The News Hour: Welfare Estimation in the Market for Local TV News*

Matthew J. Baker†
Hunter College and the Graduate Center, CUNY
Lisa M. George‡
Hunter College and the Graduate Center, CUNY

November 19, 2015

Abstract

We estimate the welfare consequences of local news broadcasting decisions in advertiser-funded television, a question that has played a central role in US regulation of media markets. We treat station broadcasting decisions as the outcome of a discrete game in which stations choose programming to maximize advertising revenue, which depends upon viewership. Using program-level data on television viewing and advertising prices during the 5-8 p.m. evening news hours, we find that local news is substantially under-provided relative to the viewer optimum. Counterfactual simulations suggest that welfare loss arises in part from the higher value advertisers place on entertainment viewers in a two-sided market framework. However in many markets additional local news broadcasting would increase joint station revenues as well as viewing, indicating that classic business stealing also plays a role in welfare outcomes.

*This paper was written with financial support from the Research Foundation of the City University of New York. We thank Tobias Klein, Lapo Filistrucci and seminar participants at the US DOJ Anti-Trust Division, Media Economics Workshop, Hunter College, Indiana University, the CUNY Graduate Center, the University of Connecticut, and Fordham University for helpful comments.

†Matthew J. Baker, Department of Economics, Hunter College, 695 Park Ave., New York, NY 10065. Email: `mjbaker@hunter.cuny.edu`.

‡Lisa M. George, Department of Economics, Hunter College, 695 Park Ave., New York, NY 10065. Email: `lisa.george@hunter.cuny.edu`.

1 Introduction

The relationship between market competition and television programming decisions has interested economists since the very start of the TV era when Steiner (1952) and later Beebe (1977) outlined theoretically how competitive broadcasters might duplicate mass content and offer too little niche programming. More recently, theoretical work on two-sided markets by Anderson and Gabszewicz (2006) shows how advertiser funding can further distort programming away from viewer preferences. The resulting concern that competitive markets might under-provide local news has motivated strong and long-standing regulation of local television markets in the United States. These regulatory requirements have remained largely unchanged despite dramatic changes in news delivery brought about by cable television, the internet and mobile platforms.

Yet despite deep policy and research interest in local news, very little empirical evidence speaks to welfare outcomes associated with television broadcasts. We offer an empirical analysis of oligopolistic competition among broadcasters in a two-sided market framework. Our approach is to model station programming decisions as the outcome of a discrete game of complete information played by rival stations. We estimate the model using program-level data on television viewing and advertising prices. Counterfactual simulations allow us to examine the consequences of programming outcomes to viewers, advertisers and stations.

After estimation, we simulate the model to determine the program patterns that maximize viewer utility, joint station revenues, and aggregate advertiser surplus. We find substantial under-provision of local news compared to the allocation that maximizes television viewing. Simulations show the average shortfall to be 10.2 half-hour local news broadcasts per market each day during the evening news hours of 5:00-8:00 p.m., or about 1 half-hour broadcast per local station. Re-allocation of programming to the viewer optimum would increase total television viewing by about 9% over the 5:00-8:00 p.m. news hours on average, with local news viewing increasing from about 5% of households to 8.5% of households. Most of the estimated shortfall occurs later in the broadcast day, during the 7:00-8:00 p.m. hour.

We study two sources of welfare loss to consumers. The first arises from divergent interests of viewers and advertisers in a two-sided market. Even though total viewing would increase with more local news broadcasts, stations chose not to provide it because advertisers will pay more for entertainment viewing, especially in later time slots. Overall, the observed broadcast pattern more closely resembles the pattern that maximizes advertiser surplus than viewership: our simulations indicate that the observed number of local news broadcasts is closer to the advertiser optimum than viewer optimum in 90% of markets.

The second source of welfare loss arises from station competition in a classic business-stealing environment. Our simulations indicate that the observed number of local news broadcast is in fact less than the number that would maximize station revenue and advertiser surplus in 83% of the sample markets. That is, profits for both advertisers and stations would rise if stations differentiated programming in several time slots rather than jointly broadcasting entertainment programming, but no single station has an incentive to do so. The average shortfall from the station optimum is about half the viewer shortfall, 4.7 half-hour broadcasts versus 10.5 broadcasts. Fully collusive programming would increase advertising prices (and hence station revenue) by $1.36 (12%) per advertising second per station, or just shy of $50,000 per market on average. Advertiser surplus would increase even more at the collusive outcome, about $655,200 per market on average. Gains are highest in larger, wealthier markets.

We take a game-theoretic approach to estimation. Because local television markets are characterized by small numbers of competing broadcasters, station broadcasting decisions are interdependent. We model local television programming as the outcome of a discrete game of complete information. We estimate station viewership and advertising revenue functions that allow interdependence in stations payoffs, then control for possible equilibrium selection using a game-theoretic econometric model to control for this interdependence.

Our viewership and advertising price models extend the approach of Berry and Waldfogel (2004). To study viewership, we formulate a nested multinomial logit model of television program demand. Our advertising price model is based on a log-linear demand for viewer impressions by advertisers, which changes functional form with program type. This specification allows the possibility that advertisers value viewer impressions differently for different types of programs, which we in fact find to be the case.[1]

We use game-theoretic techniques both in estimating the model and exploring its properties in welfare simulations. Using a method based on simulated maximum likelihood, we first estimate stations' reduced-form payoffs from alternative programming

[1]See George and Hogendorn (2012) for a theoretical treatment of advertising context.

choices. Owing to the multiplicity or cohesiveness problem (see Ciliberto and Tamer (2009), we adopt an approach to estimation inspired by Bajari et al. (2010b), and similar to Zhu et al. (2009) and Ellickson and Misra (2008). Our method allows for studying markets with varying (and sometimes fairly large) numbers of players, in which players have large-dimensional choice sets. Because we are interested in using our model for simulation of counterfactuals, it is important that it does a good job fitting the data. Following Bajari et al. (2010b), we estimate the model using the Markov chain Monte Carlo approach described in Chernozhukov and Hong (2003). In addition to the practical reasons suggested by Chernozhukov and Hong (2003) for turning to this approach to estimation, an added benefit is that we can build additional degrees of flexibility into the model (i.e., nesting different sorts of models in one estimation sweep). The estimation method then picks out, in some sense, the nature of the model that best captures the data.

Our research contributes empirical evidence to an expansive theoretical literature on inefficiencies in differentiated product markets, which can arise in both the number and mix of products. Most familiar is the potential for excess entry, which can occur if products offered by entrants are close substitutes for existing varieties so that new entrants divert consumers from existing options.[2] Also well known are inefficiencies in the product mix, which can arise when firms face incentives to cluster in regions of product space with high demand, or to excessively differentiate in order to sustain high prices.[3] Taken together, this literature demonstrates that under a range of consumer preference distributions and timing assumptions, product location choices can fall well short of first-best. Location models without prices, such as the median voter result of Downs (1957), suggest even more pronounced distortions.

An important class of differentiated product markets, namely advertiser-funded media, are also two-sided. Product positioning in two-sided markets has not been extensively studied theoretically, but conceptually any divergence between the marginal value of differentiated products to consumers versus advertisers introduces the potential for distortions in the product mix. Heterogenous valuations for media types among advertisers can arise from the demographic mix on the consumer side, but also through

[2]Models of entry and competition in this spirit begin with Chamberlin (1933) and are extended in important contributions by Spence (1976a), Spence (1976b), Dixit and Stiglitz (1977) and Sutton (1991).

[3]A large literature starting with Hotelling (1929)) and developed by d'Aspremont et al. (1979) documents inefficiencies in the location choices of firms.

better alignment with advertised products or affective context more amenable to persuasion. The source of inefficiency matters in the television market, since distortions driven by competition are likely more readily tackled with market structure regulation while distortions arising from two-sided market tradeoffs suggests a need for subsidies or other price remedies.

Few attempts have been made to empirically estimate inefficiencies in differentiated product markets, especially two-sided ones. A notable exception is Berry and Waldfogel (2004) study of excess entry in radio broadcasting. The symmetry assumptions in their model for the entry game precluded analysis of welfare consequences of station format choices. More recently, Berry et al. (2015) tackle the question of excess entry when products are differentiated. Filistrucchi et al. (2012) estimate merger effects in the Dutch newspaper markets.

More generally, our results suggest that product positioning can play an important role in platform competition. Benchmark models of two-sided markets such as Armstrong (2006) as well as more recent contributions such as Weyl (2010) and Weyl and White (2011) emphasize the balance of prices charged to each side. Product targeting has not been well explored, and our results suggest that the product mix can play both a strategic role in platform competition and alter the welfare tradeoffs from one side of the market to the other. While these effects are likely to be strongest in markets like broadcast television where there are no viewer prices, they can operate in any differentiated product setting.

The paper proceeds as follows. Section 2 relates our research to the literature and provides background on local television news markets. Section 3 describes programming, viewing and advertising data during the evening news hours and section 4 provides a preliminary outline of potential welfare losses in program choice. Section 5 outlines our econometric model and specification. Section 6 presents our simulation and estimation procedure and reports parameter estimates. Section 7 evaluates our simulation results and tackles the welfare analysis. Section 8 concludes the paper. We provide detail on our estimation methods in an appendix to the main text.

2 Background and Literature

2.1 Literature

Our study is informed by a well-developed literature on variety and consumption in media markets. Much of this literature explicitly or implicitly considers the relationship between product variety and market size. Overall, the positive relationship between market size, available variety and consumption has been demonstrated in radio (Waldfogel, 2003), newspapers George and Waldfogel (2003) and entertainment television Waldfogel et al. (2004). This literature suggests that the welfare implications of variety are particularly important for minority taste groups, as larger minority populations are generally found to increase per capita consumption among these groups. We might expect similar effects to operate in local television news markets, and one contribution of this paper is to document the effect of market size and the distribution of tastes on the supply of local news programs and local news viewing.

Our paper also informs the debate on localism, which constitutes one of the three principles of media regulation in the US and is the subject of an interesting literature on the competition between national and local media products. George and Waldfogel (2006) documents that the national expansion of the *New York Times* attracted highly educated readers away from local media, triggering repositioning in local newspapers. George (2008) documents the effect of the spread of the internet on the composition of the local newspaper audience. Anderson et al. (1997) offer a theoretical framework for thinking about competition and welfare when national and local media compete. This literature is driven by the intuition that firms producing national products can spread fixed quality investments over a larger market than local producers. Since most of the expansion in both news and entertainment programming associated with improvements in television technology has been national, this mechanism might be expected to operate in television markets. Our estimates of the substitutability of national and local television news speak to this point.

From a welfare perspective, the substitutability of local for outside products matters most when local news generates positive behavioral externalities that are lost when national or "non-news" media are privately preferred. Demand-side externalities are now well documented: George and Waldfogel (2008) show that the *New York Times* expansion differentially reduced turnout in local elections among readers targeted by the Times. Gentzkow (2006) documents significant changes in local political participation during the expansion of television in the 1950's. Oberholzer-Gee and Waldfogel (2009) media show a relationship between Spanish-language news programming and voter turnout. Evidence suggests that spillovers are not limited to voting. Strömberg (2004) shows a relationship between public spending and radio access in the 1920's, and Snyder Jr. and Strömberg (2010) describe effects of local media markets on political competition. In the context of this paper, we would like to know, for example, whether public affairs programming on cable networks is a strong substitute for local news. More generally, our model allows us to identify under-provision and estimate the cost of policy intervention to correct it.

The paper also offers unique evidence of how advertising can shape media markets through station programming choice. Most theoretical and empirical research in this area follows the contribution of Anderson and Coate (2005) in emphasizing the role of advertising minutes and associated viewer disutility in welfare tradeoffs.[4] But in television, long-term contracts and fixed program length limit the role of advertising time as a strategic variable: *Friends* is 22 minutes long and the eight minutes of advertising allowed in its half-hour time slot cannot be readily altered in the short term. We focus instead on the imperfect substitutability of programming from the perspective of advertisers. Advertiser preferences for programming can arise from different valuations of audience demographics, from different valuations for the same audience in alternative contextual environments, or from closer targeting to advertised products. We do not in this paper distinguish among the sources of heterogenous valuation, but doing so is a promising avenue for future research.[5] The model developed in this paper demonstrates how scale parameters of an advertiser profit function can impact program choice by stations and resulting advertiser welfare.

From a methodological perspective, our empirical approach builds upon the rapidly growing work that incorporates ideas from game theory into the process of model estimation.[6] The literature on estimating discrete games centers on using information

[4]See Wilbur (2008) for an empirical example in television.

[5]The potential for overall differences in the marginal value of programming to viewers and advertisers has long been noted in the empirical media literature (e.g. Berry and Waldfogel (2004) but lacks comprehensive theoretical treatment. The potential importance of advertiser heterogeneity in scale in advertising markets is studied in George (2009), which measures the role of television in the decline of local beer. Most applications of the location choice literature to media markets considers political targeting, with no direct role for product advertisers. An exception is **?**, which considers advertising directly, but considers only advertisers distributed in a left-right political space.

[6]Early literature includes Reiss and Spiller (1989), Bresnahan and Reiss (1991), and Bjorn and Vuong (1984). For recent reviews, see Berry and Reiss (2007), (Ackerberg et al., 2007), Bajari et al.

revealed in market outcomes to estimate profit functions in the absence of information about profits. Our data, however, has detailed post-outcome information in the form of both viewership and advertising expenditure, which constitute the bulk of station revenue. In this regard, our methods relate more closely to the work of Zhu et al. (2009) and Zhu et al. (2009). One challenge that we have to deal with in estimating the model is that broadcasting decisions are complicated, and stations must decide on a sequence of broadcasts to offer, where broadcasts may be complementary. Consideration of broadcast sequences, or menus, greatly expands the strategy space for each station, and makes it difficult to check equilibria for multiplicity. We present a practical approach for dealing with large strategy spaces in estimation.

2.2 Local Television Background

There has never been truly free entry in local television. Since the start of the broadcast era, television stations have been licensed by the FCC to broadcast programming over specific portions of the frequency spectrum. Because broadcast signals can interfere with each other, the number of stations in any particular region is limited by the technology available to utilize spectrum. To accommodate these technological capacity constraints, the Federal Communications Commission in the 1950's allocated three stations in the largest markets, fewer in smaller cities, setting the stage for the three-network regime that dominated television through the 1980's. The limited number of stations in small cities gave rise early on to concerns about monopoly provision and under-supply of local programming, especially local news. But at the same time, the number of broadcast stations licensed in very large markets was not much larger than the number in small markets, leading to a wide disparity in the number of stations per capita across the US. The potential for under-provision was thus a subject of concern even in large markets.

Entry barriers for local stations need not translate into restricted entry for local news, as stations have many scheduling options to satisfy demand. However in practice, the limited number of station licenses in each market did likely limit local news programming. Local station license-holders negotiate contracts with national networks to carry network programs. Station scarcity meant substantial rents paid by networks to local license holders for airing national programs. The opportunity cost of forgoing national entertainment programming in favor of additional local news broadcasts has

(2010a), and Bajari et al. (2010b).

thus always been very high. These opportunity costs were highest in the largest, most constrained markets with the greatest number of viewers per station. As a result, the amount of local news programming during the broadcast era did not vary substantially across markets.

The spread of cable television dramatically lowered entry barriers for national programming. By both offering an alternative outlet for network entertainment and diverting viewers from local stations, the spread of cable reduced the networks willingness to pay for placement on local stations. This effectively lowered the opportunity cost of airing local news, and is likely the reason that more local news programming is broadcast today than in the broadcast and early cable eras. Cable expansion has led to entry of some stations carrying local and regional news, and we will consider the effect of these stations in our analysis. But limits to "must carry" rules combined with cable system maps that do not fully coincide with broadcast geography have limited these local stations to the very largest market.

In sum, economic theories of differentiated product markets justify policy attention to potential inefficiencies in the supply of local news, but offer little practical guidance on where to look for distortions with a reduced-form approach. Our structural model can both measure the extent of under- or over-provision and pinpoint its causes. This approach can also uncover the demographic or demand characteristics associated with inefficiencies.

