
Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

# A Two-Sided, Empirical Model of Television Advertising and Viewing Markets

Kenneth C. Wilbur,

MARKETING SCIENCE
Vol. 27, No. 3, May–June 2008, pp. 356–378
ISSN 0732-2399 | EISSN 1526-548X | 08 | 2703 | 0356

DOI 10.1287/mksc.1070.0303




# A Two-Sided, Empirical Model of Television Advertising and Viewing Markets

Kenneth C. Wilbur
Marshall School of Business, University of Southern California, 3660 Trousdale Parkway, ACC 301L,
Los Angeles, California 90089, kwilbur@usc.edu

For marketers, television remains the most important advertising medium. This paper proposes a two-sided model of the television industry. We estimate viewer demand for programs on one side and advertiser demand for audiences on the other. The primary objective is to understand how each group's program usage influences the other group.

Four main conclusions emerge. First, viewers tend to be averse to advertising. When a highly rated network decreases its advertising time by 10%, our model predicts a median audience gain of about 25% (assuming no competitive reactions). Second, we find the price elasticity of advertising demand is −2.9, substantially more price elastic than 30 years ago.

Third, we compare our estimates of advertiser and viewer preferences for program characteristics to networks' observed program choices. Our results suggest that advertiser preferences influence network choices more strongly than viewer preferences. Viewers' two most preferred program genres, Action and News, account for just 16% of network program hours. Advertisers' two most preferred genres, Reality and Comedy, account for 47% of network program hours.

Fourth, we perform a counterfactual experiment in which some viewers gain access to a hypothetical advertisement avoidance technology. The results suggest that ad avoidance tends to increase equilibrium advertising quantities and decrease network revenues.

*Key words*: advertising; broadcasting; demand estimation; empirical industrial organization; endogeneity; entertainment marketing; media; television; two-sided markets
*History*: This paper was received on December 21, 2005, and was with the authors 8 months for 3 revisions; processed by Anthony Dukes.

## 1. Introduction

Television remains the single most important advertising medium. It absorbs about a quarter of all advertising expenditures in the United States. Mean household television viewing increased every year between 1996 and 2006, from 7:11 to 8:14 hours per day.[1]

Despite television's importance, some fundamental questions about the television industry remain unanswered. Econometric models of the industry have typically failed to acknowledge its two-sided nature. Networks compete to attract viewer attention and then sell that attention to advertisers. A network must recruit both sides of the market to be successful. Advertisers and marketers have long understood the two-sided nature of the industry, but theoretical treatments have only recently modeled the interplay of advertiser and viewer preferences. Yet there has been no empirical study of the industry that considers cross-group externalities: the effect the number of advertisements has on audience size, and the effect audience size has on advertiser demand.

This paper proposes a two-sided, empirical model of the television industry. The model is estimated using data on television network audience ratings[2] in 50 geographically defined television markets, advertising quantities and prices per half hour, program characteristics, and market demographics. The estimation results are used to address the following questions.

(1) *Are viewers averse to television advertising*? If the answer is yes, networks face opposing incentives to recruit the two sides of the market: Ad sales impose negative external effects on viewers and may cause audience losses. Our estimation results suggest a 10% decrease in advertising level[3] typically increases

[1] Television Bureau of Advertising (2005).

[2] "Rating" or "audience rating" are industry terms for the fraction of the potential audience that watched a program. "Share" or "audience share" is the fraction of viewers watching television who watched a program. We use the terms "rating," "audience rating," and "market share" interchangeably. We do not refer to audience share.

[3] We use the terms "advertising quantity" and "advertising level" interchangeably.

audience size by about 25% for a highly rated network (assuming competing networks do not react).

This result shows that it is necessary to control for advertising levels to obtain unbiased viewer demand estimates. When viewers are ad averse, a popular program will contain relatively more ads so the estimate of its popularity would be biased downward if advertising time is unobserved. This issue is similar to a situation in which demand for a consumer product is estimated in the absence of price data.

(2) *How competitive is the market for television audiences, and what drives advertiser demand for audiences*? Marketers spend significant fractions of their advertising budgets in television advertising markets, but the academic literature has seldom examined the determinants of advertiser demand. Knowledge of aggregate advertiser demand and market conditions may guide individual advertisers' audience purchasing strategies.

The estimates indicate that, holding audience constant, a 10% increase in advertising price decreases quantity of ads demanded by 29%. The television advertising market has become substantially more price elastic over the past 30 years (Bowman 1976).

(3) *Whose preferences have a greater effect on network program selection: viewers' or advertisers'*? Networks choose their programs to match viewers with advertisements. It may be that some programs better suit viewer tastes than advertiser tastes or vice versa. Advertiser and viewer preferences can be compared to network programs and scheduling choices to learn whether networks favor one group's preferences over the other.

The results suggest that advertiser preferences exert a stronger influence on network programming choices than viewer preferences. Viewers most prefer to watch Action and News programs. Advertisers most prefer to buy time during Comedy and Reality programs. Yet Comedy and Reality shows constitute 48% of network programming, while Action and News programs account for just 16%. Analyzing either side of the market in isolation would suggest networks were failing to satisfy their customers' tastes.

(4) *What can be discerned about the equilibrium effects of advertisement-avoidance technology on advertising quantities*? Several media articles (e.g., Garfield 2005) and books have predicted catastrophic effects on the television industry from digital video recorder proliferation. We perform a counterfactual experiment to provide educated speculation about such effects. We combine viewer and advertiser demand estimates with a structural model of network competition. We then analyze a counterfactual experiment to simulate the effects of growing advertisement avoidance on the industry. The results suggest that ad-avoidance technology will lead to increasing advertising levels and falling network revenues.

The next section explains how this paper contributes to the literature on television viewing and advertising markets. Section 3 describes the model of viewer utility, advertiser demand, and network supply of advertisements. Section 4 presents the data and discusses endogeneity and estimation. Empirical results are in §5 and the counterfactual results are in §6.

## 2. Relevant Literature

This paper's contribution is in its estimation of demand on both sides of the television industry. The importance of this two-sided approach is highlighted when viewer and advertiser preferences are compared to networks' actual programming choices. This analysis includes the first effort to measure the sensitivity of television audience size to the amount of time devoted to national advertising within a program. It also contains new evidence about the determinants of advertising demand and network program selection.

This paper extends previous work on viewer demand estimation by considering advertising time as analogous to the "price" of consuming a nominally free television program. The discrete choice literature on viewer demand for television programs originates in Rust and Alpert (1984). In the recent literature, Shachar and Emerson (2000) added interactions between viewer and program characteristics to the model. They find that program cast demographics are good predictors of program audience demographics. Estimated viewer preferences have been used to calibrate models of optimal program scheduling (Danaher and Mawhinney 2001, Goettler and Shachar 2001), to investigate sources of viewing persistence (Anand and Shachar 2004), to address whether advertising is persuasive or informative (Anand and Shachar 2005), and to estimate preference interdependence among groups of viewers (Yang et al. 2006). Several studies have investigated viewer switching during commercials (Siddarth and Chattopadhyay 1998, Zufryden et al. 1993), but no previous analysis has considered that a program's aggregate time devoted to advertising may directly affect viewer utility.

Two-sided models of media markets are relatively new to the academic literature. They are generally characterized as models in which platforms enable interactions between distinct groups of agents. The platform's goal is to get the various sides "on board" (Rochet and Tirole 2003). The pioneering treatments in the two-sided markets literature are Armstrong (2006), Caillaud and Jullien (2001, 2003), and Rochet and Tirole (2006). For advertisement-supported media, the most important insight from the two-sided markets literature is that ad prices reflect both the value of reaching a given audience and the marginal effect of the ad sale on the total size of that audience.

This paper provides evidence to support assumptions made by several recent theoretical analyses of two-sided media industries. Anderson and Coate (2006) show that media markets may supply too many ads because broadcasters fail to take into account advertising disutility imposed on nonswitching viewers. There may also be too few ads when competing programs are close substitutes and viewers are very ad averse. Dukes and Gal-Or (2003) model interactions between ad-averse viewers, broadcasters, and advertisers, and also consider the effects of informative advertising on competition in product markets. They show that when increased advertising leads to better-informed consumers, product market competition intensifies. Networks and advertisers can then benefit from exclusive advertising contracts. Liu et al. (2004) consider the effects of competition between television networks on network incentives to invest in program quality. They show that increased network competition can lead to diminished program investments and lower viewer welfare.

A few recent studies consider the effect of advertising on consumer utility in various media. Kaiser and Wright (2006) find that readers of women's magazines value advertisements. Depken and Wilson (2004) examine a large number of magazine categories and find substantial heterogeneity. They find reader utility from ads is positive in some categories and negative in others. Rysman (2004) finds that consumer utility from yellow pages directories increases with the number of pages of advertisements.

The advertising side of the television industry has received less attention than the viewing side. Crandall (1972) and Bowman (1976) estimated advertiser demand for television audiences but neither paper used data on ad quantities. Two recent working papers, Goettler (1999) and Wildman et al. (2004), estimate the relationship between audience size and advertisement price but neither paper uses data on advertising quantities. The present analysis is the first to directly estimate the price elasticity of advertising demand while accounting for audience size effects.

## 3. A Model of Television Advertisers, Networks, and Viewers

This section describes our model of viewer utility, advertiser demand, and network supply of television commercials. The viewer and advertiser models presented in §§3.1 and 3.2 are used in the estimation. The model of network behavior presented in §3.3 is used in conjunction with parameter estimates in two ways. First, its implications are used to make inferences about missing data (tune-in levels[4] are unobserved). Second, it underpins the counterfactual experiment presented in §6.

### 3.1. Viewers

Discrete choice models have been used to describe television viewers' typical behavior of watching one network at a time. In this paper, we use a random-coefficients logit as this best suits the available data. The primary benefits of this model are its extensive controls for viewer heterogeneity and its reliance on consumer characteristics to identify substitution patterns.

Each viewer $i$ in city $m$ is assumed to watch one of $J-1$ broadcast television networks (networks are indexed by $j$), to watch some other television channel (denoted option $J$), or to engage in some nontelevision pursuit at each half hour $t$. Let viewer $i$'s utility from watching network $j$ in city $m$ at time $t$ be

$$u_{imjt} = q_{jt}\alpha^*_{im} + x_{mjt}\beta^*_{im} + \xi_{jt} + \eta_{mjt} + \varepsilon_{imjt}, \tag{1}$$

where $q_{jt}$ is the number of seconds of advertising on network $j$ during half hour $t$; $x_{mjt}$ is a vector that includes the observable characteristics of the show on network $j$ at time $t$ (e.g., genre), audience flow effects,[5] and market, day, and time dummies; $\alpha^*_{im}$ and $\beta^*_{im}$ are viewer $i$'s taste parameters; $\xi_{jt}$ captures mean tastes for the unobserved characteristics of the program airing on network $j$ during time $t$; $\eta_{mjt}$ measures a deviation from mean tastes for unobserved show characteristics common to viewers in city $m$ at time $t$;[6] and $\varepsilon_{imjt}$ is viewer $i$'s idiosyncratic taste for network $j$'s time $t$ program. $q_{jt}$ does not vary across markets due to television signal distribution technology.

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

$$\begin{bmatrix}\alpha^*_{im}\\ \beta^*_{im}\end{bmatrix} = \begin{bmatrix}\alpha\\ \beta\end{bmatrix} + \Pi D_{im} + \Sigma v_{im},$$
$$D_{im} \sim P^*_{Dm}(D),\ v_{im} \sim P^*_v(v), \tag{2}$$

where $\alpha$ and $\beta$ are mean tastes for ad quantity and program characteristics, $D_{im}$ is a $d\times 1$ vector of viewer demographic characteristics (income, age, and age$^2$), $P^*_{Dm}(D)$ is the market-specific joint distribution of viewer demographics, and $v_{im}$ is a $k\times 1$ vector

[4] Tune-ins are advertisements for future network programs. People in the television industry call them promos.

[5] Byzalov and Shachar (2004) present evidence that state dependence plays an important role in television viewing choices. Our data do not allow us to observe viewer state dependence or estimate a dynamic model of program consumption, so we use network $j$'s audience size in market $m$ at times $t-1$ and $t+1$ to control for audience flow. The industry terms for these audiences are lead-in and lead-out.

[6] We assume the effects of unobserved regional differences in speech and culture on viewer preferences for unobserved show characteristics to be the primary determinants of $\eta_{mjt}$.

of unobserved preference heterogeneity (where $k = \dim(\beta^*_{im}) + 1$). We make the standard assumption that the components of $v_{im}$ are distributed normal and are independent across individuals and markets. $\Pi$ is a $k \times d$ parameter matrix that measures how tastes for program characteristics and advertising vary with observed viewer demographics, and $\Sigma$ is a $k \times k$ diagonal matrix that measures the relative importance of unobserved viewer preference heterogeneity.

It is expected that increasing levels of ads will reduce audience size. Two underlying behavioral mechanisms are consistent with this hypothesis. First, it might be that viewers are sufficiently experienced in television consumption to predict accurately whether the utility of watching a given program, net of the advertising level it contains, is better than their next best alternative. Second, each marginal unit of advertising time could provide the viewer with a potential stimulus to leave the audience. Audience losses would then increase with advertising time.

Note that negative viewer utility of advertising does not necessarily mean that viewers "dislike" advertising. Rather, it is a statement that viewers have a relative preference for watching programs over watching advertising. If it were not the case that viewers prefer programs to advertisements, it might be difficult to explain the existence of programs.

If viewer $i$ watches a nonbroadcast network (option $J$), her utility is given by:

$$u_{imJt} = x_{mJt}\beta_i + \xi_{Jt} + \eta_{mJt} + \pi_J D_{im} + \sigma_J v_{im} + \varepsilon_{imJt},$$

where $x_{mJt}$ contains audience flow, market, day, and time effects, $\xi_{Jt}$ is the mean value of the best available nonbroadcast network at time $t$, and $\eta_{mJt}$ is a deviation from mean preferences for nonbroadcast TV networks shared by viewers in city $m$ at time $t$.

