be endogenous. This lack of correlation might be a consequence of the current targeting practice, and implies the potential for improvement. In addition, the between-household variance of advertising exposure is lower than the within-household variance. For the 22 brands in the sample, the mean and median portions of the total variance between households are respectively 20.1% and 15.9%, with the lowest being 4.0% and the highest being 47.9%.

The parameters related to household segmentation are identified through differences in household response to marketing mix variables and prices.

D.3 Results

Our focus pertains to advertising effects.[1] To assess whether or not advertising affects sales, we plot the histogram of the asymptotic t-statistics of estimated advertising coefficients across all categories, segments and campaigns (see Figure D.1). Of note, advertising effects are statistically small, but when they are distinguishable from zero, they are mostly positive. To assess the external validity of the findings, we calculate the percentages of estimated advertising effects that are significant at different p-values, and compare them with the percentages reported by existing review papers on advertising effects. Table D.1 indicates the results are generally in line with previous findings. Of note, the methodology used in these studies is different from ours: they directly measure sales response to advertising spending at the aggregate level in field experiments, while we control for price, promotion, and brand loyalty in a disaggregate model.

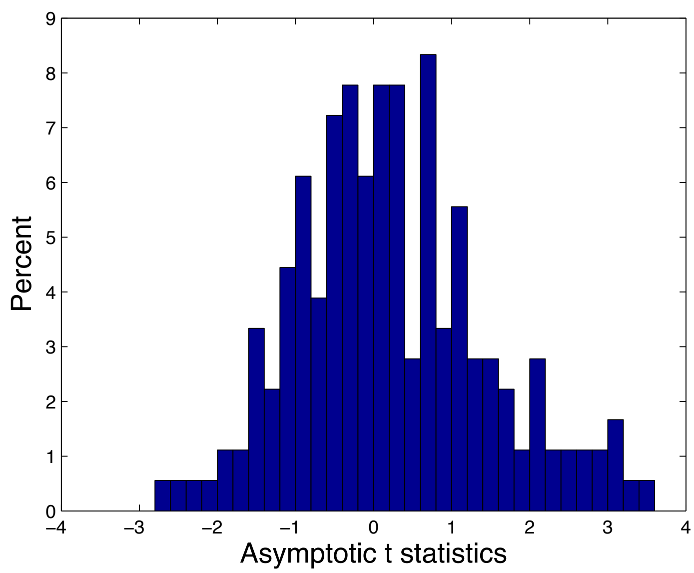


FIGURE D.1: Distribution of the Asymptotic t-Statistics of Estimated Advertising Coefficients (by Advertising Campaign and Household Segment)

[1] Detailed estimation results are presented in §D.4.

Table D.1: Percent of Significant Advertising Effects

	P=0.05	P=0.2	P=0.4
Eastlack and Rao (1989)	24%		
Hu et al. (2007)		39% before 1995; 45% after 1995	
Lodish et al. (1995)			55% for new products; 36% for established brands
This study	15%	21%	36%

To assess the magnitude of advertising effects that are significant (at the level of 10%), we simulate the effect of removing them on sales. We measure both short-term and long-term effect. The short-term effect is obtained by assuming all variables except advertising exposure to the given campaign stay unchanged. The long-term effect takes into account not only the immediate effect of eliminating the advertising campaign, but also the carry-over effect of altered choice through the purchase event feedback measure (whether the household purchased the brand in the previous category purchase).[2]

Figure D.2 depicts the results in terms of percentage changes in market share. The length of the black segment of each bar represents the short-term effect, and the total length of the bar represents the long-term effect. As shown in the figure, the change is negative for the majority (18 out of 22) of household segment - campaign pairs. Among such pairs, both the short-term and long-term changes in own market share vary from -5.5% to -0.4%.

