Advertising Rates, Audience Composition, and Competition in the Network Television Industry*

Ronald L. Goettler†
University of Chicago

February 9, 2012

*I would like to thank Mike Eisenberg, Greg Kasparian, and David Poltrack of CBS and Mark Rice of Nielsen Media Research for their help in obtaining the data for this study.

†E-mail: goettler@uchicago.edu

Abstract

We estimate the relationship between advertising prices on prime-time network television and audience size and composition. Hedonic analysis reveals a price premium for (a) viewers from large audiences, (b) viewers 35–49 years old, and (c) viewers from homogeneous audiences. We analyze the optimal scheduling of prime-time shows using the viewer-choice model of Goettler and Shachar (2001) to predict show ratings for each demographic segment. Optimal schedules increase revenue by about 17 percent for each network when competitors' schedules are fixed. The gains are similar in equilibrium of the scheduling game because the increased use of counter-programming and nightly homogeneous-programming benefits the networks as a whole at the expense of non-viewing and non-network alternatives.

KEYWORDS: advertising, hedonic analysis, spatial competition, network television, optimal scheduling, Nash equilibrium

1 Introduction

Most people think of the broadcast networks as being in the entertainment business. Entertainment, however, is merely an input to their primary business of selling access to audiences to companies and organizations seeking exposure. In 2010, spending on advertisements in the United States reached $131.1 billion, about half of which was spent on television ads. The broadcast networks received $25 billion and have increased ad revenue over recent decades despite the drop in their Nielsen ratings from 45 percent in 1985 to 26 percent in 2008 as viewers migrated to cable programming and non-viewing activities such as web-surfing.[1] One reason for these gains is the increase in the demand for advertising as the economy expands and companies increasingly seek to establish brand awareness. Another reason is that, despite shrinking audiences, network television remains the best option for advertisers seeking immediate exposure to large swaths of the U.S. population: the broadcast networks are the only consistent providers of captivated audiences that exceed 10 million viewers. Moreover, shows that deliver such large audiences command higher costs per thousand viewers (CPM) in the ad market.

Nearly as important as audience size is the demographic composition of the viewers. Marketers typically want to reach younger viewers who are more impressionable and viewers who have disposable income or are responsible for household purchases. Accordingly, advertisers are willing to pay a premium to advertise on shows attracting such consumers.

The first contribution of this paper is to estimate, for prime-time network television, the relationship between ad prices and audience size and composition.[2] We estimate hedonic pricing equations using the *Household & Persons Cost Per Thousand* data from Nielsen Media Research, and find a price premium for (a) viewers from large audiences, (b) viewers 35–49 years old, and (c) viewers from homogeneous audiences. Using magazine

[1] Advertising data are from *Kantar Media* reported by *Advertising Age* on *http://www.adage.com.* The historical Nielsen ratings are reported in Gorman (2010) "Where Did The Primetime Broadcast TV Audience Go?"

[2] Fisher, McGowan, and Evans (1980) estimate the audience-revenue relationship for local television stations to assess the revenue effects of changing cable television regulations.

advertisement prices, Koschat and Putsis (2002) also find a premium for middle-aged and homogeneous readership audiences. However, they find ad prices to be concave with respect to circulation, whereas we estimate a convex relationship between ad prices and audience size. This convexity is consistent with the broadcast networks having increased market power when selling commercial time on shows with large audiences because no other media can deliver such audiences. More importantly, this convexity can affect firms' programming and scheduling strategies.

The second goal of this paper is to determine equilibrium network strategies given the estimated relationship between ad prices and audience size and demographics. In the short run, a network's strategic choice corresponds to the scheduling, or sequential positioning, of its stock of shows. In the long run, by canceling some shows and purchasing or developing others, a network can change the set of shows available to schedule.[3] Because programming cost data are not available, researchers have focused on the short-run strategy of show scheduling, which is closer to being cost neutral. Rust and Eechambadi (1989) and Kelton and Schneider (1993) find networks can significantly increase average prime-time ratings by changing their schedules. Goettler and Shachar (2001) confirm this finding, though with more moderate gains ranging from 6 to 16 percent. One possible explanation for the finding of sub-optimal schedules is that these studies assume a network's objective is to maximize ratings, not profits. To assess whether the assumption of ratings maximization drives the discrepancies between observed and recommended schedules, we compare equilibrium schedules under both profit maximization and ratings maximization.

Equilibrium schedules might differ under ratings maximization versus profit maximization for two reasons. First, the value of additional viewers differs across shows, given the nonlinear hedonic pricing model for commercial time. For example, we estimate

[3]We assume the number of commercials aired per hour is fixed. Until 1981, the National Association of Broadcasters limited commercial time during prime time to 6 minutes per hour. However, in 1981, this restriction was declared a violation of antitrust laws and the limit was dropped. Since then, prime-time commercial time has steadily increased, reaching 15.35 minutes per hour in 1996, according to the Commercial Monitoring Report. Evaluating the equilibrium number of commercials per hour is an interesting topic we leave for future research.

that an additional million viewers increases ad revenue by $86,974 for each half-hour of *Roseanne*, a popular show in the 1990s with desirable audience demographics, compared to an increase of only $50,649 per half-hour of *Hat Squad*, a less popular show with less-desirable audience demographics. A network would therefore be willing to sacrifice viewers from a show with low marginal revenue for a smaller gain in viewers of a show with high marginal revenue. Such an exchange could be achieved, for example, by scheduling the latter show in a relatively more favorable time slot against weak competitors or following a strong lead-in show. Discrepancies between schedules that maximize ratings instead of profits can also reflect costs of adjusting schedules. When setting schedules, the appeal of new shows is unknown. This uncertainty can result in ex-post sub-optimality that might persist if adjustment costs are high.

We therefore assess whether the ratings gains found in previous studies can be attributed to modeling firms as maximizing ratings instead of maximizing profits. We find schedules that maximize ad revenue net of schedule adjustment costs are similar to schedules that maximize ratings. This finding suggests ratings-increasing strategies, such as counter-programming, placing strong shows at 8:00 or 9:00, and avoiding head-to-head battles among networks' best shows, are consistent with revenue maximization.[4] Moreover, the revenue gains from better implementing these strategies are sufficiently high that adjustment costs cannot explain why firms continue with sub-optimal schedules. Without switching costs, the equilibrium revenue gains for each network range from 11 to 18 percent. With switching costs, fewer changes occur and the gains range from 7 to 9 percent. Hence, we conclude the ratings gains Goettler and Shachar (2001) obtain and the revenue gains we find in this paper indeed suggest networks' sub-optimal behavior. The gains primarily result from more effective counter-programming by scheduling news magazines prior to 10:00 and sitcoms after 10:00. Although the networks have abandoned the rule-of-thumb that news magazines should only air from 10:00 to 11:00, they continue to avoid airing sitcoms past 10:00.

[4] All times we report refer to p.m. EST, as our focus is prime-time programming.

The paper proceeds as follows. In the next section, we discuss the market for commercial time on network television, estimate the relationship between ratings and advertisement revenue, and discuss potential implications of this relationship for network strategies. In section 3, we compute best-response schedules and Nash equilibria of the scheduling game and analyze the strategic behavior of the television networks under various specifications of the payoff function. Section 4 concludes.

2 Network Ratings and Advertisement Revenue

We describe the advertising market for network television and present our data in section 2.1. We then estimate a hedonic model for ad prices in section 2.2 and discuss its implications for the networks' strategies in section 2.3.

2.1 The Network Advertising Market and Data

Typically 70 to 80 percent of network commercial time is sold in the *up-front* market in May for the upcoming television season commencing in early September. The remainder is sold in the *scatter* market during the season, occasionally hours before the show airs. Networks use unsold time to promote the their shows or provide public service messages. Contracts specify the prices for the commercial time and minimum guaranteed ratings, as measured by Nielsen Media Research.[5] Often the guaranteed ratings correspond to particular demographic segments. For example, advertisers seeking young viewers typically receive guarantees for viewers aged 18–49. Networks also occasionally guarantee ratings particular to gender or household income. When shows do not attain the guaranteed ratings, the networks provide "make-goods" to compensate their short-changed advertisers. These make-good commercials, however, typically air on less popular shows and do not fully compensate advertisers.[6]

[5]The accuracy of these ratings has been the subject of much debate. Nonetheless, the Nielsen ratings are the best available and serve as the industry standard.