The utility of the nontelevision option (option 0) is:

$$u_{im0t} = \xi_{0t} + \eta_{m0t} + \pi_0 D_{im} + \sigma_0 v_{im} + \varepsilon_{im0t},$$

where $\xi_{0t}$ and $\eta_{m0t}$ are normalized to zero ($\xi_{jt}$ and $\eta_{mjt}$ are identified relative to this normalization) and $\pi_0 D_{im} + \sigma_0 v_{im}$ is interpreted as a fixed effect that measures the time-invariant component of viewer $i$'s value of the nontelevision option.

We assume viewers act to maximize utility. Thus, the set of demographics and preferences that leads viewer $i$ in city $m$ to watch network $j$ at time $t$ is:

$$A_{mjt} = \{(D_{im}, v_{im}, \varepsilon_{im\cdot t}) \mid u_{imjt} > u_{imkt}, \forall k \neq j\},$$

where $\varepsilon_{im\cdot t} = [\varepsilon_{im0t}, \ldots, \varepsilon_{imJt}]$. If the idiosyncratic error terms are distributed identically and independently, the rating of network $j$ in market $m$ at time $t$ is given by:

$$s_{mjt} = \int_{A_{mjt}} dP^*_\varepsilon(\varepsilon)\, dP^*_v(v) dP^*_{Dm}(D).$$

We assume the idiosyncratic error terms $\varepsilon_{imjt}$ are distributed i.i.d. Type I extreme value and integrate out over them in the standard fashion. Thus, the predicted rating can be rewritten as:

$$s_{mjt} = \int_{A_{mjt}} dP^*_\varepsilon(v, D)\, dP^*_v(v)\, dP^*_{Dm}(D), \tag{3}$$

where $dP^*_\varepsilon(v, D)$ is the standard Multinomial Logit market share function. Equation (3) will be used to estimate viewer demand for television programs. The right-hand side of Equation (3) is integrated over a large number of dimensions and does not have a closed-form solution, so simulation will be used to approximate it.

### 3.2. Advertisers

We model advertisers through their aggregate demand for advertising on a given program. Advertiser demand for a particular television audience is influenced by many factors including audience size, viewer demographics, and program characteristics that influence the efficacy of the program's advertising message delivery. We assume that aggregate inverse demand for advertising on a given program $s$ is given by:

$$p_s = q_s\lambda_q + V_s\lambda_V + d_s\lambda_d + x_s\lambda_x + \phi_s, \tag{4}$$

where $p_s$ is the price of an ad during show $s$, $q_s$ is the show's ad level, $V_s$ is the number of viewers watching show $s$, $d_s$ is a vector of viewer demographics, $x_s$ represents program characteristics that affect advertising effectiveness, the $\lambda$'s are advertiser preference parameters, and $\phi_s$ is an error term. Possible sources of error include unobserved audience demographics and measurement error in ad price.

The drawback of assuming that Equation (4) is linear is the lack of clarity in the underlying assumptions about advertiser preferences and behavior and the attendant risk of specification error.[7] However,

[7] Derivation of advertiser demand from first principles would be preferable but is complicated by an assignment problem in the matching of advertisers and audiences. To illustrate, consider a single advertiser with utility function $\sqrt{\sum_{s\in S} V_s}$, where $S$ is the set of shows purchased by the advertiser. Assume there are two audiences available: Audience $A$ consists of 9 viewers and audience $B$ consists of 16 viewers. Viewers are homogeneous and audiences do not overlap. If the advertiser purchases only audience $A$, its willingness to pay for $A$ is 3 ($=\sqrt{9}$). If the advertiser purchases only audience $B$, its willingness to pay for $B$ is 4 ($=\sqrt{16}$). If the advertiser purchases both audiences, its willingness to pay for $B$ is $B$'s marginal contribution to total advertiser utility: 2 ($=\sqrt{16+9} - \sqrt{9}$). The assignment problem is that the advertiser's willingness to pay for each audience depends on whether the advertiser purchases the other audience. There have been some recent advances in estimating many-to-many matching problems (Fox 2007) but given the focus of this paper and the technical challenges involved in many-to-many matching estimation, we restrict our approach to the assumption of a reduced-form demand function.

as shown in §5.3, the assumed functional form is found to explain 87% of the variation in advertisement prices.[8]

### 3.3. Networks

The complexity of networks' strategic interactions requires strong assumptions about their behavior. Due to the strength of these assumptions and the limitations of the data, the model of network competition is not used to estimate demand parameters. This model is used only to infer missing tune-in levels and for the counterfactual experiment described in §6.

We assume networks compete in three stages. In the first stage, networks choose their programs; in the second stage, networks schedule their programs; and in the third stage, networks set ad quantities.[9,10] The focus here is on competition in the final stage, in which program costs are sunk and program schedules have been finalized.

The data do not record tune-ins but tune-ins account for about 25% of nonprogram material[11] so it is important to account for them. We assume networks set advertisement and tune-in levels to maximize current-period advertising revenues. A higher number of tune-ins in the current period will inform viewers of upcoming programs, but tune-ins may reduce current audience size. Those audience losses reduce current-period advertising revenues.

We assume that a network's benefit from a viewer's tune-in exposure is constant at $\tau_s$ during each time show $s$ airs. $\tau_s$ is the marginal effect of exposure to one tune-in on the probability any given viewer will watch the advertised show, times the network's profit from the increase in the advertised show's audience.

[8] We tested several alternate specifications. Model fits and qualitative results were similar across all models tested.

[9] The timing of the game is consistent with reality. The first stage takes place before the up-front market, when networks renew returning shows and buy new ones. The second stage takes place at the start of the up-front, when networks announce their program schedules. The third stage occurs during the remainder of the up-front and scatter markets. Yet it should be noted that networks replace and reschedule some shows during the season.

[10] The ad-quantity-setting assumption is standard in the literature. Networks can set the number of minutes of advertisements in the short run by editing the program appropriately. The up-front contains considerable uncertainty about advertising demand and protracted price negotiations in which advertisers are known to pay nonuniform prices; this information seems inconsistent with an assumption that networks set a single price per show. Anderson and Coate (2006) show that ad price and ad quantity games yield identical results when viewers pay no subscription fees. If networks set ad prices, the empirical implication for this paper is that the inferred tune-in levels presented in §5.5 will be too small. The interested reader is referred to Crampes et al. (2005).

[11] American Association of Advertising Agencies and Association of National Advertisers (AAAA/ANA 2002).

If we denote the number of tune-ins aired during show $s$ as $r_s$, the network's stage-three profits can be written as the sum of its advertising revenues and tune-in benefits during the show:

$$\pi_j = \max_{\{q_s, r_s\}_{s\in S_j}} \sum_{s\in S_j} [q_s p_s + r_s V_s \tau_s], \tag{5}$$

where $S_j$ is network $j$'s catalogue of shows.

Substituting ad demand (4) into Equation (5) and differentiating with respect to $q_s$ and $r_s$ yields

$$p_s + q_s\lambda_q + q_s\lambda_V \frac{\partial V_s}{\partial q_s} + r_s \frac{\partial V_s}{\partial q_s}\tau_s = 0, \quad \text{and} \tag{6}$$

$$q_s\lambda_V \frac{\partial V_s}{\partial r_s} + V_s\tau_s + r_s\frac{\partial V_s}{\partial r_s}\tau_s = 0.^{12} \tag{7}$$

The first two terms in Equation (6) are similar to marginal revenue in any monopoly or oligopoly first-order condition. The third term captures the two-sided nature of the market, the decrease in advertisement price through the audience loss ($\partial V_s/\partial q_s$) engendered by commercial sales. The fourth term is the marginal effect of the network's advertising sales on its gross tune-in benefit. The first-order condition taken with respect to $r_s$ contains similar logic. The first term is the marginal effect of a tune-in on advertising revenues, and the second two terms are the network's marginal benefit of airing a tune-in.

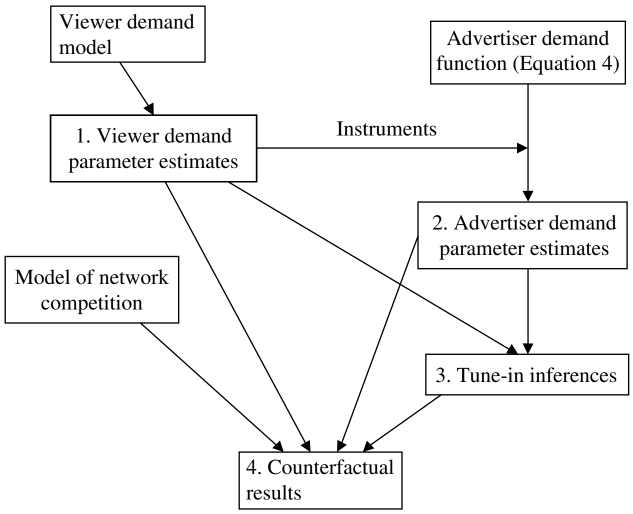
We make the additional assumption that viewers are equally averse to advertisements and tune-ins. Tune-ins are more likely to be relevant to the mean viewer than advertisements. They are also repeated more frequently within short periods of time, so they are accordingly more likely to wear out. This assumption is very strong, but it is used only to infer unobserved tune-in levels and there is no apparent alternative. It is not used in the estimation. It implies that the marginal effect of ad sales on audience size is equal to that of airing tune-ins: $\partial V_s/\partial q_s = \partial V_s/\partial r_s$. Equations (6) and (7) then imply a third relationship:

$$p_s + q_s\frac{\partial p_s}{\partial q_s} = V_s\tau_s. \tag{8}$$

Equation (8) says that, holding audience size constant, the network will choose advertising and tune-in levels to equate the marginal benefit from each activity. Equations (7) and (8) will be used in conjunction with advertiser and viewer demand parameter estimates to make inferences about $\tau_s$ and $r_s$.

[12] A computer simulation was conducted to investigate the existence, stability, and number of equilibria of this oligopoly game. Assuming $\partial V/\partial q$ is strictly negative, two equilibria exist: One is an interior, stable equilibrium in which all networks play nonzero advertising levels. In the other equilibrium, network advertising quantities approach infinity and audience sizes approach zero.

**Figure 1** **Layout of Empirical Strategy**




## 4. Data, Endogeneity, and Estimation

An overview of the empirical strategy is depicted in Figure 1. The model has three sets of primitives: assumptions about viewers, advertisers, and networks. The first stage of estimation uses assumptions about viewers to estimate the viewer demand model. The second stage uses assumptions about advertisers in conjunction with viewer demand-side instruments to estimate advertiser preferences.[13] In the third stage, the two sets of parameter estimates are used to calibrate the model of network competition. Network first-order conditions provide a means to infer the unobserved tune-in levels. In the final stage, all of the assumptions, results, and missing data inferences are used in a counterfactual to speculate about the effects of advertisement-avoidance technology.

This section describes the data, viewer demand-side endogeneity and estimation, advertiser demand-side endogeneity and estimation, and the inference procedure for missing data. The counterfactual is discussed in §6 and in the Technical Appendix that can be found at http://mktsci.pubs.informs.org.

### 4.1. Data

The model is estimated using data on television audiences, viewer demographics, advertisements, and program characteristics from four sources. The unit of observation is a network/market/day/half hour. For each observation, the data report seconds of national advertising, the average cost of a 30-second commercial, the characteristics of the program, and household audience ratings in each of 50 television markets.

The sample includes the programs and ads aired by the six most-watched U.S. broadcast television networks (ABC, CBS, FOX, NBC, UPN, and WB) in the first two hours of prime time, Monday to Friday, during the May "sweeps" period (April 24–May 21) in 2003.[14] We describe each component of the data in turn, report descriptive statistics, and conclude the section by describing some limitations of the data set.

Audience data come from Nielsen Media Research (NMR) "Viewers in Profile" reports covering the 50 largest U.S. Designated Market Areas (DMAs). More than 90% of U.S. television households are represented in the sample. NMR collected household audience data 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 network affiliate.

Two important characteristics of the audience data merit discussion. First, 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. (The data are reported at the half hour level as the average of the two fifteen-minute blocks in each half hour.) It is therefore possible for a household in Nielsen's sample to watch a program, avoid commercials perfectly with a remote control, and still be counted among the program's audience. This would suggest the household is extremely ad averse, but its ad aversion would not be detectable in the data. To the extent this occurs, the estimation may understate viewers' true ad disutility. However, the main purpose here is to measure how Nielsen's audience measurements change with advertising levels, since those measurements were the currency of the advertising marketplace in 2003.

Second, Nielsen excluded digital video recorder (DVR) households from its sample in 2003. The Yankee Group estimated DVR penetration at 2% of U.S. households in mid-2003. Therefore, viewer demand parameters should be interpreted as describing thes tastes of the remaining 98% of U.S. television households.

We can contrast the viewing data used in this paper with data sets used in prior literature in order to aid interpretation of the estimation results. Several previous studies have used individual-level viewing data generated by "peoplemeters." Our data set is relatively coarse by comparison; it consists of the aggregated choices of households in 50 geographic markets.

[13] The two sides of the model are estimated separately because of the large computational costs of the viewer model and a data manipulation required to estimate the advertiser model. Berry and Waldfogel (1999) find, in a similar context, that simultaneous estimation results are "nearly identical" to separate estimation results.

[14] Saturdays and Sundays were excluded because UPN did not broadcast on either day and WB did not broadcast on Saturdays. The time slot 10:00–11:00 P.M. was excluded because FOX, UPN, and WB did not broadcast at that time. Two nights with unpredictable ad levels were also excluded: Thursday, May 1, which contained a presidential address and Thursday, May 15, which included a live basketball game.

It is also worth considering that the households that accept peoplemeters may differ from the households that accept audimeters and diaries. Diaries require viewers to manually record the programs they view each week, while peoplemeters require viewers to "log in" via remote control once or twice per hour. The relative disadvantage of the current data set is that it is not a panel of individual viewers, so individual viewer persistence is not directly observed. The relative advantage of this data set is its size. The sample on which it is based is large (NMR sampled 1,000–1,500 households in each of the 50 DMAs) and it spans four weeks and six broadcast networks.