[2] Specifically, the steps below are taken to compute the long-term effect. First, for each household, we compute the new purchase probability for each brand in the first observed shopping trip, and simulate a purchase decision based on these probabilities. Then, using the new loyalty measure derived from the first purchase decision, we compute this household's purchase probabilities in the second shopping trip. This process is repeated until the last observed shopping trip, resulting in a new purchase sequence for this household. The entire process is repeated 100 times for each household, and the new market share is the average of the simulated market shares under the 100 simulations.

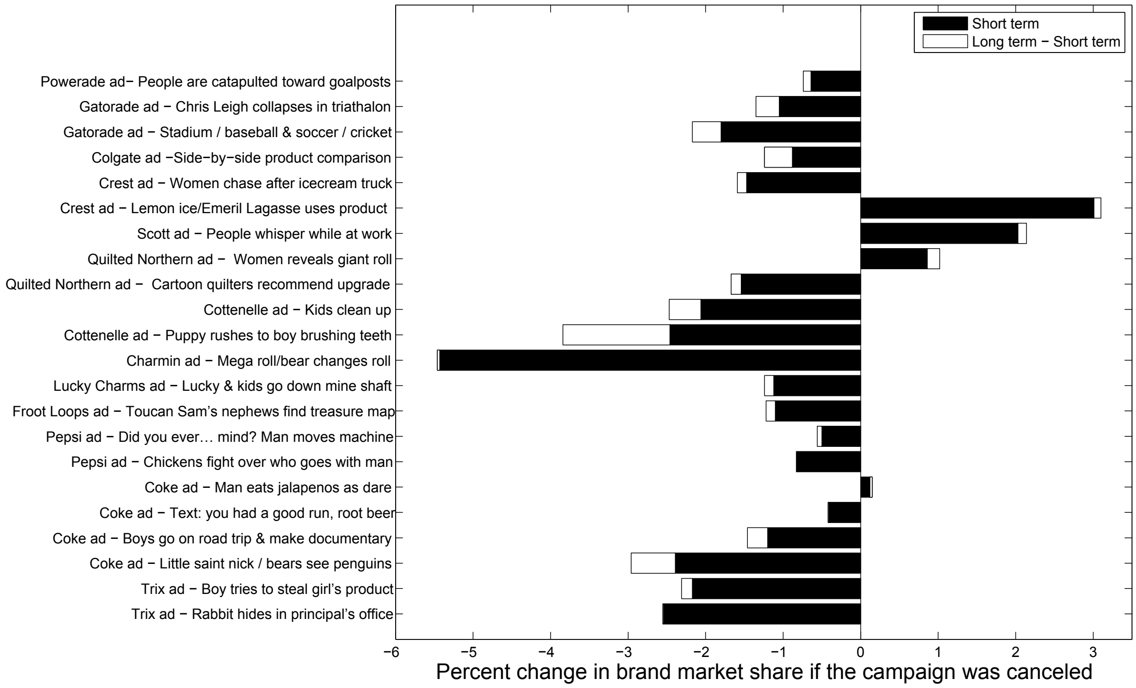


FIGURE D.2: Percentage Change in Own-Brand Market Share Following Elimination of an Advertising Campaign

Overall, we conclude that the advertising effects are small as in the previous literature (e.g., Sethuraman et al. (2011)). However, the effects vary across household segments and campaigns, suggesting the potential to enhance the efficiency of advertising spend.

D.4 Detailed Estimation Results of the Sales Response Model

D.4.1 Children's Yogurt Category

Variable	Segment 1	Segment 2
Price	-0.969 ***	-1.711 ***
	(0.252)	(0.338)
Promotion	-0.382	0.316 *
	(0.249)	(0.196)
Loyalty	1.983 ***	0.445 *
	(0.269)	(0.245)