[6]Rather than specifying the total price for airing a commercial on a particular show, the contracts could specify a per-viewer (or per viewer with desired demographics) rate that would then be converted

If commercial prices merely reflected guaranteed ratings, then the relationship between advertisement revenues and guaranteed ratings would be the relationship of interest. According to network executives contacted for this research, commercial prices reflect expected audiences, and the demographics of these expected audiences, not the lower guaranteed ratings of the contracts. Because we do not possess data on the *expected* ratings, we must estimate the relationship between *realized* ratings and advertising revenues. Assuming the expected ratings to be unbiased, we interpret the error term in the econometric models in section 2.2 as the difference between expected and realized ratings.

The data are from the *Household & Persons Cost Per Thousand* monthly publications by Nielsen Media Research for September through December, 1992. These publications list for each network show the cost of airing a 30-second commercial during its broadcast and its estimated audience sizes for the following categories: Households, Total Adult Women, Women aged 18–34, Women aged 18–49, Women aged 25–54, Total Adult Men, Men aged 18–34, Men aged 18–49, Men aged 25–54, Teens (aged 12–17), and Children (aged 2–11).[7] ABC, CBS, NBC, and Fox provided the commercial-costs data to the Nielsen monitoring service. We derive the estimated audience sizes from the Nielsen Television Index (NTI) sample, which provides minute-by-minute records of viewers' choices using the People Meter device. Because the publication is monthly, all data are averages over the month.

Though the data contain observations for all network shows, we only use shows airing past 6:00 because the ratings data are often missing for earlier shows, and our application involves scheduling evening shows. We also exclude 173 shows that aired only once during the period because many of these observations appear to be outliers. For all specifications, we reject the Wald test of equality of estimates using shows with

into a total price *after* Nielsen reports the show's audience size. While this alternative avoids the risk of paying for viewers who are not delivered, the uncertain total payment complicates the implementation of ad campaigns on fixed budgets. Given the absence of such contracts, the nuisance of a variable total payment must exceed the concern that some shows will deliver fewer viewers than expected.

[7]Despite being from 1992, the data are relevant because Nielsen still reports these same demographic categories and prime-time television has changed little. The networks still deliver a mix of news, comedy, drama, and reality programming to attract a broad swath of mainstream America. The most significant change is the nature of reality content, which is now more contest oriented.

multiple airings and using shows aired once. Excluding these 173 observations leaves 437 observations to estimate the models in the next subsection.

2.2 A Model of Ad Prices

The top panel in Figure 1 reveals a linear relationship between the logarithm of each show's audience size and the logarithm of the price of a 30-second commercial during that show. Whereas the log-log model, with an R-squared of 0.8159, explains costs per commercial well, audience size explains only 5.8 percent of the variation in costs per viewer, depicted in the bottom panel. We show viewer demographics are the primary driver of this variation, which spans .19 to 1.32 cents per viewer.

To precisely measure the effect of demographics and audience size on advertising prices, we estimate a reduced-form hedonic pricing equation. Hedonic analysis assumes the value of a product can be implicitly decomposed into valuations for separately observed characteristics (Griliches 1967, 1971, and Rosen 1974). Estimating the hedonic pricing equation yields implicit market prices for each characteristic. Rosen (1974) and Epple (1987) emphasize the difficulty of consistently estimating the structural supply and demand equations using hedonic analysis. However, we only need the reduced-form implicit prices: our counterfactual analysis of show scheduling does not affect equilibrium prices, because changing schedules only changes quantities marginally, relative to the ad market as a whole.

Although the log-log plot in Figure 1 establishes a nonlinear relationship between ad prices and audience size, the degree of curvature is not obvious.[8] The Box-Cox (1964) model provides a simple method for estimating the degree of nonlinearity inherent in the relationship between variables of interest. We define the transformation

$$x^{(\lambda)} \equiv \begin{cases} \log(x) & \text{if } \lambda = 0 \\ \frac{(x^{\lambda}-1)}{\lambda} & \text{otherwise} \end{cases} \tag{1}$$

[8]Note the log-log model allows costs per viewer to vary with audience size. For example, if the slope exceeds unity then costs per viewer increase with audience size.

for any variable x. Linear and log transformations are special cases that correspond to $\lambda = 1$ and $\lambda = 0$, respectively. We use this transformation to model ad prices as

$$P_s^{(\lambda)} = \alpha + D_s\gamma + Z_s^{(\lambda)}\beta + \epsilon_s \ , \tag{2}$$

where P_s is the price per 30-second commercial aired during show s, Z_s is the vector of explanatory variables to be transformed, and D_s is the vector of explanatory variables not to be transformed. The only explanatory variable the results suggest should be transformed is audience size (i.e., number of viewers) for show s. Thus, we define Z_s to be audience size.[9] The following demographic variables comprise D_s: the fraction of the audience aged 12–17 years old, the fraction aged 18–34, the fraction aged 35–49, the fraction aged 50 and older, the fraction of the audience that is female, the square of the fraction that is female, and the standard deviation of the ages of the viewers comprising an audience.

Table 1 presents estimates of four models: (1) the Box-Cox model using all the variables, (2) the log-log model using all the variables, (3) the Box-Cox model using only the significant variables, and (4) the log-log model using only the significant variables. First note we reject the log-log specifications using likelihood ratio tests of the restriction that $\lambda = 0$ in the Box-Cox specification. The test statistics are 14.85 and 15.33, yielding p-values less than .001. We also reject the (unreported) linear models with test statistics exceeding 280 for the restriction that $\lambda = 1$. Despite rejecting the log-log model, we report its estimates, because it is widely used and easier to interpret than the Box-Cox model. For example, a coefficient on audience size exceeding 1 implies a convex relationship between ad prices and audience size, resulting in a higher CPM for viewers in large audiences.

Another notable finding is that viewers aged 12–17 or 18–34 are no more valuable to advertisers than viewers aged 2–11 (the excluded age category). These data are from 1992, at which time the broader age category of 18–49 was considered by network strategists to be the most desirable. Our estimates indicate that the upper end of this category was the driving force behind the value of delivering 18- to 49-year-olds to advertisers. The

[9]Dividing Z_s by the U.S. population yields the familiar Nielsen rating for the show.

attractiveness of the 35- to 49-year-olds may reflect the higher income of older viewers. If we were to separately condition on income, we suspect the coefficient on "Fraction aged 18–34" would increase, relative to the coefficient on "Fraction aged 35–49." Unfortunately, the income data are not available. Also note the negative impact on ad price when shows target viewers 50 years of age and older.

The estimates corresponding to the standard deviation of viewers' ages and the gender variables reveal that more homogeneous audiences command higher prices. This result is consistent with the emphasis on targeted marketing by advertisers. Most advertised products are relevant to particular segments of the population. Accordingly, advertisers are willing to pay a higher CPM for shows that deliver the target segments in high concentrations. Ad prices therefore decrease in the dispersion of viewers' ages and are quadratic in the fraction of viewers that are female. This quadratic relationship has a minimum at 0.51 and is depicted for the show *Melrose Place*, as an example, in Figure 2. The fraction of women in audiences ranges from 0.34 to 0.70, contributing to the significant variation in prices (and prices per viewer) for audiences of the same size.[10]

Table 2 reports estimates for two models of prices per viewer. The estimates accord with the price-per-commercial models in Table 1: the signs are all the same and the same variables are statistically significant. Also, the positive coefficient for audience size (in millions of viewers) confirms the premium paid for large audiences, as implied by the estimated convexity between ad price and audience size.

Three factors contribute to the convex relationship between ad prices and audience size. First, advertising on top-rated shows provides a "halo" effect: advertised products benefit from being associated with high-quality shows. Second, advertising campaigns can attain high *reach* (i.e., share of target segment exposed to the ad at least once) more easily by advertising on shows with the largest audiences. Overlap among audiences implies an ad aired on two shows with 10 million viewers each will reach fewer than 20

[10]Koschat and Putsis (2002) investigate the benefits to magazine publishers of unbundling access to demographic segments of a magazine's readership. The ability to expose television viewers of a given show to demographic-specific or viewer-specific ads is limited.

million unique viewers—many fewer if the advertiser buys time on similar shows to target specific demographic segments. Third, the networks have some market power when selling time on top shows, because cable channels or other media outlets rarely deliver audiences over 10 million. When selling time on less popular shows, however, the networks face many close substitutes, both on television and in other media. An FCC study concluded radio, magazine, newspaper, billboard, and alternative television advertising constrain the prices of network advertising such that prices for network ads reflect competitive forces (FCC, Network Inquiry Special Staff, 1980).

