Modeling the Effects of Advertisement-Avoidance Technology on Advertisement-Supported Media

Kenneth C. Wilbur
Ph.D. Student
Department of Economics
University of Virginia
P.O. Box 400182
Charlottesville, Virginia 22904-4182
Phone: 434-466-4853
Fax: 434-982-2904
Email: Wilbur@Virginia.edu
Homepage: www.People.Virginia.edu/~kcw2w

October 2004
Preliminary and Incomplete
Comments Welcome
Abstract

Technology occasionally changes media consumers’ control over the amount of advertising to which they are exposed. Advertisers and media firms will benefit from an empirical methodology to forecast the degree to which such technologies will affect their business.

This paper introduces such a methodology. I develop a game-theoretic model of advertiser-supported media markets: media platforms (e.g., television networks) compete both for media consumers (viewers) and for advertisers. The profit-maximizing network sets its ad level to trade off marginal advertising revenue with the value of its last ad’s marginal audience loss.

I estimate the model’s structural parameters, using television advertisement and audience data from four sources. I then re-specify the model to account for the effects of advertisement-avoidance technology, and solve the re-specified model to make predictions about new equilibrium advertisement prices, ad quantities, and audience sizes.

Preliminary results indicate that viewers are averse to advertisements, and that cost per thousand viewers decreases with the number of ads sold, and increases with the size of the audience.

Key Words: Advertising; Broadcasting; Demand Forecasting; Endogeneity; Structural Modeling; Television; Two-Sided Markets; Technology
"The television business cannot exist, unless consumers are willing to give time for marketers." –Jamie Kellner, Chairman/CEO, Turner Broadcasting System

“People like commercials.” –Stanley Hubbard, CEO, Hubbard Broadcasting, responding to the threat of digital video recorders

1. Introduction

Once again, advertisement-avoidance technology threatens commercial television. A new device, the Digital Video Recorder (DVR), enables viewers to zip through commercials more easily than ever. In the extreme, theory predicts that if every viewer had a DVR and used it to skip ads, television advertisements would be worthless, and advertisers would be deprived of their most dependable means of reaching large audiences.

Have we heard this tale before? Similar arguments were used to predict the death of commercial television after the introduction of the remote control and the videocassette recorder (VCR). Yet, despite ever-diminishing audiences, US broadcast networks’ advertising revenues rose from $9.96 billion in 1990 to $15 billion in 2002. And while DVRs will increase commercial avoidance, they are also likely to increase total television viewing.

It is clear that whether, and how much, DVR proliferation harms television networks and advertisers is an empirical question. In this paper, I propose a method to answer such a question. I develop a structural model of a two-sided industry, in which media platforms compete both for media consumers and advertisers. I then show how to estimate the model’s parameters and forecast the effects of commercial-avoidance technology. To illustrate the

1 Quoted by CNN.com, December 5, 2002.
2 Quoted by Broadcasting & Cable, April 21, 2003, page 20.
3 A.k.a. TiVo, a.k.a. Personal Video Recorder (PVR).
4 “Zipping” is an industry term for fast-forwarding past an ad. “Zapping” is changing channels to avoid an ad.
5 Source: TVB.org; figures exclude WB and UPN. Revenue figures are nominal, but growth has outpaced inflation.
method, I use viewing and advertising data for US broadcast television networks. (Throughout
the paper, I will refer to media platforms as “networks” and media consumers as “viewers”).

To model television viewers, I continue the discrete-choice tradition of the literature (e.g.,
Shachar and Emerson, 2000; Rust, Kamakura, and Alpert, 1992) by using a mixed logit model.
The advantages of this approach are its detailed modeling of unobserved viewer heterogeneity,
and its flexible, realistic substitution patterns. To model supply and demand interactions
between networks and advertisers, I develop a simultaneous-equations framework. After
estimating viewer demand for television programs and advertiser demand for television
audiences, I re-specify the model to account for the effects of advertisement-avoidance
technology. Then, given plausible beliefs about unobserved parameters, I solve the re-specified
model to predict new equilibrium ad levels, ad prices, and audience sizes.

This work extends Gupta, Jain, and Sawhney (1999), who introduced a method to model
the evolution of markets with indirect network externalities, by introducing a method to model
the evolution of two-sided markets. To the best of my knowledge, this is the first paper to model
the effects of a commercial-avoidance technology on an advertisement-supported media
industry. In the application of the method, I contribute new empirical results to the literature on
audience forecasting, and advance our understanding of the market for television advertisements.

The use of a structural model provides a sound basis for predictions of the effects of
advertisement-avoidance technology. Advertising levels, ad prices, and audience sizes are

---

6 Internet websites are also affected by ad-avoidance technology; for example, the Google toolbar for Internet
Explorer blocks most “pop-up” ads. One might suspect that any advertiser-supported, digital communications
medium could be affected by commercial-avoidance technology at some point.

7 Advertisement quantity, ad level, and clutter are synonyms. In this context, all refer to the amount of time the
network gives to commercials during a program. (In other contexts, clutter can be used to refer to the number of
commercials, as opposed to total commercial time, during a program).
outcomes of market processes, and the modeling of these processes is what enables me to make predictions about how advertisement-avoidance technology will change agents’ behavior.\(^8\)

In the next section, I review DVR technology, the television industry, and some related academic literature. I then describe the model of viewer utility, advertiser demand, and network supply of advertisements. Section four discusses the data and estimation strategy, and section five will report the empirical results (when they are finished). In section six, I show how to re-specify and solve the estimated model to account for ad-avoidance. The paper concludes with a summary, some implications for managers, and directions for future research.

2. Background

In this section, I describe DVR technology, how DVR proliferation affects television industry economics, recent trends in advertising clutter, the market for TV advertisements, and the relevant academic literature.

2.1. What is a DVR?

It is common to hear anecdotes of DVR owners who claim the device has changed their lives,\(^9\) while some non-owners seem unimpressed by its features. A DVR is basically a VCR with a hard drive—meaning that many of its functions have been widely available for years—but some subtle advantages seem to add up to a big difference.

---

\(^8\) An alternative to the use of a structural model would be to use surveys of current DVR users to predict the actions of future DVR users. For example, one recent survey found that 69% of DVR owners say they fast-forward through 80% or more of commercials when watching recorded programs. (Greenspan, 2004) The problem with this alternative is that these surveys are obviously biased by sample selection: current DVR owners are “early adopters,” and as such, presumably different from the majority of the viewing population.

\(^9\) As one TiVo owner was quoted in the New York Times: “Before we got the TiVo, my son was getting C's and D's in school because he was staying up late to watch his shows and going to school half-awake”...[Now he is getting enough sleep and his grades have risen to A's and B's]...“We watch TV together as a family after dinner... And my son even has enough time to get a job. So it's improved his sense of the value of time. And it's improved my relationship with him.” (March 18, 2004) 97% of TiVo owners say they would recommend TiVo to a friend. (TiVo.com)
Unlike a VCR, a DVR is “always on” and continuously stores up to the previous 30 minutes (or more) of TV programming. Live TV can always be paused or rewound. Many DVR users delay watching TV until after their program of interest begins, so they can fast-forward through the commercials.

The DVR has several other advantages over the VCR. Fast-forwarding is quicker: a two-minute commercial break can be skipped in as little as five seconds. DVRs store programs digitally, so their owners never have to purchase, rewind, insert, remove, or store bulky videotapes. And while VCRs are notoriously difficult to program, DVR users can record programs by selecting them from an on-screen program guide. Some DVRs even include novel program search capabilities, such as the ability to record all programming credited to a specific actor or director.

Until recently, DVRs were confined to standalone, difficult-to-install boxes, available only from consumer electronics retailers and satellite television providers. Currently, DVRs are being integrated into other consumer electronics devices, like computers\textsuperscript{10} and Digital Video Disc (DVD) players and recorders. In the near future, DVR-equipped set-top boxes will be available to most digital cable subscribers for a monthly fee.