With this background, we turn to our model of viewing, advertising and program choice.

3 Programming, Viewing and Advertising Data

We estimate our model with programming, viewing and advertising data in the 101 largest Designated Market Areas (DMA) in the US over four weeks in February 2010. We work with a single, averaged observation for each station in each of the six half-hour time slots in the 5:00-8:00 p.m. time period. We focus on this time period because it brackets the traditional early evening "news hours" when local stations have discretion over programming. Earlier in the day, a large fraction of the population does not watch television, while later in the day, prime time programming is dictated by national affiliation for many stations.

Raw viewing data come from Nielsen, which records the number of households viewing television during each fifteen-minute time slot each day on all local stations and the 100 largest cable stations. We work with the share of households viewing television, averaging viewing shares to the half-hour level to match with program time slots.

To construct our working data, we merge viewing data to program schedules from Nielsen and Kantar Media and categorize programming as local news, national news, general entertainment and cable entertainment. This categorization allows us to focus on the choice faced by local stations, to broadcast local news or entertainment programming, while preserving national news and cable entertainment broadcasting as choices for viewers. This simplification of program types also allows us to think intuitively about exogenous constraints faced by stations, which must broadcast national news at times specified by broadcast network affiliates, and also the cable broadcast schedules that local stations must take as given. Categorization also reduces the complexity of the estimation. Since we will consider broadcast menus as strategies, even with the assumption that local stations choose only local news or entertainment programming, across six time slots there are possibly $2^6 = 64$ potential strategies per station. With five stations in a market, the estimation must consider 64^5 potential profiles.

While most local stations broadcast a mix of program categories over the evening time slots we study, we observe little variation in programming schedules across the four weeks. This likely reflects that stations decide on a daily programming schedule and adhere to it for an extended period of time. We also observe that stations' category choices do not vary substantially across weekdays. Because of this, we further average our viewing data for each station over the four weeks in our sample. In our estimation we thus working with a single, averaged observation for each station-timeslot from 5:00-8:00 p.m.

Station and broadcast counts per market for small, medium and large markets are summarized in table 1. Small DMA's have an average of 7.4 local stations with 7.2 local news broadcasts per evening, while the largest markets average 15.6 local stations with an average of 14.6 local news broadcasts per evening. The number of major national cable stations available also increases with market size, ranging from an average of 64.7 in small markets to 94.2 in large markets. The average number of local stations in each market is 9.9. Note that we include local cable stations in our local station category, but for brevity refer simply to cable stations rather than national cable stations.

Table 2 shows an analogous breakdown with viewing data. In this table, the share of households viewing television is summed over timeslots and stations in each market, then the totals are averaged over each market category. Both overall and local news viewing declines steeply with market size. Local news viewing averages 7% in small markets while only 1% in the largest markets. Total viewing also declines considerably with market size, falling from 42% in small markets to 8% in the largest markets. The totals in this table likely reflect both a larger share of households watching television but also each viewing household watching during more timeslots each evening.

We characterize the typical television market in the early evening in table 3, which shows the average number of broadcasts and viewing share for local news, national news, broad-based entertainment on local stations and other entertainment on cable stations. Local news broadcasts are heavily clustered on the early timeslots, falling off considerably at 7:00 p.m. Total viewing increases over the evening, but local news viewing declines, both overall and per broadcast.

Advertising data for local stations come from Kantar Media and are at the advertisement level. We average prices per second across advertisements in each half hour timeslot to correspond to program times, then average again over weeks to match with our viewing data. We do not observe cable advertising prices. To convert advertising prices into revenue, we assume 10 minutes of advertising time per half hour timeslot. [7]

Table 5 reports average prices per second for local news, national news and general entertainment programming in each time slot. Prices in this table are averaged across all stations in a market, then averaged across markets. Prices per second generally rise through the evening, reflecting higher total viewing. National news prices are about 20% higher than local news prices on average, though because national networks dictate scheduling the price difference is not central to decision making. Across the evening, the local news and entertainment categories available for station choice are very similar, at $9.61 and $9.89 for the evening.

Table **??** reports average prices per 1,000 viewers per second, a closer analog to the "Cost per Mille (CPM)" figures used most often for cross-media comparison. As above, prices per viewer in this table are averaged across all stations in a market, then averaged across markets. As above, prices per viewer rise through the evening. Early in the evening, prices per viewer for entertainment on local stations are higher than localnews prices per viewer, but the pattern reverses after the 6:00 timeslot. Few news

[7] Although we observe some variation in the number of advertising seconds in our data, some of this variation is spuriously related to missing advertisements.

broadcasts are shown after 7:00 p.m., so the values in these cases reflect small numbers of programs.

Figures 2 and 1 provide a visual description of the broadcast and viewing information in the tables. Local news is disproportionately broadcast in the 5:00 and 6:00 timeslots, followed by national news at 5:30 and 6:30, with viewership of these two programming categories showing a similar pattern. As the broadcast evening nears prime time, news broadcasting drops off and is replaced by entertainment programming and viewership. Owing to the viewership pattern, advertising prices per second tend to be higher for local news early in the evening, and lower later in the evening. While we show simple averages across major US markets, it turns out that these patterns are consistent whether we categorize markets by size, geography, or time zone.

Our model of station decisions focuses on choice of a broadcasting sequence - i.e., in which, if any, time slots to offer local news. It is thus useful to think about broadcasts as a sequence of programs. Along these lines, table 6 tabulates the most common programming sequences offered by local stations. The table shows that general entertainment programming (represented by "o") is the most common station choice in 34% of timeslots. This is the typical strategy of small local stations. The next most common configurations show stations offering local news in either the two or three earliest slots in the day, followed by national news, then general entertainment programming. We use a sequence plot to represent common schedules visually in 8. The plot groups the most common sequences together in bars, which gives a visual impression of broadcast counts over time. We will return to this plot when discussing optimal scheduling.

Having described station programming in general terms, we turn to a discussion of potential inefficiencies that might arise through program choice in television markets. We then turn to outlining the model that allows us to measure welfare loss.

4 Program Choice and Welfare

We conceptualize a market in which stations make programming decisions to attract viewers according to a simultaneous, non-cooperative game. While we ultimately will model station choices over broadcast schedules subject to constraints, the basic welfare issues are clear from simple two-by-two games.

As a backdrop for the model, suppose that stations wish to attract viewers because more viewers means higher advertising revenue. Advertisers, however, may pay differentially for viewers of different types of programming for any number of reasons. Viewer utility is proportional to total viewership, so that viewers watch more television when doing so makes them happier than the alternative. Then, the socially most desirable programming configuration from the perspective of television viewers is the one that produces the most aggregate viewership.

Table 7 presents two examples, developed from our model, where two stations must choose whether to broadcast local news or entertainment programming. The top portion of each panel shows viewing associated with the program choices; the middle panel shows advertiser surplus; and the bottom shows station revenue. In the left hand panel, the optimal programming combination from the perspective of viewers is for one station to broadcast local news and the other station to broadcast entertainment programming. Total viewership is 0.078, which is greater than it would be if both stations broadcast local news (0.048), or if both stations broadcast entertainment programming (0.070). One interpretation of this viewing outcome is that viewers have a taste for variety. There is little substitution between programming types; viewers watch their preferred program or nothing at all.

The center and lower panels show advertiser surplus and station revenue from the same programming choices. Since revenues guide station decisions, the Nash equilibrium occurs with both stations opting for entertainment programming. Thus, in equilibrium the two stations provide less variety in programming than viewers would like. Advertisers also prefer this outcome, as the aggregate advertiser surplus - the value advertisers get net of the costs of advertising - is 52, as opposed to 48 when more variety is offered to consumers. The outcome with both stations broadcasting entertainment also maximizes joint station profits. Thus, the only dimension to the welfare loss in this case is that viewers get less variety than optimal. Advertisers and stations cannot improve on their outcome unilaterally or through collusion.

A second set of potential payoffs are shown on the right side of the table. Viewers again have a preference for variety, with total viewing highest with differentiation. However in this case, both advertisers and stations also see higher total surplus or revenue with differentiation. Competition for more profitable programming leads to a Nash equilibrium where both stations broadcast entertainment programming. This outcome involves welfare losses to all market participants.

Both examples were constructed using formulas that follow from our model. For the left panel, parameters are chosen such that: 1) viewer utility from watching local news is lower than that of other programming; 2) local news broadcasts are more substitutable for one another than entertainment broadcasts; and 3) advertisers are willing to pay the same amount per viewer for local news viewers and entertainment viewers. The parameter choices in the second example maintains the assumption that entertainment programs are less substitutable for each other than local news broadcasts. But instead of differential utility for viewing and common advertiser valuations, the second panel parameters assume viewer utility is constant across program types but advertiser valuations for local news viewers is less than for entertainment viewers. Our results suggest that both of these configurations of viewer, advertiser, and station interests reflects actual outcomes in many markets.

In both of these examples, stations decisions are not the ones preferred by viewers, but they are optimal for advertisers and the stations themselves. However, it might be the case that stations would be jointly better off differentiating, but asymmetric payoffs meant that no station individually has an incentive to be the one providing local news rather than entertainment programming (or vice versa). Our estimates and simulations will allow us to characterize welfare losses.[8]

Other sorts of interest conflicts may arise from different dimensions of the variety problem, or the way in which news viewing enters the consumers utility function. For example, viewers might be less inclined to watch local news on a station later in the broadcasting period because another station already broke the news earlier in the day. In this instance, stations might wish to broadcast their news before other stations, leading to all stations (inefficiently) broadcasting news as early as possible.[9]

With this overview of potential welfare tradeoffs, we turn to our model of news broadcasts.

5 Model

We balance several objectives in modeling television programming choice, viewership and advertising. The primary function of the model is to simulate and predict counterfactual outcomes such as viewership shares and advertising revenue under different programming choices. Therefore, the predictive ability of the model is of primary im-

[8]It also bears mentioning that there are situations in which there are no conflicts of interest per se but there may be a multiplicity problem. That is, stations provide variety, but the wrong stations provide the broadcasts of each type.

[9]The dynamics of business stealing are explored theoretically in Ted Bergstrom and Bills (1995).

portance. That said, we would also like to set up the model so that parameter estimates have useful interpretations. The model should also capture important aspects of news broadcasting and viewership; breaking news might attract more viewers, or viewers might prefer to watch local news before national news. Finally, the model should capture key welfare tradeoffs, such as whether advertisers pay more for different types of viewers, and how substitutable broadcasts of different types are for viewers.

We employ a nested multinomial logit structure to model viewing and a log-linear model of advertiser revenues. In both cases, estimated parameters provide direct insight into the magnitude and direction of the important governing effects of the model. A nested multinomial logit model has several attributes which make it desirable for modeling television viewing. Berry (1994) showed how a NMNL can be estimated using a linear estimating equation when the dependent variable is a market share. Linearity eases inclusion of fixed or random effects, which allows us to thoroughly control for unobserved station and market characteristics. A NMNL is also based on a simple and tractable random utility model, simplifying calculation of viewer utility. Finally, market shares deriving from a NMNL always fall in the $(0, 1)$ range and always sum to unity across the market. This feature is important in our setting, as many shares are close to zero, and we seek to estimate accurately how one station's broadcasting decision impacts the viewership of other stations.

Ackerberg (2006) has noted that the NMNL requires very restrictive assumptions on substitution patterns and errors. Fortunately, Ackerberg (2006) shows how the NMNL can be made more flexible by including additional terms. To err on the side of caution, we include some additional terms in our NMNL model, and also include a variety of market, time, and station specific effects in the model.

5.1 Viewing

We allow for the possibility that a representative viewer may first choose a type of programming then choose a particular broadcast of the chosen type. For example, the viewer first might choose to watch local news, and then select a broadcast on one of the stations broadcasting local news. The program types we consider are local news, national news, general entertainment and cable entertainment programming. We denote this viewer choice set as $B = \{l, n, o, c\}$. Further, we use $S_b, b \in B$ to denote the set of stations broadcasting a program of type $b \in B$. We use $b_i \in B$ to denote the broadcast type of station i.

Following Berry (1994), write a station i's viewership share in a particular time slot as:

$$s_i = s_{i|b} s_b \tag{1}$$

Where $s_{i|b}$ is station's i's share of stations with broadcast type b, and s_b is the broadcast group share. If $u_k(b_k)$ denotes an index for the utility a representative viewer gets from watching station k when k broadcasts programming of type b_k. In a nested multinomial logit model, a substitutability parameter $\mu_b \in [0, 1]$ is used to characterize the degree to which items within a given type substitute for one another. As in Berry (1994), define:

$$D_b = \sum_{k \in S_b} e^{\frac{u_k(b_k)}{1-\mu_b}} \tag{2}$$

Then, the components of s_i in (1) can be written as:

$$s_{i|b} = \frac{e^{\frac{u_i(b_i)}{1-\mu_b}}}{D_b}, \quad s_b = \frac{D_b^{1-\mu_b}}{\sum_{k \in B} D_k^{1-\mu_k}} \tag{3}$$

As μ_b approaches one, items within a class become very close substitutes, as small differences in utility lead to large differences in within-group shares. At the same time, μ_b close to one implies that the group share does not change much as more broadcasts of a given type are added. Goods remain substitutes as μ_b approaches zero. However, within-group choices in this case are governed by a regular multinomial logit in which items can be viewed as being compared pairwise. We follow the usual approach of normalizing utility by setting the utility from the null alternative - not watching television - equal to zero, so $u_0 = 0$, and $D_0 = 1$.

Berry (1994) describes how to cast the model in linear form. Since $s_0 = \left(\sum_{k \in B} D_k^{1-\mu_k}\right)^{-1}$, (3) and (1) can be used to get:

$$\ln s_i - \ln s_0 = \frac{u_i(b_i)}{1-\mu_b} - \mu_b \ln D_b \tag{4}$$

From equation (3), we have:

$$\ln D_b = \frac{u_i(b_i)}{1-\mu_b} - s_{i|b} \tag{5}$$

Plugging (5) into (4) and simplifying gives:

$$y_i = \ln s_i - \ln s_0 = u_i(b_i) + \mu_b \ln s_{i|b} \tag{6}$$

Another advantage of the share-form nested multinomial logit model is straightforward calculation of the total expected utility of the representative consumer from viewing. McFadden (1980) showed that the NMNL derives from a utility function with the form:

$$U = \ln \left[\sum_{k \in B} D_k^{1-\mu_k} \right] \tag{7}$$

Equation (7) provides an alternative derivation of shares based on application of Roy's identity (differentiation of (7) with respect to $u_i(b_i)$ yields s_i). For our purposes, the important thing about expression (7) is that it links observed or counterfactual viewership shares to utility. Since $s_{i|b} = \frac{s_i}{s_b}$, utilities can be expressed using (6) as follows:

$$u_i(b_i) = \ln \left[\left(\frac{s_i}{s_0} \right)^{1-\mu_b} \left(\frac{s_b}{s_0} \right)^{\mu_b} \right] \tag{8}$$

Plugging (8) into (2) results in:

$$D_b = \sum_{k \in S_b} \frac{s_i}{s_0} \left(\frac{s_b}{s_0} \right)^{\frac{\mu_b}{1-\mu_b}} \tag{9}$$

Since $s_b = \sum_{k \in S_b} s_k$, D_b in (9) simplifies to:

$$D_b = \left(\frac{s_b}{s_0} \right)^{\frac{1}{1-\mu_b}} \tag{10}$$

Inserting this last result into the utility function (17), we have:

$$U = \ln \left[\sum_{k \in B} \frac{s_k}{s_0} \right] = -\ln s_0 \tag{11}$$

The last part of (11) follows from the fact that the sum of group shares must sum to one. Hence, the utility gains from viewership vary in inverse proportion with the fraction of non-viewers.