2.3 Implications for Scheduling Strategies

Equations (1) and (2) imply show s has an expected price of

$$P_s = \mathrm{E}\left[\left(1 + \lambda(\alpha + D_s\gamma + Z_s^{(\lambda)}\beta + \epsilon_s)\right)^{1/\lambda}\right] , \tag{3}$$

where $\mathrm{E}[\cdot]$ denotes the expectations operator. Table 3 presents expected advertisement revenue generated by three shows with different audience sizes and compositions. We use Monte Carlo integration using 50,000 draws of ϵ_s to compute the expectation in equation (3). The expected gains in revenue due to an additional million viewers appears in the lower portion of the table.

The importance of viewer demographics is evident when comparing *Hat Squad* to *Melrose Place*. Advertising on *Melrose Place* costs (and is predicted to cost) more than 50 percent more than advertising on *Hat Squad* despite the fact that the latter show has 22 percent fewer viewers. The younger and more homogeneous audience of *Melrose Place* compared to *Hat Squad* drives this premium: the *Melrose Place* audience is 64.7 percent female with 7.3 percent aged 50 years or older, whereas *Hat Squad*'s audience is 58.7 percent women with 53.9 percent aged 50 and older. Accordingly, *Melrose Place* yields almost twice as much money per viewer as does *Hat Squad*. Although the audience demographics for *Roseanne* are not quite as good as those for *Melrose Place*, reaching 32.9 million viewers triggers a large convexity-based premium that enables *Roseanne* to nearly match the per-viewer price of *Melrose Place*.

As illustrated in the last three rows of Table 3, the marginal value of additional viewers differs across shows. Per commercial, an additional million viewers is worth \$5,065 for *Hat Squad*, \$8,953 for *Melrose Place*, and \$8,697 for *Roseanne*. A network can therefore increase its revenues while lowering its average viewership across shows. Time slots vary in the degree of competition from other shows or non-viewing activities (e.g., sleep). By switching an unpopular show in a less competitive time slot with a popular show from a competitive time slot, the network can sacrifice ratings from a low-marginal-revenue show to increase the ratings of a high-marginal-revenue show. For example, suppose *Melrose Place* had a difficult time slot and *Hat Squad* had an easy time slot on the same network such that switching the two shows would result in a gain of a million viewers for the former show and a loss of a million for the latter. Using the predicted gains from Table 3, the change in weekly revenue for the network would be $2(89,529 - 50,649) =$ \$77,760 because each show is an hour long. Over a 20-episode season, the revenue gain would exceed \$1.5 million. This example demonstrates that maximizing revenue (or profits) might lead to schedules that do not maximize ratings.

The previous example is illustrative but does not account for effects of schedule changes on the ratings of the other shows on the nights involved in the switch. Such effects exist because each time slot is linked through viewer persistence, as discussed in section 3.1. Nonetheless, the general strategy of airing shows with high marginal returns in favorable time slots is implied by the models in Table 1. For example, this strategy suggests the networks should place their strongest shows at 9:00 P.M.—the most popular time slot for watching television during which competition from the non-viewing option is lowest.[11] Second, the networks should avoid airing their best shows against other networks' best shows because top shows generate high prices per viewer. For instance, NBC dominated Thursday-night television during the 1990's with little competition from the other networks who instead aired their top shows on other nights.

[11]Goettler and Shachar (2001) show the high ratings at 9:00 reflect reduced utility from non-viewing activities, in addition to high-quality shows being aired at 9:00. One possible reason for the increased viewing at 9:00 is the desire to relax before bedtime by watching television.

Of course, many other factors also influence the networks' scheduling strategies. The networks often avoid competing for the same types of viewers by *counter-programming*—airing shows with characteristics different from those of the other shows in the same time slot. Each network also tends to air similar programs in sequence in an effort to continue to serve its viewers from the previous period. This strategy is known as *homogeneous programming*. A third strategy is to air strong shows at 8:00 to secure a large audience that, due to viewer persistence, will likely stay tuned for most of the evening. We assess the extent to which these strategic considerations manifest themselves in the actual schedules the networks air, and compare the results to schedules that maximize either ratings or ad revenue.

3 Strategic Network Behavior

To analyze competition among the broadcast networks, we must specify the strategy space and each network's payoff function. Although profit maximization is the typical firm objective, revenue maximization is consistent with profit maximization if costs do not vary over choices in the strategy space. By focusing on firms' scheduling strategies, programming costs are constant across candidate strategies since the network airs the same set of shows regardless of the schedule.[12]

In section 3.1, we present our method for computing payoffs for alternative schedules. In section 3.2, we analyze scheduling when schedule changes are costless, and compare our findings to those obtained in Goettler and Shachar (2001) for networks that maximize ratings. In section 3.3, we consider optimal schedules when changing the schedule is costly. In section 3.4, we conclude the analysis by computing Nash equilibrium of the scheduling game, both with and without adjustment costs.

[12] If we had data detailing the costs of purchasing (or developing) shows, we could also analyze competition in the longer run in which the networks can change the set of shows they air. In the long run, assuming the networks maximize ratings may be inappropriate given the huge sums of money required to purchase hit shows, such as *E.R.*, for which NBC paid Warner Brothers up to $13 million per episode (Vogel 2010).

3.1 Predicting Ratings and Revenues for Alternative Schedules

To assess payoffs for scheduling strategies, we must predict shows' demographic-specific ratings for counterfactual schedules. Previous studies use a variety of models to predict ratings. Horen (1980) and Kelton and Schneider (1993) use linear aggregate ratings models to facilitate the use of integer programming to find schedules that maximize weekly ratings. The drawback of these models is they do not identify the interwoven effects of switching costs, show characteristics, and viewer heterogeneity. To treat the network's scheduling problem as an integer programming task, these studies assume each show's contribution to the weekly ratings is independent of its preceding and following shows. This assumption contradicts evidence that the number of viewers staying tuned to a network during a show change depends on the similarity of the two shows. Over the week of November 9, 1992, an average of 56 percent of a prime-time show's viewers watched the end of the previous show on the same network. This lead-in effect, however, ranges from 32 percent to 81 percent and tends to be higher the more similar a show is to its lead-in show (Goettler and Shachar, 2001).

To account for such linkages, we use a structural discrete-choice model in which viewers choose from among a set of shows possessing various show characteristics. A new schedule corresponds to a rearrangement of the show-specific characteristics, that are either estimated or specified *a priori*. Obviously, a model will accurately predict ratings only if these show characteristics effectively characterize the shows. Indeed, one of the difficult aspects of analyzing the strategic behavior of the television networks is identifying suitable show characteristics. The simplest approach is to categorize each show *a priori*, perhaps as a comedy, drama, or news show. Such labels, however, are ineffective: shows in the same category often have striking differences, and shows in different categories often have similarities. Rust and Eechambadi (1989) analyzes scheduling issues using an augmented version of the Rust and Alpert (1984) model that assigns shows *a priori* to one of several categories. Alternatively, show characteristics have been estimated by Gensch and Ranganathan (1974) using factor analysis and by Rust, Kamakura, and Alpert (1992) using multidimensional-scaling methods. Neither of these studies estimates show characteristics simultaneously with the other determinants of viewer behavior, such as costs of switching channels. Goettler and Shachar (2001) use a structural discrete-choice model to simultaneously estimate all components of the choice model, allowing for both structural state-dependence and unobserved heterogeneity to explain the observed persistence in viewers' choices.[13] Because their estimated model corresponds to the data-generating process itself, it is appropriate for evaluating counterfactual experiments, such as introducing new schedules. The other models, however, are subject to the Lucas (1981) critique because their estimates of show characteristics are based on reduced-form methods and would likely differ under different schedules. Thus, we employ the model of viewer choice that Goettler and Shachar (2001) estimate using a panel data set from Nielsen Media Research of 3,286 viewers' choices during each prime-time quarter-hour over the week of November 9, 1992.

The model is a factor-analytic logit model (Elrod and Keane 1995) with state dependence. The factor-analytic structure explains variation across shows in the lead-in effect and state dependence explains why its minimum is well-above zero.