Since the device’s 1999 introduction, DVR penetration stands at about 5% of US households, and several research firms expect it to increase rapidly. Nielsen Media Research has told clients that DVR ownership rates could reach 10% by the end of 2005, and Forrester Research recently predicted 41% penetration by 2009. If the DVR follows a diffusion curve similar to the VCR, it could be poised for explosive growth: after its 1978 introduction, VCR penetration rose from 3%, in 1982, to 58% in 1988.\textsuperscript{11}

\textsuperscript{10} Microsoft has included DVR functionality in its Windows XP Media Center Edition operating system.
\textsuperscript{11} Interestingly, VCR penetration fell for the first time ever in 2004, from 91.5% to 90.8%. (TVB.org)
2.2 Economics of Television Advertisements.

Television networks operate in a two-sided market. Networks offer viewers “free” programs in exchange for their attention, and sell that attention to advertisers. A viewer’s true cost of consuming a television program is the opportunity cost of her time spent watching commercials.

The structure of the industry suggests that the average viewer normally has better uses for her time than watching advertisements. (Otherwise, networks might have little incentive to air programs between their commercial breaks, and we might not observe so much viewer zapping of advertisements.) Thus, the television network faces a tradeoff: additional commercials may increase revenues, but will induce some viewers to leave the audience. The profit-maximizing network will sell ads until the marginal revenue of its last ad sold equals the value of the audience losses accruing to that ad.

Because viewers pay for programs by watching commercials, avoiding ads with a DVR is equivalent to a fall in price. This fall in price should produce a positive income effect (viewers will watch more TV because avoiding ads increases their total stock of free time) and a positive substitution effect (viewers will watch more TV, because it takes less time than before, relative to other activities).

The effect of DVR proliferation on clutter is less clear. Increasing scarcity of non-ad-avoiding viewers might lead networks to compete more fiercely for audience share by lowering their ad levels. On the other hand, some DVR users will fast-forward past ads, rather than

---

12 Also called a multi-sided platform industry, a two-sided market is one in which two groups of participants interact via competitive intermediaries. Examples of two-sided markets include software (operating systems coordinate users and software developers), payment mechanisms (credit card companies coordinate retailers and consumers), and academic journals (journals coordinate authors and readers). For a general treatment of two-sided markets, see Rochet and Tirole (2001) or Armstrong (2004).
leaving the audience; and smaller audience losses mean smaller disincentives for additional ads.

It is not clear, *ex ante*, which effect will dominate.

2.3. Advertising Clutter: Current Data and Recent History

A common misconception holds that advertisement quantities are relatively constant across networks and time periods. In truth, there is substantial variation: in my sample, the average difference between the maximum and minimum network ad levels, within a prime-time half-hour, is 2:49 minutes.¹³ (For comparison, the mean ad level is 5:15 minutes.) A similar pattern emerges from data on total non-program minutes reported in the 2001 Television Commercial Monitoring Report (TCMR).¹⁴ Programs that are more attractive to viewers, relative to within-time-period competition, typically contain more ads.

Ad levels once were fixed. Until 1982, broadcasters limited themselves to six minutes of non-program content per prime-time hour. Though tacitly supported by the Federal Communications Commission, this practice was discontinued after the US Department of Justice brought an antitrust suit against the industry. (Owen and Wildman, 1992) While the US does not regulate prime-time advertising levels, many other nations do, including the UK, France, Germany, and Japan. (Motta and Polo, 1997)

After the antitrust ruling, network advertising levels rose slowly, from an average of 3:24 minutes of national ads per half-hour in 1982, to 3:39 in 1988 (Owen and Wildman, 1992). As cable competitors eroded broadcast network audiences during the 1990’s, ad levels rose more rapidly: average commercial time (including local ads) climbed from 4:51 minutes per half-hour in 1991 to 6:02 minutes in 2001. (TCMR, 2001)

2.4. The Market for Television Advertisements

---

¹³ Prime time is defined as 8:00pm-11:00pm EST, Mon-Saturday, and 7:00pm-11:00pm EST, Sunday.
¹⁴ Non-program minutes include ads, “promos” for network programs, and public service announcements.
The market for television advertisements is characterized by imperfect information, price discrimination, and bundling of program audiences.

Networks present their new and returning programs to advertisers in May of each year. During the next few weeks, advertisers review scripts and watch clips in an effort to predict new shows’ audiences, while networks gather data on advertisers’ budgets. Each network then produces a “rate card,” its starting point for price negotiations with advertisers; negotiations typically last about four weeks. The market usually “breaks” in early summer, with networks selling about 70-80% of the coming year’s commercial time within a few weeks. This period is known as the upfront market.

Ads unsold during the upfront may be sold any time before the broadcast airs, on the scatter market. Networks guarantee minimum audiences for ads sold during the upfront, while ad prices in the scatter market are more closely linked to the audience delivered.

Networks engage in price discrimination, charging higher prices to more-profitable advertisers, and offering volume discounts. Audiences are typically bundled across days, time periods, and quarters. Ads are purchased in 10-, 15-, 30-, or 60-second increments.

2.5. Relevant Literature

This work is related to a very large literature in advertising, economics, and marketing, but its closest cousin is Gupta, Jain, and Sawhney (1999). Gupta, et al, show how to model the evolution of, and product diffusion in, markets related by indirect network externalities (INE). They apply the method to the digital television industry: digital television purchases affect the

---

15 When a program fails to deliver its minimum guaranteed audience, the network compensates its advertisers with make-goods, free commercials during other programs. (Unfortunately, I do not observe which ads in my sample are make-goods.)
16 Networks nearly always bundle larger audiences with smaller, a practice known as packaging.
17 There are many interesting issues to be investigated in this market, but data limitations will force me to assume networks set a uniform price per show.
18 INE exist when changes in demand or supply of one product affect demand or supply of a related product.
number of programs provided in digital formats, and program availability drives demand for
digital televisions. The authors show that ignoring INE biases predictions of future television set
sales.

Gupta, et al, introduce a tractable and useful model, but assume television manufacturers’
decisions are exogenously determined. I advance the literature by modeling entities that compete
in multiple markets, and partially internalize indirect network externalities between these
markets.19

This work contains the first empirical, microeconomic study of the market for television
advertisements in which advertising levels are endogenously determined. Two related papers are
Goettler (1999) and Kieshnick, McCullough, and Wildman (2003); both estimate the relationship
between audience size and advertisement price while, for lack of data, holding ad quantities
constant. Goettler finds a convex relationship between ad price and audience size, and shows
that audience demographics are important determinants of ad price; Kieshnick et al. show that
networks charge different implicit prices for different types of viewers. A third paper, Silk,
Klein, and Berndt (2002) uses a translog model to estimate own- and cross-price elasticities for
eight advertising media, based on time-series data. They report that aggregate demand for
advertising for most media, including network television, is price-inelastic.

Another related literature is the set of papers that examines the causes and effects of
viewers’ zapping of television commercials, and its relation to brand purchase behavior.
Members of this set include Olney, Holbrook and Batra (1991), Danaher (1995), Siddarth and

19 The difference between a two-sided market and two markets related by INE is the existence of an entity that
partially internalizes the INE that relate the two markets. The two sides of a two-sided market are necessarily
markets related by INE, but two markets related by INE do not necessarily constitute a two-sided market.
While advertisers and marketers have long understood the two-sided nature of the television industry, economic models of two-sided markets have only recently begun to appear in the literature. Two such models are presented by Anderson and Coate (forthcoming) and Dukes and Gal-Or (2003). In both papers, television networks coordinate advertisers and viewers. The former paper analyzes welfare in the industry, finding, notably, that monopoly channel ownership may result in greater welfare than competitive ownership, and that welfare can lessened when networks have the ability to charge their viewers. Dukes and Gal-Or (2003) also model interactions between advertisers and viewers in the product market, and show that, when increased advertising leads to better-informed consumers and greater product-market competition, networks can increase their profits by selling exclusive advertising rights.

3. A Model of the Television Industry

This section describes the model of viewer utility, advertiser demand, and network supply of television commercials.