Audience demographic data were collected from the U.S. Census for each Consolidated Metropolitan Statistical Area corresponding to a DMA. The demographics used are viewer age and household income, as these seem most likely to influence advertiser demand.

Data on ad quantities and prices come from TNS Media Intelligence/CMR. For each program in the sample, we observe the amount of time given to national advertising and the estimated price of a 30-second commercial during the program. The estimated prices were recorded from network reports of "the estimated cost of a 30-second spot" after the program aired. These data are commonly used by advertisers and agencies to budget for future advertising campaigns.[15]

Program characteristics were recorded by the author from videotapes of network programming made during the sample period. The videotapes were supplemented with data from a website, *TVTome.com*. Observable program characteristics include genre, thematic elements, main and supporting characters' demographics (including gender, race, age, and family structure), Program Age, setting, and current and past Emmy nominations.

National advertisement time exhibits substantial intratemporal variation in the sample. The mean ad level per half hour is 5:15 minutes and the average difference between the maximal and minimal network ad levels within a half hour is 2:49 minutes. Programs that are more attractive to viewers, relative to within-time-period competition, typically contain more ads. Broadcasters had limited themselves to three ad minutes per half hour until an antitrust suit in 1982 prohibited the practice, but nonprogram time has more than doubled since then.

[15] There is evidence that ad transaction prices are not uniform across advertisers. Advertisers can secure lower prices by procuring quantity discounts (Auletta 1992) and by using strong media negotiators (Bloom 2005). Networks keep actual transaction prices strictly confidential to avoid weakening their negotiating positions.

**Table 1 Descriptive Statistics: Ad Price, Ad Quantity, CPM, and Audience Size by Network and Day**

| | All Nets | ABC | CBS | FOX | NBC | UPN | WB |
|---|---|---|---|---|---|---|---|
| Adv. price (per 30 seconds; $000) | 147 (125) | 125 (40) | 179 (114) | 241 (186) | 212 (125) | 55 (26) | 71 (25) |
| Adv. seconds | 315 (74) | 367 (83) | 306 (55) | 284 (66) | 312 (81) | 323 (68) | 299 (58) |
| Cost per thousand viewers ($) | | | | | | | |
| Monday | 23 (6) | 20 (4) | 29 (7) | 27 (7) | 25 (7) | 19 (2) | 19 (2) |
| Tuesday | 26 (9) | 27 (5) | 14 (4) | 37 (6) | 34 (6) | 25 (3) | 19 (6) |
| Wednesday | 26 (9) | 25 (6) | 26 (13) | 29 (7) | 29 (8) | 21 (6) | 25 (7) |
| Thursday | 24 (11) | 25 (13) | 30 (13) | 15 (7) | 35 (7) | 19 (3) | 18 (4) |
| Friday | 19 (7) | 16 (3) | 20 (10) | 26 (8) | 20 (3) | 11 (1) | 23 (3) |
| Mon.–Fri. average | 24 (9) | 22 (7) | 23 (11) | 28 (9) | 28 (8) | 19 (6) | 21 (5) |
| Audience size (000 households) | | | | | | | |
| Monday | 5,846 (2,386) | 5,779 (1,275) | 8,862 (2,369) | 6,119 (1,723) | 6,928 (1,569) | 2,699 (250) | 4,690 (549) |
| Tuesday | 6,376 (3,087) | 5,299 (546) | 8,291 (551) | 11,303 (2,932) | 5,703 (1,509) | 2,768 (819) | 4,895 (670) |
| Wednesday | 6,613 (4,036) | 6,678 (1,134) | 5,409 (589) | 12,574 (4,984) | 8,551 (2,552) | 3,260 (725) | 3,210 (1,362) |
| Thursday[1] | 6,723 (4,659) | 4,281 (1,336) | 13,098 (2,547) | 4,269 (1,150) | 12,737 (828) | 3,807 (334) | 2,144 (411) |
| Friday | 4,206 (1,734) | 5,722 (549) | 5,610 (932) | 4,131 (916) | 5,574 (901) | 1,940 (290) | 2,260 (248) |
| Mon.–Fri. average | 5,868 (3,294) | 5,693 (1,182) | 7,716 (2,825) | 8,058 (4,573) | 7,361 (2,752) | 2,793 (791) | 3,584 (1,379) |

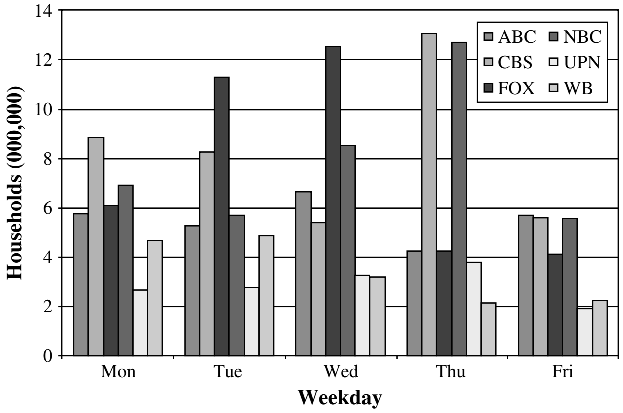
*Note.* Numbers in parentheses are standard deviations. All figures are averages over weeks and half hours, 8–10 P.M.
[1] Thursday, May 1, and Thursday, May 15, were removed from the sample.

Table 1 reports the means and standard deviations of ad price, ad quantity, cost per thousand households (CPM), and audience size by network and day. 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. FOX and NBC shared the highest average CPM ($28), while UPN charged the lowest ($19). FOX attracted the largest average audience followed by CBS, NBC, ABC, WB, and UPN.

Figure 2 shows average nightly audiences by network and day. The three highest-rated networks' nightly audiences varied considerably over days of the week. FOX dominated Tuesday and Wednesday with its *American Idol* franchise but was near average otherwise. CBS and NBC both attracted large audiences on Thursday night. The bottom three networks' audiences showed much less variation, with standard deviations less than half as large as CBS, FOX, and NBC.

Figure 3 shows how networks' prices per viewer varied over the course of the average week. NBC was

**Figure 2** Average Nightly Audience Size by Network

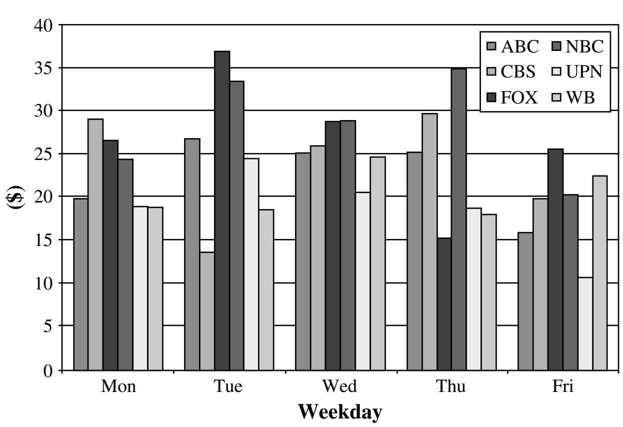



the only network to consistently charge in the upper echelon of per-viewer prices. CBS's take per viewer dipped precipitously on Tuesday while FOX's CPM fell dramatically on Thursdays. The three lowest-rated networks all charged per-viewer prices with lower means and less variation.

Table 2 lists program characteristics, their definitions, and descriptive statistics weighted by total half hours of programming in the sample. The table shows that the two most commonly programmed genres, by far, were Scripted Comedy and Psychological Drama. African-Americans appeared in prominent roles in nearly half of all shows while other minorities were Cast in leading roles less frequently (9%). The average program in the sample was 3.3 years old, was nominated for 1.4 Emmys in 2003, and had previously been nominated for 3.6 Emmys.

Table 3 shows the correlations between program audience size, advertising time, ad price, and genre. Ad price and audience size are highly collinear, with a correlation of 0.81. Advertising time is slightly negatively correlated with audience size (−0.06). Ad time is also negatively correlated with ad price (−0.10). Some of the genre correlations are interesting. Scripted Comedy is the only genre positively correlated with advertising time but it is not correlated with audience size or ad price. The Reality genre, by contrast, was the most positively correlated with ad price (0.26); it was positively correlated with audience size (0.14) but was negatively correlated with ad time (−0.1).

**Figure 3** Average Nightly Cost Per Thousand Households by Network




It is important to note that time given to local advertising and national tune-ins remains unobserved. We can gauge the effects of the omissions by referring to the 2001 Television Commercial Monitoring Report (TCMR) (AAAA/ANA 2002), which reports detailed information on nonprogram material in one week of data in November 2001. It shows that the average half hour of programming contained 8:22 minutes of nonprogram material, of which 4:52 minutes were national advertisements, 2:03 minutes were national tune-ins, and 1:24 minutes were local advertising time.[16] Local ad time is determined by long-term contracts between networks and affiliates and is nearly uncorrelated with national advertising time. In the TCMR sample, this correlation was −0.09. Networks do not set local advertising time based on contemporaneous program quality, so its omission is unlikely to bias estimates of viewer responsiveness to advertising.

The TCMR data suggest the absence of tune-in data may be more important. The program dummies $\xi_{jt}$ will capture the mean effect of tune-in levels across a program's episodes. This suggests a possible bias in programs' estimated attractiveness to viewers. Still, it seems likely that program characteristics will be stronger drivers of estimated program quality than tune-in levels.

### 4.2. Viewer Demand Endogeneity and Estimation

In the estimation of the viewer demand model presented in §3.1, mean tastes for unobserved program characteristics $\xi_{jt}$ and market-specific deviations from mean tastes for unobserved program characteristics $\eta_{mjt}$ are not observed by the econometrician. Yet these data are likely to be known by the television network and taken into account when it sets its advertising quantity $q_{jt}$. To avoid bias in the advertising response parameters, program dummies are used to estimate $\xi_{jt}$. Thus, the correlation between $q_{jt}$ and $\xi_{jt}$ exists between observed variables rather than between an observed variable and an error term.

There is the further concern that the network has knowledge of the market-specific tastes for unobserved program characteristics $\eta_{mjt}$ and uses that

[16] Local advertising time within a national broadcast is constant across affiliated stations. Affiliates vary in how much time they sell to local advertisers and how much they use to promote their own programs.

**Table 2 Variable Definitions and Descriptive Statistics**

| Variable name | Description | Mean | St. dev. |
|---|---|---|---|
| Adv seconds | Seconds of national advertisements aired during the program | 308 | (75) |
| Scripted comedy | =1 if the show is a scripted comedy | 0.31 | (0.46) |
| Action drama | =1 if the show is a scripted drama that contains action scenes | 0.10 | (0.30) |
| Psych drama | =1 if the show is a scripted drama that does not contain action scenes | 0.26 | (0.44) |
| Reality | =1 if the show is unscripted | 0.16 | (0.36) |
| News | =1 if the show is a news program or a newsmagazine | 0.06 | (0.24) |
| Movie | =1 if the show is a movie | 0.11 | (0.31) |
| African-American | =1 if at least one African-American main character | 0.44 | (0.50) |
| Other nonwhite | =1 if at least one nonwhite, non-African American main character | 0.09 | (0.29) |
| MC < 18 | =1 if at least one main character is under 18 | 0.20 | (0.40) |
| MC 18–34 | =1 if at least one main character is between 18 and 34 years old | 0.67 | (0.47) |
| MC 35–49 | =1 if at least one main character is between 35 and 49 years old | 0.62 | (0.49) |
| MC 50+ | =1 if at least one main character is over 50 years old | 0.17 | (0.38) |
| Married | =1 if at least one main character is married to another character | 0.19 | (0.40) |
| Single parent | =1 if at least one main character is single and has children | 0.10 | (0.30) |
| Female only | =1 if none of the main characters are male | 0.09 | (0.29) |
| Male only | =1 if none of the main characters are female | 0.23 | (0.42) |
| 50+% nonwhite | =1 if 50% or more of the show's cast is nonwhite | 0.20 | (0.40) |
| 25+% nonwhite | =1 if 25–49% of the show's cast is nonwhite | 0.39 | (0.49) |
| 10+% nonwhite | =1 if 10–24% of the show's cast is nonwhite | 0.55 | (0.50) |
| 50+% female | =1 if 50% or more of the show's cast is female | 0.46 | (0.50) |
| 25+% female | =1 if 25–49% of the show's cast is female | 0.74 | (0.44) |
| House | =1 if the show contains scenes set in a character's house | 0.23 | (0.42) |
| Apartment | =1 if the show contains scenes set in a character's apartment | 0.06 | (0.23) |
| Workplace | =1 if the show contains scenes set in a business or workplace | 0.34 | (0.48) |
| Outdoors | =1 if the show contains outdoor scenes | 0.49 | (0.50) |
| Studio | =1 if the show contains scenes set in a TV studio | 0.20 | (0.40) |
| Cop | =1 if the show has some law enforcement element | 0.10 | (0.31) |
| Sci–fi | =1 if the show contains elements of science fiction (i.e., Star Trek) | 0.06 | (0.24) |
| Supernatural | =1 if the show contains supernatural elements (i.e., angels, witchcraft) | 0.07 | (0.26) |
| Age | # of years since the show's debut | 3.28 | (3.5) |
| 2003 Emmy noms | 2004 Emmy nominations | 1.4 | (3.1) |
| Past Emmy noms | All pre-2004 Emmy nominations | 3.6 | (11.9) |
| ABC, CBS, FOX, NBC, UPN, WB | Network-specific dummy variables | | |
| Mon, Tue, Wed, Thu, Fri | Day-specific dummy variables | | |
| 1st HH, 2nd HH, 3rd HH, 4th HH | Dummy variables for 1st half hour of prime time, 2nd half hour of prime time, 3rd, 4th, etc. | | |

knowledge in setting its advertising quantity. There are two reasons this could be the case. First, larger markets' tastes matter more because large markets contain more viewers. Second, even after controlling for audience size and demographics (like age and income), the network might have preferences over the geographic distribution of the viewers in an audience.