Variable	Segment 1	Segment 2
Yoplait Trix	-1.663 ***	-1.952 ***
	(0.558)	(0.629)
Yoplait Go-Gurt	0.142	-0.593
	(0.664)	(0.787)
Dannon Danimals	-3.095	-3.207 ***
	(0.578)	(0.566)
Trix ad - Rabbit hides in principal's office	0.015	0.214 **
	(0.186)	(0.102)
Trix ad - Boy tries to steal girl's product	0.218	0.525 **
	(0.267)	(0.229)
Trix ad - Others	-0.507 *	0.040
	(0.271)	(0.042)
Go-Gurt ad - Nintendo sweeptakes/ kids at lunch table	0.008	-0.028
	(0.136)	(0.232)
Go-Gurt ad - Scooby-Doo! mystery tubes	0.143	0.002
	(0.119)	(0.197)
Go-Gurt ad - Others	0.002	0.057
	(0.085)	(0.037)
Danmial ads	-0.381	-0.002
	(0.676)	(0.152)
Segment size	0.231	0.769
Log likelihood	-1803	
BIC	3855	

Note: * indicates significance at 10% level; ** indicates significance at 5% level; and *** indicates significance at 1% level.

D.4.2 Children's Cereal Category

Variable	Segment 1	Segment 2
Price	-1.376 ***	-1.169 ***
	(0.084)	(0.101)
Promotion	0.493 ***	0.802 ***
	(0.026)	(0.026)
Loyalty	1.576 ***	1.262 ***
	(0.103)	(0.108)
Quaker Cap'n Crunch	-1.977 ***	-4.616 ***
	(0.224)	(0.268)
Kelloggs Froot Loops	-1.615 ***	-5.114 ***
	(0.244)	(0.293)

Variable	Segment 1	Segment 2
Kelloggs Frosted Flakes	-2.659 ***	-5.088 ***
	(0.213)	(0.238)
General Mills Lucky Charms	-1.756 ***	-4.660 ***
	(0.261)	(0.309)
General Mills Cinnamon Toast Crunch	-2.159 ***	-4.972 ***
	(0.270)	(0.299)
Other brands	0.951 ***	-1.887 ***
	(0.220)	(0.253)
Cap'n ad - Cap'n brings cereal to children at zoo/hiking boys	0.126	0.303
	(0.181)	(0.198)
Cap'n ad - Cap'n drives boat into Karate class	0.044	0.230
	(0.239)	(0.230)
Cap'n ad - Other ads	0.142	-0.077
	(0.099)	(0.177)
Froot Loops ad - Toucan Sam's nephews find treasure map	0.303 **	-0.458
	(0.147)	(0.898)
Froot Loops - Other ads	-0.459	-0.345
	(0.273)	(0.396)
Frosted Flakes ad - Tony and boys play soccer on the beach	0.136	-0.860
	(0.256)	(0.906)
Frosted Flakes ad - Tony & kids play soccer/baseball	-0.091	0.320
	(0.617)	(0.291)
Frosted Flakes ad - Other ads	-0.176	0.021
	(0.212)	(0.069)
Lucky Charms ad - Animated children climb mountain	-0.241	0.102
	(0.436)	(0.394)
Lucky Charms ad - Lucky & kids go down mine shaft	0.344 *	0.275
	(0.208)	(0.372)
Lucky Charms ad - Other ads	0.060	0.296
	(0.410)	(0.310)
Cinnamon Toast Crunch ad - Girls say why DJ loves cereal	-0.218	-0.090
	(0.178)	(0.164)
Cinnamon Toast Crunch ad - Other ads	-0.708	-0.468
	(0.443)	(0.464)
Segment size	0.290	0.710
Log likelihood	-11118	
BIC	22720	

Note: * indicates significance at 10% level; ** indicates significance at 5% level; and *** indicates significance at 1% level.