3.1. Viewers

A long line of papers in the marketing literature estimates viewer demand for television programs, including (but not limited to) Gensch and Shaman (1980), Rust and Alpert (1984), Rust and Eechambadi (1989), Rust, Kamakura and Alpert (1992), Tavakoli and Cave (1996), Shachar and Emerson (2000), Danaher and Mawhinney (2001), Goettler and Shachar (2001), and Kim (2002). A television viewer typically watches one program at a time; reflecting this, most researchers use discrete choice models to estimate viewer demand. I follow the tradition of the literature by using a mixed logit to model television viewers.
Each viewer $i$, in city $m$, is assumed, at each time $t$, to either watch one of $J$ “inside”
television networks (which are indexed by $j$), to watch some other television channel, or to
engage in a non-television pursuit, like sleeping or reading. Let viewer $i$’s indirect utility from
watching network $j$ at time $t$ be

$$u_{ijt} = q_{jt} \alpha_i^* + [x_{jt}, c_{ij,t-1}, c_{ij,t+1}] \beta_i^* + \xi_{jt} + \eta_{mjt} + \epsilon_{ijt},$$

(1)

where $q_{jt}$ is the number of seconds of advertising on network $j$ during time period $t$, $x_{jt}$ is a
vector of the observable characteristics of the show on network $j$ at time $t$, $c_{ij,t-1}$ and $c_{ij,t+1}$ indicate
whether the viewer watched network $j$ in the preceding or subsequent time periods, $\alpha_i^*$ and $\beta_i^*$
are viewer $i$’s taste parameters (defined below), $\xi_{jt}$ is a program-specific fixed effect that
captures mean tastes for unobserved show characteristics, $\eta_{mjt}$ measures the market-specific
deviation from mean tastes for unobserved show characteristics, and $\epsilon_{ijt}$ is viewer $i$’s
idiosyncratic taste for network $j$’s time-$t$ program.

To define viewer $i$’s taste parameters, let

$$\begin{pmatrix} \alpha_i^* \\ \beta_i^* \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi D_{jm} + \Sigma v_i, \quad D_{jm} \sim P_{D_{jm}}^*(D), \quad v_i \sim P^*(v),$$

(2)

where $\alpha$ and $\beta$ are mean tastes for ad level and show characteristics; $D_{jm}$ is a vector of
observed viewer demographic characteristics (income, gender, age, and age$^2$), and $P_{D_{jm}}^*(D)$ is the
market-specific joint distribution of viewer demographics; and $v_i$ is a vector of unobserved
demographics. $\Pi$ is a parameter matrix that measures how tastes for program characteristics

---

20 Numerous studies, including Goettler and Shachar (2001) and Shachar and Emerson (2000), present evidence of
state dependence in television viewing choices. While I do not observe individual data, I do observe channel $j$’s
audience size in market $m$ at times $t-1$ and $t+1$, and will use these data to control for audience flow effects.

21 $\eta_{mjt}$ captures the effect of regional differences in speech and culture on viewer preferences for unobserved show characteristics.
vary with observed viewer demographics, and $\Sigma$ is a diagonal matrix that measures the importance of unobserved viewer heterogeneity in viewing decisions.

If viewer $i$ watches a non-network channel (option “non”), her utility is given by

$$u_{i,non,t} = \xi_{non,t} + \eta_{m,non,t} + \pi_{non}D_{im} + \sigma_{non}v_i + \xi_{i,non,t},$$

where $\xi_{non,t}$ is the mean value of watching the best available “outside” network at time $t$, and $\eta_{m,non,t}$ is the market-specific deviation from that mean.

The indirect utility of the non-television option is

$$u_{i,0t} = \xi_{0t} + \eta_{m0t} + \pi_{0}D_{jm} + \sigma_{0}v_i + \xi_{i0t},$$

where $\xi_{0t}$ and $\eta_{m0t}$ are normalized to zero (and the $\xi_{jt}$’s, $\xi_{non,t}$’s, $\eta_{mjit}$’s, and $\eta_{m,non,t}$’s are identified relative to this normalization), and $\pi_{0}D_{jm} + \sigma_{0}v_i$ is essentially a fixed effect that measures the time-invariant component of viewer $i$’s value of the outside option.

If viewers know network program characteristics and ad levels,\(^{22}\) and act to maximize utility, then the set of demographics and preference shocks that leads viewer $i$ in city $m$ to watch network $j$ at time $t$ is

$$A_{mjt} = \{(D_{jm}, v_i, \xi_{jt}) | u_{ijt} > u_{ikt}, \forall k \neq j \},$$

---

\(^{22}\) Though common in the literature, the assumption that viewers know what is on TV is obviously unrealistic, and especially undesirable in the context of advertising levels. Moshkin and Shachar (2002) present evidence that viewers face search costs when switching channels, and Anand and Shachar (2004) show that audience flow can be explained with imperfect viewer information sets. Both papers use panel data on viewing choices and Bayesian techniques to update viewers’ beliefs about program choices. Because I do not observe individual viewing data, I follow Rust, Kamakura, and Alpert (1992) and Shachar and Emerson (2000) in maintaining the traditional assumption of perfect information.

There are two related justifications for this assumption. The first is that each ad presents a viewer with an opportunity, or a stimulus, to switch channels (and not switch back), so more ads decrease the likelihood a viewer will watch a given channel at a given time. The second is motivated by the observation that most television viewers are experts, having consumed thousands of hours of network programs. While viewers might not be consciously aware of networks’ offerings, they may have been conditioned over a long period of time to anticipate networks’ programs and ad levels fairly well.
where $\varepsilon_{ijt}$ is a vector of the $\varepsilon_{ijt}$'s. If the idiosyncratic error terms are distributed $i.i.d.$, the viewing share of network $j$ in market $m$ at time $t$ is given by

$$s_{mjt} = \int dP^*_x(\varepsilon)P^*_v(\nu)dP^*_D(D).$$

Equation (3) will be used to estimate viewer demand for television programs.

### 3.2. Advertisers and Networks

Two aspects of the television advertising industry complicate the modeling of advertiser preferences. The first is the combinatorial nature of television audiences: to achieve an advertiser’s audience reach and frequency goals, there will usually be several—and possibly a large number—combinations of audiences available. The second difficulty comes from the complex nature of substitute/complement relationships among program audiences. When an advertiser’s benefit from advertising is nonlinear in advertising exposures, it is difficult to predict which programs are complements, and which are substitutes, even for a single advertiser.

These are interesting and difficult problems, but to deal with them effectively, the researcher requires knowledge of each advertiser’s audience purchases. Such data can be difficult to obtain, so it is useful to think about how to estimate demand without them. To this end, I posit a simultaneous equations framework for advertisement supply and demand.

---

23 I make the common assumptions that the $\varepsilon_{ijt}$'s are distributed $i.i.d.$ Type I Extreme Value, and that each element of $\nu_j$ is distributed standard normal.

24 There are several reasons to expect such nonlinearity, including consumption externalities, increasing or decreasing marginal costs, or advertising effectiveness that varies with viewer exposures to the ad.

25 To illustrate this complexity, consider a firm whose production technology exhibits increasing returns to scale, and assume the firm buys several advertisements. Presumably, any unpurchased ads are substitutes for those that were purchased. But the second advertisement purchased is likely to complement the first: the additional sales generated by the second ad lead to lower marginal production costs, hence raising the firm’s benefit from purchasing the first ad.

26 I am also working on estimating a fully microfounded model of advertiser preferences that deals with the problems described above. For a description of this model, please see Wilbur (2003).
Industry practice among advertisers is to base audience purchase decisions on available programs’ costs per thousand viewers (CPM). This leads me to assume that aggregate demand for any program is given by the product of its audience size and the market’s valuation of the average viewer in its audience. Thus

\[ p_s = V_s f_s , \]  

(4)

where \( p_s \) is the price of an ad during show \( s \); \( V_s \) is the number of viewers watching show \( s \) (which depends on the show’s ad level, \( q_s \), and the ad quantities of its within-time-period competition for viewers, as predicted by the viewer demand model); and \( f_s \) is the price of reaching the average viewer in audience \( s \). I further assume that per-viewer ad price is linear in program and audience characteristics,

\[ f_s = [q_s, V_s, d_s, x_s] \lambda + \phi_s , \]  

(5)

where \( d_s \) is a vector of demographic characteristics of the average member of audience \( s \); \( x_s \) are show characteristics that affect advertising effectiveness or correlate with unobserved audience characteristics; \( \lambda \) is a vector of advertiser preference parameters; and \( \phi_s \) represents advertisers’ valuation of unobserved audience characteristics. Audience size is included in \( f_s \) to account for the “large-audience premium,” the notion that larger audiences contain harder-to-reach viewers.\(^{27}\)

Next, I turn to assumptions about network supply of television commercials. I assume networks compete in three stages. In the first stage, each network chooses its programs; in the second stage, networks schedule their programs; and in the third stage, networks sell

\[ \text{It should be noted that the “large-audience premium” is implied by the structure of the model of viewer demand. High-quality shows will attract greater numbers of viewers with high values of the outside option; these are the same viewers who are typically thought of as being harder to reach, because they watch less television than average.} \]
advertisements. In this paper, I focus on competition in the final stage, in which program costs are sunk, and program schedules have been finalized. Stage-three profits are given by

$$\pi_j = \max \sum_{s \in S_j} q_s (p_s - mc_s),$$

(6)

where $S_j$ is network $j$’s catalogue of shows, and $mc_s$ is the marginal cost of selling an ad during show $s$.