To correct for the first concern, we replace $\eta_{mjt}$ in Equation (1) with $\tilde{\omega}_m \eta_{mjt}$, where $\tilde{\omega}_m \equiv \omega_m/((1/M) \cdot \sum_{n=1}^{M} \omega_n)$, $\omega_m$ is the number of households in market $m$, and $M$ is the number of markets in the sample ($M = 50$). With this correction, the program-specific fixed effect is estimated as the mean effect of unobserved program characteristics across all households in the sample rather than the mean effect across DMAs in the sample. The data are defined at the market level rather than the household level, so this correction is necessary to ensure that the effects captured by $\xi_{jt}$ and $\eta_{mjt}$ are consistent with their definitions in §3.1.

There is no parsimonious strategy to control for the second concern listed above. We acknowledge this as a potential objection but it does not seem to be a first-order issue. Individual advertisers are certain to have preferences over the geographic distribution of audience members but if those prefer-

**Table 3 Correlations between Audience Size, Ad Price, Ad Seconds, and Program Genres**

| | Audience size | Adv. price | Adv. seconds |
|---|---|---|---|
| Audience size | 1.00 | | |
| Adv. price | 0.81 | 1.00 | |
| Adv. seconds | −0.06 | −0.10 | 1.00 |
| Comedy | 0.00 | 0.02 | 0.44 |
| Action | 0.04 | 0.02 | −0.19 |
| Drama | −0.12 | −0.16 | −0.23 |
| Reality | 0.14 | 0.26 | −0.10 |
| News | 0.00 | −0.09 | −0.13 |
| Movie | −0.05 | −0.07 | −0.02 |

*Sources.* Nielsen Media Research; TNS Media Intelligence/CMR.

ences are very strong, they are likely to buy audiences in the "spot" television market (advertising on local stations) in place of national advertising. Additionally, geographic preferences are likely to vary across advertisers but network incentives are influenced by the cumulative preferences of all advertisers in the market.

We now turn to the details of viewer demand estimation. The estimation routine is similar to that introduced by Berry, Levinsohn, and Parkes (BLP 1995) and described in detail by Nevo (2000, 2001). The main idea behind this estimation routine is to numerically solve the audience rating functions defined by Equation (3) for programs' mean utility levels and to use these imputed mean utilities in a moment condition. Use of this algorithm confers two primary benefits. First, it defines the objective function as a smooth function of the parameters, which reduces simulation error. Second, it significantly reduces the number of parameters to be estimated nonlinearly. This greatly speeds computation.[17]

New notation is needed to facilitate explanation of the estimation routine. Let $N$ be the number of market/network/day/half hour market shares observed in the sample, and let $P$ be the number of programs in the sample. Let $H$ be a $N \times P$ matrix of program dummies and $\psi$ be the $P \times 1$ vector of mean program utilities that are constant across viewers and time periods in which the program airs. Let two partitions of $x_{mjt}$ be labeled $x_{jt}$, which includes program dummies, and $\tilde{x}_{mjt}$, which contains advertising level, audience flow effects, the last half hour dummy, and day-, time-, and market-specific fixed effects. Label two partitions of $\beta$ as $\beta_1$, which interacts with $x_{jt}$, and $\beta_2$, which interacts with $\tilde{x}_{mjt}$. Denote the set of parameters to be estimated using the Generalized Method of Moments (GMM) as $\theta = \{\psi, \alpha, \beta_2, \Pi, \Sigma\}$. Define the mean utility that viewers in city $m$ derive from watching network $j$ at time $t$ as $\delta_{mjt} = \psi_p + q_{jt}\alpha + \tilde{x}_{mjt}\beta_2 + \tilde{\omega}_m \eta_{mjt}$, where $\psi_p$ is the *non-ad mean utility* (NAMU) of the program on network $j$ at time $t$. $\psi_p$ will later be used as an instrument to identify endogenous advertiser demand parameters. We now describe the estimation procedure.

We use simulation to approximate the integral in Equation (3). We draw simulated viewer demographics $D_{im}$ from the market-specific nonparametric marginal distributions in the census microdata, and draw $\nu_{im}$ from a standard multivariate normal distribution. The predicted audience rating of network $j$ at time $t$ in market $m$ is, then, the fraction of the simulated viewers in that market for which network $j$'s time $t$ program maximizes utility, conditional on the time $t$ programs aired by all networks and a guess of the parameter set $\theta$.

We use the contraction mapping suggested by BLP to solve for the $J$ vector of mean utilities $\tilde{\delta}_{m\cdot t}(\theta)$ that, for a given value of $\theta$, equates predicted market shares to observed market shares in market $m$ at time $t$,

$$s_{m\cdot t}(\tilde{\delta}_{m\cdot t}(\theta)) = S_{m\cdot t}.$$

The error term is the market-specific deviation from mean tastes for network $j$'s time $t$ broadcast, written as $\tilde{\eta}_{mjt}(\theta) = (1/\tilde{\omega}_m)(\tilde{\delta}_{mjt}(\theta) - \psi_p - q_{jt}\alpha - \tilde{x}_{mjt}\beta_2)$.

Next, we construct the moment conditions $EX'\tilde{\eta}(\theta) = 0$, where $X$ is a matrix of instruments defined below, and $\tilde{\eta}(\theta)$ is an $N$ vector of the $\tilde{\eta}_{mjt}(\theta)$'s. The GMM estimate of $\theta$ is:

$$\hat{\theta} = \arg\min_{\theta} \tilde{\eta}(\theta)' X A^{-1} X' \tilde{\eta}(\theta),$$

where $A$ is a positive-definite weighting matrix.[18]

We use the minimum-distance procedure suggested by Nevo (2000) to decompose the program-level mean utilities $\hat{\psi}$ into the taste parameters associated with observed program characteristics ($\beta_1$) and unobserved program characteristics ($\xi$). Let $X_p$ be a $P \times K$ matrix of program characteristics and let $\hat{\psi}$ be the $P \times 1$ vector of estimated mean program utilities. Then, since $\psi = X_p\beta_1 + \xi$, the estimate of $\beta_1$ is computed as $\hat{\beta}_1 = (X_p'\hat{\Omega}^{-1}X_p)^{-1}X_p'\hat{\Omega}^{-1}\hat{\psi}$ and $\hat{\xi} = \hat{\psi} - X_p\hat{\beta}_1$, where $\hat{\Omega}$ is the estimated covariance matrix of $\hat{\psi}$. This minimum-distance procedure is analogous to a GLS regression wherein the dependent variable is the set of estimated program mean utilities, the independent variables are the programs' observed characteristics, and the number of observations is equal to the number of programs in the sample.

The columns of the $X$ matrix must all be orthogonal to the vector of markets' deviations from mean program utility $\eta$. The obvious candidates for inclusion in $X$ are the program dummies. These are valid instruments because the inclusion of $\tilde{\omega}_m$ ensures that the mean program utilities are mean independent of $\eta$. $X$ also includes the market, day, and time dummies, and interactions between some elements of $x_{mjt}$ and moments of the market-specific distributions of viewer demographics. For each variable $k$ whose effect on utility varies with consumer demographics, and for each observable viewer demographic $d$,

[17] The method traditionally used to estimate random coefficient logit models is to define the objective function as the difference between the predicted market shares and the observed market shares. The advantages of the BLP estimation algorithm listed here are relative to the traditional estimation method.

[18] The most efficient choice of $A$ is the covariance matrix of the moments. We follow the standard method of setting $A = X'X$ to obtain a consistent estimate of $\theta$. We then use this estimate to construct a consistent estimate of the asymptotically efficient weighting matrix $EX'\tilde{\eta}(\hat{\theta})\tilde{\eta}(\hat{\theta})'X$, which we use to obtain the final estimate of $\theta$.

$X$ includes $X_d^k$, an $N$ vector whose $n$th element is $x_{jt}^k \mathrm{E}d_{im}$, where $\mathrm{E}d_{im}$ is the expected value of viewer demographic $d$ in market $m$. Note that the audience flow effects can not be used, as they are highly likely to be correlated with market-specific tastes for unobserved program characteristics. To meet necessary identification conditions, we also include an $N \times J$ matrix $Q$ whose $n$th row contains the time $t$ ad levels of the associated network and its $J-1$ competitors. Inclusion of advertising levels in $X$ is justified by the discussion at the beginning of this subsection.

We impose some zero restrictions on $\Pi$ and $\Sigma$ to limit the number of parameters that enter the GMM objective function nonlinearly.[19] Three regressors are assumed to interact with viewer demographics: the outside option dummy, the nonbroadcast network television dummy, and the network's ad quantity. The first two variables were chosen to improve the reliability of predicted substitution patterns among various options, and the third variable improves the reliability of predicted audience changes resulting from varying levels of advertisements.

### 4.3. Advertiser Demand Endogeneity and Estimation

Advertiser demand parameters are estimated using instrumental variables. A limited-information approach is used in place of a full-information approach, based on the network supply model presented in §3.3, for two reasons. First, the assumptions underlying the model in §3.3 are strong and potentially unreliable. Second, there are no good instruments available for the unobserved elements in the networks' first-order conditions. We describe here the endogeneity concerns in advertisement demand estimation, the instruments used, a necessary data manipulation, and the estimation strategy.

The error in the advertisement demand equation $\phi_s$ is assumed to primarily reflect unobserved program and audience characteristics that influence advertiser demand for ads.[20] For example, $\phi_s$ might reflect viewers' level of "engagement" with the program, or it might represent the fraction of audience members who drink cola. The network is likely to have partial knowledge of $\phi_s$ and to take it into account when setting its advertising level. Thus, $q_s$ could be correlated with $\phi_s$ and, because audience size depends on $q_s$, $V_s$ could also be correlated with $\phi_s$. (Note, this second correlation is due solely to the direct dependence of $V_s$ on $q_s$. There is no obvious reason to think that audience size is systematically related to advertiser preferences for unobserved program and audience demographics.)

Instruments are required to obtain unbiased estimates of the advertiser demand parameters in Equation (4). Instruments must meet two requirements for validity: They should be correlated with $V_s$ and $q_s$, and not correlated with advertiser preferences for unobserved audience demographics $\phi_s$. We use the common strategy of using program characteristics as instruments. We use two sets of program characteristics as instruments: the observed program characteristics thought to influence advertiser demand $x_s$, and the non-ad mean utility (NAMU) of the program and its within-time-period competition estimated by the viewer demand model. (In the notation of §4.2, the NAMU is $\hat{\psi}_p$ or $X_p\hat{\beta}_1 + \hat{\xi}$.) These instruments meet the first requirement for validity—they enter directly into the audience rating functions of Equation (3), so they are correlated with $V_s$. They are correlated with $q_s$ because a network's optimal advertising level depends partly on its audience size, as illustrated by Equations (6) and (7). These instruments can also be assumed to meet the second requirement—that they are not correlated with advertiser preferences for unobserved program and audience characteristics. NAMU is the mean utility of program consumption across all viewers, which influences advertiser preferences only through the program's audience size. And the program characteristics $x_s$ are included in ad demand (4) so that $\phi_s$, by definition, consists of advertiser preferences for those program and audience characteristics that are not included in $x_s$.

Before estimating advertiser demand parameters, we modify the data to account for how the market operates. Audiences are sold and prices are reported at the program level but networks distribute ads within a program based on strategic considerations. Longer programs tend to contain more ads in their latter stages, as viewers are relatively more "captive" at that point and the program is less likely to attract new viewers from competing networks. This phenomenon is likely to be an important source of variation in the advertisement price/quantity relationship. To account for it, we average each program's audience characteristics over the half hours in which the program aired, so that the unit of observation is a network/day/program, rather than a network/day/half hour.

To estimate advertisement demand parameters, we interact the instruments described above with the ad demand function residual and use GMM to solve the moment conditions. Let $Z_s$ be a vector that includes the set of instruments for show $s$ described above,

[19] If there are $\bar{d}$ elements in $D_i$, each observable program characteristic whose effect on utility varies with viewer demographics adds $\bar{d}+1$ nonlinear parameters. This is a problem because computation time increases at an increasing rate with the number of parameters that enter the objective function nonlinearly.

[20] $\phi_s$ could also be assumed to reflect measurement error in advertisement prices. This ad price data is often used by media buyers to plan future purchases, so any measurement error is presumably small.

and let $Z$ be a matrix constructed by stacking the $Z_s$'s. The moments, then, are $\mathrm{E}Z'\phi = \mathbf{0}$, where $\phi$ is a vector of the $\phi_s$'s. The GMM estimates of advertisement demand and supply parameters are:

$$\hat{\lambda} = \arg\min_{\lambda} \phi' Z B^{-1} Z' \phi,$$

where $B$ is the covariance matrix of the moments. Following standard practice, we first set $B = Z'Z$ to obtain a consistent estimate of the asymptotically efficient weighting matrix and then use $\hat{B}$ to re-estimate the model and obtain the final results.

### 4.4. Network Supply Inferences

Demand parameter estimates are used in conjunction with networks' first-order conditions to infer unobserved tune-in levels and benefits.[21] The basis for this procedure is the structural model of network competition described in §3.3, which stipulates that a profit-maximizing network will air tune-ins and ads so as to equate its marginal benefit from each activity. Solving Equation (8) for $\tau_s$,

$$\hat{\tau}_s = \frac{p_s + q_s \hat{\lambda}_q}{V_s}. \tag{10}$$

We can use $\hat{\tau}_s$ to make inferences about tune-in levels. From either of the network's first-order conditions (Equations (6) and (7)), it can be shown that:

$$\hat{r}_s = \frac{q_s \lambda_V (\partial V_s/\partial q_s) + V_s \hat{\tau}_s}{-(\partial V_s/\partial q_s)\hat{\tau}_s}. \tag{11}$$

We report test results in §5.5 that indicate the tune-in inferences are reliable.

## 5. Empirical Findings

This section reports and discusses estimation results. It begins with an estimation of a Multinomial Logit model (MNL) of viewer demand for television programs, then presents the results from the full Random Coefficients Logit (RCL). Advertiser demand parameter estimates, comparisons of advertiser and viewer preferences for program characteristics to networks' observed programming choices, and a discussion of tune-in inferences then follow.