5.2 Advertising

We follow Berry and Waldfogel (1999) in modeling advertising revenue. The idea is to model station advertising revenues so that a measure of consumer's surplus can be computed, where in our application the consumers are advertisers. Suppose advertisers have diminishing willingness to pay for additional viewers, and that price-per-viewer is set competitively. Suppose further that price-per-viewer is of an exponential form:

$$ppv_j = \tilde{K} v_j^{\eta-1} \tag{12}$$

where $\eta \in (0,1)$. Multiplying ppv in (12) by the number of viewers v_j then gives an expression for how advertising revenues depend upon total viewership:

$$r = ppv_j * v_j = \tilde{K}_j v_j^{\eta}$$

That is, price-per-viewer multiplied by the number of viewers gives total advertising revenues (per unit time). Taking logs gives the linear relationship:

$$\ln r_j = K_j + \eta \ln v_j \tag{13}$$

Equation (13) is a log-linear equation describing how total advertising revenues depend upon viewership. As was the case with the viewership model, one convenience is that a linear estimating equation can be matched quite flexibly to data through inclusion of fixed and/or random effects.

We can use the willingness to pay function (12) to study advertiser's surplus, and how it depends on viewership. Integrating (12) with respect to total viewers gives $\frac{\tilde{K}}{\eta} v^{\eta}$, which is the gross "utility" advertisers derive from contact with v viewers. Subtracting from this r - what is paid by advertisers to attract the viewers - gives advertiser's surplus as:

$$AS_j = \frac{\tilde{K}_j}{\eta} v_j^{\eta} - \tilde{K}_j v_j^{\eta} = \frac{1-\eta}{\eta} \tilde{K}_j v_j^{\eta} \tag{14}$$

Thus, advertisers' surplus amounts to weighting observed revenues in a way that depends upon the elasticity of willingness to pay with respect to viewers. By allowing the parameters of (13) to vary with broadcast type, one can allow for advertisers to weight and value viewers of different sorts of programming differently.

6 Estimation

The viewership estimating equation (6) and the advertising revenue equation (13) form the core of our model. The rest of the model controls for the impact of strategic, optimizing behavior on the part of stations in estimating these two equations. Our assumption is that stations make programming choices to maximize advertising revenue across the early evening, which entails attracting viewership. We detail how this assumption impacts estimation after describing how we supplement (6) and (13) to increase flexibility.

6.1 Viewing

The viewership data covers a large number of stations across markets spanning multiple time slots. Accordingly, we index viewing shares by station i, market m, and time slot t. We break the viewer utility index $u_i(b_i)$ into a component depending on observable covariates, market-time random effects, station specific random effect, program type, and an unobserved idiosyncratic error term. With these modifications, equation (6) becomes:

$$\underbrace{\ln s_{jmt} - \ln s_{0mt}}_{y_{jmt}} = \mu_b \ln s_{jmt|s_{bmt}} + \underbrace{\underbrace{\alpha_b^v + X_{jmt}^v \beta_v + \zeta_b^v \ln(1+n_b)}_{z_{jmt}^v} + \underbrace{\omega_j^v + \omega_{mt}^v + \varepsilon_{jmt}^v}_{e_{jmt}^v}}_{u_{jmt}} \tag{15}$$

The superscript v is used to distinguish terms in the viewership equation from similar terms in the advertising revenue equation. As we detail when describing estimation results, the vector of covariates X_{jmt}^v includes elements designed to capture lead-in effects and other features of television broadcasting. The term $\zeta_b^v \ln(1 + n_b)$ is a way of including a function of the number of broadcasts of a given type, which Ackerberg (2006) suggests should be included in the model to make errors flexible. The underbraces in (15) indicate how viewing utility depends upon a systematic component z and a random component e, which contains station, market-time, and idiosyncratic components. We re-introduce y_{jmt} from equation (6) as a simplifying notation for the right hand side of (15).

As outlined above, we model programming as nested into the four categories of general entertainment, national news, local news, and cable entertainment. We adopt an error structure with market-time, station, and idiosyncratic random effects. We add to the estimation equation independent variables that capture intuition and have strong predictive power. These include dummy variables for programming category, and also 1) a set of variables that capture lead-in effects, 2) a set of variables that capture a "breaking news" effect, and a set of variables that allows for additional flexibility in the error terms. We have also observed that viewership shares appear to be systematically smaller in larger markets, so we also include a variable capturing market size.

We include lead-in effects in the viewership model by including lagged viewership shares, and a set of dummies that indicate whether a local news broadcast follows another local news broadcast, a local news broadcast follows a national news broadcast, a national news broadcast follows a local news broadcast, or a national news broadcast follows another national news broadcast. We interact the lagged viewership share with these dummies. The idea behind these variables is that people might have beliefs or expectations about the correct sequence within which to ingest news, in addition to capturing the standard inertia in channel viewership.

We also interact local news and national news with the cumulative share of local and national news viewed in a market up to the current time. This is how we allow for a "breaking news" effect in the model, in that people might be less likely to watch news later in the day if news has already been broadcast earlier on some other station.

6.2 Advertising

We break the constant term K_j in (13) into a component depending on observable covariates, market-time random effects, and station random effects, as we did with the viewership model. We allow for the elasticity of revenue with respect to viewership to depend upon the type of programming. This allows advertisers to weight additional viewers of different types of programming differently. The revenue equation becomes:

$$\ln r_{jmtb} = \underbrace{\alpha_b^p + \eta_b \ln v_{jmtb}}_{z_{jmt}^p} + \underbrace{\omega_j^v + \omega_{mt}^v + \varepsilon_{jmt}^p}_{e_{jmt}^p} \tag{16}$$

6.3 Maximizing advertising revenue

Stations choose broadcast lineups to maximize advertising revenue. In our model, stations choose over program types throughout the early evening into prime time. Of course, maximizing revenues across the evening is different than maximizing revenues obtained in a given half-hour time slot. This is because of viewer lead-in effects, which might be program type-specific.[10] For example, people may be more likely to watch national news after a local news broadcast, so a station may prefer to broadcast local news prior to national news even though other programming would give more viewership and revenue in the current time slot.

Of the four types of programming categories ($B = \{l, e, n, c\}$), only "local news (l)" and "general entertainment (e)" can be selected by local stations. Local stations view cable schedules as fixed and are obligated by national contracts to broadcast national news at certain times. Hence, when a station is free to choose programming, it makes a choice from a binary local choice set, $B' = \{l, e\}$. National news is a fixed broadcast from the perspective of the station, around which other broadcasts must be arranged.

While the binary nature of programming choice is a simplification, it still admits a large number of broadcast menus for local stations over the six early-evening time slots. If a station does not broadcast national news at all, it is making a binary l, e programming decision in six time slots in our observation period. This creates $2^6 = 64$ possible programming sequences.

Expanding previous notation a bit, we write u_i in (15) as:

$$u_i = u(b_i, z_i, e_i, \theta_v) \tag{17}$$

That is, viewer utility depends upon broadcast type and parameters θ_v, and has both systemic and idiosyncratic components. The NMNL creates a viewership share from the utility indices of all firms in the market, which we may alternatively write as:

$$s_i = s_i(u_i, u_{-i}, \theta_v) \tag{18}$$

Similarly, we expand on notation and write our revenue equation (16) as:

$$r_i = \tilde{r}_i(v_i, b_i, \theta_p) \tag{19}$$

[10]The importance of lead-in effects in television viewing is well documented, see for example Esteves-Sorenson and Perretti (2012), and Wilbur (2004).

By virtue of equation (18), (19) can be rewritten:

$$r_i = \tilde{r}_i(Ms_i(u_i(b_i, z_i, e_i), u_{-i}(b_{-i}, z_{-i}, e_{-i}), \theta_v), b_i, e_i^p, \theta_p) \tag{20}$$

Expression (20) reflects that advertising revenue depend upon programming choices and viewing, which in turn depends upon the viewership model. It is helpful to write (20) in a more streamlined form:

$$r_i = r_i(u_i, u_{-i}, b_i, e_i^p, \theta) \tag{21}$$

The convenience of (21) is that it provides a recipe for computing counterfactual advertising revenues. One first obtains a counterfactual utility index u_i', and uses this to compute a counterfactual viewership share s_i' using (18). Then, the share can be inflated into viewership and inserted into the revenue function to obtain hypothetical revenues from broadcasting different sorts of programming.

To introduce a time component to the model, let $b_i = \{b_{it}\}_{t=1}^T$ now denote station i's "broadcast lineup," and let b_{-i} denote the lineups of other stations in the market. Let $\mathcal{B}_i$ denote all of station i's possible lineups. Revenue maximization by i implies that station i has chosen its revenue maximizing lineup from available lineups, given the broadcast lineups of other stations. Aggregate revenue is:

$$R_i(u_i(b_i), u_{-i}(b_{-i}), \theta) = \sum_{t=1}^{T} r_{it}(u_{it}(b_{it}), u_{-it}(b_{-it}), \theta) \tag{22}$$

Under the maintained assumption that stations engage in Nash competition in broadcast choices to maximize profits, we require that the broadcast menu chosen by a station maximizing a station's revenues, given the broadcast menus chosen by all other stations. Using (22), this means that:

$$\begin{aligned} R_j(u_j(b_j, z_j, e_j), u_{-j}), \theta) \geq R_{jm}(u_j(b_j', z_j', e_j'), u_{-j}), \theta) \\ \forall b' \in \mathcal{B}_j, \forall j \in L \end{aligned} \tag{23}$$

Equation (23) reflects the idea that the broadcast sequence actually chosen by station j maximizes revenues, given the broadcast sequences chosen by other stations. The condition $\forall j \in L$ means that the condition applies to all the local stations making programming decisions in the market.

Condition (23) implies that each station is choosing a broadcast menu to maximize revenues given the decisions of other stations, so that (23) means observed programming in a market is a Nash equilibrium.

6.4 Likelihood

We now describe the likelihood function built around the viewership and revenue equations, introducing additional details as needed. The contribution of each market m to the likelihood function can be broken into four components: the viewership model, the advertising revenue model, the revenue maximization condition, and then a correction term necessitated by the assumption that outcomes are consistent with the outcome of a game. The last term derives from condition (23) not fully determining a unique market outcome, as pointed out by Ciliberto and Tamer (2010) and Bajari, Hong, and Ryan (2011). This "cohesiveness" problem has to be addressed in the likelihood function.

The likelihood contribution of a market as a function of its four components is:

$$L_m = L_{vm} L_{pm} L_{rm} L_{cm} \tag{24}$$

The first two components of the log-likelihood are straightforward. The latter two are somewhat simple conceptually, but induce a degree of complexity in estimation.

6.4.1 Viewership Likelihood

To form an expression for L_{vm}, rewrite equation (15) as:

$$e^v = y - \mu_b \ln s_{.|b} - \alpha_b - X^v \beta_v \tag{25}$$

Which expresses the joint likelihood of y in terms of e^v. Accordingly, we have:

$$L_{vm} = f_v(e^v, X_v, \theta_v) \left| \frac{de^v}{dy} \right| \tag{26}$$

where $\left|\frac{de^v}{dy}\right|$ is the Jacobian determinant of e^v with respect to y. In the appendix, it is shown that this determinant is:

$$\left|\frac{de^v}{dy}=\right|\prod_{b\in B}(1-\mu_b)^{N_b-1}$$

That is, to compute the jacobian determinant, one simply counts the number of broadcasts of each type, subtracts one from this count, and the raises the terms $(1-\mu_b)$ to the power N_{b-1}. Under the assumption that the terms are jointly normally distributed according to the covariance matrix Ω_v, where Ω_v is built to take into account common market-time and station-level random effects.

6.5 Revenue Likelihood

For L_{pm}, we rewrite equation (16):

$$e^p = \ln r - \alpha_b^p - \eta_b \ln v \tag{27}$$

As the transformation from dependent variable to error term for (27) is unitary, we have simply

$$L_{pm} = f_p(e^p, X_p, \theta_p) \tag{28}$$

We again assume that e^P is normally distributed with covariance matrix Ω_p, where Ω_p considers market-time and station-level random effects.

6.6 Profit Maximization

Profit maximization requires that we calculate:

$$L_{rm} = \int_{e'v_j, e'p_j \in \Gamma_j} R_j \geq R_j(u_j(b_j', z_j', e_j^{'v}), u_{-j}, e_j^{'p}), \theta) \qquad \forall b' \in \mathcal{B}_j, \forall j \in L \tag{29}$$

That is, a term calculating the probability that observed advertising revenues from observed broadcasting decisions given parameter values and data must be added to the likelihood. While conceptually clear, condition (29) obviously presents some computational challenges, as the region Γ_j is quite complicated. How we addressed these challenges will be discussed in the estimation section of the paper.

6.7 Nash Equilibrium

To close the empirical model, we also need a condition describing an equilibrium selection mechanism. That is, conditional on (29), which means that observed broadcast menus across the market *could* be an outcome, what are the chances that they actually *will* be a market outcome? We write this condition as:

$$L_{cm} = P(b|R \geq R(u(b', z', e^{'v}), e^{'p}, \theta) \tag{30}$$

We now give an overview as to how the model likelihood $L = \prod_{m=1}^{M} L_m$, where L_m is comprised of (25), (28), (29), and (30), is actually maximized to produce estimates of model parameters θ. Once we have these parameters, we can then recover implied random effects and use the estimated model to conduct counterfactuals.

6.8 Estimation Details

In this section we stress the key aspects of estimating the model, referring to an online appendix for the finer details. While most of the model can be estimated without difficulty using standard maximum likelihood methods, selection effects and game-theoretic aspects of the problem render estimation a bit more difficult. Selection effects arise because stations choose the observed broadcast menu because it is expected to generate higher profits than alternative broadcast menus. Game-theoretic concerns arise because stations make their revenue maximizing broadcasting decisions in consideration of the broadcasting decisions of other stations.

Both of these aspects of the problem require consideration of a variety of counterfactuals and asking questions about what prices and viewership would be if stations had followed different courses of action. Because these counterfactuals are computationally costly, we adopt a simulation-based approach to estimation that follows Ackerberg (2010), and also borrows elements from Bajari, Hong, and Ryan (2010), and Chernozhukov and Hong (2003). Our approach is to estimate a simple, preliminary model, and then based on this model simulate counterfactual shares and revenues consistent with the observed broadcasts in each market constituting a Nash equilibrium. We then check these simulated market outcomes for other equilibria and correct the simulated likelihood for multiplicity or cohesiveness (see, e.g., Ciliberto and Tamer (2010)). We then weight the simulations according to their likelihood value, and reestimate the model. These final estimates are then taken to be the right ones - fully corrected for multiplicity and selection.

In the course of estimation, we restrict all of the multinomial substitutability parameters to the $(0, 1)$ range. One advantage of the MCMC method is that in the event that a parameter tends to either its upper or lower limit, estimation can proceed without creating problems created by, say, numerical differentiation.

While the details are discussed in the appendix, some additional aspects of estimation should be mentioned. For one, since we consider broadcast menus as strategies, even under our assumption that each station may select only local news or entertainment programming, across six time slots there are possibly $2^6 = 64$ potential strategies per station. If there are five stations in a market, there are 64^5 potential profiles that must be checked. The computational costs of checking each strategy so thoroughly is prohibitive, so we employ what we believe to be a novel markov-chain-monte-carlo approach to finding additional equilibria of the game. The whole model was estimated using methods described in Chernozhukov and Hong (2003), and Baker (2014).

As described in the appendix, estimation proceeds by piecing together each of the four components of the likelihood function. While there are closed-form expressions for the observed viewership share and the observed price, closed forms are impossible to calculate for the revenue maximizing condition and the nash equilibrium condition. For each market and station that we assume is choosing a broadcast menu, we therefore simulate $S = 25$ drawn counterfactual errors, where errors are drawn so that the observed revenue and viewership outcome produces the highest revenue over the course of the evening, given the broadcast menus of other stations. We then adjust likelihood by observing whether or not another equilibrium outcome is consistent with any of our draws. The details as to how we performed this drawing, and some of the nuances involved, are described in the appendix. To get starting values for the model, we first fit the model in simple linear form - i.e., without the correction for the possibility of multiple equilibria or without assuming what is observed is a revenue-maximizing outcome on behalf of stations. These starting values provide a point of contrast with our eventual model estimates.