Before deriving the network’s first-order condition, it is necessary to consider whether it has market power. There are at least two reasons to believe it does: governmental spectrum regulations and cable system capacity constraints limit entry of new networks, and advertisers have few good substitutes for large television audiences. I proceed with the assumption that networks act as ad-quantity-setting oligopolists, and use the network’s first-order condition to define its advertisement supply function. Substituting ad demand into equation (6) and differentiating with respect to $q_s$ yields

$$p_s + q_s \left[ V_s \frac{\partial f_s}{\partial q_s} + \frac{\partial V_s}{\partial q_s} f_s \right] = mc_s.$$  

(7)

The first term in equation (7) is the network’s ad price, its incentive to sell commercials; the second term is the decrease in its ad price resulting from its final commercial sale ($\frac{\partial f_s}{\partial q_s}$ is expected to be negative); and the third term is the value of the audience loss accruing to the last ad sold ($\frac{\partial V_s}{\partial q_s}$ is expected to be negative; this audience loss will be predicted by the model of

---

28 Each stage corresponds clearly to reality. The first stage takes place before the upfront, when networks renew returning shows, and buy new ones; the second stage takes place when networks announce their program schedules at the start of the upfront, and the third stage occurs during the remainder of the upfront and scatter markets.

29 The ad-quantity-setting assumption simplifies the model, relative to an ad-price-setting alternative: if networks set ad prices, then the model would have to account for feedback effects between the advertiser demand and viewer demand. The observation that networks charge different prices to different advertisers lends credence to the ad-quantity-setting assumption.
viewer demand described in the section 3.1.) Equations (4) and (7) form the advertisement supply-and-demand framework used to estimate advertiser and network parameters.

4. Data and Estimation

This section describes the components of the dataset, the estimation routines, and the instruments used to estimate viewer and advertiser demand parameters.

4.1. Data

The model is estimated using data on television programs, advertisements, and audience characteristics from four sources. The sample includes all shows aired on the six most-watched US broadcast television networks,30 between 8:00 p.m. and 10:00 p.m., Monday to Friday, April 24 to May 21, 2003. I exclude two nights in which live broadcasts took place, since these programs curtail networks’ ability to set their ad levels.31

Audience share data come from Nielsen Media Research (NMR) “Viewers in Profile” reports covering the 50 largest US Designated Market Areas; more than 90% of US television households are included in the sample. Audience estimates are measured in households, which NMR collected in each market with a sample of “audimeters,” set-top boxes that record viewing choices and transmit data to Nielsen via telephone lines. The data report audiences by half-hour, day, and channel.32,33

---

30 The networks in the sample are ABC, CBS, FOX, NBC, UPN, and WB. All networks aired national programs during all time periods in the sample.
31 There were two programs broadcast live during the sample: a presidential address, on Thursday, May 1, from 9:00-9:30 p.m.; and a NBA basketball game, on Thursday, May 15, from 8:00-10:00 p.m.
32 If a household watched a program for five not-necessarily-consecutive minutes during a 15-minute period, Nielsen includes that household in the program’s audience. In spite of this definition of “viewing,” Nielsen data are the industry standard, as it produces the best (and, usually, only) audience measurements available.
33 Nielsen excluded DVR households from its sample, classifying them as “technically difficult.” In mid-2003, DVR penetration was estimated at about 2% of US households. (Yankee Group, 2003) Nielsen will start including DVR households in its audience samples in 2005. (MediaDailyNews, 3/4/2004)
The Nielsen audience data were only available in printed reports, so I used a scanner to transfer the data into a machine-readable format. This process resulted in numerous errors, which I corrected both by hand and with computer programs. Before starting the estimation, I manually checked all of the data against the original source to reduce the likelihood of error induced by the scanning process.

Audience demographic data were collected from US Census 1% IPUMS files for each Consolidated Metropolitan Statistical Area corresponding to a geographic television market.

Television advertisement data were purchased from TNS Media Intelligence/CMR. For each program in the sample, I observe an estimated ad price, and the start time and duration of each national advertisement during the show. Price data was collected by TNSMI/CMR as the network-reported “estimated cost of a 30-second commercial” during the show.

Table 1 reports the means and standard deviations of ad price, ad quantity, cost per thousand viewers, and audience size. Some interesting results emerge. ABC aired the most national advertising during the sample, averaging 6:07 minutes of national ads per half-hour, while FOX aired the least, 4:44 minutes per half-hour. The highest average nightly CPM in the sample was $37, charged by FOX for its Tuesday-night lineup (which included American Idol, the highest-rated program in the sample); the second-highest CPM was NBC’s Thursday-night price of $35 per thousand viewers. FOX and NBC consistently received the highest CPM ($28), while UPN charged the least ($19). The networks with the largest average audiences were FOX, CBS, and NBC; ABC was fourth, and WB came fifth.

Program characteristics were recorded from videotapes of network programming made during the sample period. Observable program characteristics include genre, thematic elements,
main and supporting characters’ demographics\textsuperscript{34} (including gender, race, age, and family structure), program age,\textsuperscript{35} setting, and current and past Emmy nominations. The videotapes were supplemented with data from internet websites like \textit{tvtome.com}. Table 2 lists the program characteristics, their definitions, and descriptive statistics.

4.2. Viewer Demand Estimation

I estimate viewer demand parameters using the error structure and estimation algorithm introduced by Berry, Levinsohn, and Pakes (1995; hereafter, “BLP”). The main idea of this estimation routine is to numerically solve the market share functions defined by equation (3) for programs’ mean utility levels, and to use these imputed mean utilities in a moment condition. Its principal benefits are to reduce simulation error, to speed computation by reducing the number of parameters to be estimated nonlinearly, and to allow for gradient-based numerical optimization by defining the objective function as a smooth function of the parameters. The innovation of the BLP estimation algorithm is that, rather than defining the error term as the difference between predicted and observed market shares, it defines the error as the difference between the mean utility predicted by the model and the mean utility implied by the assumed equality of the predicted and observed market shares. The complication is that, for each guess of the parameters, implied mean utility levels must be computed to construct the error.

While I do not observe individual viewing choices (and hence do not observe $c_{ij,t-1}$ or $c_{ij,t+1}$), I do see each channel’s audience size in each market in neighboring time periods. These audience sizes should correlate with the probability that $c_{ij,t-1} = 1$ or $c_{ij,t+1} = 1$ for any individual viewer in market $m$. I therefore define $x_{mijt}$ as a vector of networks’ local-market audience sizes

\textsuperscript{34} A main character is defined as a character on whom major plotlines are based.
\textsuperscript{35} For those movies that were not originally produced for TV, age is defined as the number of years since the movie’s theatrical opening.
in neighboring time periods, and include it in viewer i’s utility from option j at time t in place of $c_{ij,t−1}$ and $c_{ij,t+1}$. To facilitate the presentation of the estimation strategy, I partition the program characteristics parameter vector $\beta$ into $\beta_1$, which interacts with program characteristics $x_{j,t}$, and $\beta_2$, which interacts with $x_{m,j,t}$.