### 5.1. Viewer Demand: Multinomial Logit Results

The Multinomial Logit is a special case of the full Random Coefficients Logit model wherein parameter matrices $\Pi$ and $\Sigma$ are restricted to zero. While the MNL has unrealistic substitution patterns, its ease of computation makes it a good tool for comparing results across multiple specifications.

In the specifications reported below, advertising quantity ($q_{jt}$) enters viewer utility linearly. We estimated several alternate models but found no evidence that ad levels have nonlinear effects on utility. Nor was there any evidence that the number of commercial breaks affects program audience size (controlling for ad quantity). We therefore maintain the assumption that viewers' marginal utility from advertising is linear in advertising quantity.[22]

Table 4 reports estimation results from eight MNL specifications. Parameter estimates are computed from ordinary least squares (OLS) regressions of transformed log ratings, $\log(s_{mjt}) - \log(s_{m0t})$, on alternate mean utility specifications. The specifications are permutations that include program-specific fixed effects, market dummies, and market-size indices $\tilde{\omega}_m$. Program dummies control for correlation between ad levels and unobserved program quality; market dummies approximate the way the full model controls for unobserved viewer heterogeneity (since simulated viewer demographics are drawn from market-specific distributions); and market-size indices control for the effects of market size on networks' ad quantity decisions.

The MNL model fits the data very well; the Adjusted $R^2$ ranges from 0.79 to 0.91. This high degree of fit suggests that there is not much unobserved viewer heterogeneity for the random coefficients to explain.

Table 4 shows that the point estimate of advertising time on viewer utility is negative across all specifications. Inclusion of program dummies has the effect of making this point estimate significant and quadrupling it. This is consistent with standard models of consumer demand. Failure to control for unobserved product characteristics biases the price coefficient toward zero. These findings and the improved fit of the model validate the endogeneity controls described in §4.2. They show that networks know their programs' qualities and take them into account when setting advertising levels, so a model that ignores this strategic behavior will yield biased parameter estimates.

The finding that higher aggregate advertising levels are associated with smaller program audiences is intuitive. However, it contrasts with Kaiser and Wright's (2006) finding that women's magazine readership grows with the number of advertising pages the magazines contain. The difference in the sign of the effect might be attributable to the two mediums'

[21] The sample of videotapes was incomplete so tune-in levels could not be recorded from the tapes. Thirty seven percent of network program hours in the sample were available on tape.

[22] The audience data's relative insensitivity to viewer zapping of commercial breaks is discussed in §4.1. This is a specification test, not a formal test of whether the number of ad breaks affects viewer zapping.

**Table 4 Multinomial Logit Viewer Demand Parameter Estimates**

| | (i) | (ii) | (iii) | (iv) | (v) | (vi) | (vii) | (viii) |
|---|---|---|---|---|---|---|---|---|
| Advertising seconds | −5.9E−5 | −2.2E−4* | −5.7E−5 | −2.2E−4* | −5.4E−5 | −2.2E−4* | −4.9E−5 | −2.1E−4* |
| | (4.7E−5) | (4.6E−5) | (4.3E−5) | (4.2E−5) | (4.7E−5) | (4.6E−5) | (4.3E−5) | (4.1E−5) |
| Lead-in audience | 5.133* | 4.756* | 4.735* | 4.288* | 5.293* | 4.920* | 4.832* | 4.380* |
| | (0.083) | (0.081) | (0.077) | (0.075) | (0.085) | (0.084) | (0.078) | (0.075) |
| Lead-out audience | 4.649* | 4.478* | 4.193* | 3.958* | 4.864* | 4.736* | 4.170* | 3.950* |
| | (0.083) | (0.082) | (0.078) | (0.075) | (0.084) | (0.082) | (0.077) | (0.074) |
| Nonbroadcast network | −39.564* | −2.868* | −41.838* | −2.571* | −38.855* | −2.942* | −41.755* | −2.547* |
| | (3.378) | (0.027) | (3.125) | (0.030) | (3.406) | (0.026) | (3.093) | (0.025) |
| Last half hour[a] | −4.6E−3 | 3.9E−3 | −8.1E−3 | 2.0E−3 | 2.0E−3 | 0.010 | −3.4E−3 | 5.6E−3 |
| | (0.010) | (0.011) | (9.3E−3) | (9.8E−3) | (0.010) | (0.011) | (9.1E−3) | (9.7E−3) |
| Includes program dummies? | — | yes | — | yes | — | yes | — | yes |
| Includes market dummies? | — | — | yes | yes | — | — | yes | yes |
| Weighted by market size? | — | — | — | — | yes | yes | yes | yes |
| Adjusted $R^2$ | 0.791 | 0.813 | 0.821 | 0.844 | 0.864 | 0.876 | 0.897 | 0.911 |

*Notes.* Reported here are OLS regressions wherein the dependent variable is $\log(s_{mjt}) - \log(s_{m0t})$. Number of observations: 23,588. All specifications included day (Tues, Wed, Thurs, Fri) and half hour (8:30, 9:00, 9:30) dummies. When program dummies are not included, observable program characteristics (genre, setting, Main Character and Cast demographics, age, and current and past award nominations) and a constant are included.

* = Significant at the 1% level.

[a]Indicates whether a program was in its last half hour only for those programs that lasted 60 minutes or longer.

varying levels of consumer control over their advertising exposures. Magazine readers are free to choose how much time they spend with each ad based on how much they value it. Television viewers have less perfect control over their advertising exposure. An alternate explanation may be found in the level of audience heterogeneity. If television audiences are heterogeneous relative to magazine readerships, the ads provided may be less appealing to the average audience member.

Audience flow effects and the last half hour dummy control for viewing persistence. Numerous studies find evidence of this phenomenon. Among these, Shachar and Emerson (2000) find that viewer switching costs grow as they become more experienced with a program and are higher when they have fewer new programs to sample on competing networks. Moshkin and Shachar (2002) find that uncertainty reduction through program consumption plays a larger role in explaining viewing persistence than state dependence. Zhou (2004) shows that networks tend to air more and longer commercial breaks toward the end of a popular program. Table 4 shows that lead-in and lead-out effects are positive and significant. The last half hour dummy is not significant, likely due to small audience losses at the end of a program being offset by smaller numbers of viewers tuning in (Shachar and Emerson 2000). The level of aggregation of viewers in the current data set makes this model a less reliable judge of the presence of state dependence than other studies that observe individual viewer behavior. The interested reader is referred to Moshkin and Shachar (2002).

## 5.2. Viewer Demand: Random Coefficients Logit Results

Here, we present estimation results of the full Random Coefficients Logit model of viewer demand. The model fit the data quite well, with a pseudo $R^2$ of 0.86 and average relative error of 0.22.

Table 5a shows the random utility parameters estimates. Consistent with the MNL, the mean effect of advertising seconds is negative and significant. The heterogeneity parameters associated with tastes for advertising are not significant, indicating that

**Table 5a Random Utility Parameter Estimates**

| | Means ($\beta$'s) | Standard deviations ($\sigma$'s) | Interactions with demographic variables ($\pi$'s) | | |
|---|---|---|---|---|---|
| | | | Income | Age | Age[2] |
| Advertising seconds | −1.5E−3** | −0.478 | −4.1E−5 | −0.092 | 5.8E−4 |
| | (2.1E−4) | (9.269) | (4.9E−5) | (1.799) | (0.019) |
| Nonbroadcast network | −10.024** | −0.501 | 7.1E−14 | 0.258 | −6.0E−3 |
| | (1.387) | (0.935) | (1.1E−4) | (0.389) | (4.6E−3) |
| TV off | 0[a] | 1.819** | 0[b] | 0[b] | 0[b] |
| | | (0.479) | | | |
| GMM objective | | 8.120 | | | |
| $\chi^2$ Goodness of fit (degrees of freedom) | | 1.00E+10 (163) | | | |
| Pseudo $R^2$ | | 0.8588 | | | |

*Notes.* Number of observations: 23,588.

**Significant at the 1% confidence level.

[a]The mean utility of "TV off" is normalized to zero. It is not separately identified from the mean utilities of the "inside" options.

[b]Interactions between "TV off" and interactions with consumer demographics are, in practice, very difficult to estimate separately from the market-specific fixed effect. They are set to zero.

**Table 5b Nonrandom Utility Parameter Estimates**[a]

| Variable | Estimate (std. error) | Variable | Estimate (std. error) |
|---|---|---|---|
| Lead in | 3.144 (2.152) | Main characters: male only[b] | 0.112** (0.037) |
| Lead out | 20.917** (6.055) | Cast: 50+% nonwhite[b] | −0.062 (0.044) |
| Last half hour[c] | 0.026 (0.038) | Cast: 25+% nonwhite[b] | −0.046 (0.040) |
| Tuesday | −0.439 (0.507) | Cast: 10+% nonwhite[b] | −0.094** (0.047) |
| Wednesday | 0.176 (0.099) | Cast: 50+% female[b] | −0.046 (0.038) |
| Thursday | 0.322** (0.136) | Cast: 25+% female[b] | 0.060 (0.047) |
| Friday | 0.772** (0.287) | Setting: house[b] | 0.151** (0.040) |
| 2nd prime time half hour | −0.246** (0.092) | Setting: apartment[b] | 0.042 (0.057) |
| 3rd prime time half hour | 0.124** (0.042) | Setting: workplace[b] | −0.112** (0.027) |
| 4th prime time half hour | 0.105 (0.070) | Setting: on location[b] | −0.073** (0.026) |
| Constant[b] | 1.664 (6.644) | Special[b] | 0.080 (0.041) |
| Genre: scripted comedy[b] | −0.214** (0.028) | Cop[b] | 5.8E−3 (0.039) |
| Genre: action drama[b] | 0.112** (0.037) | Sci-fi[b] | 0.286** (0.065) |
| Genre: reality[b] | −0.105** (0.037) | Supernatural[b] | 0.035 (0.040) |
| Genre: news[b] | 0.046 (0.053) | SeasonFirstAired[b] | −2.9E−3 (3.3E−3) |
| Genre: movie[b] | −0.203** (0.047) | 2003 Emmy nominations[b] | −0.042** (5.5E−3) |
| Main characters: African-American[b] | 0.139** (0.039) | Past Emmy nominations[b] | 0.012** (1.1E−3) |
| Main characters: other nonwhite[b] | −0.032 (0.040) | ABC[b] | 0.177** (0.042) |
| Main characters: <18 years old[b] | 0.027 (0.045) | FOX[b] | 0.040 (0.038) |
| Main characters: 18–34 years old[b] | 0.144** (0.031) | NBC[b] | −0.026 (0.036) |
| Main characters: 35–49 years old[b] | 0.043 (0.026) | UPN[b] | −0.125** (0.049) |
| Main characters: married[b] | −0.052 (0.038) | WB[b] | 0.020 (0.050) |
| Main characters: single parent[b] | −0.012 (0.031) | Minimum distance $\chi^2$ | 1.23E+07 |
| Main characters: female only[b] | −0.035 (0.035) | Minimum-distance pseudo $R^2$ | 0.6447 |

**Significant at the 1% confidence level.
[a]Parameters are GMM estimates, except where noted.
[b]Minimum-distance estimate.
[c]Defined only for those programs whose duration exceeds 30 minutes.

market-level differences in age and income do not affect audience responsiveness to advertising.

The mean taste for nonbroadcast network programming is negative. This shows that the "TV off" option is generally preferred to nonbroadcast options like cable networks. There are two reasons for this. First, the TV off market share exceeds the nonbroadcast TV market share in 77% of the market/time periods in the sample. Second, advertising quantities on cable networks are not observed, so their effect is captured by the nonbroadcast TV fixed effects. The random utility parameters associated with the nonbroadcast TV option were not significant.

Table 5a indicates that unobserved heterogeneity plays a significant role in tastes for the nontelevision option. The other eight random utility parameter estimates are not significant. Consideration was given to removing the random effects from the demand model. The following null hypotheses were tested: $\Pi = 0$, $\Sigma = 0$, and $\{\Pi = 0\} \cap \{\Sigma = 0\}$. Newey-West "D" tests and Wald tests reject all hypotheses at the 1% confidence level. Yet some parameter estimates are imprecisely estimated. The random utility parameter estimates were therefore not used to construct unobserved audience demographics $d_s$.

Table 5b reports the effects of audience flow, program, and time characteristics. Many studies have found that audience flow effects are very strong predictors of audience ratings. Consistent with that research, the lead-out effect is very large, positive, and significant. The lead-in effect is not significant but this is probably due to its high correlation with the lead-out effect.

Day effects indicate that viewers exhibit a statistically significant preference for watching television on Thursday and Friday nights. The large Friday effect contrasts with the broadcast networks' tendency to air low-quality programs on Friday nights; this is explored further in §5.4.

The half hour parameter estimates indicate that the 8:00–8:30 P.M. (EST) time block is preferred to the 8:30–9:00 P.M. block. The 9:00–9:30 P.M. block is preferred to both of the preceding half hours. These results agree with Goettler and Shachar (2001) who find that television utility peaks in the first quarter of each hour.

Program genre is a powerful predictor of audience size. Viewer preference order (using point estimates) is Action Drama, News, Psychological Drama, Reality, Movie, and Scripted Comedy. News is the only genre whose effect is not significantly different from Psychological Drama.

Main character demographics influence audience size. Viewers prefer programs that include African-American and 18- to 34-year-old main characters. Main characters in other age groups and of other minority races have no significant effect on audience size. Main character marital status does not have a significant effect on audience size. Programs without female main characters fared slightly better, but exclusion of male main characters has no significant effect.