One thing that is worth mentioning before discussing model results is that we do not rely on a classical estimation method. This is for practical as well as conceptual reasons.

Because we have an interest in generating a predictive model, we have included a variety of different sorts of controls, substitution parameters, and random effects. It is bound to be the case (and indeed, may even be of some interest) that some of these effects are not well-identified. With this in mind, we employ a Markov-chain Monte Carlo method of estimation, as it experiences none of the problems a direct optimization method might encounter when faced with an overidentified model (i.e., missing derivatives). Details of how the draw occurs are also described in the appendix; readers with general interest should consult Chernozhukov and Hong (2003) for theory, and Baker (2014) for some details of practical application.

6.9 Estimation Results

Estimation results are presented in table 8. For purposes of comparison, we present both the starting values and the final estimation results. Starting values were obtained by ignoring selection and multiplicity and estimating viewership and pricing models separately by maximum likelihood.

The top portion of the table reports program substitutability parameters from the viewer model. The estimated values are all close to zero, indicating that when the model is fully saturated with market-time and station fixed effects as well as broadcast counts, the nesting parameters are not that important in describing viewing choice. In other words, the nested logit adds little over a standard multinomial logit in viewer program choice. In practical terms, this indicates that viewers tend to look at the spectrum of broadcasts and make a choice rather than first choosing a category. [11]

The program type dummies show that local stations typically get more viewers than the control group (cable stations), and that this effect is even bigger for local news broadcasts.

The Dynamics portion of the table shows a series of controls for lagged viewing by program type. Overall, the magnitude and statistical significance of these estimates indicates that the dynamics matter – early viewing spills over into later viewing, and local news viewing especially bumps national news viewing. The results are reminiscent of Wilbur (2008), which emphasizes dynamics in viewing.

The negative cumulative viewing estimates in the fourth section of the table indicate

[11] Our result contrasts markedly with the estimates in Berry and Waldfogel (2004) for radio, where broadcast substitutability parameters approach unity revealing that categories are highly differentiated from the listener perspective.

a substantial breaking news effect for local news. Higher local news viewing earlier in the evening is associated with lower local news viewing in the current timeslot. This result suggests one reason why local news broadcasts are clustered early in the evening. This effect does not operate for national news (the parameter estimate is close to zero) indicating a reason that local news typically precedes national news.

In terms of the variances of the random effects, a substantial share of variation in viewership is more or less split equally between station and market-time fixed effects, with the contribution of idiosyncratic errors being smaller.

Estimates for advertising prices are shown in the Revenue portion of the table. The elasticity estimates indicate that the demand for local news viewers is modestly more inelastic than for entertainment viewers (0.6 vs. 0.7), controlling for station and market-time fixed effects. In other words, an increase in the number of viewers raises the advertising price per second more for entertainment broadcasts than for local news broadcasts. However, the modest difference in the estimates indicates that the number of viewers is more important in determining station price than the type of viewer. In our welfare analysis this will be evident again in examining how much welfare loss is due to two-sided market tradeoffs versus business stealing. The table also indicates that advertiser demand for national news viewers is substantially less elastic than for local news viewers, 0.3 vs. 0.6, however this likely reflects the small number of similar stations that broadcast national news.

In terms of random effects, the station, the market-time, and the idiosyncratic components of viewership contribute equally to variation in revenue across stations.

Overall the coefficient patterns in the equation-by-equation and selection models are similar, with two key differences. The substitutability parameters for local news and entertainment are substantially lower in the selection model, indicating that program categories are less important in choice than in the naive approach. Ignoring selection would also underestimate the importance of dynamics in program scheduling.

With these basic parameter estimates, we turn to the counter-factual simulations that allow for welfare analysis.

7 Welfare Analysis

The goal of this section is to evaluate and compare the three program configurations that maximize viewer welfare, advertiser surplus and station revenues. We make three comparisons. First, we assess generally how far the observed program configuration lies from the configuration that maximizes viewing (viewer optimum), the configuration that maximizes advertiser surplus (advertiser optimum) and the configuration that maximizes station revenue (station optimum). We then compare the viewing optimum to the advertiser optimum, which shows the potential tradeoffs between viewer and advertiser welfare in the context of a two-sided market. Finally, we compare the program configuration that maximizes station revenue to the observed outcome, which provides a sense of the extent to which classic business stealing reduces welfare in relation to a fully collusive outcome. With these comparison's, we proceed to evaluate the characteristics of markets most likely to experience welfare loss from sub-optimal program choices and discuss policy remedies best suited to the competitive environment.

To conduct our simulations and counterfactual analysis, we first estimate the random effects for each station and each market-time block. With these terms, we then simulate the model 25 times. After simulating the model, we calculate market equilibria for each simulation, and then redraw from the simulations so as not to overweight simulations with multiple equilibria.

Because the number of possible sequences and strategy profiles for each station and market is large, it is impractical to search through them all when trying to maximize viewership, joint station profits, or advertisers' surplus. We therefore use a a Markov-chain Monte Carlo approach to maximize shares. We begin by randomly perturbing a station's broadcast profile, check to see that the proposed change in strategy increases the quantity of interest, and proceed in this fashion until further guesses fail to find an improvement. For our evaluation, we average the results across the 25 simulations.

An overall summary of the welfare calculations is presented in table 9. The top portion of the table shows the optimal number of local news and entertainment broadcasts from the viewer, advertiser and station perspectives along with observed outcomes. (We show national news and cable broadcasts for completeness, but these are taken as given by stations so do not vary across profiles.) Both broadcast and viewing are summed over the entire evening in each market, so the viewing totals are best viewed as program-views per capita.

The first row indicates that the average market shows 10.2 local news broadcasts per night. The program allocation that maximizes total viewership would include 20.7 local news broadcasts on average, approximately double. Total viewing at the viewer optimum would increase from 0.297 to 0.325, about 10%. As would be expected, local news viewing increases substantially, from 0.049 to 0.084, almost double. Local news viewing increases more than entertainment and national news viewing falls, suggesting additional local news broadcasts would increase viewing overall rather than simply reallocate viewing across programming.

The third and fourth data columns report the program allocation and outcomes associated with the advertiser and station optimum. Interestingly, the number of local news broadcasts at the advertiser and station maximum is higher than observed levels by about 4 broadcasts, about 45%, though not as high as at the viewer maximum. Local news viewing would be about 5% higher at the station maximum, and entertainment viewing would increase as well.

The bottom two rows of the table provide a measure of advertiser surplus and station revenue at the different program allocations. The reported surpluses and prices are averaged rather than summed across markets in this table. Total surplus and revenue can be approximated by multiplying the averages by 3600, or 600 seconds of advertising per timeslot times 6 timeslots.[12] The table indicates that a shift to the viewer optimal allocation would reduce station revenue by 11%, while a shift to the station optimum would increase revenue by 12%. Advertiser surplus would increase substantially with a re-allocation of programming.

The table reveals some interesting welfare results. There clear is evidence of a two-sided market tradeoff in product choice. In dividing approximately 52 program timeslots on average, viewers prefer 21 local news and 31 entertainment broadcasts while advertisers prefer 15 local news and 37 entertainment broadcasts. The difference of 6 broadcasts (12%) reflects the size of the two-sided market distortion. However, the fact that observed broadcasts deviate from even the station maximum indicate that business stealing occurs in the market. The difference between the station optimum of 15 local news and 37 entertainment broadcasts and the observed outcome of 10 and 42 suggest a mis-allocation of 5 broadcasts, almost as large as the two-sided market distortion.

[12]Recall that we observe very little variation in advertising minutes in our data, with an average of about 10 minutes per 30 minute timeslot.

The price and surplus simulation results allow us to quantify the cost of better satisfying viewers. A move from the observed profile to the viewer equilibrium would decrease station revenue by \$1.25 x 3600 or about \$4,500 per station. With an average of about 10 stations per market, this is \$45,000 per market on average. The shift would increase viewership by 0.028 from the observed to the viewer optimum, which implies a cost of approximately \$16,000 to stations to increase total viewing by one percentage point. We cannot, however, estimate whether the benefits of this shift outweigh the cost, as we cannot with our model quantify the relative value of broadcasts to viewers.

The role of business stealing can perhaps be seen best with a market example. Table 10 shows the program allocation in Lexington, Kentucky during the 7:00-7:30 timeslot. All of the local stations broadcast entertainment during this period. The viewer optimum includes one local news broadcast. If WLEX were to switch to local news broadcasting, station revenue would fall. However gains to competitors broadcasting entertainment would rise substantially more than WLEX would decline, increasing total revenues in the market by about 8%. Both local news and entertainment viewing would also increase.

Timing plays an important role in broadcast outcomes. Table 11 shows observed and optimal broadcasts and viewing by timeslot, averaged across markets. The differences between the observed and viewer optimum are small early in the evening, increasing later, with the largest gap in the 7:00-8:00 timeslot heading into prime time. The differences between observed and optimal local news and entertainment viewing also diverge as the evening progresses. The differences can perhaps be captured best visually. Figure 3 shows a sequence plot that shows observed broadcast patterns over the evening in relation to the viewer, advertiser and station optimum. The value of more differentiated programming across the evening can be seen in the multicolored bars of the optimal allocation relative to the solid areas on the observed outcome.

Table 12 provides information on cost tradeoffs at the timeslot level. Switching program configurations to the viewer maximum reduces the price per second that stations earn by 1-4% and increase viewing by a commensurate amount. In proportional terms, the lowest cost for largest gains is in the middle to late time period.

To better understand possible drivers of inefficiencies, we regress deviations from optimal outcomes on market characteristics. Results are shown in table 13. The first column shows the deviation of observed outcomes from the viewership maximizing profile. The dependent variable is the difference between the viewer maximizing viewing share and the observed viewing share, expressed as a percentage. Larger, wealthier, younger markets see the largest deviation from the viewer optimum in equilibrium. The results also indicate that a larger black fraction is associated with a smaller gap and a larger hispanic population with a larger gap. The results support the marketing notion that advertiser's care more and thus perhaps impose a greater influence in these markets.

The second column and third columns shows the difference between the viewership maximizing profile and the advertiser optimum, a measure of the potential for advertiser bias. Column two reports the difference in viewing and column three in advertising prices, again expressed as a percentage. Market size is positively correlated with the gap in viewership, indicating large markets show the greatest potential for distortion. Larger markets also have a larger gap (more negative) in advertising prices between the viewer and advertiser optimum. An implication of this result is that the scope for two-sided market distortions away from viewer preferences toward those of advertisers is greatest in large markets. Wealth is not associated with a greater viewer distortion, but is associated with a greater (more negative) gap in advertiser prices. Expressed another way, the potential for advertiser bias is highest in large, wealthy markets.

The fourth column speaks to business stealing effects. The dependent variable in this specification is the difference between the station revenue maximizing allocation and the observed equilibrium, again expressed in percentage terms. Larger markets have the largest gap, experiencing the largest losses from business stealing. This is consistent with standard results that suggest coordinated action among firms is more difficult with more market participants, though in this case coordinated action would improve welfare.

More broadly, we find that the observed number of local news broadcasts is below the station optimum in 84 of 101 markets. Advertising prices, which convert directly to station revenues, are in all markets below what would be expected under a fully collusive outcome. Observed prices are even lower than predicted under the viewer maximum in 13 markets. Taken together with the results in 13, the welfare estimates suggest that coordination among stations would increase welfare.

8 Conclusions

We estimate the welfare consequences of local news broadcasting decisions in advertiser-funded television, finding a substantial shortfall in local news provision relative to the viewer optimum. The shortfall is greatest late in the evening news hours from 7:00-8:00 p.m. Higher advertiser value for entertainment programming during these timeslots explains some of the shortfall. However we also find the number of local news broadcasts to be less than the revenue maximizing allocation for stations, suggesting that classic business plays a role in local television broadcast decisions. The two different mechanisms speak to the role of policy interventions. While subsidies might be used to increase incentives for stations to broadcast local news, alliances and mergers that allow stations to internalize business stealing incentives might prove a less costly alternative.

A Estimation details

In this section we discuss the details of our econometric model and how we estimate it. Our basic method is maximum likelihood, and indeed, most of the likelihood function can be written in closed terms. A few aspects of estimation present some difficulties which require simulation-based methods. Combining the terms in (25), (28), (29), and (30), results in an expression for the log-likelihood contribution of a market to overall likelihood as:

$$\ln L_m = \ln L_{vm} + \ln L_{pm} + \ln L_{rm} + \ln L_{cm} \tag{31}$$

The first two components of L_m on the right-hand side of (31) are the viewership model and the advertising revenue model. The latter two parts effectively enforce some constraints on the model; the term $\ln L_{rm}$ imposes that observed revenues be maximal given the decisions of other stations, while the term $\ln L_{cm}$ imposes cohesiveness on the model. That is, given a set of error terms and parameters, it is possible that there are other equilibria. This term corrects for this possibility.

As described previously, the nested multinomial logit model of viewership share results in share expression that can be written as:

$$\underbrace{\ln s_{jmt} - \ln s_{0mt}}_{y_{jmt}} = \mu_b \ln s_{jmt|bmt} + \underbrace{\alpha_b^v + \zeta_b^v \ln(1+n_b) + X_{jmt}^v \beta_v}_{U_{vjmt}} + \underbrace{\omega_j^v + \omega_{mt}^v + \epsilon_{jmt}}_{e_{jmt}^v} \tag{32}$$

In the underbraces, we have indicated how the above can be broken into a dependent variable, which can be explained by a substitution component, an observable component, and an unobserved component which has an error structure allowing for idiosyncratic effects, in which there are market-time, station, and idiosyncratic components.