The model consists of (a) show characteristics in a four-dimensional latent attribute space, (b) the distribution of viewers' most preferred locations (ideal points) in this space, (c) a vertical characteristic (quality) for each show, (d) switching costs (when transitioning across channels or across viewing/non-viewing options), (e) the utility from not watching television, and (f) the utility from watching non-network programming. The switching costs, ideal points, utility from not watching television, and utility from non-network programming all vary across demographic segments defined by age, gender, household income, and other measures. Shows that appeal to the same viewers (i.e., have high joint audiences) are estimated to be near each other in the latent attribute space so that consumers with ideal points near these similar locations are likely to watch both of them. See Appendix A for model specification and estimation details.

[13] Dubé, Hitsch, and Rossi (2010) illustrate the importance of allowing for heterogeneous preferences when estimating state dependence.

For a given schedule, we use predicted ratings for each demographic segment to construct Z_s and D_s for each show, which translate to revenues using equation (3).

3.2 Optimal Scheduling

Network strategists and executives actively debate the scheduling of their shows. For the most part, they rely on their intuition and insights from aggregate and demographic-specific Nielsen ratings. Unfortunately for the networks, to precisely account for the many factors influencing viewer behavior, such as the lead-in effect, show competition, and viewer heterogeneity, one needs the individual-level Nielsen data.[14] Disentangling the interaction of these factors is the attraction of the model we use to evaluate alternative schedules.

Each network's best-response schedule maximizes its payoff holding the other networks' schedules fixed. We employ the "iterative improvements" approach of combinatoric optimization to find an approximate solution to the discrete optimization that yields the optimal schedule. As detailed in Appendix B, we search for payoff-improving swaps of blocks of programming of various lengths and from various initial schedules.

Table 4 presents elements of the best-response schedules for each of the big three networks for the week of November 9, 1992, using ad revenue as the payoff function. Each network obtains significant revenue gains—17.7 percent for ABC, 16.8 percent for CBS, and 16.5 percent for NBC.[15] These percentage gains translate into \$6.1 million to \$7.1 million per week, assuming 10 minutes of network commercial time each hour.[16] Multiplying by 52 yields annual gains ranging from \$317 million to \$369 million per year. Also note the percentage gains in ratings are similar to the revenue gains.

What types of scheduling strategies generate these gains in revenues and ratings?

[14]The lead-in effect refers to the tendency for shows to have high ratings if the preceding (or lead-in) show had obtained a high rating.

[15]In 1992, Fox only broadcasted shows on Wednesday, Thursday, and Friday, and only from 8:00–10:00. As such, we focus on the three major networks.

[16]The 15.35 minutes of commercials time reported in the introduction includes time sold by the affiliates and time used to promote the network's programs.

The lower half of Table 4 provides summary measures of several strategies the networks implement. We measure counter-programming (CP) as the average distance from the competition in each quarter-hour. Two measures reflect the degree of homogeneous programming for network j: nightly-homogeneity (NH) is the average distance between shows' attribute-space locations z_{jt} for shows airing the same night, and sequential-homogeneity (SH) is the average distance between a network's z_{jt} in sequential periods. Three other strategies pertain to the placement of quality or power shows. Often a network airs "power on the hour" because more viewers have just finished watching a show, and are willing to switch channels, than at the half-hour when many viewers are in the middle of an hour-long show. A network also tends to air its power early in an effort to build a large audience it can retain with the help of switching costs and inertia. Both "power on the hour" and "power early" are captured by $\overline{\eta}_t$, the average over days of each time slot's show quality, denoted η_{jt} in Goettler and Shachar (2001). The third power-placement strategy involves not airing one's best shows against other strong shows with (relatively) similar z characteristics. We call this strategy "power counter" (PC) and measure its implementation by the average over the week of the ratio $RR_{jt}/RR_{\hat{j}t}$, where $\hat{j}$ refers to the closest competitor (in the latent attribute space) to j at time t and RR_{jt} is the *relative rating* defined as the rating for j at t divided by the average rating for j over the week.

We expect optimal schedules to have higher CP (than the actual schedule), higher PC, lower NH and NS, and higher $\overline{\eta_t}$ on the hours and early in the night. Table 4 reveals that the ratings-maximizing schedule indeed implements each of these strategies to a greater degree than observed in the actual schedules.

Though not reported in the table, the ratings-maximizing schedule implements these strategies to a similar degree as the revenue-maximizing schedules. In fact, the ratings- and revenue-maximizing schedules are the same for NBC and are similar for CBS and ABC. The reason for the strong similarity is that ratings maximization also provides the incentives related to audience size and composition (discussed at the end of section 2.3). For example, audience composition is already important under ratings maximization because ratings are higher if shows that target a specific demographic segment are scheduled on the same night. Also, avoiding "power wars" by placing hit shows in relatively easy slots is important to both ratings maximization and revenue maximization. This analysis demonstrates the simpler objective of ratings maximization is consistent with revenue maximization, when assessing scheduling strategies.

3.3 Costs of Adjusting Schedules

The predicted revenue gains computed above are upper-bounds if costs to implementing the best response schedule are non-trivial. Such costs could be substantial because many schedule changes are needed to move from the original schedules to the best-response schedules: 26 changes for ABC, 20 for CBS, and 18 for NBC.[17]

To account for adjustment costs in the iterative-improvements algorithm, we compute best-response schedules assuming networks only swap one programming block with another if revenue increases by at least 2 percent. This required percentage gain translates to a cost of about \$38.7 million for ABC (0.02 $\times$ \$124,061 per ad for ABC $\times$ 20 ads per hour $\times$ 15 prime-time hours per week $\times$ 52 weeks per year). Because the primary cost of changing the schedule is the opportunity cost of commercial time used to promote the new schedule, we can view this cost as paying for 312 commercials to promote the new schedule. Because this cost estimate is high, we are erring on the side of being too conservative with respect to the recommended schedule changes. Using these costs, ABC would optimally implement three changes in its schedule for a revenue gain of 10.2 percent, CBS would optimally implement two changes for a gain of 9.7 percent, and NBC would optimally implement two changes for a gain of 12.3 percent.

Hence, the networks appear to be scheduling sub-optimally, even when we consider possible costs of schedule changes. Although these gains are non-trivial, they are modest compared to the predicted gains reported in previous studies of network television scheduling. For example, Rust and Eechambadi (1989) find a 78 percent improvement in

[17] We define a single schedule change as a swap of one continuous block of shows with another continuous block, regardless of the number of shows in each block.

NBC's schedule using a stochastic, heuristic approach to finding the optimal schedule.

3.4 Nash Equilibrium

A natural question is whether the gains under autarkic optimal scheduling will persist in equilibrium. A possible scenario is that strategic responses from the other networks will erase the gains. We therefore compute Nash equilibria of the scheduling game. To find an equilibrium, we cycle through the four networks, individually implementing their best-response schedules holding the other schedules fixed. A schedule is a Nash equilibrium if no network has an incentive to change unilaterally. The order in which the networks hypothetically implement their best-response schedules marginally influences the equilibrium payoffs, but not in any predictable manner. Although one might expect first-movers to attain the highest gains in payoffs, their gains are often not the highest.

Despite the possibility that no pure-strategy equilibrium exists, we always find an equilibrium in fewer than four rounds. The equilibrium we find depends on the assumed sequence of play in the equilibrium search. Fortunately, the usefulness of analyzing the equilibrium is not in pinpointing the exact equilibrium schedule the networks should play. Rather, we focus on strategic behavior that is common across all equilibria.

The most important finding is that the strategic responses from other networks do not erode away the gains from unilateral optimization. In each equilibria we find, the ratings and payoffs of the big three networks (ABC, CBS, and NBC) increase, though by less than the gains associated with the best-response schedules (holding other networks' schedules fixed). When ABC moves first, the percentage gains in revenue are 17.6 percent for ABC, 10.9 percent for CBS, and 12.3 percent for NBC. When costs of schedule changes are included in the payoff function as discussed above, revenues increase by 9.1 percent for ABC, 7.0 percent for CBS, and 7.2 percent for NBC.[18]

As expected, the same strategies responsible for the gains under the autarkic sce-

[18]The gains are similar when other networks move first. For example, when CBS moves first the percentage revenue gains without costs are 14.4, 10.1, and 15.5, respectively, for ABC, CBS, and NBC. With costs, the percentage gains are 8.5, 7.7, and 12.1, respectively.

nario are at work in equilibrium. Furthermore, we again find these strategies are the same regardless of whether each network maximizes ratings or revenue. That is, the networks increase their revenues and ratings by increasing their use of counter-programming, homogeneous programming, and airing strong shows on the hour and early in the night.

