Merger enforcement in two-sided markets *

Przemysław Jeziorski †

July 3, 2010

Abstract

This paper studies mergers in two-sided markets by estimating a structural supply and demand model and performing counterfactual experiments. The analysis is performed on data for a merger wave in U.S. radio that occurred between 1996 and 2006. The paper makes two main contributions. First, I identify the conflicting incentives of merged firms to exercise market power on both sides of the market (listeners and advertisers in the case of radio). Second, I disaggregate the effects of mergers on consumers into changes in product variety and changes in supplied ad quantity. I find that firms have moderate market power over listeners in all markets, extensive market power over advertisers in small markets and no market power over advertisers in large markets. Counterfactuals reveal that extra product variety created by post-merger repositioning increased listeners' welfare by 1.3% and decreased advertisers' welfare by about \$160m per-year. However, subsequent changes in supplied ad quantity decreased listener welfare by 0.4% (for a total impact of +0.9%) and advertiser welfare by an additional \$140m (for a total impact of -\$300m).

*I would like to thank: Lanier Benkard and Peter Reiss, Ilya Segal, Alan Sorensen, Benjamin Van Roy and Ali Yurukoglu, and participants of numerous seminars.

†Department of Economics, Johns Hopkins University

Between 1996 and 2006, the U.S. radio industry experienced an unprecedented merger wave. This merger wave was prompted by the 1996 Telecommucation Act, which raised ownership caps in local markets and abolished cross-market ownership restrictions. At the height of merger activity, about 30% of stations changed ownership each year and about 20% changed programming format. In this paper, I use this merger wave to study the consequences of consolidation in two-sided markets. I make two main contributions. First, I identify conflicting incentives for stations to exercise market power on both sides of the market (in the case of radio, the two sides are advertisers and listeners). In particular, I separate the impact of consolidation on listener and advertiser surplus. Second, I decompose this impact into two parts: changes in product variety, and market power. In particular, I evaluate whether extra variety can mitigate the negative effects of a decrease in competition. Similar issues arise in other two-sided markets such as credit cards, newspapers or computer hardware. The framework proposed in this paper can be easily adjusted to analyze any of these industries.

In two-sided markets, firms face two interrelated demand curves from two distinct types of consumers. These demands give merging firms conflicting incentives because exercising market power in one market lowers profits in the other market. In the case of radio, a company provides free programming to listeners but draws revenue from selling advertising that is priced on a per-listener basis. In the listener market, a merged firm would like to increase post-merger advertising because it captures some switching listeners. This advertising decreases the welfare of listeners and increases the welfare of advertisers. However, from the perspective of the advertising market, the merged firm would like to supply less advertising, which has the exact opposite impact on listener and advertiser welfare. The firm's ultimate decision, which determines the impact of consolidation on the welfare of both consumer groups, depends on the relative demand elasticities in the two markets.

In this paper, I separately estimate elasticities for both groups using a structural model of the demand and supply of radio programming and advertising. Using those estimates, I perform counterfactual policy experiments that quantify the impact of consolidation on listener and advertiser surplus. I find that market power on the listener side is similar across geographical markets. In contrast, the amount of market power on the advertiser side depends on market population. In particular, firms have a considerable control over advertising price in smaller markets; however, they are price takers in larger markets. Consequently, mergers result in firms lowering advertising quantity in small markets (less than 500 thousand people) by about 13%, which leads to a 6% per-listener increase in ad prices. Mergers increase listener surplus by 2.5% but at the same time decrease advertiser surplus by $235m per year. Conversely, in large markets (more than 2 million people) mergers lead to a 5.5% increase in total advertising minutes while per-listener price stays constant. This results in a 0.3% decrease in listener welfare as well as a slight decrease in advertiser welfare ($0.1m per year). The aggregate national impact of the merger wave amounted to a listener welfare gain of 1% and a $300m per year advertiser welfare loss. I conclude that listeners benefited and advertisers were disadvantaged by the 1996 Telecom Act.