One of the challenges in any empirical study of market demand is dealing with the correlation between endogenous regressors and consumer tastes for unobserved product characteristics. That correlation arises here because television programs have some quality components that are either not observed, or not quantified, by the econometrician (e.g., sex appeal), but are known by the network; and the network takes these quality components into account when it sets its advertisement level. Common strategies for dealing with this endogeneity include the use of instrumental variables and/or brand-specific fixed effects. As explained by Nevo (2000), the latter treatment is preferred, though it is not always feasible.

Two characteristics of this study enable me to employ program-level fixed effects to control for the correlation of ad level with unobserved program characteristics. The first is the structure of the industry: network affiliates are contractually bound to air the network’s programs and advertisements without modification, so program characteristics and ad levels do not vary from city to city within a given time period. The second is the richness of the data. For each network/time period in the sample, I observe audience shares in 50 metropolitan areas, so I have enough observations to identify the program-level fixed effects.

To facilitate explanation of the estimation routine, let $N$ be the number of city/day/time/channel market shares observed in the sample (including market shares for the non-broadcast-network television option), and let $P$ be the number of programs in the sample. Let $H$ be a $NxP$ matrix of program-level dummy variables, and let $\psi$ be the $Px1$ vector of mean
utilities that are constant across viewers and time periods in which the program airs. Finally, I denote the parameter set \( \theta = \{\psi, \alpha, \beta_2, \Pi, \Sigma\} \), and I define the mean utility that viewers in city \( m \) derive from watching network \( j \) at time \( t \) as
\[
\delta_{mjt} = \psi_p + q_j \alpha + x_{mjt} \beta_2 + \eta_{mjt},
\]
where \( \psi_p \) is the mean utility of the program on network \( j \) at time \( t \).

The integral that defines the predicted market share function (equation 3) has no closed form solution, so I use simulation to approximate it. For each simulated viewer \( i \) in market \( m \), I draw \( D_{i|m} \) from the joint distribution of viewer demographics defined by the IPUMS data; I draw \( \nu_{i} \) from a standard multivariate normal distribution; and I integrate over the assumed distribution of the \( \epsilon_{ijt} \)'s in the normal fashion. The predicted market share of network \( j \) at time \( t \) in market \( m \) is then the fraction of the simulated viewers in market \( m \) for whom network \( j \)'s time-\( t \) program is utility-maximizing.

I use the numerical technique suggested by BLP to solve for the \((J+1)\)-vector of mean utilities \( \tilde{\delta}_{m:t}(\theta) \) that, for a given value of \( \theta \), equate predicted market shares to observed market shares in market \( m \) at time \( t \),
\[
S_{m:t}(\tilde{\delta}_{m:t}(\theta)) = S_{m:t}.
\]
I then define the error term, the market-specific deviation from mean tastes for network \( j \)'s time-\( t \) broadcast, as
\[
\tilde{\eta}_{mjt}(\theta) = \tilde{\delta}_{mjt}(\theta) - \psi_p - q_j \alpha - x_{mjt} \beta_2.
\]

Next, I construct the moment conditions, \( EX'\tilde{\eta}(\theta) = 0 \), where \( X \) is a matrix of instruments defined below, and \( \tilde{\eta}(\theta) \) is a vector of the \( \tilde{\eta}_{mjt}(\theta) \)'s. The Generalized Method of Moments (GMM) estimate of \( \theta \) is
\[
\hat{\theta} = \arg\min_{\theta} \tilde{\eta}(\theta)' X B^{-1} X' \tilde{\eta}(\theta),
\]
where $B$ is a positive-definite weighting matrix. Any choice of $B$ will produce a consistent estimate of $\theta$, but the most efficient is a consistent estimate of the covariance matrix of the moments. I set $B = X'X$ to obtain an initial estimate of $\theta$, which I then use to construct an estimate of the asymptotically efficient weighting matrix, $E X' \tilde{\eta}(\theta) \tilde{\eta}(\theta)^T X$. I then use the new weighting matrix to obtain the final estimate of $\theta$.

After estimating the program-level mean utilities $\hat{\psi}$, I use the minimum-distance procedure suggested by Nevo (2000) to disentangle the taste parameters associated with the observed program characteristics ($\beta_i$) from the unobserved program characteristics ($\xi$). Let $X_p$ be a $P \times K$ matrix of program characteristics (where $K \leq P$), and let $\psi$ be a $P \times 1$ vector of mean program utilities. Then, since $\psi = X_p \beta_1 + \xi$, parameter estimates are given by

$$\hat{\beta}_i = \left(X_p' \hat{\Omega}^{-1} X_p\right)^{-1} X_p' \hat{\Omega}^{-1} \hat{\psi}$$

and $\hat{\xi} = \hat{\psi} - X_p \hat{\beta}_1$, where $\hat{\Omega}$ is the estimated variance-covariance matrix of $\hat{\psi}$.

It remains to discuss the contents of the $X$ matrix and parameter restrictions. All instruments included in $X$ must share the characteristic that they can be assumed to be orthogonal to the vector of market-specific deviations from mean program utility. The argument for each of the instruments will be the same: all are variables set by the networks at the national level and do not vary across markets. Thus, because all networks broadcast to a large number of markets, the effect of any single market’s deviation from mean tastes on the network’s national decision is negligible.

The obvious candidates for inclusion in $X$ are the program-level dummy variables, $H$. I also include a $N \times J$ matrix $Q$, whose $n^{th}$ row contains the associated network’s time-$t$ ad level in the first column, and the time-$t$ ad levels of each of its $J-1$ competitors in the adjoining columns.
I also include in $X$ interactions between show characteristics and moments of the market-specific distributions of demographics. For each program characteristic $k$ in $X_p$ whose effect on utility varies with consumer demographics, and for each observable viewer demographic $d$, I include $X^k_d$, a $N \times 1$-vector whose $n^{th}$ element is $x^k_p Ed_{im}$, where $Ed_{im}$ is the expected value of viewer demographic $d$ in market $m$. These interactions are equal in number to the number of parameters to be estimated in $\Sigma$. (Note that the audience flow effects, $x_{mjt}$, cannot be used as instruments here, as they are highly likely to be correlated with market-specific tastes for unobserved program characteristics.)

To reduce what could become an extremely costly nonlinear minimization, I impose further zero-restrictions on $\Sigma$ and $\Pi$ by limiting the number of observable program characteristics that interact with viewer demographics. If there are $\#d$ elements in $D_{lim}$, each observable program characteristic whose effect on utility varies with viewer demographics adds $\#d+1$ parameters to the nonlinear optimization. I therefore limit the number of regressors that interact with viewer demographics to three: the viewer’s taste for the outside option, her taste for non-broadcast-network television, and her disutility of advertisements. (I choose the first two variables to improve the reliability of predicted substitution patterns among various options, and the third variable to improve the reliability of predicted audience changes resulting from varying levels of advertisements.)

4.3. Advertiser Demand Estimation

To estimate advertisement demand and supply parameters, I interact instruments with the residuals derived from the simultaneous-equations framework laid out in section 3.2 to construct moment conditions, which I solve using GMM.
Audiences are sold and prices are reported at the program level, but networks distribute ads within a program based on strategic considerations. To help correct for this, I average all program and audience characteristics over half-hours, so the unit of observation is a network/day/program.

The demand residual is the effect of unobserved audience characteristics on ad price per viewer, and can be derived from equations (4) and (5):

$$\phi_s = \frac{p_s}{V_s} - [q_s, V_s, d_s, x_s] \lambda .$$

To construct the supply residual, I assume the marginal cost of selling an ad during show $s$ is $mc_s = \sigma + \kappa_s$, where $\sigma$ is an intercept, and $\kappa_s$ is a show-specific component of cost. From equation (7), then,

$$\kappa_s = p_s - q_s \left[ V_s \frac{\partial f}{\partial q_s} + \frac{\partial V}{\partial q_s} f_s \right] - \sigma .$$

Audience size, audience demographics, and their derivatives are endogenous, since they depend partially on ad quantity, which is determined within the model. To control for this endogeneity, I use two sets of instruments. The first is the non-ad, mean utility derived by viewers from the program, and its within-time-period competitors. (In the viewer demand notation introduced previously, the non-ad, mean utility for show $s$ is the corresponding element of $\hat{\psi}$.) These instruments are correlated with $q_{jt}$ and the endogenous regressors because they are major determinants of audience size, demographics, and their derivatives, and $q_{jt}$ depends on this latter set of variables. The non-ad, mean utilities can reasonably be assumed to be independent of the errors; $\phi_s$ represents advertisers’ valuation of the average audience member’s unobserved demographic characteristics, whereas the mean utilities capture the effects of program
characteristics on viewer utility. \(\text{namu}_s\) can also be assumed to be orthogonal to \(\kappa_s\), since this latter variable is most likely to reflect the degree of difficulty of editing a program to accommodate a changing level of commercials, which is not likely to be related to the utility that viewers derive from the program. The second set of instruments consists of the exogenous regressors, \(x_s\).