**Table 5c Market-Effect Estimates**

| Market | Size (000 households) | Estimate (std. error) | Market | Size (000 households) | Estimate (std. error) |
|---|---|---|---|---|---|
| New York, NY | 9,343 | 2.98** (0.80) | San Diego, CA | 1,004 | 0.00 (0.38) |
| Los Angeles, CA | 6,834 | −0.08 (0.58) | Hartford and New Haven, CT | 1,566 | 2.72** (0.32) |
| Chicago, IL | 3,782 | 3.00** (0.54) | Charlotte, NC | 1,248 | 0.17 (0.26) |
| Philadelphia, PA | 3,748 | 0.33 (0.63) | Raleigh-Durham (Fayetteville), NC | 1,324 | 3.03** (0.28) |
| San Francisco-Oakland-San Jose, CA | 3,866 | 3.22** (0.48) | Nashville, TN | 1,023 | 0.04 (0.38) |
| Boston, MA (Manchester, NH) | 2,354 | 0.07 (0.45) | Milwaukee, WI | 957 | 2.87** (0.50) |
| Dallas-Ft. Worth, TX | 2,611 | 2.82** (0.38) | Cincinnati, OH | 1,309 | −0.11 (0.40) |
| Washington, D.C. (Hagerstown, MD) | 3,598 | 0.27 (0.68) | Kansas City, MO | 1,091 | 2.78** (0.49) |
| Atlanta, GA | 2,246 | 3.14** (0.41) | Columbus, OH | 1,111 | −0.06 (0.34) |
| Detroit, MI | 2,481 | 0.36 (0.33) | Greenville-Spartanburg, SC Asheville, NC, Anderson, SC | 1,124 | 3.22** (0.41) |
| Houston, TX | 2,009 | 2.92** (0.42) | Salt Lake City, UT | 780 | −0.35 (0.58) |
| Seattle-Tacoma, WA | 1,788 | 0.02 (0.31) | San Antonio, TX | 862 | 2.59** (0.77) |
| Tampa-St. Petersburg, FL | 1,837 | 0.09 (0.59) | Grand Rapids-Kalamazoo-Battle Creek, MI | 950 | −0.25 (0.77) |
| Minneapolis-St. Paul, MN | 1,908 | 0.17 (0.27) | West Palm Beach-Ft. Pierce, FL | 1,390 | 3.15** (0.46) |
| Cleveland-Akron (Canton), OH | 2,380 | 3.12** (0.48) | Birmingham (Anniston and Tuscaloosa), AL | 831 | 0.26 (0.60) |
| Phoenix (Prescott), AZ | 1,930 | 0.02 (0.51) | Norfolk-Portsmouth-Newport News, VA | 690 | 2.94** (0.52) |
| Miami-Ft. Lauderdale, FL | 1,981 | 2.83** (0.35) | New Orleans, LA | 731 | −0.07 (0.35) |
| Denver, CO | 1,991 | −0.08 (0.21) | Memphis, TN | 886 | 2.68** (0.58) |
| Sacramento-Stockton-Modesto, CA | 2,328 | 3.20** (0.36) | Buffalo, NY (including Canadian audiences) | 795 | 1.84** (0.55) |
| Orlando-Daytona Beach-Melbourne, FL | 1,688 | 0.25 (0.60) | Oklahoma City, OK | 768 | 2.69** (0.37) |
| Pittsburgh, PA | 1,656 | 2.60** (0.58) | Greensboro-High Point-Winston Salem, NC | 973 | 0.05 (0.54) |
| St. Louis, MO | 1,349 | 0.04 (0.39) | Harrisburg-Lancaster-Lebanon-York, PA | 974 | 2.48** (0.50) |
| Portland, OR | 1,229 | 2.99** (0.53) | Providence, RI-New Bedford, MA | 2,138 | −0.26 (0.45) |
| Baltimore, MD | 2,611 | 0.17 (0.28) | Albuquerque-Santa Fe, NM | 659 | 2.87** (0.73) |
| Indianapolis, IN | 1,319 | 2.88** (0.30) | | | |

**Significant at the 1% confidence level.

Program setting and thematic elements also influence viewer choices. Viewers most prefer programs with House settings, followed by Apartment, Studio, Outdoors, and Workplace. Workplace scenes may be unpopular because they interfere with viewers' recreational use of television, while House and Apartment scenes may be popular for the opposite reason. Three of four setting effects are significant but they are relatively small. Of the thematic effects, Sci-fi has a large and positive impact on viewership but Cop and Supernatural are not significant.

Past Emmy award nominations increase a program's attractiveness to viewers. Emmy nominations for 2003 were not announced until after the sample concluded and actually correlate negatively with program viewership. This suggests that Emmy

**Table 5d Median Own- and Cross-Advertising-Level Audience Elasticities[a]**

| Network | ABC | CBS | FOX | NBC | UPN | WB |
|---|---|---|---|---|---|---|
| ABC | −5.68 | 0.22 | 0.20 | 0.29 | 0.31 | 0.26 |
| CBS | 0.74 | −2.55 | 0.28 | 0.39 | 0.38 | 0.30 |
| FOX | 0.71 | 0.32 | −2.49 | 0.38 | 0.40 | 0.31 |
| NBC | 0.64 | 0.28 | 0.26 | −2.59 | 0.38 | 0.31 |
| UPN | 0.47 | 0.20 | 0.19 | 0.24 | −7.81 | 0.24 |
| WB | 0.63 | 0.26 | 0.26 | 0.33 | 0.34 | −5.63 |
| Nonbroadcast network TV | 0.20 | 0.08 | 0.07 | 0.09 | 0.10 | 0.08 |
| Outside option (TV off) | 0.70 | 0.20 | 0.24 | 0.24 | 0.64 | 0.55 |

[a]Table entry $i, j$ reports the median-estimated elasticity of network $i$'s national audience given an unanticipated 10% decrease in network $j$'s observed advertising level. (If rival networks anticipate the change in $q_i$, they will change their ad quantities in the same direction and the elasticity will be smaller in absolute value.) Medians are over days and half hours.

nominations play an important role in signaling program quality to both viewers and networks. The negative effect of current-year Emmy nominations likely reflects unfavorable timeslots and tune-in support given to unproven new shows.

Table 5c reports market-specific fixed effects. The excluded DMA is Louisville, Kentucky. Television utility varies considerably across markets and is significantly different from the excluded DMA in 24 of 49 markets. The highest point estimates are found in diverse markets such as San Francisco, West Palm Beach, and Atlanta. There are no clear correlations between television utility and DMA size or location.

Table 5d shows median own- and cross-advertising elasticities by network. Audience sizes are responsive to advertising levels. For example, if CBS unilaterally decreases the advertising levels in all of its programs by 10%, its median audience gain would be about 25% (assuming no competitive reactions). In general, more highly rated networks gain viewers at a slower rate from falling advertising levels. The three lowest rated networks have the most elastic audience demand but audience elasticity is not uniformly related to network audience ratings. ABC had a higher average program rating than the WB but a more elastic viewer demand. This is probably because the WB tended to provide narrower niche programming than ABC.

Higher rated networks' audiences were generally less responsive to advertising quantity. This is because programs are differentiated vertically as well as horizontally. Programs with high levels of vertical differentiation garner larger, more diverse audiences and have fewer good substitutes. This reasoning may explain why all six broadcast networks are more likely to lose viewers to the nontelevision option than to the nonbroadcast network option. Cable programs tend to exhibit high levels of horizontal differentiation and low levels of vertical differentiation, so they may be poor substitutes for broadcast network programs.

The results in Table 5d should be read with three caveats in mind. First, ABC, UPN, and WB audiences were smaller on average than those of CBS, FOX, and NBC, so percentage changes are based on different bases across networks. Second, these are point elasticities and would not be constant if calculated across the range of possible changes in advertising levels. Third and most importantly, these are estimated responses to deviations from equilibrium. Network advertising levels are strategic complements, so any anticipated deviation from equilibrium would invite a competitive response. Such a response would result in a smaller change in audience than the estimated response to unilateral deviations shown in the table.

### 5.3. Advertiser Demand Parameter Estimates

Advertisement demand parameters were estimated using the instruments and the procedure described in §4.4. The advertisement demand regression includes program characteristics that can be presumed to influence audience receptivity to advertisements or to proxy for unobserved demographics valued by advertisers.[23] These are:

- Network dummies: to account for the varying ability of networks to bundle desirable program audiences with smaller audiences.
- Genre: viewer mood at the time of exposure to advertising affects ad message processing (Goldberg and Gorn 1987).
- Main Character and Cast demographics: Shachar and Emerson (2000) found that people prefer to watch programs about characters demographically similar to themselves, so cast demographics should correlate with audience demographics.
- A "special" dummy: networks typically air irregularly-scheduled programs in place of their weakest programs.[24]
- Thematic elements: these may correlate with viewer psychographics, which advertisers use to target ad messages.
- Program Age: networks typically renew their best programs, so Program Age reflects viewers' and advertisers' past appraisals of a show.
- Award nominations and past award nominations: these seem likely to indicate the degree to which viewers are emotionally involved in a show and to correlate with desirable viewer demographics.

[23] We had hoped to construct audience demographics $d_s$ using the random coefficients estimated by the viewer demand model, but the imprecision of the random utility parameter estimates made these estimates unreliable.

[24] Specials in the program sample include *Miss Dog Beauty Pageant* and *Married! with Children Cast Reunion*.

**Table 6 Advertisement Demand Parameter Estimates**

| Variable | Estimate (std. error) | Variable | Estimate (std. error) |
|---|---|---|---|
| Advertising seconds | −1,302.4** (127.1) | Main char: single parent | 7,144.7** (500.8) |
| Audience size | 19.4** (6.3) | Cast: 50+% nonwhite | −88,517.6** (691.7) |
| Network: ABC | −14,307.7** (1,679.5) | Cast: 25+% nonwhite | 36,147.8** (631.1) |
| Network: FOX | −15,960.4** (662.2) | Special | −36,644.9** (842.9) |
| Network: NBC | −16,261.2** (681.3) | Theme: cop | 6,678.1** (791.2) |
| Network: UPN | −7,228.7** (1,094.7) | Theme: sci-fi | −60,251.8** (1,132.4) |
| Network: WB | −57,648.9** (667.5) | Theme: supernatural | 4,119.5** (685.9) |
| Genre: scripted comedy | 38,966.3** (466.0) | SeasonFirstAired | 110.6 (71.3) |
| Genre: action drama | −72,658.6** (778.8) | 2003 Emmy nominations | 12,949.7** (99.0) |
| Genre: reality | 73,264.1** (619.8) | Past Emmy nominations | 229.4** (18.3) |
| Genre: news | −17,757.6** (865.7) | Day: Tues | 3,152.7** (577.5) |
| Genre: movie | −20,272.5** (1,517.0) | Day: Wed | 7,340.9** (603.0) |
| Main char: African-American | 1,058.8** (519.1) | Day: Thurs | 41,766.7** (573.9) |
| Main char: other nonwhite | 52,127.7** (711.4) | Day: Fri | 12,303.1** (564.7) |
| Main char: married | 16,958.8** (532.7) | Constant | 83,812.4** (13,486.1) |
| Pseudo $R^2$ | 0.8747 | Average relative error | 0.248 |

*Notes.* Number of observations = 262. Dependent variable is advertisement price.

• Weekday: consumers often shop on the weekend and have limited memories, so advertisers prefer that their commercials be aired later in the week (Auletta 1992).

Table 6 reports advertiser demand parameter estimates. The model fit the data well, with a pseudo $R^2$ of 0.87 and an average relative error of 0.248.

The direct effect of advertising quantity on ad price was negative and significant and implies a mean price elasticity (holding audience size constant) of −2.9.[25] This is substantially more elastic than previous findings; Crandall (1972) estimated a price elasticity of −0.45 and Bowman (1976) found elasticities (using two different specifications) of −0.73 and −0.92. These differences in results are likely due to increased competition: the number of broadcast networks increased from three to six; cable networks and other media entered into competition for television advertising dollars; and broadcasters are no longer allowed to collude by setting a cap on advertising minutes.

The effect of audience size on ad price is positive and significant, with a mean elasticity of 0.83. This figure is substantial given the large variation in audience size. It underscores the importance of considering both sides of the television industry in this paper—that ad revenues are dependent on audience size just as audience size is highly responsive to advertising level.

The genre parameter estimates show that advertisers value Reality programs most, followed by Scripted Comedy, Psychological Drama, News, Movie, and Action Drama. Reality programs receive, *ceteris paribus*, $146,000 more per spot than Action shows, a difference that is larger than the mean ad price. This helps to explain the rapid proliferation of Reality programs after their introduction in the mid-1990s.

Why do Reality programs earn so much more than other shows? Their distinguishing characteristics are unscripted Action, nonprofessional Cast members, competitive themes, and increased ability to accommodate product placement. It might be that advertisers are seeking to take advantage of companion advertising to complement their product placement or to "jam" rival advertisers' product placements. It also could be that viewers identify with the nonprofessional casts, enhancing their receptivity to ad messages. This question deserves further study. Comedies earn more than average because they generate positive feelings among viewers which reduces resistance to persuasion and increases liking, a feeling that can be transferred to advertising (Goldberg and Gorn 1987).

Cast demographics also affect ad prices. Programs featuring African-American main characters earned slightly more than average while programs featuring other minorities earned far more than average. This latter finding might be due to the relative scarcity of nonwhite and non-African-American actors on television. Shows with married or single-parent main characters earned more than average. Advertisers pay significantly less for programs featuring 50% or greater minority Cast representation, but programs with 25%–50% minority casts charge a premium. The first result likely reflects the audience demographics of 50+% minority Cast programs, since nearly all of these shows feature African-American casts and African-American viewers tend to earn less than average.

Program thematic elements influence advertiser demand. "Cop" shows earn slightly more than average as do programs with Supernatural elements. However, science fiction programs earn $60,000 less per spot

[25] The bounds of the 95% confidence interval for this statistic are −2.4 and −3.5.

than do other shows. This is perhaps due to incongruities between program and advertising content.

Irregularly scheduled programs earn less per ad but there does not appear to be any advertising premium associated with Program Age. The effect of current-year Emmy nominations (which were not announced until after the sample was complete) was quite large, suggesting that advertisers generally prefer to place messages in highly engaging programs. Past Emmy nominations increased advertiser demand but their effect was much smaller. These observations indicate evidence of a "halo of quality" effect conferred by highly engaging programs on their advertising.