My work is related to several theoretical papers studying complexity of pricing strategies in two-sided markets. The closest studies related to this paper are: Armstrong (2006), Rochet and Tirole (2006), Evans (2002) and Dukes (2004). The general conclusion in this literature is that using a standard supply and demand framework of single-sided markets might be not sufficient to capture the economics of two-sided markets. Additionally, there have been several empirical studies on this topic. For example Kaiser and Wright (2006), Argentesi and Filistrucchi (2007) and Chandra and Collard-Wexler (2009) develop empirical models that recognize the possibility of market power in both sides of the market. They use a form of the Hotelling model proposed by Armstrong (2006) to deal with product heterogeneity. I build on their work, incorporating recent advances in the literature on demand with differentiated products. This allows me to incorporate richer consumer heterogeneity and substitution patterns (e.g. Berry, Levinsohn, and Pakes (1995), and Nevo (2000)) that are necessary to capture complicated consumer preferences for radio programing. Moreover, I supplement reduced form results on market power with out-of-sample counterfactuals that explicitly predict changes in supplied ad quantity and consumer welfare.

The second contribution of this paper is the decomposition of the total impact of mergers on consumer surplus into changes in product variety and effects of exercising extra market power from joint ownership. This exercise is motivated by the fact that in most cases consumers have preference for variety, so it is possible that extra variety created by mergers might mitigate the negative effects of extra market power. In order to verify the above claim, I quantify consumers' value for extra variety and compare it to the loss in surplus coming from the extra market power.

This approach relates to Kim, Allenby, and Rossi (2002), who compute the compensating variation for the changes of variety in tastes of yogurt and Brynjolfsson, Hu, and Smith (2003) who do the same for the variety of books offered in on-line bookstores. These papers assume away the fact that changes in variety will be followed by readjustments in equilibrium prices. In this paper, taking their analysis one step forward, I incorporate such strategic responses by performing counterfactual experiments in which new equilibrium prices are computed.

Berry and Waldfogel (2001) and Sweeting (2008) document that the post-1996 merger wave resulted in an increase in product variety. I investigate their results using a structural utility model and find that extra variety leads to a 1.3% increase in listener welfare. However, because product repositioning softened competition in the advertising market and caused some stations to switch to a “Dark“ format [1], advertiser welfare also decreased by $147m per year. Additionally, I find that product ownership consolidation and repositioning are followed by advertising quantity readjustments. I estimate, that this effect leads to a 0.3% decrease in listener welfare and an additional $153m decrease in advertiser welfare. The two effects combined total to $1% increase in listener welfare and $300m decrease in advertise welfare. While extra variety mitigates the negative effects of mergers on listeners, it strengthens the negative impact on advertisers.

This paper is organized as follows. Section 2 outlines the questions investigated in the paper in a formal way and describes the structural model of the industry. Section 3 contains the description of the data. Estimation techniques used to identify the parameters of the model are described in Section 4. Results of the structural estimation are presented in Section 5. Section 6 describes the results of counterfactual experiments. Robustness checks of different modeling assumptions are contained in Section 7. Section 8 provides the conclusion.

1 Radio as a two-sided market

The radio industry is an example of a two-sided market (other examples include advertising platforms, credit cards or video games). Such markets are usually characterized by the existence of three types of agents: two types of consumers and a platform provider. What distinguishes this

[1] When in “dark” format, the station holds the frequency so that other stations cannot use it. “Dark” stations typically do not broadcast or broadcast very little non-commercial programming.

setup from a standard differentiated product oligopoly is that the platform provider is unable to set prices for each type of consumer separately. Instead, the demand curves are interrelated through a feedback loop in such a way that quantity sold to one consumer determines the market clearing price for the other consumer. In this subsection I argue that this feedback makes it complicated to determine whether the supplied quantities are strategic substitutes or complements (as defined in Bulow, Geanakoplos, and Klemperer (1985)). This creates important trade-offs in the case of a merger and affects the division of surplus between both types of consumers. The remainder of this subsection discusses this mechanism in detail using the example of radio; however, the discussion applies to the majority of other two-sided markets.