Let \(\text{namu}_s\) be a vector of non-ad, mean utilities for the programs that compete for viewers at the same time as show \(s\); let the set of instruments for show \(s\) be represented by vector \(Z_s = [x_s \text{namu}_s \text{namu}_s]\); and let \(Z\) be a matrix constructed by stacking the \(Z_s\)’s. Then the moments are as \(EZ'\phi = 0\) and \(EZ'\kappa = 0\), where \(\phi\) and \(\kappa\) are vectors of the \(\phi_s\)’s and \(\kappa_s\)’s, respectively. The initial GMM weighting matrix is defined as \(W = [Z'Z \otimes I_2]\), and the GMM estimates of advertisement demand and supply parameters are

\[
\{\hat{\lambda}, \hat{\sigma}\} = \arg\min_{\{\lambda, \sigma\}} \begin{bmatrix} Z'\phi \\ Z'\kappa \end{bmatrix} W^{-1} \begin{bmatrix} Z'\phi \\ Z'\kappa \end{bmatrix}.
\]

As with the viewer demand estimation, I use \(W\) to obtain an initial, consistent estimate of the parameters. I then compute the asymptotically efficient weighting matrix, and re-estimate the model to obtain the final results.

5. Empirical Results

I have recently estimated the model on a subsample of the data, but I have not yet settled on final specifications of the demand functions, nor finished a full analysis of the results. I present some preliminary findings concerning the more interesting parameter estimates.

5.1. Viewer Demand
The mean effect of advertising on utility is negative. The model predicts that a 30-second increase over the observed advertising level will decrease the mean audience by 84,000 households (1.4%).

I find that higher-income viewers dislike ads less than lower-income viewers, and women dislike ads less than men. Advertising disutility is convex in age, and reaches a minimum at 42, indicating that viewers in their early 40’s are the most likely to leave a program’s audience in response to additional commercials.

5.2. Advertisement Demand

I find that per-viewer ad price decreases with the number of times an audience is sold; on average, one additional 30-second ad lowers CPM by $1.80. Cost per viewer is increasing in total audience size: an extra million viewers raises a show’s CPM by $0.54.

In the preliminary demand specification, I defined $d_s$ to be a vector containing the mean income and age of audience $s$. I found that CPM is decreasing in both mean viewers’ age and income. The second result was unexpected, and could be explained by advertiser preferences that depend more on age than on income, combined with a strongly negative correlation between viewers’ age and income. I expect that a better definition of $d_s$—such as including the portion of viewers who fit into categories like “young, high-income, male,” “old, low-income, female,” etc.—will yield more intuitive findings.

CPM also varies with observed program characteristics. Comedy and Reality shows fetched the highest prices per viewer, all else equal; Action and News programs earned the least. Shows that featured married characters earned more than average, and those featuring single parents earned less. Irregularly scheduled programs earned less per viewer, as did younger
shows, shows with fewer past Emmy nominations, and shows with minority cast representation exceeding 50%.

6. Predicting the Effects of Advertisement-Avoidance Technology

In this section, I show how to re-specify and solve the model to account for the effects of ad-avoidance technology on viewer utility and advertiser demand. When finished, I will report new predicted equilibrium ad quantities, ad prices, and audience sizes, given plausible assumptions about unobserved parameter values.

6.1. Model Respecification

Proliferation of advertisement-avoidance technology will change agents’ behavior in the model in several ways. Ad-avoiding viewers’ advertisement disutility will be reduced, and audiences will contain (and therefore advertisers will purchase) ad-avoiders. The unobserved parameters are:

- $\gamma_1$: Ad-avoiders’ proportional reduction in ad nuisance
- $\gamma_2$: The proportion of ad-avoiders in the viewing population
- $\gamma_3$: Advertisers’ proportional valuation of an ad-avoider, relative to a non-ad-avoider

In what follows, I denote ad-avoiding viewers with a superscript $a$, and non-ad-avoiders with a superscript $n$. I rewrite ad-avoider $i$’s utility of watching network $j$ at time $t$ in city $m$ as

---

36 To make explicit the ad-avoidance technology that I am modeling here, I assume that an ad-avoiding viewer can either view or record one network per half-hour, and that users must fast-forward through advertisements (as opposed to being able to skip them entirely). Alternative technologies can be modeled using appropriate re-specifications; the point here is to illustrate the method of model re-specification for a particular technology.

37 $\gamma_3$ is likely to be bounded by zero and one. The main reason to believe $\gamma_3$ might be positive is that when a viewer uses a DVR or similar technology to fast-forward past an ad, she watches the screen intently, to avoid overrunning the return of the program. Even at high speed, she is likely to be able to identify the advertised brand, product, or campaign; and if she finds the commercial interesting, she may stop fast-forwarding to view it. One recent survey found that 15% of DVR users “always,” and another 52% “sometimes” notice TV ads while fast-forwarding (Mandese, 2004); another found that 54% of DVR users say they have replayed or paused television commercials to learn more about an advertised product. (Greenspan, 2004)
\[ u_{ijt}^n = q_j \hat{\alpha}_i \gamma_i + x_j \hat{\beta}_i + \hat{\xi}_{jt} + \hat{n}_{mjt} + \epsilon_{ijt}, \quad (1') \]

where \(\hat{\alpha}, \hat{\beta}, \hat{\xi}_{jt},\) and \(\hat{n}_{mjt}\) are viewer demand model estimates. Ad-avoidance also should affect the utility of the non-broadcast-network-television option, as most viewers engaging in this option are likely to be watching advertisement-supported cable networks. I assume that an ad-avoider’s utility from option non at time \(t\) increases by the same amount as the average network option at the same time:

\[
u_{t,non}^u = \xi_{non,t} + \hat{n}_{m,non,t} + \hat{\pi}_{non} D_{ijm} + \hat{\sigma}_{non} V_i + \epsilon_{i,non,t} - \hat{\alpha}_i (1 - \gamma_1) \frac{1}{J} \sum_{j=1}^{J} q_{jt}.
\]

The utility from the non-television option, and type-\(n\) viewers’ utility from any of their options, remain unchanged. I construct predicted market share functions for each type of viewer, \(\hat{s}_{mjt}^a\) and \(\hat{s}_{mjt}^n\), in the same manner as described in section 3.1.

Network \(j\)’s total audience at time \(t\) depends on its share of each type of viewer, and the proportion of viewers of each type (\(\gamma_2\) and \(1 - \gamma_2\)). To facilitate exposition in what follows, I assume that ad-avoidance technology is distributed randomly among viewers in the population. Thus

\[
V_{jt} = \sum_m [(1 - \gamma_2) \hat{s}_{mjt}^a + \gamma_2 \hat{s}_{mjt}^a] N_m
\]

where \(N_m\) is the number of viewers in city \(m\).