Thursday programs command a $41,766 premium over Monday programs. Next came Friday with a $12,303 premium, followed by Wednesday (a $7,340 premium), and Tuesday (a $3,152 premium). Consumers often save shopping trips (e.g., autos or movies) for the weekend, so advertisers seek to send messages on Thursdays. Advertisers' relative preference for Friday over Wednesday is smaller but this does not conform to conventional wisdom; we discuss these results further below.

### 5.4. Comparing Advertiser and Viewer Demand Parameter Estimates

It is interesting to compare advertiser and viewer preferences for program characteristics to network programming decisions. Viewer preferences reflect the entertainment value of program characteristics while advertiser preferences are more likely to measure the effect of program characteristics on advertising delivery. Network revenues depend on getting both sides of the market on board, so networks must take the preferences of both sides into account when they acquire and schedule programs.

Consider program genre. Viewers' most preferred genre was Action Drama with a 95% confidence interval of $[0.04, 0.18]$, followed by News ($[-0.06, 0.15]$), Psychological Drama (restricted to 0), Reality ($[-0.18, -0.03]$), Movie ($[-0.30, -0.11]$), and Scripted Comedy ($[-0.27, -0.16]$). It is striking that viewers' three most preferred genres account for only 42% of network schedules. The explanation becomes clear when it is recognized that advertiser genre preferences are nearly opposite those of viewers. Advertisers most prefer to buy time during Reality programs ($[72049, 74479]$), followed by Scripted Comedy ($[38053, 39880]$), Psychological Drama (restricted to 0), News ($[-19454, -16061]$), Movie ($[-23246, -17299]$), and Action Drama ($[-74212, -71159]$). It becomes clear why Reality and Scripted Comedy programs account for 47% of network programming: They command large premiums over other genres. Analyzing either side of the market in isolation might suggest that networks were failing to satisfy their customers' tastes.[26]

It is also interesting to consider the interplay between consumer and viewer preferences across days of the week. Viewers' preferred evening for television is Friday ($[0.21, 1.33]$), followed by Thursday ($[0.06, 0.59]$). Advertisers' most preferred night is Thursday ($[40642, 42892]$), followed by Friday ($[11196, 13410]$), Wednesday ($[6159, 8523]$), Tuesday ($[2021, 4285]$), and Monday (restricted to 0). The estimation results suggest that networks schedule increasingly strong programs as the week progresses, peaking on Thursday before falling sharply on Friday. The average program NAMU on Monday was $-0.20$, $-0.15$ on Tuesday, $-0.14$ on Wednesday, $-0.07$ on Thursday, and $-0.28$ on Friday.

It is easy to see why network competition for viewers has historically been fiercest on Thursday night. Advertisers are willing to pay more on Thursday nights, while viewers are more likely than average to watch television.

Friday night presents a less intuitive picture. This is viewers' most preferred night to watch television, as freedom from going to work or school the next day allows many viewers extra leisure time; but networks air low-quality programs. Advertisers are even willing to pay slightly more for Friday night slots than for nights earlier in the week. There are several strategic factors that may explain broadcast networks' paucity of strong programming on Friday nights. First, viewers as a group are more likely to watch on Friday nights but any individual viewer may be more likely to have occasional social opportunities that would preempt involvement with a regular television series.[27] Second, the value of tune-ins is lower on Friday nights. With the networks' strongest programs on Thursdays, tune-ins are more likely to be valuable on Mondays, Tuesdays, and Wednesdays. Third, broadcast networks face stronger-than-average competition from cable networks on Friday nights. For example, the USA and Sci-Fi cable networks usually debut original programming on Fridays. Cumulative broadcast audiences fell from 45.8% to 29.3% from Thursday night to Friday night, but nonbroadcast networks' cumulative audience rating rose from 22.4% to 28.9%. This final explanation is perhaps the most compelling. Flint (2006) points out that networks scheduled strong programs like *Dallas*, *Miami Vice*, and *Dukes of Hazzard* on Friday nights in the 1980s before cable networks offered strong competition.

[26] These results are suggestive but not conclusive. There may also be unobserved cost differences across types of program that influence networks' actions.

[27] This is supported by the observation that nonserial shows (e.g., movies) are often programmed on Friday nights.

**Figure 4** **Histogram of Imputed Tune-In Benefits (Per 1,000 Households)**

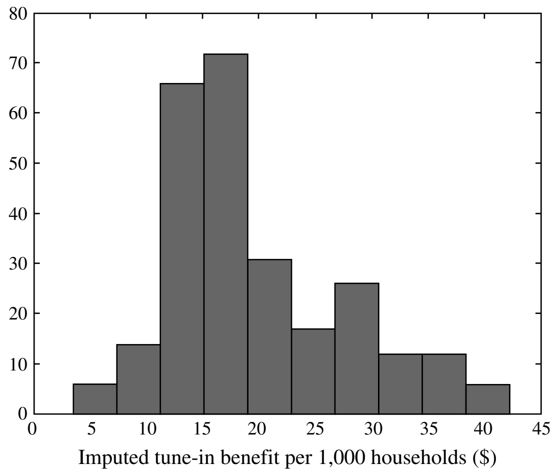
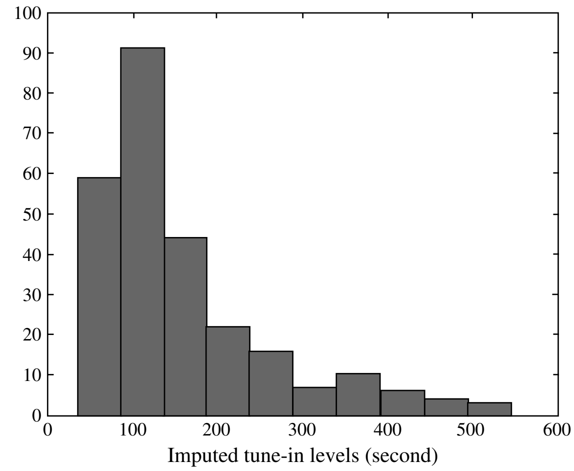



**Figure 5** **Histogram of Imputed Tune-In Levels**




This discussion has focused on genre and weekday effects. Further insights can be drawn by comparing Tables 2, 5b, and 6. For example, viewers prefer Sci-fi programming but advertisers do not, so Sci-fi accounts for just 6% of network program hours.

### 5.5. Network Tune-In Inferences

We use advertiser demand parameter estimates in conjunction with assumptions about network behavior to make inferences about unobserved tune-in levels in the manner described in §4.4. Figure 4 contains a histogram of $\hat{\tau}_s$, the imputed network tune-in benefits (per thousand households) across shows. $\hat{\tau}_s$ was unrestricted but the imputations were strictly positive, with a mean of \$19.50. The minimum tune-in benefit was \$3.16 for the WB's *Reba* on Friday night at 9:00; this show aired on a low-rated network on the night before that network goes dark. The maximum tune-in benefit was on NBC's Wednesday-night *Law & Order*; this show aired at 9:00 p.m. on the night before the network's "Must See TV" Thursday night lineup. Other programs have similarly intuitive places in the distribution of tune-in benefits, with programs on higher rated networks having higher per-viewer tune-in benefits.

We checked the reasonableness of the imputed tune-in benefits by measuring their correlation with the mean program utilities estimated by the viewer demand model. This potential correlation was the reason we did not use the structural model of network competition to estimate advertiser demand parameters. The correlation has the hypothesized sign and turns out to be quite large: 0.48.[28]

[28] This correlation is significant at the 1% confidence level.

Next, we turn to inferences about tune-in levels. Figure 5 depicts the histogram of imputed tune-in levels. The distribution appears to be approximately lognormal. Like imputed tune-in benefits, imputed tune-in levels were unrestricted but found to be strictly positive. The programs with the highest inferred tune-in levels are the season finales of *American Idol* (FOX) and *Everybody Loves Raymond* (CBS). The lowest tune-in inferences corresponded to second-hour, Friday-night programs on low-rated networks: *Reba* and *Grounded for Life* on the WB and movies on UPN.

To check tune-in inference reasonableness, we compared the nightly aggregate tune-in inference to the aggregate tune-in levels observed in a randomly selected 20% sample of the available day/network videotapes. The inferred tune-in levels were regressed on their corresponding observed tune-in levels without a constant. The result was a regression coefficient of 0.91, a standard error of 0.16, and an $R^2$ of 0.80. (If the inferences were perfect, the coefficient in this regression would be 1.) This check suggests the tune-in inferences are quite reliable.

## 6. Counterfactual Experiment: Ad-Avoidance Technology Proliferation

We seek to gain insight into the effects of advertisement-avoidance technology (AAT) proliferation on equilibrium advertising time. AAT may increase or decrease equilibrium ad levels. To understand how AAT can increase ad quantities, note that viewer channel switching is networks' primary incentive to keep advertising levels low. AAT users fast-forward past ad messages rather than switching channels, so it might be that AAT's primary effect is to reduce network audience losses from advertising. Falling disincentives to advertise would lead networks to increase their advertising quantities. Rises in ad time would make AAT more valuable to ad-averse viewers, leading to mutually reinforcing rises in AAT penetration and advertising time.

The other possibility may arise if AAT penetration's primary effect is to lower advertiser willingness to pay for viewers using AAT, since those viewers would presumably be fast-forwarding past most ads. This would make non-AAT users more scarce, and this increased scarcity could lead to higher advertising prices. Networks might respond by competing more intensely for non-ad-avoiding viewers. This competition would take the form of lower advertising levels. Falling ad levels could dampen the advertisement-avoidance benefits of AAT ownership and therefore slow its rate of growth.

We use a counterfactual experiment to gain insight into how AAT may affect the industry. We test the sensitivity of ad levels to a hypothetical AAT within the model of network competition using demand parameter estimates from both sides of the market and tune-in inferences. We report predicted equilibrium ad quantities, ad revenues, and audience sizes, given assumptions about the effects of ad-avoidance technology on viewer and advertiser behavior.

We consider an AAT that allows each viewer to view or record one program per half hour and that gives all ad-averse television viewers an identical, proportional reduction in ad disutility. We do not consider the effects of AAT use on network program scheduling or quality investments. The exercise here is a counterfactual: How would market equilibria in May 2003 have been different if $x$% of viewers had access to the assumed ad-avoidance technology? The results should be interpreted as educated speculation rather than as prediction.

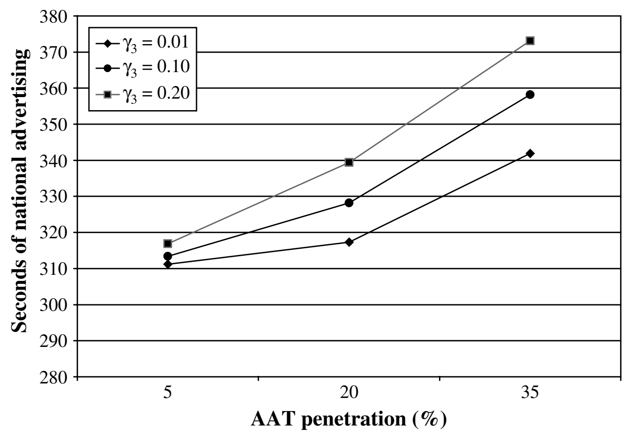
The effects of the hypothetical AAT are governed by three parameters and which viewers have the technology (the "technology distribution rule"). The parameters are:

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

The Technical Appendix that can be found at http://mktsci.pubs.informs.org describes in detail how the model's primitives are respecified to account for AAT and shows how to solve for the resulting equilibrium.

There is no published research to guide the selection of values for the unobserved parameters, and it is not possible to estimate them from the available data.

**Figure 6** **Mean Equilibrium Advertising Seconds per Network Half Hour, by AAT Penetration (Assuming $\gamma_1 = 0.667$ and Most Ad-Averse Viewers have AAT)**




We therefore use ranges of values that seem reasonable and report the sensitivity of the counterfactual results to the assumptions used. We assume that the proportional reduction of advertising disutility resulting from AAT use is 0.67, that $\gamma_3$ is near zero,[29] and that AAT is first adopted by the most ad-averse viewers in the population. The counterfactual results are most sensitive to assumptions about $\gamma_2$ and $\gamma_3$, so we report predictions for various combinations of those parameters.[30]

Figure 6 shows mean predicted advertising time (advertising seconds plus tune-in seconds) for various assumptions about $\gamma_2$ and $\gamma_3$.[31] An increase in AAT penetration from 5% to 35% increases equilibrium advertising time by about 14%. This indicates the attenuation in audience sensitivity to advertising outweighs the viewer-scarcity effect described above. Ad time also increases with advertiser valuations of AAT-using viewers. When ad skippers are more valuable, networks have an incentive to sell more ads to take advantage of ad skippers' less elastic viewing demand.

Figure 7 shows how mean audience size and effective audience size change with AAT penetration.

[29] Wilbur (2007) reviews research on advertising exposure and finds three reasons that advertisers might attach some value to ad-skipping viewers. First, viewer learning has been shown to increase with advertising exposure speed. Second, there is evidence that advertising can have latent effects on consumers' consideration sets, even when consumers do not remember the advertising. Third, the heightened attention required to fast-forward past ads has been linked to increased consumer awareness and recall.

[30] The model's predictions change very little with assumptions about $\gamma_1$ and the technology distribution rule because these two assumptions primarily affect viewers with AAT. These viewers are not highly valued by advertisers so their actions have little effect on networks' strategies.

[31] Predictions are based on the first week of data due to large computational costs.