In the case of radio there are three types of agents: radio stations, listeners, and advertisers. Radio stations provide free programming for listeners and draw revenue from selling advertising slots. First, consider the demand curve for radio programming. The listener market share of the radio station j is given by

$$r_j = r_j(q|s, d, \theta^L) \tag{1.1}$$

where q is the vector of advertising quantities, s are observable and unobservable characteristics of all active stations, d are market covariates and θ^L are parameters of the listener demand. Since radio programming is free, there is no explicit price in this equation. However, because listeners have disutility for advertising, its effect is similar to price, i.e. $\frac{\partial r_j}{\partial q_j} < 0$.

The market clearing price of an advertising slot in station j depends on the amount of advertising supplied and the number of listeners to station j. Therefore, the inverse demand curve for advertising slots is

$$p_j = p_j(q, r_j(q)|s, d, \theta^A) \tag{1.2}$$

where θ^A are parameters. Note that the advertising quantity affects the advertising price in two ways: directly through the first argument and indirectly through the listener demand feedback loop (the second argument).

Suppose for now that each owner owns a single station and there is no marginal cost (I relax these assumptions later). In equilibrium, each radio station chooses their optimal ad quantity, keeping the quantities of the other stations fixed, i.e.

$$\max_{q_j} p_j(q, r_j(q)|q_{-j})q_j \tag{1.3}$$

In contrast to a differentiated products oligopoly, the firm has just one control (ad quantity) that determines the equilibrium point on both demand curves simultaneously. The first order conditions for profit maximization are given by

$$\frac{\partial p_j}{\partial q_j} q_j + \frac{\partial p_j}{\partial r_j} \frac{\partial r_j}{\partial q_j} q_j + p_j = 0$$

The important fact is that this condition shares features with both the Cournot and Bertrand models. On the one hand, the first term represents the direct effect of quantity on price, and it is reminiscent of the standard quantity setting equilibrium (Cournot). On the other hard, the second component represents the listener feedback loop and is reminiscent of the price setting model (Bertrand), because ad quantities function like prices in the demand for programming.

In order to determine the impact of a merger on the equilibrium ad quantities supplied we need to know if they are strategic complements or substitutes. The duality described in the previous paragraph make it ambiguous. This is because in the Cournot model quantities are strategic substitutes and in the differentiated product Bertrand model prices are strategic complements. Without knowing the relative strengths of the direct effects and the feedback loop, we cannot conclude whether a merger leads to an increase or decrease in ad quantity on the margin. Moreover, in the borderline case in which the effects cancel each other, a merger does not effect quantity at all; in this case, even though companies have market power over both consumers, they are unable to exercise it. Measuring these effects is critical for predicting the split of surplus between advertisers and listeners. When the direct effect is stronger, mergers lead to contraction in the ad quantity supplied and higher prices. This will benefit listeners but hurt advertisers. However, if the feedback loop is stronger than the direct effect then merger leads to more advertising and lower prices, benefiting advertisers and hurting listeners.

Because the theory does not give a clear prediction about the split of surplus, I investigate this question empirically using a structural model. In the remainder of this section I put more structure on equations (1.1), (1.2) and (1.3), enabling separate identification of both sets of demand elasticities. I discover the relative strength of the direct and feedback effects and perform counterfactuals that quantify the extent of surplus reallocation.