This is a useful way of writing the viewership equation because we can now write the viewership model in the form (see equation (6)):

$$y_{jmt} = \mu_b \ln \left[\frac{e^{\frac{U_{vjmt}+e^v_{jmt}}{1-\mu_b}}}{\sum_{b_k \in b} e^{\frac{U_{vkmt}+e^v_{kmt}}{1-\mu_b}}} \right] + U_{vjmt} + e^v_{jmt} \tag{33}$$

While Berry (1994) advocates using instrumental-variables methods to tackle this problem, we take a more direct approach and compute the jacobian determinant of the transformation implied by (32) and adding this on to the likelihood. Write the density of y, $f(y)$, as:

$$f(y) = f(e)|\frac{de}{dy}| = f(y) = f(e)\frac{1}{|\frac{dy}{de}|}$$

This jacobian is fairly straightforward to calculate using expression (33). Expanding and simplifying, we have:

$$y_{jmt} = \frac{1}{1-\mu_b}(U_{vjmt} + e^v_{jmt}) - \ln \left(\sum_{b_{kmt} \in b} e^{\frac{U_{vkmt}+e^v_{kmt}}{1-\mu_b}} \right) \tag{34}$$

Differentiating (34) with respect to e^v_{kmt}, and using the fact that i's share within the group is:

$$s_{imt|bmt} = \frac{e^{\frac{U_{vimt}+e^v_{imt}}{1-\mu_b}}}{\sum_{b_{kmt} \in b} e^{\frac{U_{vkmt}+e^v_{kmt}}{1-\mu_b}}}$$

This results in the following partial derivatives:

$$\begin{aligned} \frac{\partial y_{imt}}{\partial e_{jmt}} &= \tfrac{1}{1-\mu_b} - \tfrac{\mu_b}{1-\mu_b} s_{j|bmt} & & i = j \\ &= -\tfrac{\mu_b}{1-\mu_b} s_{j|bmt} & & i \neq j, i, j \in b_m t \\ &= 0 & & \text{otherwise} \end{aligned} \tag{35}$$

Accordingly, for a given market and timeslot, we have a block matrix in which stations are grouped according to broadcast type:

$$\mathbf{J} = \begin{bmatrix} \mathbf{J_l} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J_e} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{J_n} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{J_c} \end{bmatrix}$$

Each block is a square matrix of dimension N_b, with shape.

$$\mathbf{J_b} = \frac{1}{1-\mu_b} \begin{bmatrix} 1-\mu_b s_{1|bmt} & -\mu_b s_{2|bmt} & \cdots & -\mu_b s_{N_b|bmt} \\ -\mu_b s_{1|bmt} & 1-\mu_b s_{2|bmt} & \cdots & -\mu_b s_{N_b|bmt} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_b s_{1|bmt} & \mu_b s_{2|bmt} & 1-\mu_b s_{N_b|bmt} & \end{bmatrix} \tag{36}$$

Since the matrix is of block form, we can calculate the absolute value of the determinant as:

$$|\mathbf{J}| = |\mathbf{J_l}||\mathbf{J_e}||\mathbf{J_n}||\mathbf{J_c}|$$

Since $\mathbf{J_b}$ in (36) can be written as:

$$\mathbf{J_b} = \frac{1}{1-\mu_b}(\mathbf{I} + \mathbf{u_b}\mathbf{v_b}')$$

Where:

$$u_b = \begin{bmatrix} -\mu_b \\ -\mu_b \\ \vdots \\ -\mu_b \end{bmatrix}, \quad v_b = \begin{bmatrix} s_{1|bmt} \\ s_{2|bmt} \\ \vdots \\ s_{N_b|bmt} \end{bmatrix}$$

It follows the determinant of J_b is:

$$|\mathbf{J_b}| = \left(\frac{1}{1-\mu_b}\right)^{N_b} (1 + \mathbf{u_b}'\mathbf{v_b})$$

Since:

$$1 + \mathbf{u_b}'\mathbf{v_b} = 1 - \mu_b \sum_{i=1}^{N_b} s_{i|bmt} = (1-\mu_b)$$

because the sum of shares within group must add to one, we find that:

$$|\mathbf{J_b}| = \left(\frac{1}{1-\mu_b}\right)^{N_b-1}$$

Therefore, the log of the reciprocal of this determinant must be added to the likelihood. Baltagi shows how the density of errors with a two-way random effect assuming normally distributed error terms can be computed (Baltagi, 2008, p. 42). Putting his expression in context, and adding on the jacobian term, we have the log-likelihood as:

$$\ln L_{vm} = \text{constant} - \frac{1}{2}\ln|\mathbf{\Omega_v}| - \frac{1}{2}e_j'\mathbf{\Omega_v}^{-1}e_j + \sum_{t=1}^{6}\sum_{bmt\in B}(N_{bmt}-1)\ln(1-\mu_b) \quad (37)$$

Where

$$\mathbf{\Omega_v} = \sigma^2_{mtv}(\mathbf{I_{mtv}} \otimes \mathbf{J_{sv}}) + \sigma^2_{sv}(\mathbf{J_{mtv}} \otimes \mathbf{I_{sv}}) + \sigma^2_{smtv}\mathbf{I_{mtv}} \otimes \mathbf{I_{sv}}) \quad (38)$$

and I_{mtv}, I_{sv} are identity matrices of dimension mtv and sv, where mtv are the total number of viewership shares across the market, and sv are the number of observed shares per station across the market (i.e., $sv = 6$). Similarly, J_{mtv}, J_{sv} are matrices of ones of dimensions mtv and sv. Note that the Jacobian transform that must be used is akin to multiplying the the substitution term by the number of observations less one. Also, note that the inclusion of the term puts downward pressure on the likelihood, as, absent other concerns, the likelihood would be maximized by setting $\mu_k = 0$.

The price equation is:

$$\ln r_{jmtb} = \alpha^p_b + \eta_b \ln v_{jmtb} + \omega^p_s + \omega^p_{mt} + \epsilon_{jmt}$$

This can be written as:

$$\underbrace{\omega^p_s + \omega^p_{mt} + \epsilon_{jmt}}_{e^p_{jmt}} = \ln r_{jmtb} - \alpha^p_b - \eta_b \ln v_{jmtb} \quad (39)$$

As before, this two-way random components structure has a likelihood mirroring that described in (40) and (41), which can be written as:

$$\ln L_{pm} = \text{constant} - \frac{1}{2}\ln|\Omega_p| - \frac{1}{2}e^{p\prime}\Omega_p^{-1}e^p \quad (40)$$

Where

$$\Omega_p = \sigma^2_{mtp}(I_{mtp} \otimes J_{sp}) + \sigma^2_{sp} J_{mtp} \otimes I_{sp}) + \sigma^2_{smtp} I_{mtp} \otimes I_{sp}) \tag{41}$$

and I_{mt}, I_s are identity matrices of dimension mtp and sp, where mtp is the number of revenue observations across all time slots in a market, and sp is the number of observations per station (i.e., $sp = 6$). Similarly, J_{mtp}, J_{sp} are matrices of ones of dimensions mtp and sp.

So far, the likelihood has consisted of terms for which there are a closed form. The parts of the likelihood capturing the idea that stations must be choosing a profit-maximizing configuration of broadcasts across the evening, and that all broadcast configurations in a market must be Nash equilibria, are typically impossible to calculate in closed form and therefore must be simulated. The drawing is fairly straightforward, but does involve some nuances that need to be discussed.

First, in the text, we combine expressions (16) and (22) to get an expression for total revenues:

$$R_j = \sum_{t=1}^{6} \left(\alpha^p_b + \eta_b \ln v_{jmtb} + \omega^p_j + \omega^p_{mt} + \omega_{jmt} \right)$$

Using the viewership equation in the text, and the means of expressing it we developed above, we have total viewership as market population M_m multiplied by the viewership share, which can be written as:

$$v_{jmtb} = M \frac{e^{\frac{U_{vjmt} + \omega^v_j + \omega^v_{mt} + \epsilon_{jmt}}{1-\mu_b}}}{\sum_{k \in b} e^{\frac{U_{vkmt} + \omega^v_k + \omega^v_{mt} + \epsilon_{kmt}}{1-\mu_b}}} \frac{\sum_{k \in b} e^{\frac{U_{vkmt} + \omega^v_k + \omega^v_{mt} + \epsilon_{kmt}}{1-\mu_b}}}{1 + \sum_{k \in b} \left(\sum_{k \in b} e^{\frac{U_{vkmt} + \omega^v_k + \omega^v_{mt} + \epsilon_{kmt})}{1-\mu_b}} \right)^{1-\mu_b}} \tag{42}$$

Beginning with draws for market-time and station effects, we can then compute *actual* viewership errors $\hat{e}^v_{jmt}$. To get shares due to one firm deviating, while all others hold decisions constant, we draw a new set of viewership errors $\tilde{e}^v_{jmt}$, and use these to predict counterfactual shares for the whole market, for each broadcast menu that a firm might follow. This gives us a set of hypothetical viewership patters for the deviating firm (in fact, for all firms) That can be used in the revenue equation as v_{jmtb}. Simulating all potential viewerships involves computing viewership patterns for all possible combinations of observed and hypothetical actions and error terms. While this makes things tricky, we can recursively draw the error terms by working backwards.

To illustrate, at the very end of prime time it must be the case that:

$$\begin{aligned} R_j &= \sum_{t=1}^{5} \alpha_b^p + \eta_b \ln v_{jmtb} + \omega_j^p + \omega_{mt}^p + \omega_{jmt} \\ &\quad + \alpha_b^p + \eta_b \ln v_{jm6b} + \omega_j^p + \omega_{m6}^p + \omega_{jm6} \\ &\qquad \geq \\ R_j^{(6)} &= \sum_{t=1}^{5} \alpha_b^p + \eta_b \ln v_{jmtb} + \omega_j^p + \omega_{mt}^p + \omega_{jmt} \\ &\quad + \alpha_b^p + \eta_b \ln v_{jm6b} + \omega_j^p + \omega_{m6}^p + \tilde{\omega}_{jm6} \end{aligned} \tag{43}$$

Which reduces to:

$$\eta_b \ln v_{jm6b} + \omega_{jm6} \geq \eta_b \ln \tilde{v}_{jm6b} + \tilde{\omega}_{jm6} \tag{44}$$

Or

$$T_6 = \eta_b \left(\ln v_{jm6b} - \ln \tilde{v}_{jm6b} \right) + \omega_{jm6} \geq \tilde{\omega}_{jm6} \tag{45}$$

Since by assumption $\omega_{jmt} \sim N(0, \sigma^2_{vjmt},\ \tilde{\omega}_{jm6})$ can be drawn from a right-truncated normal distribution $N(0, \sigma^2_{vjmt}; T_6)$. This gives a counterfactual error term for period 6. Backing up one period, we now have the requirement that the observed schedule generate the greatest possible revenue of *all potential schedules*, that might be followed from t forward. So:

$$\begin{aligned} R_j &= \sum_{t=1}^{6} \exp(\alpha_b^p + \eta_b \ln v_{jmtb} + \omega_j^p + \omega_{mt}^p + \omega_{jmt}) \\ &\qquad \geq \\ R_j^{(56)} &= \sum_{t=1}^{4} \exp(\alpha_b^p + \eta_b \ln v_{jmtb} + \omega_j^p + \omega_{mt}^p + \omega_{jmt}) \\ &\quad + \max \Big[\exp(\alpha_b^p + \eta_b \ln v_{jm5b} + \omega_j^p + \omega_{m5}^p + \tilde{\omega}_{jm5}) \\ &\quad + \exp(\alpha_b^p + \eta_b \ln v_{jm6b} + \omega_j^p + \omega_{m6}^p + \tilde{\omega}_{jm6}), \\ &\quad \exp(\alpha_b^p + \eta_b \ln v_{jm5b} + \omega_j^p + \omega_{m5}^p + \tilde{\omega}_{jm5}) \\ &\quad + exp(\alpha_b^p + \eta_b \ln v_{jm6b} + \omega_j^p + \omega_{m6}^p + \tilde{\omega}_{jm6}) \Big] \end{aligned} \tag{46}$$

Which becomes:

$$
\begin{aligned}
R_j &= \sum_{t=1}^{6} \exp(\alpha_b^p + \eta_b \ln v_{jmtb} + \omega_j^p + \omega_{mt}^p + \omega_{jmt}) \\
&\qquad \geq \\
R_j^{(56)} &= \sum_{t=1}^{4} \exp(\alpha_b^p + \eta_b \ln v_{jmtb} + \omega_j^p + \omega_{mt}^p + \omega_{jmt}) \\
&\quad + \max \Big[\exp(\alpha_b^p + \eta_b \ln v_{jm5b} + \omega_j^p + \omega_{m5}^p + \tilde{\omega}_{jm5}) \\
&\quad + \exp(\alpha_b^p + \eta_b \ln v_{jm6b} + \omega_j^p + \omega_{m6}^p + \tilde{\omega}_{jm6}), \\
&\quad \exp(\alpha_b^p + \eta_b \ln v_{jm5b} + \omega_j^p + \omega_{m5}^p + \tilde{\omega}_{jm5}) \\
&\quad + exp(\alpha_b^p + \eta_b \ln v_{jm6b} + \omega_j^p + \omega_{m6}^p + \tilde{\omega}_{jm6}) \Big]
\end{aligned}
\tag{47}
$$

That is, we require that the error term be such that changing so that a different item is broadcast in the 5th timeslot is less than what is observed, when the continuation strategy could be to either follow the observed broadcast in period 6, or change broadcasts in period 6 along with a change in broadcast in period 5. Out of this logic, we find an upper bound for the error term of $\tilde{\omega}_{jm5}$, T_5. So, this can be drawn accordingly.

We proceed in this fashion until we arrive at the first broadcast period, at which point we have a bound T_1. we can then calculate the density of everything as

$$
f = \frac{\phi\left(\frac{\tilde{\omega}'_{vjm1}}{\sigma^p_{jmt}}\right)}{\Phi\left(\frac{T_1}{\sigma^p_{jmt}}\right)}
$$

We now have a set of counterfactual and actual error terms for each station that is choosing a broadcast menu. The last remaining thing to do is check for other potential equilibria. This is somewhat challenging because each station has a rather large strategy space, and each market is characterized by a reasonably large number of stations. So, the strategy space for the entire broadcasting game can be of a large dimension. With 4 stations, each with 64 possible menus and two different types of broadcasts, this gives $4 * 2^6 4 = 7.3786976e + 19$ potential strategy profiles across the viewership market. It is clearly impracticable to check them all.

We therefore follow an approximate procedure, which resembles an adaptive Markov-chain Monte Carlo technique. For each station, we first pare down the strategy space to include only those that are likely to be viable alternatives. It typically happens in simulating prices and shares, that some strategies would generate significantly lower revenue than what is observed (for example, broadcasting local news in all six possible slots, or broadcasting local news at 7:30). We therefore pare down the set of potential choices to the ten best possible strategies. In the above example, this reduces the available strategy space to $10^4 = 10000$. While this is an improvement, it is still hard to check all possibilities - especially in markets in which there are as many as 10 stations. Accordingly, what we do is select a number of stations randomly, perturb their strategies using choices from the likely set of strategies, and then iterate from this point to a Nash equilibrium. If we find additional equilibria using this procedure, we make note of this. In the final assessment, we adjust the likelihood of the entire vector of error draws by $\frac{1}{N_{eq}}$, where N_{eq} is the (approximate) number of observed equilibria. For small dimensional examples, (i.e., markets with two stations), this approach produces accurate results.

As an additional note on estimation methods, because our parametric model is in some sense overidentified (i.e., we have some redundancy in controlling for heterogeneity in the model), we do not want to resort to standard simulated maximum likelihood. Instead, we rely on the methods in Chernozhukov and Hong (2003) and use an Markov-chain Monte Carlo approach to classical estimation. The practical means by which this is done are detailed in Baker (2014).

Here, we detail how we can apply an Ackerberg-Keane-Wolpin type simulation-based estimation strategy to our problem. The problem emerges because it is computationally difficult and time consuming to simulate Nash equilibria of the broadcast menu choice game, the error terms operate over a difficult-to-define region of integration, and because it is also costly to compute counterfactual shares.

We basically follow Ackerberg (2010) in our estimation strategy. The basic idea is as follows. Suppose that one wishes to calculate a likelihood function that includes data x and parameters θ, which requires some integration of an unobserved variable ϵ over a region Γ. ϵ could be, and often is, multidimensional. Thus:

$$L = \int_{\epsilon \in \Gamma} F(x, \theta, \epsilon) f(\epsilon) d\epsilon$$

The difficulty arises because $F(x, \theta, \epsilon)$ is costly and/or time-consuming to calculate. Ackerberg (2010) shows how the problem may be, in certain circumstances, recast as follows. First, if one can form an index $u(x, \theta, \epsilon)$, one can write:

$$L = \int_{\epsilon \in \Gamma} F(u(x, \theta, \epsilon)) f(\epsilon) d\epsilon$$

u is often a linear index function of the form $u = x\beta + \epsilon$, where $\beta \in \theta$. Then, one can introduce a change of variables:

$$L = \int_{u \in \Gamma'} F(u) h(u, \theta, x) du$$

Suppose that one can find a way to simulate values of u that does not depend upon θ. Then one can introduce an importance sampler as follows:

$$L = \int_{u \in \Gamma'} F(u) h(u, \theta, x) \frac{g(u, x)}{g(u, x)} du$$

Values of u can be drawn from $g(u)$, which in practice often derives from some approximate, yet reliable and easy-to-estimate model, so $g(u, x)$ is in fact $g(u, x, \theta_0)$. Then, an approximation of L is obtained using S simulated values of u:

$$L \approx \frac{1}{S} \sum_{s=1}^{S} \frac{F(u_s)}{g(u_s, x)} h(u_s, \theta, x)$$

The convenience of this last expression is clear; the problem now relies on calculating $h(u_s, \theta, x)$ instead of repeated calculation of $F(x, \theta, \epsilon)$. Essentially, simulated observations are re-weighted by the estimation procedure.

In our problem, we seek to integrate over the unobserved features of the problem, which include unobserved viewership error terms and unobserved advertising price error terms. Counterfactual viewership and prices from unobserved choices are time consuming to calculate, and checking the cohesiveness of the results (i.e., checking for multiple Nash equilibria) is also time-consuming. Accordingly, we rely on Ackerberg's (2010) procedure as follows. The idea is to first simulate viewership for a set of draws, and then simulate prices so that observed prices constitute the revenue maximizing price. Given these conditions, we check for existence of alternative Nash equilibria.