Inspecting the placement of shows in both the best-response schedules and the equilibrium schedules, we find the counter-programming and homogeneous-programming measures increase primarily by moving sitcoms past 10:00. Network strategists have avoided airing half-hour shows past 10:00. Our findings suggest this rule-of-thumb may not be optimal. At a minimum, experimenting with sitcoms after 10:00 seems warranted.

Because each network increases ratings and revenues, the gains must be achieved by pulling viewers from the non-viewing and, to a lesser extent, non-network viewing alternatives. This finding reflects the benefit to all the networks of counter-programming, along both the vertical (quality) and horizontal dimensions of show attributes. Essentially, the increased use of counter-programming enables the networks to provide programming in each time slot that appeals to more viewers. Also, the increased homogeneous programming induces viewers to stay tuned to the networks longer once they start watching.

4 Conclusion

First, we estimate the importance of audience size and demographic composition in determining prices for commercial time on the broadcast networks. We find that for a given audience size, higher prices are obtained by shows with more homogeneous viewers (as measured by age and gender) and by shows with a high percent of 35- to 49-year-old viewers. Shows with a high percent of viewers 50 years old and older generate yield lower ad prices. Furthermore, we find the price per commercial is convex in the total number of viewers. That is, cost-per-viewer (CPM) increases in the size of the audience.

Second, we combine the hedonic pricing function with the viewer-choice model of Goettler and Shachar (2001) to evaluate optimal scheduling when the networks maximize ad revenue, both with and without accounting for costs of adjusting schedules. We find the optimal schedules increase the networks' revenues and are similar to the schedules that maximize average ratings. Consequently, the ratings gains earlier studies report are not an artifact of researchers using ratings as the firm's objective function. Rather, the findings of sub-optimal scheduling suggest the networks can better implement the recognized strategies of counter-programming across networks within time slots and homogeneous programming by each network during each night.

We close by noting that, despite advertisers' preferences for reaching specific demographic segments, "addressable television advertising," in which viewers of the same show are exposed to different ads, remains in its infancy (Vascellaro 2011). Privacy concerns and the high adoption costs for addressable distribution systems are formidable barriers to growth, particularly given the uncertainty of the benefits to media companies. Theoretical studies by Iyer, et al. (2005), Gal-Or, et al. (2006), Athey and Gans (2010), and Bergemann and Bonatti (2011), among others, provide interesting insights, but do not agree on the effect of increased targeting on media firms' prices and profits. If addressable television advertising eventually takes off, we would hope to compare the implicit prices in this paper to those under the new regime.

Table 1: Estimates of Box-Cox and Log-Log Models of 30-second Commercial Prices

	Model 1	Model 2	Model 3	Model 4
λ	0.1484	0	0.1504	0
	(0.0404)		(0.0403)	
Constant	13.8109	-4.0589	14.5982	-3.9348
	(10.9132)	(1.0525)	(10.6542)	(0.7689)
Audience size$^{(\lambda)}$	0.5443	1.1881	0.5381	1.1863
	(0.1155)	(0.0227)	(0.1138)	(0.0219)
Fraction aged 12–17	1.7281	0.5755		
	(4.7007)	(0.9206)		
Fraction aged 18–34	-0.5036	-0.1594		
	(2.3926)	(0.4702)		
Fraction aged 35–49	15.3507	3.0107	14.8117	2.7225
	(7.2910)	(0.4997)	(6.9312)	(0.3032)
Fraction aged 50+	-5.1219	-0.9580	-5.4289	-1.0243
	(2.9686)	(0.3596)	(2.4701)	(0.0863)
Std. dev. of ages	-0.2434	-0.0485	-0.2495	-0.0488
	(0.1254)	(0.0127)	(0.1220)	(0.0102)
Fraction female	-84.1238	-15.3597	-86.1786	-15.4026
	(41.0696)	(2.4782)	(41.8296)	(2.3992)
(Fraction female)^2	82.8157	15.2010	84.9380	15.2769
	(40.0120)	(2.2493)	(40.7882)	(2.1764)
σ^2	1.7189	0.0668	1.7978	0.0666
R-squared		0.8919		0.8917
Log Likelihood	-4880.0806	-4887.5083	-4880.2075	-4887.8709
χ^2 for H_0: $\lambda = 0$	14.8553		15.3267	
χ^2 for H_0: $\lambda = 1$	281.5602		282.8470	

Standard errors are in parentheses.
The dependent variable (price per 30-second commercial) and audience size are transformed using equation (1).
The excluded age demographic is the fraction of viewers aged 2–11.
The .05 critical value for the likelihood ratio tests is 3.84.

Table 2: Estimates of Models of Costs Per Viewer (in cents)

	Model 1	Model 2
Constant	3.3476	3.4341
	(0.5419)	(0.3954)
Millions of viewers	0.0071	0.0070
	(0.0012)	(0.0011)
Fraction aged 12–17	0.1032	
	(0.5423)	
Fraction aged 18–34	0.0527	
	(0.2767)	
Fraction aged 35–49	1.3628	1.3271
	(0.2877)	(0.1740)
Fraction aged 50+	-0.4595	-0.5080
	(0.2131)	(0.0508)
Std. dev. of ages	-0.0235	-0.0246
	(0.0075)	(0.0059)
Fraction female	-10.4588	-10.5391
	(1.4484)	(1.4032)
(Fraction female)^2	10.0878	10.1620
	(1.3144)	(1.2726)
σ^2	0.0229	0.0228
R-squared	0.4007	0.4006

Standard errors are in parentheses.

Table 3: Expected Revenue for Three Shows with Different Audience Sizes and Demographics

	Hat Squad	*Melrose Place*	*Roseanne*
Total viewers	14,770,000	11,520,000	32,900,000
Fraction ages 35–49	0.202	0.154	0.243
Fraction ages 50+	0.539	0.073	0.179
Standard deviation of viewer ages	18.307	14.569	17.718
Fraction women	0.587	0.647	0.604
Actual price per commercial	$ 55,700	$ 90,600	$ 254,600
Predicted price per commercial	$ 60,707	$ 92,791	$ 256,551
Predicted cents per viewer per commercial	0.411	0.805	0.780
Predicted revenue per half-hour	$ 607,070	$ 927,910	$ 2,565,510
Predicted revenue per season for the show	$ 12,141,400	$ 18,558,200	$ 51,310,200
Gain from extra million viewers, per commercial	$ 5,065	$ 8,953	$ 8,697
Gain from extra million viewers, per half-hour	$ 50,650	$ 89,530	$ 86,970
Gain from extra million viewers, per season	$ 1,013,000	$ 1,790,600	$ 1,739,400

Calculations assume 10 commercials per half-hour and 20 episodes per season.
The revenue gain per commercial exceeds cents-per-viewer × one million
because of the convex relationship between ad prices and audience size.