D.4.3 Regular Cola Category

Variable	Segment 1	Segment 2	Segment 3
Price	2.971 ***	-5.759 ***	-37.292 ***
	(0.137)	(0.227)	(1.639)
Promotion	4.432 ***	0.866 ***	3.226 ***
	(0.210)	(0.130)	(0.295)
Loyalty	1.135 ***	1.092 ***	1.139 ***
	(0.129)	(0.103)	(0.172)
Coke	-14.209 ***	7.096 ***	57.928 ***
	(0.460)	(0.493)	(2.716)
Pepsi	-12.689 ***	5.842 ***	54.457 ***
	(0.406)	(0.461)	(2.571)
Coke ad - Little saint nick / bears see penguins	0.417 **	0.020	0.725 **
	(0.155)	(0.126)	(0.354)
Coke ad - Boys go on road trip & make documentary	0.411 *	-0.089	0.299
	(0.233)	(0.151)	(0.404)
Coke ad - Man steals sips of soda in store	-0.501	-0.574	0.586
	(0.693)	(0.363)	(0.742)
Coke ad - Text: you had a good run, root beer	-0.155	0.516 **	-0.847
	(0.734)	(0.246)	(1.179)
Coke ad - Man eats jalapenos as dare	0.156	0.069	-1.451 **
	(0.403)	(0.250)	(0.597)
Coke ad - Other ads	0.084	0.029	0.018
	(0.120)	(0.031)	(0.091)
Pepsi ad - Chickens fight over who goes with man	0.220	0.579 **	0.163
	(0.571)	(0.225)	(0.409)
Pepsi ad - Jimmy Fallon dances in the street	-0.637	-0.085	-0.631
	(0.851)	(0.283)	(1.128)
Pepsi ad - Mariah Carey ringtones	-0.167	0.287	-1.225
	(0.511)	(0.297)	(1.092)
Pepsi ad - V Guerrero & a rodriguez hit moon & run	-1.978	-0.447	0.279
	(1.306)	(0.400)	(0.437)
Pepsi ad - Mice watch man run through maze	-0.771	-0.313	0.253
	(0.746)	(0.436)	(0.513)
Pepsi ad - Other ads	-0.009	-0.288	0.029
	(0.140)	(0.160)	(0.132)
Segment size	0.170	0.234	0.596
Log likelihood	-4616		
BIC	9795		

Note: * indicates significance at 10% level; ** indicates significance at 5% level; and *** indicates significance at 1% level.

D.4.4 Diet Cola Category

Variable	Segment 1	Segment 2
Price	-3.327 ***	-1.876 ***
	(0.085)	(0.084)
Promotion	2.838 ***	1.809 ***
	(0.096)	(0.104)
Loyalty	1.171 ***	1.23 ***
	(0.078)	(0.074)
Coke	-0.235	-0.359
	(0.205)	(0.253)
Pepsi	-0.030	-0.090
	(0.245)	(0.262)
Coke ad - Bubbles follow man drinking product	0.093	0.083
	(0.074)	(0.121)
Coke ad - Starry eyed surprise / woman opens can	0.046	0.062
	(0.041)	(0.110)
Coke ad - Goody two shoes / woman gets haircut	0.132	-0.127
	(0.115)	(0.222)
Coke ad - Couple leaves theatre and ends date	-0.215	-0.361
	(0.144)	(0.251)
Coke ad - Other ads	0.016	0.031
	(0.024)	(0.098)
Pepsi ad - Jay Mohr makes deal with another celebrity	0.090	0.010
	(0.100)	(0.119)
Pepsi ad - Cans dance to radio in market case	-0.101	0.112
	(0.142)	(0.152)
Pepsi ad - Did you ever... mind? / Man moves machine	0.345 ***	-0.057
	(0.112)	(0.178)
Pepsi ad - Staying alive / Cindy and Carson ogle man	0.004	-0.175
	(0.116)	(0.132)
Pepsi ad - Other ads	-0.012	-0.164 **
	(0.032)	(0.069)
Segment size	0.598	0.402
Log likelihood	-9286	
BIC	18898	

Note: * indicates significance at 10% level; ** indicates significance at 5% level; and *** indicates significance at 1% level.