\(\gamma_3\) captures advertisers’ value of an ad-avoiding viewer, relative to a non-ad-avoider. For example, if the exposure of an ad-avoider to the advertiser’s message is worthless, \(\gamma_3\) will be zero; or, if such an exposure is worth half as much as the exposure of a non-ad-avoider to the

---

\(\gamma_3\) captures advertisers’ value of an ad-avoiding viewer, relative to a non-ad-avoider. For example, if the exposure of an ad-avoider to the advertiser’s message is worthless, \(\gamma_3\) will be zero; or, if such an exposure is worth half as much as the exposure of a non-ad-avoider to the

---

38 A good alternative would be to assume that the most ad-averse viewers are those who adopt the ad-avoidance technology first. I plan to report predictions made under both assumptions.
advertiser’s message, then $\gamma_3 = .5$. I use $\gamma_3$ to construct the “effective” audience size, which is the number of type-$n$ viewers in the audience, plus the number of type-$a$ viewers, weighted by advertisers’ value of this latter group. Effective audience size is

$$\tilde{V}_{jt} = \sum_m [(1 - \gamma_2) \hat{s}^n_{mjt} + \gamma_2 \gamma_3 \hat{s}^a_{mjt}] N_m.$$ 

I then aggregate shows’ ad levels and audiences over half-hours, as I did in the estimation of the viewer demand model, and rewrite price per viewer and ad demand in terms of effective audience size,

$$\tilde{f}_s = [q_s, \tilde{V}_s, d_s, x_s] \hat{\lambda} + \hat{\phi},$$

$$\tilde{p}_s = \tilde{V}_s \tilde{f}_s,$$

where $\hat{\lambda}$ and $\hat{\phi}$ are advertiser demand estimates.

Next, I use the re-specified viewer and advertiser demand functions to rewrite network $j$’s first-order condition with respect to show $s$ (originally given by equation 6) as

$$\tilde{p}_s + q_s \left[ \tilde{V}_s \frac{\partial \tilde{f}_s}{\partial q_s} + \frac{\partial \tilde{V}_s}{\partial q_s} \tilde{f}_s \right] - \tilde{\sigma} - \tilde{\kappa}_s = 0.$$ 

Finally, I solve the re-specified model for new equilibrium ad levels, ad prices, and audience sizes. For a guess of the $q_s$’s, the steps are:

1) Calculate predicted audience shares, $\hat{s}^n_{mjt}$ and $\hat{s}^a_{mjt}$, and demographic characteristics, $d_{jt}$, for all networks and time periods;

2) Calculate the effective audience size, $\tilde{V}_{jt}$, and its derivative with respect to ad level, $\frac{\partial \tilde{V}_{jt}}{\partial q_{jt}}$, for all networks and time periods;

3) Average all program-level variables over the half-hours in which they air;
4) Calculate new ad prices and derivatives, \( \tilde{p}_i \), \( \tilde{f}_i \), and \( \frac{\partial \tilde{f}_i}{\partial q_i} \); and finally,

5) Calculate each firm’s first-order condition.

I numerically optimize over the \( q_i \)’s, performing steps (1)-(5) at each guess of the equilibrium ad levels, to set networks’ first-order conditions to zero.

6.2. Results.

I have not yet finished this section.

7. Summary, Managerial Implications, and Future Research

This paper has made a first attempt to give advertisers, media companies, and regulators a rigorous method to predict the effects of advertisement-avoidance technology on an advertisement-supported media industry. The major advantages of the method presented here are its flexibility, and its reliance on a structural model to make out-of-sample predictions.

The issues examined in this paper suggest courses of action for DVR manufacturers, cable system operators, and satellite TV companies (whom I will collectively refer to as the “DVR industry”), as well as for television networks. Demand for DVRs, or monthly DVR service, depends partially on television networks’ program quality. Because advertising revenues are networks’ incentives to invest in program quality, both industries benefit from larger advertiser valuations of ad-avoiding viewers. This convergence of interests suggests that the DVR industry should not include “30-second skip” buttons, or any sort of automatic commercial avoidance feature, in its products; and that television networks should accommodate the DVR industry’s attempts to make DVR users more attractive to advertisers.\(^{39}\) Collaborative

\(^{39}\) An example of such an attempt is the TiVo “Showcase” service. When an advertiser buys a TiVo Showcase, three- or four-minute ads are downloaded to TiVo hard drives. When a TiVo user sees a “regular” TV ad for the
development of technologies that maximize advertisers’ value of ad-avoiders will produce ancillary revenue streams for the DVR industry and help to strengthen advertiser demand for TV audiences.

This research also has implications for advertisers. One implication is old: advertisers could make better audience predictions if networks informed advertisers during the upfront how many minutes of advertising would be sold during each show. Other implications are new: when planning their media buys, advertisers should consider how much they value DVR users (if at all), and take into account data on audiences’ DVR usage and ad-avoidance.

Managers in the DVR industry could combine this model with a model of product diffusion and a conjoint analysis of DVR product characteristics to forecast demand for their products, and how that demand will be affected by television networks’ actions.

Of course, this study has its limitations, some of which suggest avenues for future research. To make predictions, I have held constant many factors, including program schedules, networks’ ad scheduling strategies and, most significantly, program quality. The model could be extended to endogenize program scheduling, alternative advertisement vehicles like product placement or screen overlays, or endogenous commercial break timing. It would also be interesting to supplement the broadcast network data with cable network data to examine how the relationship between the two industries will change as DVR usage increases. Finally, it would be very interesting to model the third side of the television industry, wherein networks compete to purchase programs, to examine the effects of ad-avoidance on the quality and costs of television programs.

product in question, a signal encoded in the commercial alerts her DVR. The DVR then displays an icon on her television screen. If the viewer is interested in learning more about the product, she can press a button on the TiVo remote to play the longer commercial. At the end of the commercial, she is returned to her program.
Acknowledgements  I am indebted to many people for their helpful comments and suggestions, most especially Simon Anderson, Steven Stern, and Ron Wilcox; and also including Keith Brand, Paul Carrillo, Federico Ciliberto, Anthony Dukes, Paul Farris, John Harris, David Mick, John Pepper, Adam Rennhoff, Roland Rust, and seminar participants at the IDEI/ZEI Conference on Regulation of Media Markets. I gratefully acknowledge financial support from the Bankard Fund for Political Economy at the University of Virginia.

References


InsightExpress (2004), “


<table>
<thead>
<tr>
<th></th>
<th>All Nets</th>
<th>ABC</th>
<th>CBS</th>
<th>FOX</th>
<th>NBC</th>
<th>UPN</th>
<th>WB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adv Price (per 30</td>
<td>147</td>
<td>125</td>
<td>179</td>
<td>241</td>
<td>212</td>
<td>55</td>
<td>71</td>
</tr>
<tr>
<td>seconds; $000</td>
<td>(125)</td>
<td>(40)</td>
<td>(114)</td>
<td>(186)</td>
<td>(125)</td>
<td>(26)</td>
<td>(25)</td>
</tr>
<tr>
<td>Adv Seconds</td>
<td>315</td>
<td>367</td>
<td>306</td>
<td>284</td>
<td>312</td>
<td>323</td>
<td>299</td>
</tr>
<tr>
<td></td>
<td>(74)</td>
<td>(83)</td>
<td>(55)</td>
<td>(66)</td>
<td>(81)</td>
<td>(68)</td>
<td>(58)</td>
</tr>
</tbody>
</table>

Cost Per Thousand Viewers ($)\(^1\)

<table>
<thead>
<tr>
<th></th>
<th>Monday</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20</td>
<td>29</td>
<td>27</td>
<td>25</td>
<td>19</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4)</td>
<td>(7)</td>
<td>(7)</td>
<td>(7)</td>
<td>(2)</td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tuesday</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>14</td>
<td>37</td>
<td>34</td>
<td>25</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5)</td>
<td>(4)</td>
<td>(6)</td>
<td>(6)</td>
<td>(3)</td>
<td>(6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wednesday</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>26</td>
<td>29</td>
<td>29</td>
<td>21</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6)</td>
<td>(13)</td>
<td>(7)</td>
<td>(8)</td>
<td>(6)</td>
<td>(7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thursday</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>30</td>
<td>15</td>
<td>35</td>
<td>19</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(13)</td>
<td>(13)</td>
<td>(7)</td>
<td>(7)</td>
<td>(3)</td>
<td>(4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Friday</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>20</td>
<td>26</td>
<td>20</td>
<td>11</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(10)</td>
<td>(8)</td>
<td>(3)</td>
<td>(1)</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>All Nights</td>
<td>24</td>
<td>22</td>
<td>23</td>
<td>28</td>
<td>28</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>(9)</td>
<td>(7)</td>
<td>(11)</td>
<td>(9)</td>
<td>(8)</td>
<td>(6)</td>
<td>(5)</td>
</tr>
</tbody>
</table>

Audience Size (000)\(^2\)