Table 4: Revenue-Maximizing Schedules Compared to Actual Schedules

		ABC		CBS		NBC	
		Actual	Optimal	Actual	Optimal	Actual	Optimal
predicted revenue per ad		\$124,061	\$146,047	\$117,080	\$136,755	\$124,346	\$144,918
revenue gain per ad			\$21,986		\$19,674		\$20,572
weekly revenue gain (20 ads/hour)			\$6,595,722		\$5,902,320		\$6,171,654
percentage revenue gain			17.72		16.80		16.54
predicted weekly rating		8.55	9.81	8.74	9.78	8.32	9.56
weekly ratings gain			1.27		1.03		1.24
percentage gain			14.81		11.80		14.86
CP	$\equiv$ average($\lvert z_{jt} - z_{j't}\rvert$)	0.63	0.70	0.67	0.71	0.59	0.72
NH	$\equiv$ average($\lvert z_{jt} - z_{jt'}\rvert$)	0.55	0.43	0.54	0.37	0.59	0.39
SH	$\equiv$ average($\lvert z_{jt} - z_{j,t-1}\rvert$)	0.36	0.35	0.40	0.33	0.52	0.35
$\overline{\eta}_{8:00}$	$\equiv$ average($\eta_{j,8:00}$)	2.33	2.13	2.21	2.10	2.12	2.10
$\overline{\eta}_{8:30}$	$\equiv$ average($\eta_{j,8:30}$)	2.12	2.09	2.25	2.08	2.06	2.10
$\overline{\eta}_{9:00}$	$\equiv$ average($\eta_{j,9:00}$)	2.33	2.59	1.95	2.02	2.11	2.32
$\overline{\eta}_{9:30}$	$\equiv$ average($\eta_{j,9:30}$)	1.94	2.25	1.78	1.93	1.95	2.07
$\overline{\eta}_{10:00}$	$\equiv$ average($\eta_{j,10:00}$)	2.09	2.11	1.93	2.01	1.90	1.75
$\overline{\eta}_{10:30}$	$\equiv$ average($\eta_{j,10:30}$)	2.09	1.71	1.93	1.90	1.90	1.71
PC	$\equiv$ average($RR_{jt}/RR_{\hat{j}t}$)	1.09	1.55	1.13	1.24	1.27	1.30

The first column contains the variable names: CP for Counter-Programming, NH for Nightly Homogeneity, SH for Sequential Homogeneity, $\overline{\eta}_t$ for quality at time t, and PC for Power Counter. The variable RR_{jt} is the *Relative Rating* defined as the rating for j at t divided by the average rating for j over the week. The subscript $\hat{j}$ in the definition of PC refers to the closest competitor to j at time t. Optimal schedules maximize each network's ad revenue, holding competitors' schedules fixed.

Figure 1: Commercial Prices and Audience Size

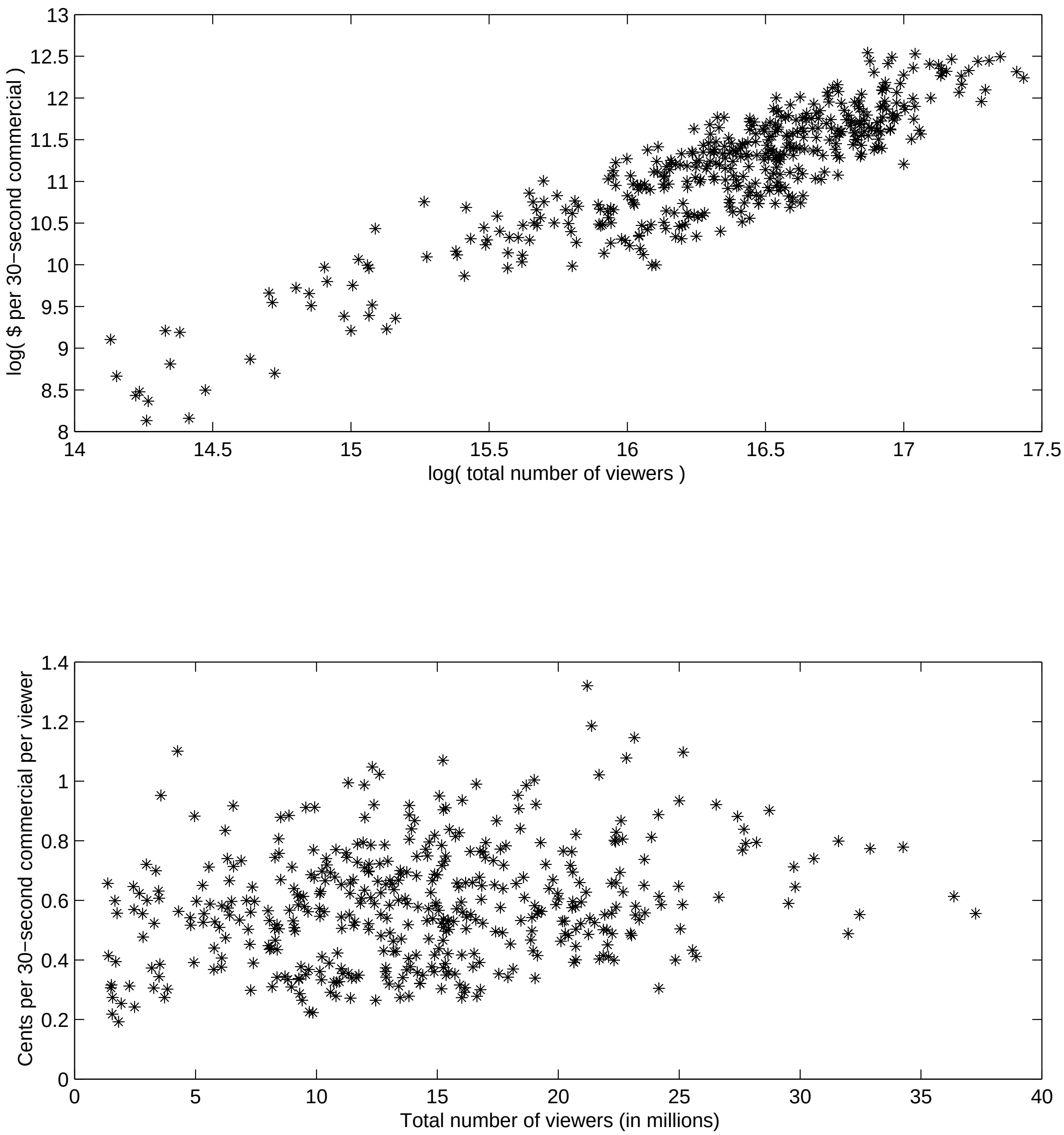
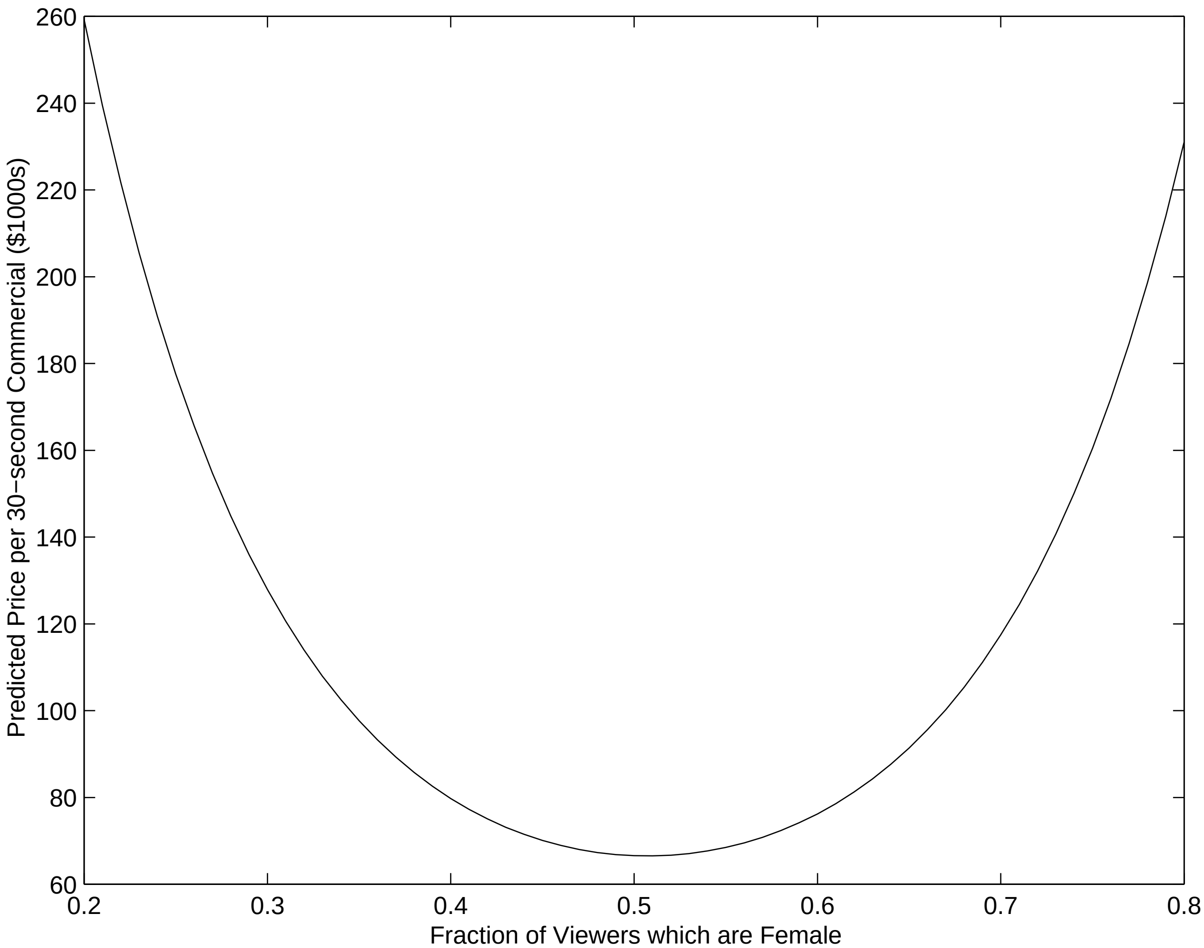