D.4.5 Sports Drink Category

Variable	Segment 1	Segment 2	Segment 3
Price	-2.926 ***	-2.309 ***	-4.363 ***
	(0.111)	(0.098)	(0.167)
Promotion	-0.539 **	1.201 ***	3.412 ***
	(0.206)	(0.100)	(0.189)
Loyalty	0.908 ***	1.331 ***	0.613 ***
	(0.168)	(0.112)	(0.156)
Gatorade	1.854 ***	1.383 ***	2.840 ***
	(0.255)	(0.231)	(0.292)
Powerade	0.551 **	0.490 **	-0.701 **
	(0.260)	(0.223)	(0.273)
Gatorade ad - Soccer team travels to compete	-0.715	0.226	-0.375
	(0.494)	(0.223)	(0.304)
Gatorade ad - Stadium / baseball & soccer / cricket	0.645 ***	0.007	-0.070
	(0.195)	(0.252)	(0.295)
Gatorade ad - Kids pretend to be professional players	-0.128	0.037	-0.190
	(0.336)	(0.165)	(0.338)
Gatorade ad - Volleyball in rain morphs into a woman	-0.232	0.331	-0.547
	(0.410)	(0.264)	(0.490)
Gatorade ad - Chris Leigh collapses in triathalon	-0.086	0.646 **	-0.011
	(0.294)	(0.347)	(0.508)
Gatorade ad - Other ads	0.001	-0.040	0.049 **
	(0.034)	(0.041)	(0.018)
Powerade - Text: sport is what you make it	-0.654	0.236	0.369
	(1.797)	(0.390)	(0.597)
Powerade - People are catapulted toward goalposts	-0.083	0.761 **	-0.436
	(0.604)	(0.413)	(0.633)
Powerade - Other ads	0.319 ***	-0.150	-0.019
	(0.090)	(0.152)	(0.085)
Segment size	0.357	0.157	0.487
Log likelihood	-5053		
BIC	10554		

Note: * indicates significance at 10% level; ** indicates significance at 5% level; and *** indicates significance at 1% level.

D.4.6 Toothpaste Category

Variable	Segment 1	Segment 2	Segment 3
Price	1.484 ***	-13.086 ***	-5.550 ***
	(0.119)	(0.527)	(0.366)
Promotion	2.831 ***	-3.205 ***	3.264 ***
	(0.212)	(0.300)	(0.232)
Loyalty	0.721 ***	0.918 ***	1.895 ***
	(0.213)	(0.234)	(0.336)
Crest	-7.973 ***	16.706 ***	2.554 ***
	(0.282)	(0.854)	(0.648)
Colgate	-7.732 ***	15.748 ***	3.206 ***
	(0.259)	(0.760)	(0.571)
Aquafresh	-8.235 ***	12.059 ***	0.830
	(0.264)	(0.663)	(0.541)
Other brands	-9.924 ***	7.525 ***	5.474 ***
	(0.439)	(0.621)	(0.583)
Crest ad - Lemon ice/Emeril Lagasse uses product	-0.341	-0.712 **	0.127
	(0.271)	(0.349)	(0.295)
Crest ad - Women chase after icecream truck	-0.107	0.295 **	0.165
	(0.347)	(0.146)	(0.299)
Crest ad - Women face appears all over town	-0.503	0.246	-0.089
	(0.541)	(0.372)	(0.740)
Crest ad - Other ads	0.019	0.015	0.008
	(0.089)	(0.011)	(0.023)
Colgate ad - Woman discusses seeing pink	0.057	-0.400	0.011
	(0.244)	(0.436)	(0.271)
Colgate ad -Side-by-side product comparison	0.174	-0.452	0.413 *
	(0.226)	(0.464)	(0.244)
Colgate ad -Man on snowboard on bathroom	-0.693	-0.154	-0.121
	(0.639)	(0.358)	(0.437)
Colgate ad - Other ads	0.071 *	-0.044	0.025
	(0.042)	(0.063)	(0.023)
Aquafresh ads	-0.522	-0.010	-0.557
	(0.623)	(0.202)	(0.643)
Segment size	0.317	0.396	0.287
Log likelihood	-4081		
BIC	8690		

Note: * indicates significance at 10% level; ** indicates significance at 5% level; and *** indicates significance at 1% level.