The details are as follows. First, note that viewership can be written as:

$$v_{jmtb} = M \frac{e^{\frac{U_{vjmt}+\omega_j^v+\omega_{mt}^v+\epsilon_{jmt}}{1-\mu_b}}}{\sum_{k\in b} e^{\frac{U_{vkmt}+\omega_k^v+\omega_{mt}^v+\epsilon_{kmt}}{1-\mu_b}}} \frac{\sum_{k\in b} e^{\frac{U_{vkmt}+\omega_k^v+\omega_{mt}^v+\epsilon_{kmt}}{1-\mu_b}}}{1+\sum_{k\in b}\left(\sum_{k\in b} e^{\frac{U_{vkmt}+\omega_k^v+\omega_{mt}^v+\epsilon_{kmt})}{1-\mu_b}}\right)^{1-\mu_b}} \quad (48)$$

We break up the viewership utility index as follows:

$$u_{vjmt} = U_{vjmt} + \omega_j^v + \omega_{mt}^v + \epsilon_{jmt}$$

After replacng U_{vjmt} with its definition, we have:

$$u_{vjmt} = \alpha_b^v + X_{jmt}^v \beta_v + \zeta_b^v \ln(1+n_b) + \omega_j^v + \omega_{mt}^v + \epsilon_{jmt}$$

In this term, we think about creating simulated viewership indexes as a three step procedure in which first market-time and station level random effects are drawn, which then imply a value of the error term ϵ_{jmt}. Conditional on the market-time and station level fixed effects, we then draw error terms ϵ'_{jmt} for unobserved programming choices and calculate hypotheticals. That is, we draw viewership utility indexes as:

$$u'_{vjmt} = \underbrace{\alpha_b^v + X_{jmt}^v \beta_v + \zeta_b^v \ln(1+n_b) + \underbrace{\omega_j^v}_{u_{v1}} + \underbrace{\omega_{mt}^v}_{u_{v2}} + \epsilon_{jmt}}_{u_{v3}}$$

Based on these draws, we then create a hypothetical share:

$$v'_{jmtb} = M \frac{e^{\frac{u'_{vjmt}}{1-\mu_{b'}}}}{\sum_{k\in b'} e^{\frac{u_{kmtb'}}{1-\mu_{b'}}}} \frac{\sum_{k\in b} e^{\frac{u_{kmtb'}}{1-\mu_b}}}{1+\sum_{k\in b}\left(\sum_{k\in b} e^{\frac{u_{kmtb}}{1-\mu_b}}\right)^{1-\mu_b}} \quad (49)$$

We then have a density for the draw as:

$$\begin{aligned} h(u'_{vjmtbs}, \theta, X) \quad &= \quad \phi(\alpha_b^v + X_{jmt}^v \beta_v + \zeta_b^v \ln(1+n_b) + \omega_{js}^v + \omega_{mts}^v, \sigma_{mt}) \times \\ & \quad\quad \phi(0, \sigma_j)\phi(0, \sigma_{mt}) \end{aligned}$$

The importance sampling weight is calculated using preliminary estimates of the parameters as:

$$g(u) = \phi(u_{v1}, \sigma_{v0})\phi(u_{v2}, \sigma_{mt0})\phi(u_{v3}, \sigma_{smt0})\phi(u'_{v3}, \sigma_{smt0})$$

Note that there are actually two individual-level densities to be drawn in this case. Once counterfactual viewership shares are created, we can now draw counterfactual advertising prices. First, inserting our hypothetical viewership draw into the pricing equation gives:

$$\ln r_{jmtb} = \underbrace{\alpha_b^p + \eta_b \ln v'_{jmtb} + \underbrace{\omega_j^v}_{u_{p1}} + \underbrace{\omega_{mt}^v}_{u_{p2}} + \varepsilon_{jmt}^p}_{u_{p3}} \tag{50}$$

The nuance here is that we have a set of bounds, as detailed previously, that the advertising price revenues have to satisfy. These give bounds for $T_{js1}, T_{js2}, \ldots, T_{js6}$. Which results in our values of u_p belonging to a truncated normal density. Hence, we have:

$$\begin{aligned} h(u'_{pjmtb}, \theta, X) &= \frac{\phi(\alpha_b^p + \eta_b \ln v'_{jmtb} + u_{p1} + u_{p2}, \sigma_{mt})}{\Phi(T_{jmtb})} \times \\ & \phi(u_{1p}, \sigma_j)\phi(u_{2p}, \sigma_{mt}) \end{aligned}$$

with important sampling weight:

$$g(u) = \phi(u_{p1}, \sigma_{p0})\phi(u_{p2}, \sigma_{pmt0})\phi(u_{p3}, \sigma_{psmt})\phi(u'_{v3}, \sigma_{psmt0})$$

As a final step, we have to calculate the number of Nash equilibria for each alternative draw, and then add a weight $\frac{1}{N_{eq}}$ to the likelihood. We do this as detailed in the text. We first perturb each strategy, and then begin a process of iterated dominance. Often, the process leads back to the original equilibrium, but in the event that it leads to some other equilibrium, we decrease the sampling weight proportionally. As a practical matter, when another equilibrium is found, it typically involves stations switching broadcast types in one or two settings.

In terms of practical details, we employ $S = 40$ for each station and market. For each draw, we check 5000 randomly drawn perturbations to the strategies to find alternative Nash equilibria. Once this is complete, we perform $T = 1000$ iterations of the MCMC estimation procedure over the full likelihood, for which we then discard the first 500 draws. After discarding the first 500 draws, we retain every fifth draw only, and then use the remaining 100 draws to compute standard errors, confidence intervals, etc. See Baker (2015) and Chernozhukov and Hong (2003) for the logic behind these choices.

Table 1: Station and Broadcast Counts per Market, by Market Category

	N	Mean	SD	5th Pct.	95th Pct.
Small Markets (Under 1 Million HH)					
- Local Stations	40	7.4	1.4	5.5	10.5
- Natl. Cable Stations	40	64.7	9.8	53.0	84.0
- Local News Broadcasts	40	7.2	1.9	5.0	10.0
Medium Markets (1-3 Million HH)					
- Local Stations	43	9.9	2.1	7.0	14.0
- Natl. Cable Stations	43	93.0	14.4	61.0	110.0
- Local News Broadcasts	43	11.2	3.0	7.0	17.0
Large Markets (Over 3 Million HH)					
- Local Stations	18	15.6	3.7	10.0	22.0
- Natl. Cable Stations	18	94.2	7.5	81.0	107.0
- Local News Broadcasts	18	14.6	3.9	9.0	22.0
All Markets					
- Local Stations	101	9.9	3.6	6.0	18.0
- Natl. Cable Stations	101	82.0	18.2	56.0	107.0
- Local News Broadcasts	101	10.2	3.9	6.0	17.0

Table 2: Total News and Television Viewing per Market, by Market Category

	N	Mean	SD	5th Pct.	95th Pct.
Small Markets (Under 1 Million HH)					
- Local News Broadcasts	40	7.2	1.9	5.0	10.0
- All TV Viewing	40	0.42	0.08	0.32	0.57
- Local News Viewing	40	0.07	0.02	0.05	0.11
Medium Markets (1-3 Million HH)					
- Local News Broadcasts	43	11.2	3.0	7.0	17.0
- All TV Viewing	43	0.27	0.08	0.15	0.42
- Local News Viewing	43	0.04	0.01	0.02	0.07
Large Markets (Over 3 Million HH)					
- Local News Broadcasts	18	14.6	3.9	9.0	22.0
- All TV Viewing	18	0.08	0.03	0.02	0.13
- Local News Viewing	18	0.01	0.00	0.00	0.02

Table 3: Average broadcast count and aggregate viewership shares by times and program.

	Local News	National News	Entertainment	Cable	Total
Broadcasts					
5:00	3.37	1.32	6.30	80.93	91.91
5:30	1.85	3.22	5.91	80.93	91.91
6:00	3.52	2.04	5.28	81.07	91.91
6:30	0.75	3.80	6.29	81.07	91.91
7:00	0.53	1.19	8.70	81.49	91.91
7:30	0.19	0.96	9.28	81.49	91.91
Total 1	10.22	12.52	41.75	486.97	551.47
Viewing					
5:00	0.0168	0.0004	0.0048	0.0201	0.0420
5:30	0.0082	0.0099	0.0051	0.0213	0.0445
6:00	0.0211	0.0008	0.0052	0.0222	0.0492
6:30	0.0016	0.0113	0.0137	0.0237	0.0504
7:00	0.0012	0.0007	0.0259	0.0272	0.0550
7:30	0.0001	0.0003	0.0274	0.0285	0.0562
Total 1	0.0489	0.0233	0.0821	0.1430	0.2973

Table 4: Average advertisement prices per second by times and program type.

	Local News	National News	Entertainment	Average
Broadcasts				
5:00	\$9.10	\$4.58	\$3.91	\$6.54
5:30	\$9.11	\$10.68	\$3.96	\$6.91
6:00	\$11.47	\$15.44	\$6.04	\$8.82
6:30	\$11.41	\$13.00	\$7.68	\$9.60
7:00	\$8.65	\$16.53	\$17.76	\$16.79
7:30	\$14.09		\$18.94	\$18.40
Total Average	\$9.61	\$11.49	\$9.89	\$11.18

Table 5: Average advertisement prices per 1,000 viewers, by times and program type.

	Local News	National News	Entertainment	Average
Broadcasts				
5:00	\$0.76	\$0.39	\$1.28	\$0.99
5:30	\$0.89	\$1.02	\$1.15	\$0.95
6:00	\$0.83	\$1.13	\$1.51	\$1.13
6:30	\$1.46	\$1.42	\$1.19	\$1.17
7:00	\$1.60	\$0.86	\$1.65	\$1.59
7:30	\$5.58		\$1.66	\$1.73
Total Average	\$0.91	\$1.10	\$1.38	\$1.26

Figure 1: Number of broadcasts of each type by timeslot (other cable programming omitted).

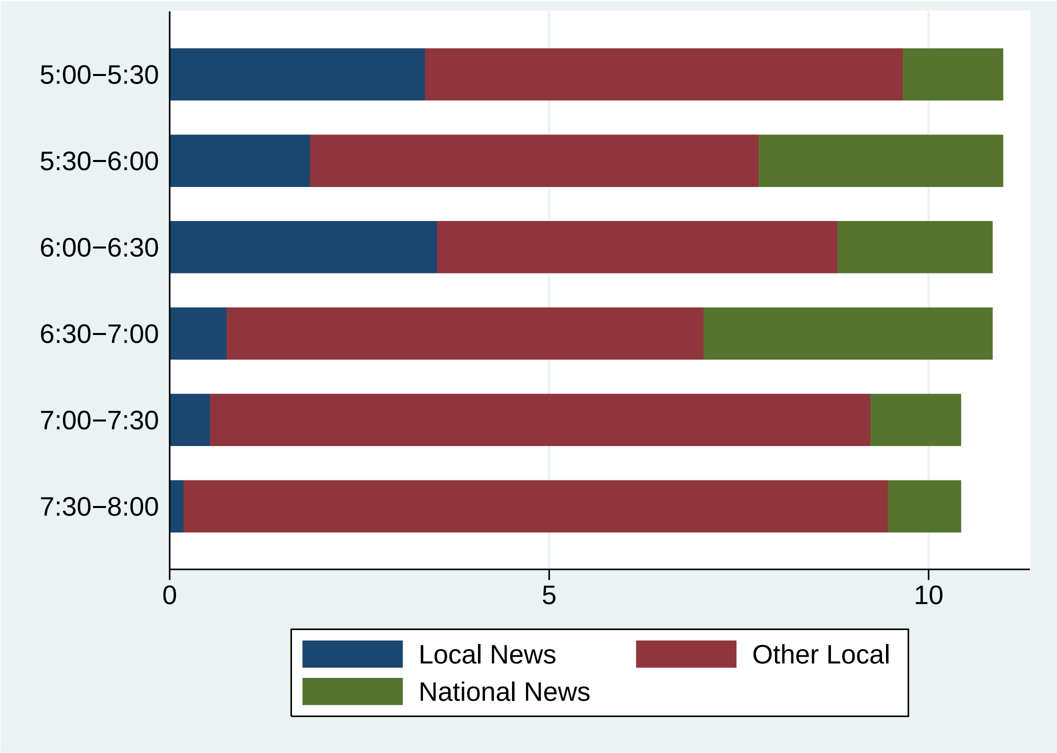


Figure 2: Average viewership shares by programming type and time slot.

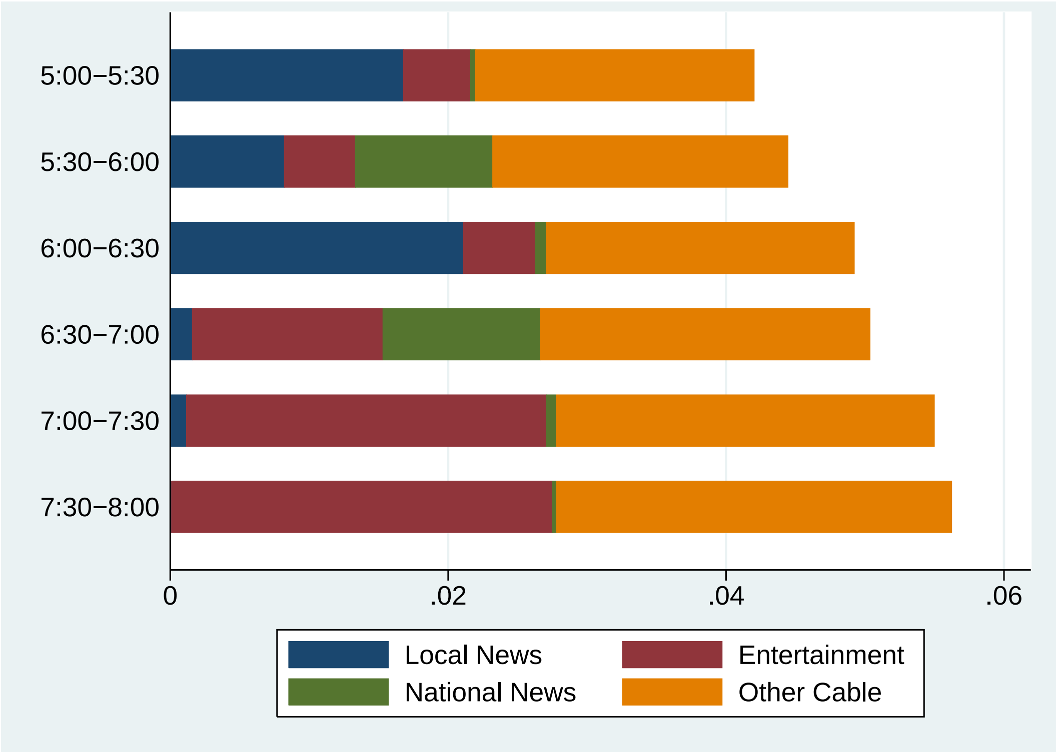


Table 6: Most common programming sequences

Programming lineups	No.	Col %	Cum %
oooooo	213	34.2	34.2
lnlooo	106	17.0	51.2
lllnoo	96	15.4	66.6
oolnoo	35	5.6	72.2
lnlloo	21	3.4	75.6
lnoooo	21	3.4	79.0
lllnlo	17	2.7	81.7
lloooo	16	2.6	84.3
llllooo	15	2.4	86.7
oooloo	10	1.6	88.3
other	73	11.7	100.0
Total	623	100.0	