<table>
<thead>
<tr>
<th></th>
<th>Monday</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5,779</td>
<td>8,862</td>
<td>6,119</td>
<td>6,928</td>
<td>2,699</td>
<td>4,690</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1275)</td>
<td>(2369)</td>
<td>(1723)</td>
<td>(1569)</td>
<td>(250)</td>
<td>(549)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tuesday</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5,299</td>
<td>8,291</td>
<td>11,303</td>
<td>5,703</td>
<td>2,768</td>
<td>4,895</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(546)</td>
<td>(551)</td>
<td>(2932)</td>
<td>(1509)</td>
<td>(819)</td>
<td>(670)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wednesday</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6,678</td>
<td>5,409</td>
<td>12,574</td>
<td>8,551</td>
<td>3,260</td>
<td>3,210</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1134)</td>
<td>(589)</td>
<td>(4984)</td>
<td>(2552)</td>
<td>(725)</td>
<td>(1362)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thursday(^3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4,281</td>
<td>13,098</td>
<td>4,269</td>
<td>12,737</td>
<td>3,807</td>
<td>2,144</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1336)</td>
<td>(2547)</td>
<td>(1150)</td>
<td>(828)</td>
<td>(334)</td>
<td>(411)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Friday</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5,722</td>
<td>5,610</td>
<td>4,131</td>
<td>5,574</td>
<td>1,940</td>
<td>2,260</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(549)</td>
<td>(932)</td>
<td>(916)</td>
<td>(901)</td>
<td>(290)</td>
<td>(248)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>All Nights</td>
<td>5,868</td>
<td>5,693</td>
<td>7,716</td>
<td>8,058</td>
<td>7,361</td>
<td>2,793</td>
</tr>
<tr>
<td></td>
<td>(3294)</td>
<td>(1182)</td>
<td>(2825)</td>
<td>(4573)</td>
<td>(2752)</td>
<td>(791)</td>
<td>(1379)</td>
</tr>
</tbody>
</table>

Numbers in parentheses are standard deviations

\(1\) averaged over half-hours, 8:00-10:00 p.m., and weeks

\(2\) measured in households; averaged over half-hours, 8:00-10:00 p.m., and weeks

\(3\) Thursday, May 1, and Thursday, May 15, were removed from the sample

Sources: Nielsen Media Research; TNS Media Intelligence/CMR
Table 2. Variable Definitions and Descriptive Statistics

<table>
<thead>
<tr>
<th>variable name</th>
<th>description</th>
<th>mean</th>
<th>st. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adv Seconds</td>
<td>Seconds of national advertisements aired during the program</td>
<td>308</td>
<td>(75)</td>
</tr>
<tr>
<td>Lead-in</td>
<td>The size of the network’s audience in the previous half-hour varies by market</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lead-out</td>
<td>The size of the network’s audience in the subsequent half-hour varies by market</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ScriptedComedy</td>
<td>=1 if the show is a scripted comedy</td>
<td>.31</td>
<td>(.46)</td>
</tr>
<tr>
<td>ActionDrama</td>
<td>=1 if the show is a scripted drama that contains action scenes</td>
<td>.10</td>
<td>(.30)</td>
</tr>
<tr>
<td>PsychDrama</td>
<td>=1 if the show is a scripted drama that does not contain action scenes</td>
<td>.26</td>
<td>(.44)</td>
</tr>
<tr>
<td>Reality</td>
<td>=1 if the show is unscripted</td>
<td>.16</td>
<td>(.36)</td>
</tr>
<tr>
<td>News</td>
<td>=1 if the show is a news program or a newsmagazine</td>
<td>.06</td>
<td>(.24)</td>
</tr>
<tr>
<td>Movie</td>
<td>=1 if the show is a movie</td>
<td>.11</td>
<td>(.31)</td>
</tr>
<tr>
<td>African-American</td>
<td>=1 if at least one African-American main character</td>
<td>.44</td>
<td>(.50)</td>
</tr>
<tr>
<td>Other Nonwhite</td>
<td>=1 if at least one non-white, non-African American main character</td>
<td>.09</td>
<td>(.29)</td>
</tr>
<tr>
<td>MC&lt;18</td>
<td>=1 if at least one main character is under 18</td>
<td>.20</td>
<td>(.40)</td>
</tr>
<tr>
<td>MC18-34</td>
<td>=1 if at least one main character is between 18 and 34 years old</td>
<td>.67</td>
<td>(.47)</td>
</tr>
<tr>
<td>MC35-49</td>
<td>=1 if at least one main character is between 35 and 49 years old</td>
<td>.62</td>
<td>(.49)</td>
</tr>
<tr>
<td>MC50+</td>
<td>=1 if at least one main character is over 50 years old</td>
<td>.17</td>
<td>(.38)</td>
</tr>
<tr>
<td>Married</td>
<td>=1 if at least one main character is married to another character</td>
<td>.19</td>
<td>(.40)</td>
</tr>
<tr>
<td>Single Parent</td>
<td>=1 if at least one main character is single and has children</td>
<td>.10</td>
<td>(.30)</td>
</tr>
<tr>
<td>Female Only</td>
<td>=1 if none of the main characters are male</td>
<td>.09</td>
<td>(.29)</td>
</tr>
<tr>
<td>Male Only</td>
<td>=1 if none of the main characters are female</td>
<td>.23</td>
<td>(.42)</td>
</tr>
<tr>
<td>50+% Nonwhite</td>
<td>=1 if 50% or more of the show’s cast is non-white</td>
<td>.20</td>
<td>(.40)</td>
</tr>
<tr>
<td>25+% Nonwhite</td>
<td>=1 if 25-49% of the show’s cast is non-white</td>
<td>.20</td>
<td>(.40)</td>
</tr>
<tr>
<td>10+% Nonwhite</td>
<td>=1 if 10-24% of the show’s cast is non-white</td>
<td>.19</td>
<td>(.40)</td>
</tr>
<tr>
<td>50+% Female</td>
<td>=1 if 50% or more of the show’s cast is female</td>
<td>.46</td>
<td>(.50)</td>
</tr>
<tr>
<td>25+% Female</td>
<td>=1 if 25-49% of the show’s cast is female</td>
<td>.31</td>
<td>(.46)</td>
</tr>
<tr>
<td>House</td>
<td>=1 if the show contains scenes set in a character’s house</td>
<td>.21</td>
<td>(.41)</td>
</tr>
<tr>
<td>Apartment</td>
<td>=1 if the show contains scenes set in a character’s apartment</td>
<td>.06</td>
<td>(.23)</td>
</tr>
<tr>
<td>Workplace</td>
<td>=1 if the show contains scenes set in a business or workplace</td>
<td>.32</td>
<td>(.47)</td>
</tr>
<tr>
<td>Studio</td>
<td>=1 if the show contains scenes set in a TV studio</td>
<td>.28</td>
<td>(.45)</td>
</tr>
<tr>
<td>Cop</td>
<td>=1 if the show has some law enforcement element</td>
<td>.10</td>
<td>(.31)</td>
</tr>
<tr>
<td>Sci-Fi</td>
<td>=1 if the show contains elements of science fiction (i.e. Star Trek)</td>
<td>.06</td>
<td>(.24)</td>
</tr>
<tr>
<td>Supernatural</td>
<td>=1 if the show contains supernatural elements (i.e. angels, witchcraft)</td>
<td>.07</td>
<td>(.26)</td>
</tr>
<tr>
<td>Age</td>
<td># of years since the show’s debut</td>
<td>3.28</td>
<td>(3.5)</td>
</tr>
<tr>
<td>2003EmmyNoms</td>
<td>2004 Emmy Nominations</td>
<td>1.4</td>
<td>(3.1)</td>
</tr>
<tr>
<td>Past Emmy Noms</td>
<td>All pre-2004 Emmy Nominations</td>
<td>3.6</td>
<td>(11.9)</td>
</tr>
<tr>
<td>ABC, CBS, FOX,</td>
<td>Network-specific dummy variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NBC, UPN, WB</td>
<td>Mon, Tue, Wed, ...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thu, Fri</td>
<td>Day-specific dummy variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1stHH, 2ndHH,</td>
<td>Dummy variables for 1st half-hour of prime time, 2nd half-hour of prime time, 3rd, 4th, etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rdHH, 4thHH</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>