1.1 Industry setup

During each period t, the industry consists of $\mathbb{M}$ geographical markets that are characterized by a set of demographic covariates $d \in \mathcal{D}_m$. Each market m can have up to $\mathbb{J}^m$ active radio stations and $\mathbb{K}^m$ active owners. Each radio station is characterized by one of $\mathbb{F}$ possible programming formats. Station formats include the so-called "dark" format when a station is not operational The set of all station/format configurations is given by $\mathbb{F}^{\mathbb{J}^m}$. Ownership structure is defined as a $\mathbb{K}^m$-element partition of station/format configuration $s^{mt} \in \mathbb{F}^{\mathbb{J}^m}$. In an abuse of notation, I will consider s^{mt} to be a station/format configuration for market m at time t, as well as an ownership partition. Each member of the ownership partition (denoted as s_k) specifies the portfolio of stations owned by firm k.

The quality of the programming of radio station j is fully characterized by a one-dimensional quality measure $\xi_j \in \Xi \subset \mathbb{R}$. The state of the industry at time time t in market m is therefore fully characterized by: a station/format configuration and ownership structure s^{tm}, vector of station quality measures ξ^{tm} and market covariates d^{tm}. In the next subsections I present a detailed model of listener demand, advertiser demand, and supply side. Throughout the description I take the triple $(s^{tm}, \xi^{tm}, d^{tm})$ as given and frequently omit market or time subscripts to simplify the notation.

1.2 Listeners

This subsection describes the details of the demand for listenership introduced in equation (1.1). The model will be a variation on the random coefficient discrete choice setup proposed by Berry, Levinsohn, and Pakes (1995).

I assume that each listener chooses only one radio station to listen to at a particular moment. Suppose that s is a set of active stations in the current market at a particular time. For any radio station $j \in s$, I define a vector $\iota_j = (0, \ldots, 1, \ldots, 0)$ where 1 is placed in a position that indicates the format of station j.

The utility of listener i listening to station $j \in s$ is given by

$$u_{ij} = \theta^L_{1i}\iota_j - \theta^L_{2i} q_j + \theta^L_3 \mathrm{FM}_j + \xi_j + \epsilon_{ji} \tag{1.4}$$

where θ^L_{2i} is the individual listener's demand sensitivity to adverting, q_j the amount of advertising,

ξ_j the unobserved station quality, ϵ_{ji} an unobserved preference shock (distributed type-1 extreme value), and finally θ^L_{1i} is a vector of the individual listener's random effects representing preferences for formats.

I assume that the random coefficients can be decomposed as

$$\theta^L_{1i} = \theta^L_1 + \Pi D_i + \nu_{1i}, \quad D_i \sim F_m(D_i|d), \quad \nu_{1i} \sim N(0, \Sigma_1)$$

and

$$\theta^L_{2i} = \theta^L_2 + \nu_{2i}, \quad \nu_{2i} \sim N(0, \Sigma_2)$$

where Σ_1 is a diagonal matrix, $F_m(D_i|d)$ is an empirical distribution of demographic characteristics, ν_i is unobserved taste shock, and Π is the matrix representing the correlation between demographic characteristics and format preferences. I assume that draws for ν_i are uncorrelated across time and markets.

The random effects model allows for fairly flexible substitution patterns. For example, if a particular rock station increases its level of advertising, the model allows for consumers to switch proportionally to other rock stations depending on demographics.

Following Berry, Levinsohn, and Pakes (1995), I can decompose the utility into a part that does not vary with consumer characteristics

$$\delta_j = \delta(q_j|\iota_j, \xi_j, \theta^L) = \theta^L_1 \iota_i - \theta^L_2 q_j + \theta^L_3 \mathrm{FM}_j + \xi_j$$

an interaction part

$$\mu_{ji} = \mu(\iota_j, q_j, \Pi D_i, \nu_i) = (\Pi D_i + \nu_{1i})\iota_j + \nu_{2i} q_j$$

and error term ϵ_{ji}.