Table 7: Welfare Loss from Station Choice - Examples

Two-Sided Market Distortion			Business Stealing Distortion		
Viewing (HH Share)			*Viewing (HH Share)*		
Station1 / Station2	Local News	Entertainment	Station1 / Station2	Local News	Entertainment
Local News	0.024, 0.024	0.033, 0.045	Local News	0.019, 0.019	0.028, 0.028
Entertainment	0.045, 0.033	0.035, 0.035	Entertainment	0.028, 0.028	0.024, 0.024
Advertiser Surplus ($ per Second)			*Advertiser Surplus ($ per Second)*		
	Local News	Entertainment		Local News	Entertainment
Local News	16, 16	19, 25	Local News	15.3, 15.3	22.8, 28.4
Entertainment	25, 19	23, 23	Entertainment	28.4, 22.8	24.2, 24.2
Station Revenue (Ad Price $ Per Second)			*Station Revenue (Ad Price $ Per Second)*		
	Local News	Entertainment		Local News	Entertainment
Local News	8, 8	9.5, 12.5	Local News	7.7, 7.7	11.4, 14.2
Entertainment	12.5, 9.5	11.5, 11.5	Entertainment	14.2, 11.4	12.1, 12.1

	Parameter	Full Model b	se	Eq-by-Eq ML b	se
Substitution					
Local news	μ_l	0.033+	0.016	0.157*	0.021
Entertainment	μ_e	0.084*	0.025	0.141*	0.010
National news	μ_n	0.098*	0.007	0.090*	0.013
Other cable	μ_c	0.000	0.001	0.000	0.000
Program_Type					
Local news dummy	α_l^v	4.881*	0.041	5.075*	0.033
Entertainment dummy	α_e^v	3.895*	0.102	3.173*	0.063
National news dummy	α_n^v	2.963*	0.071	3.030*	0.060
Dynamics					
loc. news after loc. news	β_{ll}	0.105	0.090	-0.212*	0.077
loc. news after nat. news	β_{ln}	0.369	0.198	-0.015	0.100
nat. news after loc. news	β_{nl}	1.076*	0.139	0.753*	0.085
nat. news after nat. news	β_{nn}	1.362*	0.087	1.519*	0.048
share prev. timeslot	β_s	0.235*	0.008	0.228*	0.006
loc. news after loc. news X share	β_{sll}	0.068*	0.016	0.022*	0.008
loc. news after nat. news X share	β_{sln}	0.144*	0.037	0.076*	0.014
nat. news after loc. news X share	β_{snl}	0.053	0.029	-0.006	0.011
nat. news after loc. news X share	β_{snn}	0.144*	0.009	0.162*	0.008
Breaking_News					
Cum. share local news	β_{cuml}	-11.744*	0.441	-11.802*	0.870
Cum. share nat. news	β_{cumn}	0.449	1.313	-0.456	2.084
Controls					
Market Size	β_M^v	-0.678*	0.004	-0.690*	0.003
ln(1+no. loc. news)	ζ_l	0.084	0.091	0.211*	0.043
ln(1+no. enter.)	ζ_e	0.209	0.112	0.656*	0.024
ln(1+no. nat. news)	ζ_n	0.428*	0.026	0.481*	0.054
ln(1+other cable)	ζ_c	0.720*	0.015	0.764*	0.005
constant	α^v	-0.290*	0.071	-0.554*	0.029
Rand_FX					
log(sd) station RE	$\ln \sigma_s^v$	0.790*	0.009	0.795*	0.007
log(sd) market RE	$\ln \sigma_{mt}^v$	0.942*	0.020	0.898*	0.030
log(sd) model	$\ln \sigma_{smt}^v$	0.332*	0.001	0.328*	0.001
Revenue					
Local news viewer elast.	η_l	0.596*	0.003	0.612*	0.002
Entertainment viewer elast.	η_e	0.702*	0.009	0.708*	0.001
Nat. news viewer elast.	η_n	0.297*	0.005	0.270*	0.003
Local news dummy	α_l^p	-2.708*	0.023	-3.113*	0.025
Entertainment dummy	α_e^p	-3.406*	0.080	-3.693*	0.015
Market Size	β_M^p	0.874*	0.002	0.875*	0.001
constant		-12.337*	0.035	-12.460*	0.032
Rand_FX					
log(sd) station RE	σ_s^p	0.404*	0.012	0.371*	0.014
log(sd) market RE	σ_{mt}^p	0.324*	0.009	0.311*	0.013
log(sd) model	σ_{smt}^p	0.250*	0.005	0.294*	0.005
N					
Total obs.		55698		55698	
Markets		101		101	
Stations		623		623	
Local news		1032		1032	
Other local		4217		4217	
Nat'l news		1265		1265	

Table 8: Estimation results: starting values and results from full model

Table 9: Observed vs. Optimal Total Broadcasts and Viewing

	Observed	Viewer Max	Advertiser Max	Station Max
Local News Broadcasts	10.2	20.7	14.7	14.9
National News Broadcasts	12.5	12.5	12.5	12.5
Entertainment Broadcasts	41.8	30.8	37.5	37.4
Cable Broadcasts	487.0	487.0	487.0	487.0
Evening Total Viewing	0.297	0.325	0.313	0.316
Local News Viewing	0.049	0.084	0.054	0.054
National News Viewing	0.023	0.024	0.021	0.023
Entertainment Viewing	0.082	0.075	0.099	0.100
Cable Viewing	0.143	0.142	0.142	0.142
Advertiser Surplus ($/Second)	$7.46	$19.97	$25.67	$25.67
Station Revenue ($/Second)	$11.40	$10.15	$12.77	$12.80

Table 10: Station Competition Example: Lexington, KY 7:00-7:30 p.m.

Lexington, KY							
		Broadcast		Ad Price		Viewing	
Station	Network	Observed	Station Optimum	Observed	Station Optimum	Observed	Station Optimum
WTVQ	ABC	Entertainment	Entertainment	$3.04	$3.07	0.0020	0.0021
WKYT	CBS	Entertainment	Entertainment	$12.13	$13.14	0.0119	0.0119
WLEX	NBC	Entertainment	Local News	$5.00	$4.91	0.0068	0.0071
EKYT	CW	Entertainment	Entertainment	$1.50	$1.53	0.0007	0.0007
WDKY	FOX	Entertainment	Entertainment	$3.91	$4.04	0.0070	0.0073
Market Total				$25.59	$26.69	0.0283	0.0291

Table 11: Observed vs. Optimal Broadcasts and Viewing by Timeslot

	Observed	Viewer Max	Advertiser Max	Station Max
Local News Broadcasts				
5:00	3.4	5.0	4.3	5.0
5:30	1.9	2.5	2.1	1.9
6:00	3.5	4.2	3.6	3.6
6:30	0.8	2.5	2.1	2.0
7:00	0.5	2.2	2.3	2.1
7:30	0.2	4.4	0.3	0.3
Local News Viewing				
5:00	0.017	0.017	0.011	0.017
5:30	0.008	0.008	0.005	0.005
6:00	0.021	0.023	0.018	0.018
6:30	0.002	0.008	0.007	0.006
7:00	0.001	0.010	0.010	0.007
7:30	0.000	0.019	0.001	0.001
National News Viewing				
5:00	0.000	0.000	0.000	0.000
5:30	0.010	0.010	0.007	0.009
6:00	0.001	0.001	0.001	0.001
6:30	0.011	0.012	0.012	0.012
7:00	0.001	0.001	0.001	0.001
7:30	0.000	0.000	0.000	0.000
Entertainment Viewing				
5:00	0.005	0.004	0.008	0.005
5:30	0.005	0.010	0.011	0.011
6:00	0.005	0.009	0.012	0.013
6:30	0.014	0.012	0.014	0.014
7:00	0.026	0.025	0.027	0.028
7:30	0.027	0.015	0.029	0.029
Cable Viewing				
5:00	0.020	0.020	0.020	0.020
5:30	0.021	0.021	0.021	0.021
6:00	0.022	0.022	0.022	0.022
6:30	0.024	0.024	0.024	0.024
7:00	0.027	0.027	0.027	0.027
7:30	0.028	0.028	0.028	0.028

Figure 3: Combined sequence plots

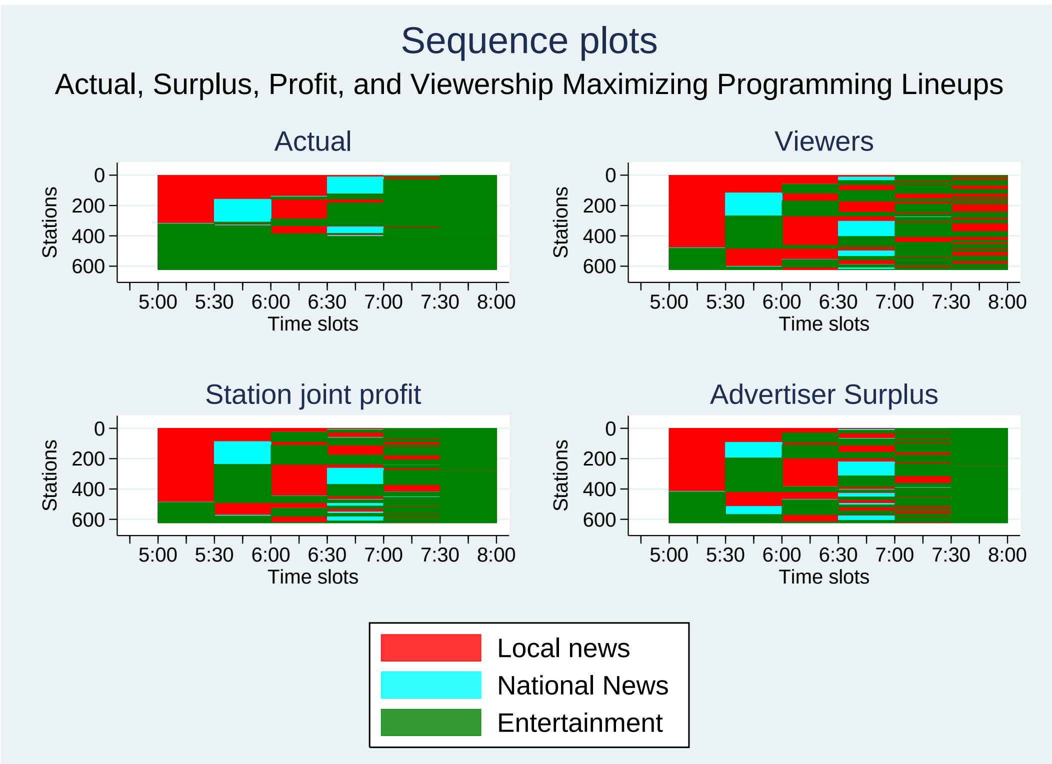


Table 12: Observed vs. Optimal Advertising Prices per Second by Timeslot

Timeslot	Observed Price	Viewer Max Price	%Difference	Observed Viewing	Viewer Max	%Difference
5:00	$6.50	$6.06	-1.30%	0.0420	0.0428	0.70%
5:30	$6.81	$6.03	-2.12%	0.0445	0.0477	1.90%
6:00	$9.00	$8.25	-1.63%	0.0492	0.0547	2.78%
6:30	$9.80	$9.17	-0.96%	0.0504	0.0553	2.70%
7:00	$17.52	$15.78	-1.67%	0.0550	0.0611	2.87%
7:30	$18.80	$15.62	-3.83%	0.0562	0.0634	3.68%

Table 13: Welfare Loss and Market Characteristics: Equilibrium Differences (%)

	Viewer Max-Observed	Viewer Max-Advertiser Max	Viewer Max-Advertiser Max	Station Max-Observed
	(1)	(2)	(3)	(4)
	(Viewing %)	(Viewing %)	(Price/Second %)	(Price/Second %)
Market HH Pop (100,000)	.022**	.007*	–.038**	.009*
	(.01)	(.00)	(.01)	(.00)
Market Black HH Share	–4.970*	–.056	1.820	–1.612
	(2.00)	(1.14)	(3.89)	(1.73)
Market Hispanic HH Share	3.148+	.071	–3.882	1.974
	(1.86)	(1.06)	(3.62)	(1.61)
Market Median Income	.023+	–.013+	–.075**	.037**
	(.01)	(.01)	(.02)	(.01)
Market Median Age	–.083*	–.055*	.184*	.060+
	(.04)	(.02)	(.08)	(.03)
Constant	4.248*	3.548**	–6.991*	–1.524
	(1.73)	(.99)	(3.37)	(1.50)
Mean Y	2.59	1.01	–4.68	2.69
N	101	101	101	101
Adj. R2	.41	.10	.42	.26

Figure 4: A depiction of the broadcasting game

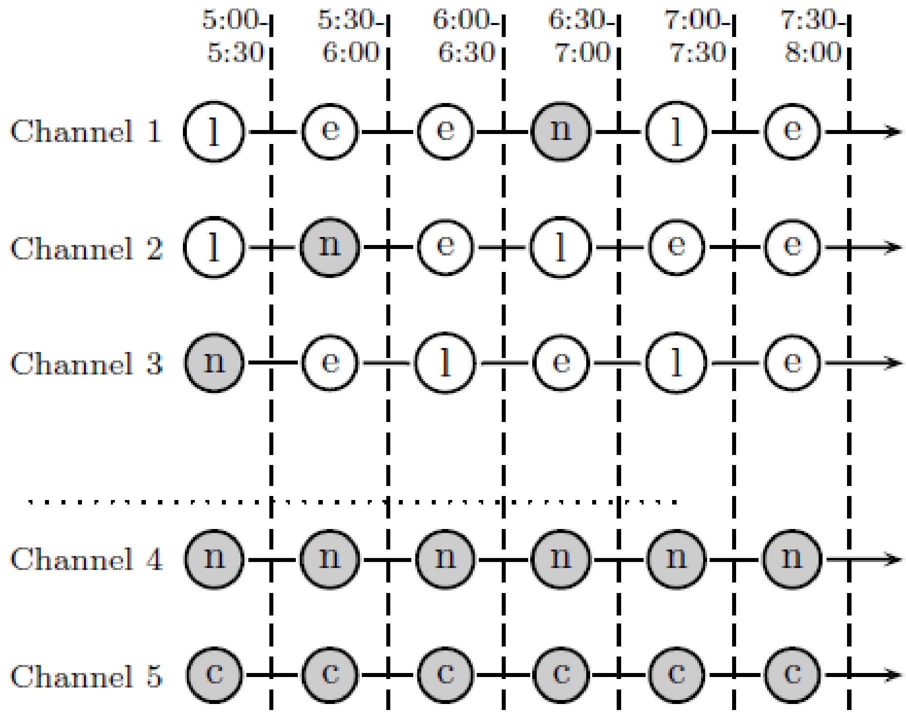


Figure 5: Actual shares and simulated shares at the viewer optimal programming configuration.

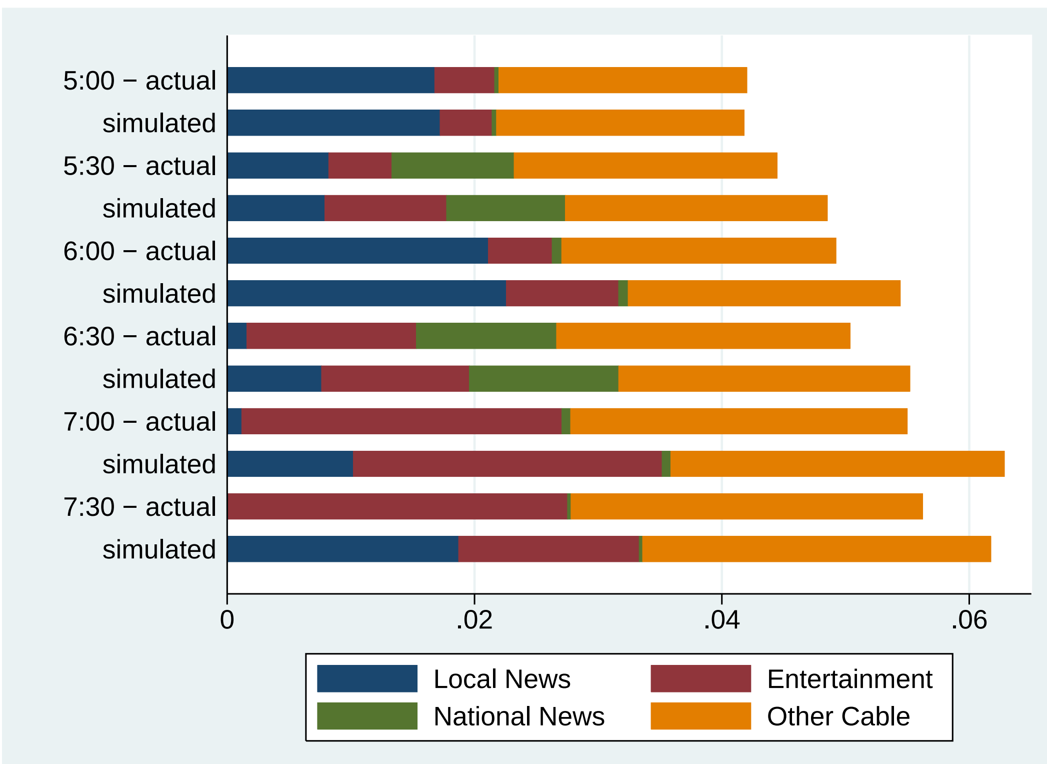


Figure 6: Actual shares and simulated shares at the station revenue maximizing configuration.