D.4.7 Bathroom Tissue Category

Variable	Segment 1	Segment 2
Price	0.831 ***	-3.087 ***
	(0.076)	(0.152)
Promotion	0.425 ***	0.633 ***
	(0.022)	(0.017)
Loyalty	2.700 ***	1.288 ***
	(0.090)	(0.071)
Charmin	-7.572 ***	-0.836 ***
	(0.211)	(0.309)
Angel Soft	-7.220 ***	-2.796 ***
	(0.181)	(0.226)
Cottonelle	-6.918 ***	-1.802 ***
	(0.180)	(0.245)
Quilted Northern	-7.242 ***	-1.178 ***
	(0.187)	(0.265)
Scott	-8.213 ***	3.368 ***
	(0.301)	(0.460)
Other brands	-5.001 ***	-1.033 ***
	(0.158)	(0.203)
Charmin ad - Mega roll/bear changes roll	0.054	0.461 ***
	(0.199)	(0.155)
Charmin ad - Side-by-side product comparison	0.360	0.426
	(0.293)	(0.269)
Charmin ad - Bears in woods/competitor comparison	0.085	-0.207
	(0.353)	(0.253)
Charmin ad - Other ads	-0.178	0.017
	(0.152)	(0.027)
Angel Soft ads	0.152	0.262 **
	(0.144)	(0.114)
Cottenelle ad - Puppy rushes to boy brushing teeth	0.618 ***	-0.079
	(0.196)	(0.228)
Cottenelle ad - Kids clean up	0.677 **	-0.003
	(0.280)	(0.233)
Cottenelle ad - Other ads	-0.067	0.011
	(0.160)	(0.066)
Quilted Northern ad - Cartoon quilters recommend upgrade	0.268	0.339 ***
	(0.198)	(0.112)
Quilted Northern ad - spot the quilters $1,000,000 game	0.282	-0.100
	(0.242)	(0.314)

Variable	Segment 1	Segment 2
Quilted Northern ad - Women reveals giant roll	-0.078	-0.390 **
	(0.752)	(0.196)
Quilted Northern ad - Other ads	-0.149	-0.004
	(0.259)	(0.058)
Scott ad - People whisper while at work	0.140	-0.975 **
	(0.185)	(0.361)
Scott ad - Delivery drivers discover soft side	0.188	0.418
	(0.186)	(0.269)
Scott ad - Other ads	-1.029	0.114 ***
	(0.891)	(0.039)
Segment size	0.265	0.735
Log likelihood	-12194	
BIC	24922	

Note: * indicates significance at 10% level; ** indicates significance at 5% level; and *** indicates significance at 1% level.

Bibliography

Anand, Bharat N., Ron Shachar. 2009. Targeted advertising as a signal. *Quantitative Marketing and Economics* **7**(3) 237–266.

Anand, Bharat N., Ron Shachar. 2011. Advertising, the matchmaker. *The RAND Journal of Economics* **42**(2) 205–245.

Ansari, Asim, Carl F. Mela. 2003. E-customization. *Journal of Marketing Research* **40**(2) 131–145.

Arcidiacono, Peter, Patrick Bayer, Jason R. Blevins, Paul B. Ellickson. 2013. Estimation of dynamic discrete choice models in continuous time with an application to retail competition. *NBER Working Paper* .

Becker, Gary S., Kevin M. Murphy. 1993. A simple theory of advertising as a good or bad. *The Quarterly Journal of Economics* **108**(4) 941–964.