Given this specification, and the fact that ϵ_{ji} is distributed as an extreme value, one can derive the expected station rating conditional on a vector of advertising levels q, market structure s, a vector of unobserved station characteristics ξ, and market demographic characteristics d,

$$r_j(q|s, \xi, d, \theta^L) = \int\int \frac{\exp[\delta_j + \mu_{ji}]}{\sum_{j' \in s} \exp[\delta_{j'} + \mu_{j'i}]} dF(\nu_i) dF_m(D_i|d)$$

1.3 Advertisers

In this subsection I present the details of the demand for advertising introduced in equation (1.2). The model captures several important features specific to the radio industry. In particular, the pricing is done on a per-listener basis, so that the price for a 60sec slot of advertising is a product of cost-per-point (CPP) and station rating (market share in percents). Moreover, radio stations have a direct market power over advertisers, so that CPP is a decreasing function of the ad quantities offered by a station and its competitors. The simplest model that captures these features and is a good approximation of the industry is a linear inverse demand for advertising, such as

$$p_j = \theta_1^A r_j \left(1 - \theta_2^A \sum_{f' \in \mathbb{F}} \omega_{ff'}^m q_{f'}\right) \tag{1.5}$$

where f is a format of station j, θ_1^A is a scaling factor for value of advertising, θ_2^A is a market power indicator and $\omega_{ff'} \in \Omega$ are weights indicating competition closeness, between formats f and f'.

The weights ω are a key factor determining competition between formats and thus market power. They reflect the fact that some formats are further and others are closer substitutes for advertisers because of differences in the demographic composition of their listeners. In principle, one could proceed by estimating these weights from the data. However, here it is not feasible to do that because the available data do not contain radio station level advertising prices. Instead, I make additional assumptions that will enable me to compute the weights using publicly available data. The reminder of this subsection discusses the formula for the weights and provides an example supporting this intuition. The formal micro-model is given in Appendix B.

Let there be $\mathcal{A}$ types of advertisers. Each type $a \in \mathcal{A}$ targets a certain demographic group(s) a. I.e. advertiser of type a gets positive utility only if a listener of type a hears an ad. Denote $r_{f|a}$ to be the probability that a listener of type a chooses format f and $r_{a|f}$ to be the probability that a random listener of format f is of type a. Advertisers take these numbers, along with station ratings r_j, as given and decide on which station to advertise. This assumption is is motivated by the fact that about 75% is purchased by small local firms. Such firms' advertising decisions are unlikely to influence prices and station ratings in the short run.

This decision problem results in an inverse demand for advertising with weights $\omega_{jj'}$, that are given by

$$\omega_{ff'} = \frac{1}{\sum_{a \in \mathcal{A}} r_{a|f}^2} \sum_{a \in \mathcal{A}} r_{a|f} \left(r_{a|f} r_{f'|a}\right) \tag{1.6}$$

The formal justification and derivation of this equation is given in Appendix B. The intuition behind it is that the total impact on the per-listener price of an ad in format f is a weighted average of impacts on the per-listener value of an ad for different types of advertisers. The weighting is done by the advertisers' arrival rates, which are equal to the listeners' arrival rates $r_{a|f}$. For each advertiser of type a the change of value of an ad in format f, in response to a change of total quantity supplied in format f', is affected by two things: it is proportional to the probability of correct targeting in format f, given by $r_{a|f}$, because advertisers are expected utility maximizers; and it is proportional to the share of advertising purchased by this advertiser in format f', given by $r_{f'|a}$. Assembling these pieces together and normalizing the weights to sum to 1 gives equation (1.6).

To illustrate how these weights work in practice, consider the following example. Suppose that there are only two possible formats of programming: Talk and Hits, and two types of consumers: Teens and Adults. Teens like mostly Hits format and Adults like Talk format. However, Adults like Hits more than Teens like Talk. Hypothetical numerical values of $r_{f|a}$ and $r_{a|f}$ are given in Table 1.