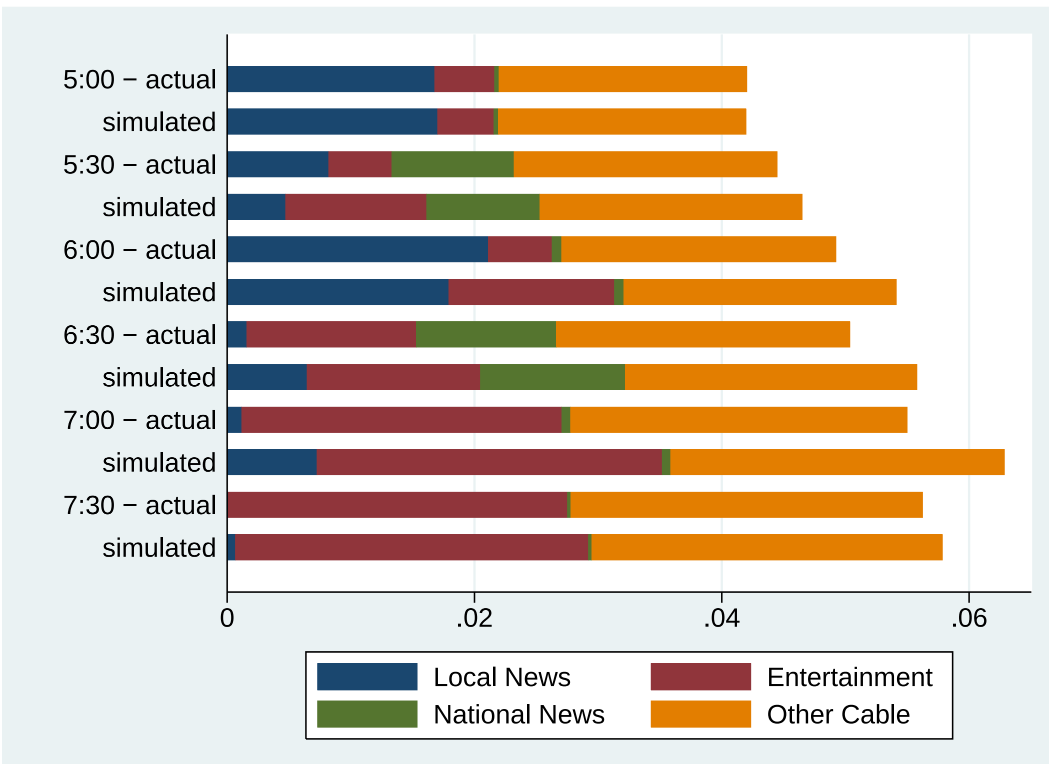


Figure 7: Actual shares and simulated shares at advertiser's surplus maximizing configuration.

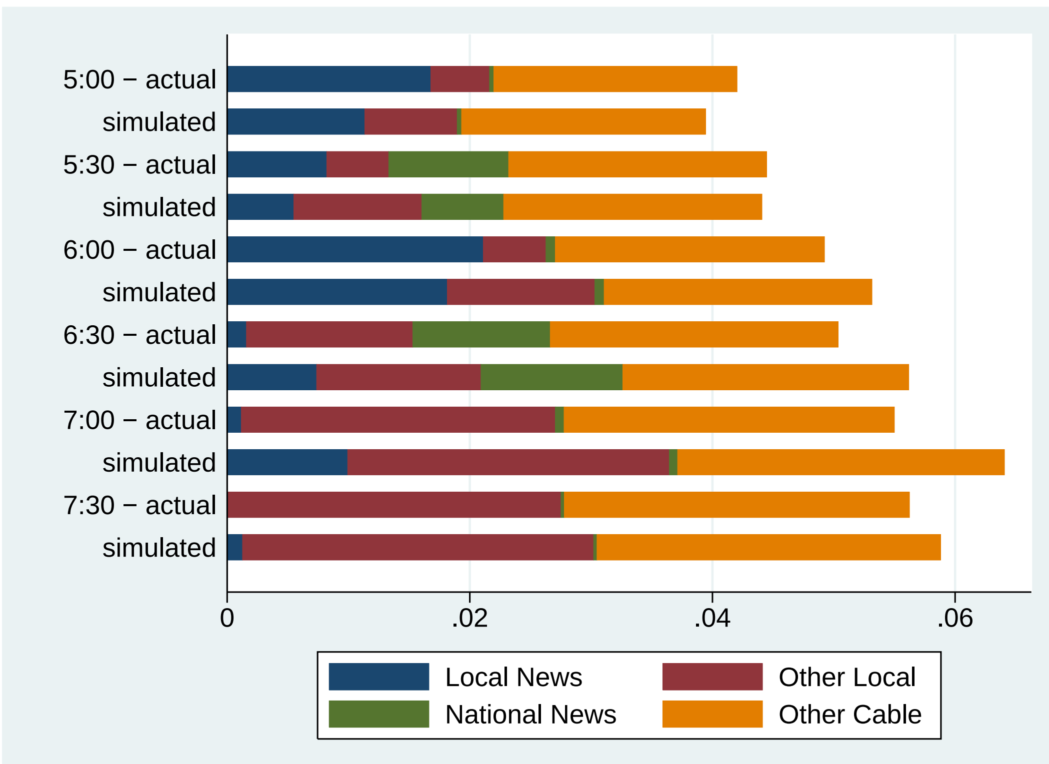
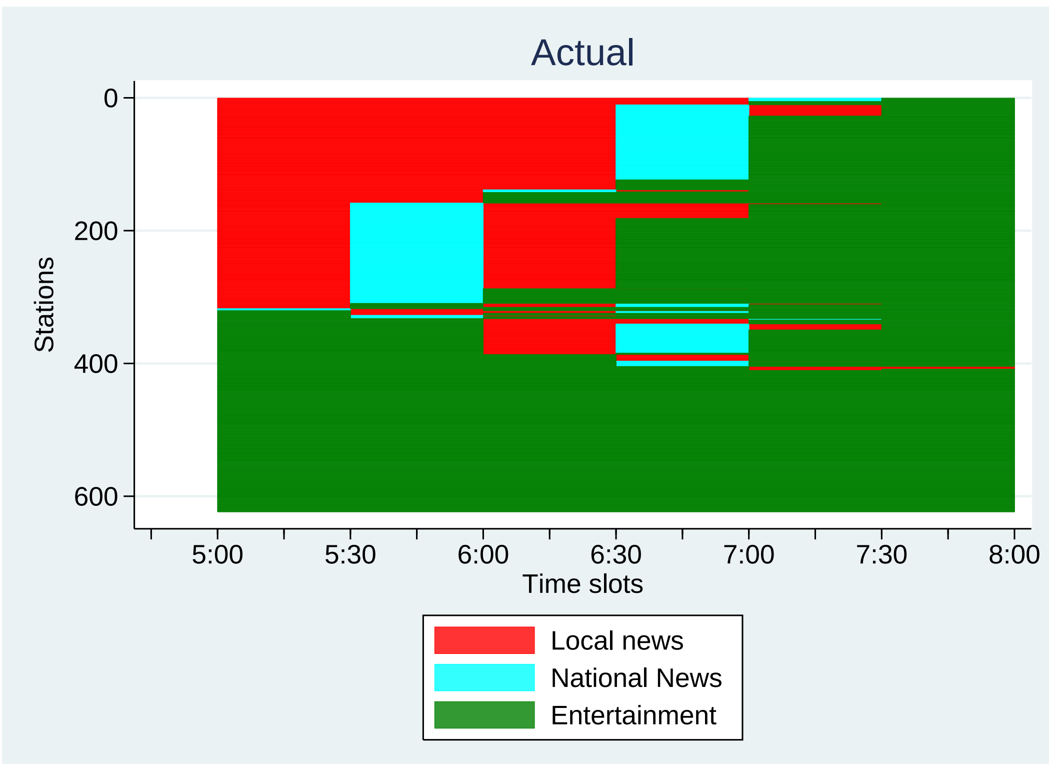


Figure 8: Sequence plot of actual programming offered



References

Ackerberg, D., Benkard, C. L., Berry, S., and Pakes, A. (2007). "Econometric tools for analyzing market outcomes." *Handbook of econometrics*, *6*, 4171–4276.

Anderson, S. P., and Coate, S. (2005). "Market provision of broadcasting: A welfare analysis." *Review of Economic studies*, 947–972.

Anderson, S. P., and Gabszewicz, J. J. (2006). "The media and advertising: A tale of two-sided markets." In V. Ginsburgh, and D. Throsby (Eds.), *Handbook of the Economics of Art and Culture*, vol. 1, chap. 18, 567–614, Amsterdam: Elsevier.

Anderson, S. P., Goeree, J. K., and Ramer, R. (1997). "Location, location, location." *Journal of Economic Theory*, *77*(1), 102 – 127.

Armstrong, M. (2006). "Competition in two-sided markets." *The RAND Journal of Economics*, *37*(3), 668–691.

Bajari, P., Hong, H., Krainer, J., and Nekipelov, D. (2010a). "Estimating static models of strategic interactions." *Journal of Business & Economic Statistics*, *28*(4).

Bajari, P., Hong, H., and Ryan, S. P. (2010b). "Identification and estimation of a discrete game of complete information." *Econometrica*, *78*(5), 1529–1568.

Baker, M. J. (2014). "Adaptive markov chain monte carlo sampling and estimation in mata." *Stata Journal*, *14*(3), 623–661(39).

Beebe, J. H. (1977). "Institutional structure and program choices in television markets." *The Quarterly Journal of Economics*, *91*(1), 15–37.

Berry, S., Eizenberg, A., and Waldfogel, J. (2015). "Optimal Product Variety in Radio Markets." Working paper ssrn 2668157.

Berry, S., and Reiss, P. (2007). "Empirical models of entry and market structure." In M. Armstrong, and R. Porter (Eds.), *Handbook of Industrial Organization*, vol. 3, chap. 29, 1845 – 1886, Amsterdam: Elsevier.

Berry, S., and Waldfogel, J. (2004). "Free entry and social inefficiency in radio broadcasting." *RAND Journal of Economics*, *30*(3), 397–420.

Bjorn, P. A., and Vuong, Q. H. (1984). "Simultaneous Equations Models for Dummy Endogenous Variables: A Game Theoretic Formulation with an Application to Labor Force Participation." Working Papers 537, California Institute of Technology, Division of the Humanities and Social Sciences.

Bresnahan, T., and Reiss, P. C. (1991). "Entry and competition in concentrated markets." *Journal of Political Economy*, *99*(5), 977–1009.

Chamberlin, E. (1933). *The Theory of Monopolistic Competition*. No. v. 38 in Harvard economic studies, vol. XXXVIII, Harvard University Press.

Chernozhukov, V., and Hong, H. (2003). "An mcmc approach to classical estimation." *Journal of Econometrics*, *115*(2), 293–346.

Ciliberto, F., and Tamer, E. (2009). "Market structure and multiple equilibria in airline markets." *Econometrica*, *77*(6), 1791–1828.

d'Aspremont, C., Gabszewicz, J. J., and Thisse, J.-F. (1979). "On Hotelling's "Stability in Competition"." *Econometrica*, *47*(5), 1145–50.

Dixit, A. K., and Stiglitz, J. E. (1977). "Monopolistic Competition and Optimum Product Diversity." *American Economic Review*, *67*(3), 297–308.

Downs, A. (1957). *An Economic Theory of Democracy*. New York: Harper and Row.

Ellickson, P. B., and Misra, S. (2008). "Supermarket pricing strategies." *Marketing Science*, *27*(5), 811–828.

Filistrucchi, L., Klein, T. J., and Michielsen, T. O. (2012). "Assessing unilateral merger effects in a two-sided market: An application to the dutch daily newspaper market." *Journal of Competition Law and Economics*.

Gentzkow, M. (2006). "Television and voter turnout." *The Quarterly Journal of Economics*, 931–972.

George, L., and Hogendorn, C. (2012). "Aggregators, search and the economics of new media institutions." *Information Economics and Policy*, *24*(1), 40–51.

George, L. M. (2008). "The internet and the market for daily newspapers." *The B.E. Journal of Economic Analysis & Policy*, *8*(1), 1–33.

George, L. M. (2009). "National television and the market for local products: The case of beer." *The Journal of Industrial Economics*, *57*(1), 85–111.

George, L. M., and Waldfogel, J. (2003). "Who affects whom in daily newspaper markets?" *Journal of Political Economy*, *111*(4), 765–784.

George, L. M., and Waldfogel, J. (2006). "The new york times and the market for local newspapers." *American Economic Review*, *96*(1), 435–447.

George, L. M., and Waldfogel, J. (2008). "National media and local political participation: The case of the new york times." In R. Islam (Ed.), *Information and public choice : from media markets to policy making*, 33–49, World Bank Washington, D.C.

Hotelling, H. (1929). "Stability in competition." *The Economic Journal*, *39*(153), 41–57.

Oberholzer-Gee, F., and Waldfogel, J. (2009). "Media markets and localism: Does local news en espaol boost hispanic voter turnout?" *American Economic Review*, *99*(5), 2120–28.

Reiss, P. C., and Spiller, P. T. (1989). "Competition and entry in small airline markets." *Journal of Law and Economics*, S179–S202.

Snyder Jr., J. M., and Strömberg, D. (2010). "Press coverage and political accountability." *Journal of Political Economy*, *118*(2), 355–408.

Spence, A. (1976a). "Product differentiation and welfare." *American Economic Review*, *66*(2), 407–14.

Spence, A. (1976b). "Product selection, fixed costs, and monopolistic competition." *Review of Economic Studies*, *43*(2), 217–35.

Steiner, P. O. (1952). "Program patterns and preferences, and the workability of competition in radio broadcasting." *The Quarterly Journal of Economics*, *66*(2), 194–223.

Strömberg, D. (2004). "Mass media competition, political competition, and public policy." *The Review of Economic Studies*, *71*(1), 265–284.

Sutton, J. (1991). *Sunk Cost and Market Structure*. Cambridge, MA: MIT Press.

Ted Bergstrom, M. C., and Bills, A. (1995). "Hotelling locational problems with directional constraints: An application to television news scheduling." *Journal of Industrial Economics*, *43*(1), 121–124.

Waldfogel, J. (2003). "Preference externalities: An empirical study of who benefits whom in differentiated-product markets." *RAND Journal of Economics*, *34*(3), 557–68.

Waldfogel, J., Holmes, T. J., and Noll, R. G. (2004). "Who benefits whom in local television markets?[with comments]." *Brookings-Wharton Papers on Urban Affairs*, 257–305.

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

Zhu, T., Singh, V., and Manuszak, M. D. (2009). "Market structure and competition in the retail discount industry." *Journal of Marketing Research*, *46*(4), 453–466.