Bronnenberg, Bart J., Jean-Pierre Dubé, Carl F. Mela. 2010. Do digital video recorders influence sales? *Journal of Marketing Research* **47**(6) 998–1010.

Clark, C Robert, Ulrich Doraszelski, Michaela Draganska. 2009. The effect of advertising on brand awareness and perceived quality: An empirical investigation using panel data. *Quantitative Marketing and Economics* **7**(2) 207–236.

Eastlack, Joseph O, Ambar G Rao. 1989. Advertising experiments at the campbell soup company. *Marketing Science* **8**(1) 57–71.

eMarketer. 2014. Advertisers blend digital and TV for well-rounded campaigns.

Esteves-Sorenson, Constanca., Fabrizio Perretti. 2012. Micro-costs: Inertia in television viewing. *The Economic Journal* **122** 867–902.

Farahat, Ayman, Michael C. Bailey. 2012. How effective is targeted advertising? *Proceedings of the 21st international conference on World Wide Web*. ACM, 111–120.

Furnham, Adrian, Barrie Gunter, Freya Richardson. 2002. Effects of product-program congruity and viewer involvement on memory for televised advertisements. *Journal of Applied Social Psychology* **32**(1) 124–141.

Furnham, Adrian, Marie-Therese Price. 2006. Memory for televised advertisements as a function of program context, viewer-involvement, and gender. *Communications* **31**(2) 155–172.

Gal-Or, Esther, Mordechai Gal-Or, Jerrold H. May, William E. Spangler. 2006. Targeted advertising strategies on television. *Management Science* **52**(5) 713–725.

Ghosh, Bikram, Axel Stock. 2010. Advertising effectiveness, digital video recorders, and product market competition. *Marketing Science* **29**(4) 639–649.

Goettler, Ronald L. 2012. Advertising rates, audience composition, and competition in the network television industry. *Working Paper* .

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

Gustafson, Paul, S. Siddarth. 2007. Describing the dynamics of attention to tv commercials: A hierarchical bayes analysis of the time to zap an ad. *Journal of Applied Statistics* **34**(5-6) 585–609.

Homer, Pamela M, Sun-Gil Yoon. 1992. Message framing and the interrelationships among ad-based feelings, affect, and cognition. *Journal of Advertising* **21**(1) 19–33.

Hu, Ye, Leonard M Lodish, Abba M Krieger. 2007. An analysis of real world tv advertising tests: A 15-year update. *Journal of Advertising Research* **47**(3) 341–353.

Kim, Jun B., Paulo Albuquerque, Bart J. Bronnenberg. 2010. Online demand under limited consumer search. *Marketing Science* **29**(6) 1001–1023.

Kitts, Brendan, Liang Wei, Dyng Au, Stefanie Zlomek, Ryan Brooks, Brian Burdick. 2010. Targeting television audiences using demographic similarity. *2010 IEEE International Conference on Data Mining Workshops (ICDMW)*. IEEE, 1391–1399.

Lee, Mei-Ling Ting, G. A. Whitmore. 2006. Threshold regression for survival analysis: Modeling event times by a stochastic process reaching a boundary. *Statistical Science* **21**(4) 501–513.

Lehmann, Donald R. 1971. Television show preference: Application of a choice model. *Journal of Marketing Research* **8**(1) 47–55.

Lodish, Leonard M, Magid Abraham, Stuart Kalmenson, Jeanne Livelsberger, Beth Lubetkin, Bruce Richardson, Mary Ellen Stevens. 1995. How tv advertising works: A meta-analysis of 389 real world split cable tv advertising experiments. *Journal of Marketing Research* 125–139.

Lovett, Mitchell, Michael Peress. 2014. Targeting political advertising on television. *Working Paper* .

Malaviya, Prashant, Jolita Kisielius, Brian Sternthal. 1996. The effect of type of elaboration on advertisement processing and judgment. *Journal of Marketing Research* **33**(4) 410–421.