In Table 1, the impact of Hits on the price of Talk is greater than the impact of Talk on the price of Hits. This is due to the fact that the quantity supplied in the Hits format affects Adult-targeting advertisers (who drive the price of the Talk format) to a much greater extent than ad quantity in Talk affects Teen-targeting advertisers (who drive the price of the Hits format). Moreover, because the weights sum up to 1, it must be that the own effect of Talk is weaker than that of Hits. This is exactly the essence of the mechanism behind Equation (1.6). More examples from the data with an extensive discussion are given in Section 4.

In the next section I will combine demand for programming and advertising to compose the profits of the radio station owners.

1.4 Radio station owners

In this subsection I will describe a profit maximizing problem for the radio station owners. It will be a version of equation (1.3) that allows for non-zero cost in selling advertising and common radio station ownership. Given the advertising quantity choices of competing owners q_{-k}, the profit of radio station owner k is given by

$$\begin{aligned}\bar{\pi}_k(q_k|q_{-k},\xi,\theta) &= \max_{\{q_j;j\in s_k\}} \sum_{j\in s_k} r_j(q|\xi,\theta^L)p_j q_j - \mathrm{MC}_j(q_j) = \\ &= \theta_1^A \max_{\{q_j;j\in s_k\}} \sum_{j\in s_k} q_j r_j(q|\xi,\theta^L)\left(1-\theta_2^A \sum_{f'\in\mathbb{F}} \omega^m_{ff'} q_{f'}\right) + \mathrm{C}_j(q_j|\theta^A,\theta^C)\end{aligned} \tag{1.7}$$

where $C_j(q_j)$ is the total cost of selling advertising. I assume constant marginal cost and allow for a firm level of unobserved cost heterogeneity η_j, i.e. $\mathrm{C}_j(q_j|\theta^A,\theta^C) = \theta_1^A[\theta^C+\eta_j]q_j$.

I assume that the markets are in a Cournot Nash Equilibrium. The first order conditions for profit optimization become

$$r_j p_j + \sum_{j'\in s_k} q_{j'}\left[\frac{\partial r_{j'}}{\partial q_j} p_{j'} - r_{j'}\theta_2^A \omega^m_{jj'}\right] - \theta^C - \eta_j = 0 \quad \forall k \text{ and } j\in s_k \tag{1.8}$$

Additionally, I assume that station unobserved quality is exogenous but serially correlated. It evolves according an AR(1) process such that

$$\xi_j^t = \rho\xi_j^{t-1} + \zeta_j^t \tag{1.9}$$

where ζ_j^t is an exogenous innovation to station quality.

2 Data description

I have constructed a panel of data on radio stations and radio station ownership merging data from two sources: *BIA Financial Network Inc.* and *the SQAD Media Market Guide.*

BIAfn provided me data on: radio station ownership, revenues, market shares and formats. The data are a 1996-2006 panel covering each radio station in the market in 2006. The data are incomplete in the sense that I do not observe all the stations that exited the market between 1996 and 2006. According to Sweeting (2007) there were only 50 stations that exited during this period, mostly due to violations of FCC regulations. Because this number is small relative to the 11,000 stations in the sample, this omission is unlikely to significantly influence the results.

The BIAfn data are supplemented with data on aggregate advertising prices. Unfortunately, price data at the station level are not available. SQAD instead provides estimates of market prices that are obtained using proprietary formulas. According to anecdotal evidence, those estimates are widely recognized as the industry standard and are the best available data on market prices. Radio market prices are reported as a Cost per Rating Point (CPP). CPP is the cost of advertising per 1 percent of listenership. SQAD provides CPP broken down into daytime and demographic categories. We will estimate station level prices from SQAD CPPs using radio station ratings that are broken down by time of day and demographics.