**Figure 7** Mean Equilibrium Audience and "Effective" Audience per Network Half Hour, by AAT Penetration (Assuming $\gamma_1 = 0.667$ and Most Ad-Averse Viewers have AAT)

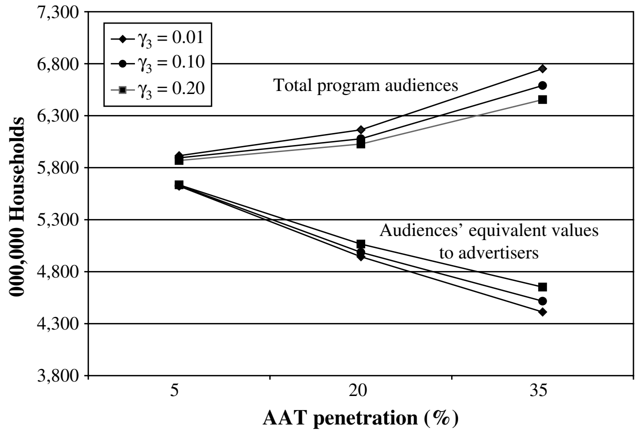
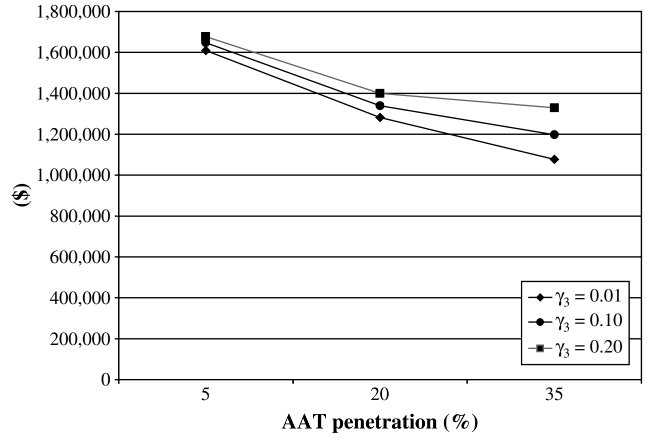



**Figure 8** Mean Equilibrium Advertising Revenues per Network Half Hour, by AAT Penetration (Assuming $\gamma_1 = 0.667$ and Most Ad-Averse Viewers have AAT)




Cumulative audience size rises as AAT proliferates, because AAT users watch more television. Those rises are negatively impacted by AAT users' value to advertisers because networks increase ad levels when ad avoiders are relatively more valuable. However, networks' "effective" audience size falls with AAT penetration. The bottom set of lines in Figure 7 shows the value of the expanded audience sizes when translated to non-ad-avoider equivalents.[32] Ad avoiders' fast-forwarded exposure to commercials represents an unavoidable loss of audience value.

Figure 8 shows that equilibrium advertising revenues fall with AAT penetration. The size of the fall depends on advertiser valuations of ad skippers. A conservative estimate (1 non-AAT user is worth 100 AAT users) indicates network revenues fall 38% when AAT use climbs to 35%. A more liberal estimate (1 non-AAT user is worth 5 AAT users) indicates network revenues fall 22%. The main implication of this result for marketers is that digital video recorders may decrease network incentives to invest in program quality.

It is important to note that we have not accounted for program scheduling; networks might respond to AAT by decreasing intertemporal competition among high-quality programs (for example, they might move some good shows to Monday or Friday night). There is also some question about the extent to which falling ad revenues will be distributed among the networks and their content providers, and the feasibility of strategic changes in advertising content and delivery.

[32] As an example, assume an audience contains 10 viewers with AAT and 10 viewers without, and that $\gamma_3 = 0.10$. The cumulative audience size is 20 but the "effective" audience size is 11, since each of the viewers with AAT is worth one-tenth as much as a viewer without AAT.

## 7. Discussion

Television networks operate in a two-sided market, choosing programs to match advertisers with viewers. This paper estimates a two-sided model of advertiser demand for audiences and viewer demand for programs to quantify the effect of each group's program usage on the other group. We find strong evidence of cross-group externalities. A 10% increase in advertising time decreases the median audience size on a highly rated broadcast network by about 25% (assuming no competitive reactions). Advertisement prices are highly responsive to audience size (elasticity of 0.8). The estimated price elasticity of advertiser demand (−2.9) indicates that the advertising market has become substantially more competitive since the 1970s (Bowman 1976).

We sought to gain some insights into network strategies by comparing advertiser and viewer demand estimates to networks' program characteristics. Viewers' two most preferred genres (Action and News) account for just 16% of network program schedules. Advertisers' two most preferred genres (Reality and Comedy) occupy 47% of network timeslots. These and other results suggest that advertiser preferences influence network program and scheduling choices more strongly than viewer preferences.

Demand estimates were combined with a structural model of network competition to perform a counterfactual experiment. The experiment evaluates the equilibrium effects of proliferation of a hypothetical advertisement-avoidance technology. This exercise offers educated speculation about the effects of digital video recorder penetration on market equilibria. The results suggest that ad avoidance increases equilibrium advertising levels and decreases network advertising revenues.

A limitation of this analysis is that the market-level audience data used in this paper mask viewer zapping behavior (though the audience ratings used as currency in the advertising market also masked viewer zapping behavior). It would be interesting to see how the results would vary by using individual viewing data and commercial minute ratings.

This paper could be extended in several interesting directions. Network program scheduling decisions could be endogenized. Individual advertisers' demand for audiences could be modeled and estimated. It would be interesting to model the third side of the television industry: network program acquisition and expenditure. And it would be interesting to consider how ad-avoidance technology penetration impacts the *value* of advertising exposures lost due to ad avoidance rather than the *number* of ads avoided.

## Acknowledgments

The author thanks the many colleagues and seminar audiences who provided helpful comments, most notably Shantanu Dutta, Paul Farris, John Pepper, S. Siddarth, Ron Wilcox, the editor, the area editor, and four anonymous reviewers. He is especially grateful to his thesis advisors, Simon Anderson and Steve Stern, for their guidance and support. Parts of this work are based on the author's 2005 dissertation at the University of Virginia. Dorothy Kinney and Linli Xu provided superlative research assistance. Financial support was provided by the Bankard Fund for Political Economy at the University of Virginia.

## References

American Association of Advertising Agencies and Association of National Advertisers. 2002. *2001 Television Commercial Monitoring Report*.

Anand, B., R. Shachar. 2004. Brands as beacons: A new source of loyalty to multiproduct firms. *J. Marketing Res.* **41**(2) 135–150.

Anand, B., R. Shachar. 2005. Advertising: The matchmaker. Working paper, Harvard Business School Strategy Unit Research Paper No. 264206, Cambridge, MA.

Anderson, S., S. Coate. 2006. Market provision of broadcasting: A welfare analysis. *Rev. Econom. Stud.* **72**(4) 947–972.

Anderson, S., A. de Palma, J.-F. Thisse. 1992. *Discrete Choice Theory of Product Differentiation*. MIT Press, Cambridge, MA.

Armstrong, M. 2006. Competition in two-sided markets. *RAND J. Econom.* **37**(3) 688–711.

Auletta, K. 1992. *Three Blind Mice*. The Easton Press, New York.

Berry, S. 1994. Estimating discrete-choice models of product differentiation. *RAND J. Econom.* **25**(2) 242–262.

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

Berry, S., J. Levinsohn, A. Pakes. 1995. Automobile prices in market equilibrium. *Econometrica* **63**(4) 841–890.

Bloom, J. 2005. You have to live with the upfront, so at least try to be smart about it. *Advertising Age* **76**(18) 28.

Bowman, G. W. 1976. Demand and supply of network television advertising. *Bell J. Econom.* **7**(1) 238–267.

Byzalov, D., R. Shachar. 2004. The risk reduction role of advertising. *Quant. Marketing Econom.* **2**(4) 283–320.

Caillaud, B., B. Jullien. 2001. Competing cybermediaries. *Eur. Econom. Rev.* **45**(4–6) 797–808.

Caillaud, B., B. Jullien. 2003. Chicken and egg: Competition among intermediation service providers. *RAND J. Econom.* **34**(2) 309–328.

Chen, Y., J. Xie. 2007. Cross-market network effect with asymmetric customer loyalty: Implications for competitive advantage. *Marketing Sci.* **26**(1) 52–66.

Crampes, C., C. Haritchabalet, B. Jullien. 2005. Advertising, competition and entry in media industries. Working paper, CESifo Working Paper Series No. 1591, Munich, Germany.

Crandall, R. W. 1972. FCC regulation, monopsony, and network television program costs. *Bell J. Econom. Management Sci.* **3**(2) 483–508.

Danaher, P. 1995. What happens to television ratings during television commercials? *J. Advertising Res.* **35**(1) 37–48.

Danaher, P., D. Mawhinney. 2001. Optimizing television program schedules using choice modeling. *J. Marketing Res.* **38**(3) 298–312.

Depken II, C. A., D. P. Wilson. 2004. Is advertising a good or a bad? Evidence from U.S. magazine subscriptions. *J. Bus.* **77**(2) S61–S80.

Dukes, A., E. Gal-Or. 2003. Negotiations and exclusivity contracts for advertising. *Marketing Sci.* **22**(2) 222–247.

Flint, J. 2006. Are Friday nights back from the dead? *Wall Street J.* (May 11). http://online.wsj.com/article/SB114713062500047187.html?mod=The+Small+Screen.

Fournier, G. M., D. L. Martin. 1983. Does government-restricted entry produce market power? New evidence from the market for television advertising. *Bell J. Econom.* **14**(1) 44–56.

Fox, J. T. 2007. Estimating matching games with transfers. Working paper, University of Chicago, Chicago, IL.

Gal-Or, E., M. Gal-Or. 2004. Customized advertising via a common media distributor. *Marketing Sci.* **24**(2) 241–253.

Garfield, B. 2005. The chaos scenario. *Advertising Age* **76**(14) 1–4.

Gensch, D. H., P. Shaman. 1980. Models of competitive television ratings. *J. Marketing Res.* **17**(3) 307–315.

Godes, D., D. Mayzlin. 2004. Using on-line conversations to study word-of-mouth communication. *Marketing Sci.* **23**(4) 545–560.

Goettler, R. 1999. Advertising rates, audience composition, and competition in the network television industry. Working paper, Carnegie Mellon University, Pittsburgh, PA.

Goettler, R., R. Shachar. 2001. Spatial competition in the network television industry. *RAND J. Econom.* **32**(4) 624–656.

Goldberg, M. E., G. J. Gorn. 1987. Happy and sad TV programs: How they affect reactions to commercials. *J. Consumer Res.* **14**(3) 387–403.

Hansen, L. P. 1982. The large sample properties of generalized-method of moments estimators. *Econometrica* **50**(4) 1029–1054.

Kaiser, U., J. Wright. 2006. Price structure in two-sided markets: Evidence from the magazine industry. *Internat. J. Indust. Organ.* **24**(1) 1–28.

Lehmann, D. 1971. Television show preference: Application of a choice model. *J. Marketing Res.* **8**(1) 47–55.

Li, X. 2005. Cheap talk and bogus network externalities in the emerging technology market. *Marketing Sci.* **24**(4) 531–543.

Liu, Y., D. Putler, C. Weinberg. 2004. Is having more channels really better? A model of competition among commercial television broadcasters. *Marketing Sci.* **23**(1) 120–133.

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

Nevo, A. 2000. A practitioner's guide to estimation of random-coefficients logit models of demand. *J. Econom. Management Strategy* **9**(4) 513–548.

Nevo, A. 2001. Measuring market power in the ready-to-eat cereal industry. *Econometrica* **69**(2) 307–342.

Newey, W. K., K. D. West. 1987. Hypothesis testing with efficient method of moments estimation. *Internat. Econom. Rev.* **28**(3) 777–787.

Olney, T. J., M. B. Holbrook, R. Batra. 1991. Consumer responses to advertising: The effects of ad content, emotions, and attitude toward the ad on viewing time. *J. Consumer Res.* **17**(4) 440–453.

Owen, B., S. Wildman. 1992. *Video Economics*. Harvard University Press, Cambridge, MA.

Rochet, J.-C., J. Tirole. 2003. Platform competition in two-sided markets. *J. Eur. Econom. Assoc.* **1**(4) 990.

Rochet, J.-C., J. Tirole. 2006. Two-sided markets: A progress report. *RAND J. Econom.* **37**(3) 645–667.

Rojas-Mendez, J. I., G. Davies. 2005. Avoiding television advertising: Some explanations from time allocation theory. *J. Advertising Res.* **45**(1) 34–48.

Roth, A. E., M. Sotomayor. 1990. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press, Cambridge, UK.

Ruggles, S., M. Sobek, T. Alexander, C. A. Fitch, R. Goeken, P. K. Hall, M. King, C. Ronnander. 2004. *Integrated Public Use Microdata Series: Version 3.0* [Machine-readable database]. Minnesota Population Center, Minneapolis, MN. Available at http://www.ipums.org.

Rust, R., M. Alpert. 1984. An audience flow model of television viewing choice. *Marketing Sci.* **3**(2) 113–127.

Rust, R., W. Kamakura, M. Alpert. 1992. Viewer preference segmentation and viewing choice models for network television. *J. Advertising* **21**(1) 1–17.

Rysman, M. 2004. Competition between networks: A study of the market for yellow pages. *Rev. Econom. Stud.* **71**(2) 483.

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

Shapiro, S., D. J. MacInnis, S. E. Heckler. 1997. The effects of incidental ad exposure on the formation of consideration sets. *J. Consumer Res.* **24**(1) 94–105.

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

Stern, S. N. 1997. Simulation-based estimation. *J. Econom. Literature* **35**(4) 2006–2039.

Television Bureau of Advertising. 2005. Universal McCann and Nielsen Media Research data. *Trends in Advertising Volume*. http://www.tvb.org.

Train, K. E. 2003. *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge, UK.

Wilbur, K. C. 2007. How the digital video recorder changes traditional television advertising. *J. Advertising* Forthcoming.

Wildman, S. S., B. D. McCullough, R. Kieschnick. 2004. A new look at the economics of markets for television ad time. Working paper, Michigan State University, East Lansing, MI.

Yang, S., V. Narayan, H. Assael. 2006. Estimating the interdependence of television program viewership between spouses: A Bayesian simultaneous equations model. *Marketing Sci.* **25**(4) 336–349.

Zhou, W. 2004. The choice of commercial breaks in television programs: The number, length, and timing. *J. Indust. Econom.* **52**(3) 315–326.

Zufryden, F., J. H. Pedrick, A. Sankaralingam. 1993. Zapping: Who does it and what is its impact on brand purchase behavior? *J. Advertising Res.* **33**(1) 58–66.