Mela, Carl F., Sunil Gupta, Donald R. Lehmann. 1997. The long-term impact of promotion and advertising on consumer brand choice. *Journal of Marketing Research* **34**(2) 248–261.

Moore, Robert S., Claire Allison Stammerjohan, Robin A. Coulter. 2005. Banner advertiser-web site context congruity and color effects on attention and attitudes. *Journal of Advertising* **34**(2) 71–84.

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

Narayanan, Sridhar, Puneet Manchanda. 2009. Heterogeneous learning and the targeting of marketing communication for new products. *Marketing Science* **28**(3) 424–441.

Nevskaya, Yulia, Paulo Albuquerque. 2013. A continuous time model of product usage: Measuring the effect of product design and rewards in online games. *Working Paper* .

Nielsen. 2013. Advertising & audiences: State of the media.

Nielsen. 2014. More of what we want: The cross-platform report.

Norris, Claire E., Andrew M. Colman. 1993. Context effects on memory for television advertisements. *Social Behavior and Personality: an international journal* **21** 279–296.

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

Rust, Roland T., Wagner A. Kamakura, Mark I. Alpert. 1992. Viewer preference segmentation and viewing choice models for network television. *Journal of Advertising* **21**(1) 1–18.

Schweidel, David A., Robert J. Kent. 2010. Predictors of the gap between program and commercial audiences: An investigation using live tuning data. *Journal of Marketing* **74**(3) 18–33.

Sethuraman, Raj, Gerard J Tellis, Richard A Briesch. 2011. How well does advertising work? generalizations from meta-analysis of brand advertising elasticities. *Journal of Marketing Research* **48**(3) 457–471.

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

Siddarth, S., Amitava Chattopadhyay. 1998. To zap or not to zap: A study of the determinants of channel switching during commercials. *Marketing Science* **17**(2) 124–138.

Teixeira, Thales S., Michel Wedel, Rik Pieters. 2010. Moment-to-moment optimal branding in tv commercials: Preventing avoidance by pulsing. *Marketing Science* **29**(5) 783–804.

Tellis, Gerard J. 1988. Advertising exposure, loyalty, and brand purchase: a two-stage model of choice. *Journal of Marketing Research* **25** 134–144.

Tellis, Gerard J, Rajesh K Chandy, Deborah MacInnis, Pattana Thaivanich. 2005. Modeling the microeffects of television advertising: Which ad works, when, where, for how long, and why? *Marketing Science* **24**(3) 351–366.

Tuchman, Anna, Harikesh S Nair, Pedro Gardete. 2015. Complementarities in consumption and the consumer demand for advertising. *Working Paper* .

Vakratsas, Demetrios, Tim Ambler. 1999. How advertising works: what do we really know? *The Journal of Marketing* **63** 26–43.

Weitzman, Martin L. 1979. Optimal search for the best alternative. *Econometrica* **47**(3) 641–654.

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

Wilbur, Kenneth C, Linli Xu, David Kempe. 2013. Correcting audience externalities in television advertising. *Marketing Science* **32**(6) 892–912.

Yang, Sha, Vishal Narayan, Henry Assael. 2006. Estimating the interdependence of television program viewership between spouses: A bayesian simultaneous equation model. *Marketing Science* **25**(4) 336–349.

Yao, Song, Wenbo Wang, Yuxin Chen. 2015. Channel search and welfare implications of commercial breaks. *Working Paper* .

Biography

Yiting Deng was born in Taiyuan, China. She earned her B.A. in Economics and B.S. in Statistics from Peking University in 2006, and her M.A. in Economics from Peking University in 2009. In 2009, she began her Ph.D. study at the Fuqua School of Business, Duke University. She received her en-route M.S. in Statistics from Duke University in 2012. During the Ph.D. program, she was nominated to be Duke's attendee at the AMA-Sheth Foundation Doctoral Consortium in 2013.