Figure 2: Impact of Female Composition for *Melrose Place*



Appendix A: Modeling Viewer Choice to Predict Ratings

We use the estimated model of Goettler and Shachar (2001) to simulate viewers' choices for each quarter-hour of prime-time programming, 8 to 11 p.m. EST, Monday through Friday, for the week of November 9, 1992. In each period t, individual i chooses from among $J = 6$ options indexed by j, corresponding to (1) TV off, (2) ABC, (3) CBS, (4) NBC, (5) Fox, and (6) non-network programming, such as cable or public television.

Individual i's utility from watching network j in quarter-hour t is

$$\begin{aligned} u_{ijt} &= \eta_{jt} - (z_{jt} - \nu_{i,z})'(z_{jt} - \nu_{i,z}) \\ &\quad + \delta_{Sample} Sample_{ijt} + \delta_{InProgress} InProgress_{ijt} \\ &\quad + \delta_{Start,i} Start_{ijt} + \delta_{Cont,i} Cont_{ijt} + \varepsilon_{ijt}, \end{aligned} \tag{4}$$

where η_{jt} is a quality attribute equally valued by all individuals, $\nu_{i,z}$ denotes viewer i's K-dimensional ideal point, and z_{jt} denotes the K-dimensional location of network j's show during period t. Because none of these parameters are observed by the econometrician, the factor-analytic structure in the first line of equation (4) is a latent attribute space. We assume preferences $\nu_{i,z}$ are constant over time and viewers know the z_{jt} and η_{jt} of all shows. We assume the unobserved $\nu_{i,z}$ are distributed $N(X_i'\Gamma_z,\ \Sigma_z)$, where Σ_z is diagonal and X_i is an L-vector of demographic characteristics (age, gender, household income and number of adults and children, urban-residence status, education of head-of-household, cable subscription level).

The second and third lines of equation (4) contain the state-dependence component of the model, along with the idiosyncratic error ε_{ijt}. The dummy variables $Start_{ijt}$, $Cont_{ijt}$, $Sample_{ijt}$, and $InProgress_{ijt}$ describe the viewer's previous choice as it relates to each of the current period's network alternatives: $Start_{ijt} = 1$ if i was tuned to network j at $t-1$ and the show on j starts in period t, $Cont_{ijt} = 1$ if i was tuned to j and the show on j is a continuation from last period, $Sample_{ijt} = 1$ if i was tuned to j and the show on j is entering its second quarter-hour and is longer than 30 minutes, and $InProgress_{ijt} = 1$

if i was not tuned to j and the show on j is a continuation from last period.

Both $\delta_{Start,i}$ and $\delta_{Cont,i}$ are functions of the demographic variables X_i:

$$\begin{aligned}\delta_{Start,i} &= X_i'\Gamma_\delta \ , \quad \text{and} \\ \delta_{Cont,i} &= \delta_{Start,i} + \delta_{Cont} \ .\end{aligned} \tag{5}$$

For parsimony, $\delta_{Start,i}$ serves as a "base" measure of persistence for viewer i, and δ_{Cont} is the incremental cost of leaving a continuing show that was watched last period. We expect $\delta_{Cont} > 0$, $\delta_{Sample} < 0$, and $\delta_{InProgress} < 0$.

The utility from a nonnetwork show has the same structure as utility from a network show. Because our data does not specify which nonnetwork channel is watched, we use the expected maximum utility over the N_i nonnetwork options available to individual i. The utility from each nonnetwork channel, indexed by $j' = 1, \ldots, N_i$, is

$$u_{ij't} = \eta_{Non} + (\delta_{Mid,i} Mid_t + \delta_{Hour,i} Hour_t)\, \mathcal{I}\{y_{i,j',t-1} = 1\} + \varepsilon_{ij't}, \tag{6}$$

where η_{Non} is common across all N_i options, $Hour_t = 1$ if t is the hour's first quarter-hour, $Mid_t = 1 - Hour_t$, $\mathcal{I}\{\cdot\}$ is an indicator function, and

$$\begin{aligned}\delta_{Mid,i} &= \delta_{Start,i} + \delta_{Mid} \\ \delta_{Hour,i} &= \delta_{Start,i} + \delta_{Hour} \ .\end{aligned} \tag{7}$$

We expect $\delta_{Hour} < \delta_{Mid}$ because more shows are continuations during the hour than on the hour. Assuming $\{\varepsilon_{ij't}\}_{j'=1}^{N_i}$ are independently distributed type I extreme value, the expected maximum is

$$u_{i6t} = \log\left[\sum_{j'=1}^{N_i} \exp(u_{ij't} - \varepsilon_{ij't})\right] + \varepsilon_{i6t}. \tag{8}$$

Substituting equation (6) into equation (8) yields

$$u_{i6t} = \eta_{Non} + \log\left[N_i - 1 + \exp\left((\delta_{Mid,i} Mid_t + \delta_{Hour,i} Hour_t)\mathcal{I}\{y_{i,6,t-1} = 1\}\right)\right] + \varepsilon_{i6t} \tag{9}$$

because $y_{i,j',t-1} = 1$ is satisfied by one j' when $y_{i,6,t-1} = 1$ and zero j' otherwise. We model $N_i = exp(\nu_{i,N})$ as unobserved heterogeneity with $\nu_{i,N} \sim N(X_i'\Gamma_N,\ \exp(X_i'\Gamma_{\sigma_N})^2)$.

Finally, utility from the $j = 1$ nonviewing option differs among individuals according to their previous choice, the time of day, the day of the week, and unobserved tastes for the outside alternative, $\nu_{i,Out} \sim N(X_i'\Gamma_{Out},\ \sigma^2_{Out})$:

$$\begin{aligned} u_{i1t} &= X_i'\Gamma_9 Hour9_t + X_i'\Gamma_{10} Hour10_t + X_i'\Gamma_{Day} Day_t \\ &\quad + \eta_{Out,t} + \delta_{Out} I\{y_{i,1,t-1} = 1\} + \nu_{i,Out} + \varepsilon_{i1t}\ , \end{aligned} \tag{10}$$

where the variables $Hour9_t$ and $Hour10_t$ indicate t is in the 9:00 to 10:00 hour and 10:00 to 11:00 hour, respectively, the variable Day_t is a vector of length five with all zeros except for a one in the current day's position, and Γ_{Day} is an L by five parameter matrix. The time slot and day effects differ across demographic segments. For example, children go to bed earlier than adults.

Goettler and Shachar (2001) estimate the model using maximum simulated likelihood and use the Bayesian information criterion to determine that the latent-attribute space has four dimensions. As in factor analysis, the attribute space can be rotated to yield interpretable dimensions. Three of the dimensions appear to represent plot complexity, degree of realism, and age (of characters and viewers). The fourth dimension is harder to interpret, but appeals to urban, educated men aged 18 to 34. See Goettler and Shachar (2001) for plots depicting the shows' estimated locations and for tables reporting estimates of the other parameters.