An observation in my data is a radio station operating in a specific half-year and in a specific market. BIAfn and SQAD use Arbitron market definitions. An Arbitron market is in most cases a county or a metropolitan area. According to the surveys conducted by CRA International (2007) for the Canadian market (which is similar to the US market): "The majority of radio advertisers are local. They are only interested in advertising in their local area since most of their customers and potential buyers live in or very near their city." In our analysis, I assume no interdependence between markets. To further assure that there is no overlap between markets, I use only the 88 market sub-selection that was developed in Sweeting (2007). Table 7 presents a list of the 88 markets, along with their populations.

To achieve a sharper identification of the random effects covariance matrix, I use listenership shares of different demographic groups in each of the formats that has been aggregated from the 100 biggest markets [2]. I observe listenership shares of different age/gender groups within each station format between 1998 and 2006, and shares for income, race and education groups between 2003 and 2006. Unfortunately, I do not observe a full matrix of market shares for all the combinations of demographic variables. For example, I do not see what the share of rock stations is among black, educated males. Instead I have shares for blacks, educated people, and males.

Table 2 contains some basic aggregate statistics about the industry. The top part of the table documents changes in concentration of radio station ownership. The average number of stations owned in our dataset grew from 4.43 in 1996 to 6.28 in 2006. This ownership consolidation resulted in growth of the market share of the 3 biggest owners (C3) from 77% in 1996 to 90% in 2006, peaking at 97% in 2000. The middle part of the table contains the average percentages of stations that switched owners and that switched formats. Between 1996 and 2000 more than 10% of stations switched owners yearly. After 2000 the number dropped to below 4%. Greater concentration activity in the 1996-2000 period was also associated with more format switching.

[2] Source: Arbitron Format Trends Report

The percentage of stations that switched format peaked in 1998 and 2001 at 13%.

3 Estimation

The estimation of the model is done in two steps. In the first step, I estimate the demand model that includes parameters of the consumer utility θ^L (see equation (1.4)) and the unobserved station quality lag parameter ρ (see equation (1.9)). In the second step, we recover parameters of the inverse demand for advertising θ^A, $w_{jj'}$ (see equation (1.5)) and cost parameters θ^C (see equation (1.7)).

3.1 First stage

This stage provides the estimates of demand for radio programming θ^L. Estimation is done using the generalized method of simulated moments. I use two sets of moment conditions. The first set is based on the fact that innovation to station unobserved quality ξ_j has a mean of zero conditional on the instruments:

$$E[\xi_{jt} - \rho\xi_{jt-1}|Z_1, \theta^L] = 0 \tag{3.1}$$

This moment condition follows Berry, Levinsohn, and Pakes (1995) and extends it by explicitly introducing auto-correlation of ξ. I use instruments for advertising quantity since it is likely to be correlated with unobserved station quality. My instruments include: lagged mean and second central moment of competitors' advertising quantity, lagged market HHIs and lagged number and cumulative market share of other stations in the same format. These are valid instruments under the assumption that ξ_t follows an AR(1) process and the fact that decisions about portfolio selection are made before decisions about advertising.

A second set of moment conditions is based on demographic listenership data. Let R_{fc} be the national market share of format f among listeners possessing certain demographic characteristics c. The population moment conditions are

$$\int_t \int_{(D_{ic}^t, m)} \int_{\nu_i} \frac{\exp[\delta_j^{mt} + \mu_{ji}^{mt}]}{\sum_{j' \in s^{mt}} \exp[\delta_{j'}^{mt} + \mu_{ij'}^{mt}]} dF(\nu_i) dF_c^t(D_{ic}^t, m) dt = R_{fc} \tag{3.2}$$

where $F_c^t(D_i, m)$ is a national distribution of people who possess characteristic c at time t. Each person is characterized by the demographic characteristics D_i and the market m they belong to.