We forecast ratings for each prime-time show during the week of November 9, 1992, given a candidate schedule Y, by aggregating predicted choices for each of the 3,286 viewers used in the estimation. For each viewer, we randomly draw an ideal point ν_i from the estimated distribution of ideal points, which is conditional on the viewer's demographics X_i. We then compute the viewer's probability of watching each show. Let $y_{i\cdot t}$ denote the response vector, such that for $j = 1, \ldots, J$, $y_{ijt} = 1$ if i chooses j at time t and $y_{ijt} = 0$ otherwise. Because the additive stochastic utility term in the model is type I extreme value, the probability of viewer i with preference vector ν_i choosing $y_{ijt} = 1$ at time t

conditional on her previous choice of $y_{i,\cdot,t-1}$ is of the convenient form

$$f(y_{ijt}=1|\hat{\theta},y_{i,\cdot,t-1},Y_{jt},\nu_i)=\frac{\exp(\bar{u}_{ijt}(\hat{\theta};y_{i,\cdot,t-1},Y_{jt},\nu_i))}{\sum_{j'=1}^{J}\exp(\bar{u}_{ij't}(\hat{\theta};y_{i,\cdot,t-1},Y_{j't},\nu_i))}\ , \tag{11}$$

where $\hat{\theta}$ is the vector of estimated parameters, and $\bar{u}_{ijt}(\hat{\theta};y_{i,\cdot,t-1},Y_{jt},\nu_i)$ is the non-stochastic component of utility for viewer i watching choice j at time t with schedule Y, conditional on having chosen $y_{i,\cdot,t-1}$ last period.

State dependence implies the i's choice in period t is conditional on her choice from the previous period. The marginal probability $s(y_{ijt}=1|\hat{\theta},Y_{jt},\nu_i)$ is therefore the probability-weighted average of the conditional probabilities in equation (11). Explicitly,

$$s(y_{ijt}=1|\hat{\theta},Y,\nu_i)=\sum_{\hat{y}_{i,\cdot,t-1}\in(1,\ldots,J)}\left[s(\hat{y}_{i,\cdot,t-1}|\hat{\theta},Y,\nu_i)\cdot f(y_{ijt}=1|\hat{\theta},\hat{y}_{i,\cdot,t-1},Y_{jt},\nu_i)\right]\ . \tag{12}$$

We could alternatively draw random logit errors and simulate the full sequence of choices. This alternative, however, would introduce simulation error.

We convert the viewer probabilities in equation (12) to expected network ratings for network j in period t by averaging $s(y_{ijt}=1|\hat{\theta},Y,\nu_i)$ over all n viewers. Letting $r_t(j;\hat{\theta},Y)$ denote the ratings for network j under schedule Y, we have

$$r_t(j;\hat{\theta},Y,(\nu_1,\ldots,\nu_n))=\frac{1}{n}\sum_{i=1}^{n}s(y_{ijt}=1|\hat{\theta},X_i,Y,\nu_i)\ . \tag{13}$$

To obtain ratings specific to a demographic segment D, as needed to predict revenues using equation (3), we sum only over consumers in the target segment:

$$r_t(j,D;\hat{\theta},Y,(\nu_1,\ldots,\nu_n))=\frac{\sum_{i=1}^{n}s(y_{ijt}=1|\hat{\theta},X_i,Y,\nu_i)\,\mathcal{I}(X_i=D)}{\sum_{i=1}^{n}\mathcal{I}(X_i=D)}\ . \tag{14}$$

Appendix B: Finding Best-response Schedules

Each network chooses a schedule from its strategy space—the set of feasible schedules for shows during the week of November 9, 1992—to maximize its objective function. Possible objective functions are profits, advertisement revenue, and average ratings. Each of these payoff functions requires predicting the ratings of the candidate schedules in the strategy space, as described in Appendix A.

The most obvious approach to finding the optimal schedule is to compute the payoff for each feasible schedule and select the schedule with the highest payoff. Each network airs about 20 prime-time shows yielding approximately $20! = 2.4 \times 10^{18}$ candidate schedules. Optimistically assuming each schedule's payoff can be computed in one second, this approach would require 77 billion years to find the optimal schedule for a single network.

We instead implement an "iterative improvements" approach of combinatoric optimization to find approximate best-response schedules. Given an initial schedule, we find and execute ratings-improving swaps of continuous blocks of shows (ranging in length from 30 minutes to 3 hours) until no more ratings-improving swaps exist. This process is sure to converge because the number of possible schedules is finite. Thus a schedule with a (weakly) maximum payoff exists. If only payoff-improving changes are executed, then in finite time, either the optimal schedule will be reached, or the process will terminate at a sub-optimal schedule that single block swaps cannot improve. The possibility of terminating at a sub-optimal schedule is the sense in which this algorithm is an approximate (or local) solution. We use the best terminal schedule obtained from starting at a set of initial schedules that includes the network's current schedule and several randomly generated ones.

The approximate best-response schedule is not unique: if the algorithm were to change the order of show-block swappings, the terminating schedule would be different. Accordingly, we focus on characteristics of the best-response schedules that are invariant to the order in which blocks are swapped.

We also considered extending the algorithm to permit combinations of two simultaneous swaps (involving 3 or 4 continuous blocks of shows). We found no improvements relative to the faster algorithm that only swaps pairs of show-blocks.

References

Athey, S. and Gans., J. (2010), "The Impact of Targeting Technology on Advertising Markets and Media Competition," *American Economic Review: Papers and Proceedings*, 100(2), 608–613.

Bergemann, D. and A. Bonatti (2011), "Targeting in Advertising Markets: Implications for Offline vs. Online Media," *RAND Journal of Economics*, 42(3), 417–443.

Box, G. and D. Cox (1964), "An Analysis of Transformations," *Journal of the Royal Statistical Society, Series B*, 211–264.

Dubé, J. P., G. Hitsch, and P. Rossi (2010), "State Dependence and Alternative Explanations for Consumer Inertia," *RAND Journal of Economics*, 41(3), 417–445.

Elrod, T. and Keane, M.P. (1995), "A Factor-Analytic Probit Model for Representing the Market Structure in Panel Data," *Journal of Marketing Research*, 32, 1–16.

Epple, D. (1987), "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products," *Journal of Political Economy*, 95(1), 59–80.

Fisher, F. M., J. J. McGowan, and D. S. Evans (1980) "The Audience-Revenue Relationship for Local Television Stations," *Bell Journal of Economics*, 11(2), 694–708.

Gal-Or, E., M. Gal-Or, J. H. May, and W. E. Spangler (2006) "Targeted Advertising Strategies on Television," *Management Science*, 52(5), 713–725.

Goettler, R. and R. Shachar (2001), "Spatial Competition in the Network Television Industry", *RAND Journal of Economics*, 32(4), 624–656.

Gorman, B. (2010, April 12), "Where Did The Primetime Broadcast TV Audience Go?", *http://tvbythenumbers.zap2it.com*.

Griliches, Z. (1967), "Hedonic Price Indexes Revisited: Some Notes on the State of the Art," *Proceedings of the Business and Economic Statistics Section*, American Statistical Association, 324–332.

Griliches, Z. (1971), "Hedonic Price Indexes of Automobiles: An Econometric Analysis of Quality Change," in Zvi Griliches (ed.), *Price Indexes and Quality Change*, Cambridge: Cambridge University Press.

Horen, J. H. (1980), "Scheduling of Network Television Programs", *Management Science*, 26, 354–370.

Iyer, G., D. Soberman, and J. M. Villas-Boas (2005), "The Targeting of Advertising," *Marketing Science*, 24(3), 461–476.

Kelton, C. M. L. and L. G. Schneider (1993), "Optimal Television Schedules in Alternative Competitive Environments," Carlson School of Management mimeo, University of Minnesota.

Koschat, M. A. and W. P. Putsis Jr. (2002), "Audience Characteristics and Bundling: A Hedonic Analysis of Magazine Advertising Rates," *Journal of Marketing Research*, 39, 262–273.

Lucas, R. E. Jr. (1981), "Econometric Policy Evaluation: A Critique," in *Studies in Business Cycle Theory*, MIT Press.

Rosen, S. (1974), "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 82(1), 34–55.

Rust, R. T. and M. I. Alpert (1984), "An Audience Flow Model of Television Viewing Choice," *Marketing Science*, 3, 113–124.

Rust, R. T. and N. V. Eechambadi (1989), "Scheduling Network Television Programs: A Heuristic Audience Flow Approach to Maximize Audience Share," *Journal of Advertising*, 18, 11–18.

Rust, R. T., W. A. Kamakura, and M. I. Alpert (1992), "Viewer Preference Segmentation and Viewing Choice Models for Network Television," *Journal of Advertising*, 21, 1–18.

Vascellaro, J. E. (2011, March 7), "TV's Next Wave: Tuning In to You," *The Wall Street Journal*.

Vogel, H. L. (2010), *Entertainment Industry Economics: A Guide for Financial Analysis*, Cambridge: Cambridge University Press.