For each time t and demographic characteristic c, I draw $\mathcal{I}$ observation pairs (D^t_{ic}, m) from the nationally aggregated CPS. Let $g = (g_1, g_2)$ represent the empirical moments and W be a weighting matrix. I estimate the model by using the constrained optimization procedure:

$$\min_{\theta^L, \xi, g} g'Wg$$

Subject to:

$$\begin{aligned}
&\hat{r}_{jmt}(q_{mt}|s_{mt}, \xi_{mt}, d_{mt}, \theta^L) = r_{jmt} \quad \forall t, m \\
&\frac{1}{\mathcal{T}\mathcal{I}} \sum_t \sum_{(D^t_{ic}, m)} \int_{\nu_i} \frac{\exp[\delta^{mt}_j + \mu^{mt}_{ji}]}{\sum_{j' \in s^{mt}} \exp[\delta^{mt}_{j'} + \mu^{mt}_{ij'}]} dF(\nu_i) - R_{fc} = g_1 \quad \forall c \\
&\frac{1}{\text{size of } \xi} Z_1(\xi - \rho L \xi) = g_2
\end{aligned} \tag{3.3}$$

where L is a lag operator that converts the vector ξ into one-period lagged values. If the radio station did not exist in the previous period, the lag operator has a value of zero. Integration with respect to demographics when calculating the first constraint is obtained by drawing from the CPS in the particular market and period. This way of integrating allows us to maintain proper correlations between possessed demographic characteristics. The same is true when obtaining the data set D_{ict}. When computing the interaction terms μ in the second constraint, I draw one vector ν_i from the normal distribution for each D_{ict}.

3.2 Second stage

The second stage of the estimation obtains the competition matrix Ω and the parameters of demand for advertising θ^A. The estimation is done separately for every market, thereby allowing for different Ω and θ^A.

To compute the matrices Ω^m for each market I use the specification layed out in section 1.3. The elements of the matrix Ω are specified as

$$\omega_{ff'} = \frac{1}{\sum_{a \in \mathcal{A}} r^2_{a|f}} \sum_{a \in \mathcal{A}} r_{a|f} \left(r_{a|f} r_{f'|a} \right)$$

following equation (1.6). The $r_{f|a}$ are advertisers' beliefs about listeners' preferences for formats. These are constant across markets. To recognize that advertisers know the demographic composition of each market I allow for market specific listener arrival rates for each format $r^m_{f|a}$. However, I assume that the advertisers compute those values by using Radio Today reports and the Current Population Survey. After computing weights, I treat Ω^m as exogenous and fixed in all of the following steps [3].

After computing matrices Ω, I estimate θ^A. Using estimates of demand for radio programming θ^L from the first stage, I compute ratings for each station conditioned on the counterfactual advertising quantities. I use the set of $3M$ moment conditions

$$E_m[\eta^m | Z_2, \theta^A, \theta^C] = 0 \quad \forall m \in \mathbf{M} \tag{3.4}$$

where the integral is taken with respect to time and stations in each market. η_j^{tm} is an unobserved shock to marginal cost defined in equation (1.5). The Z_2 are three instruments: a column of ones, the AM/FM dummy and number of competitors in the same format. They are uncorrelated with η^m under the IID assumption, but are correlated with the current choice of quantity because they describe the market structure.

We back out η_j^{tm} using FOCs for owner's profit maximization (see equation (1.7))

$$\eta_j^t = r_j^t p_j^t + \sum_{j' \in s_k^{tm}} q_{j'}^t \left[\frac{\partial r_{j'}^t}{\partial q_j^t} p_{j'}^t - \theta_{2m}^A r_{j'}^t \omega_{ff'}^m \right] - \theta_m^C \quad \forall t \in \mathbf{T}, k \in \mathbf{K}^{tm}, j \in s_k^{tm} \tag{3.5}$$

Since the equation does not depend on θ_{1m}^A, I can use it to estimate θ_{2m}^A and θ_m^C. During the estimation, I allow for a different value of marginal cost for each market. I allow for 3 different values for the slope of inverse demand depending on the population of the market (up to 500 people, between 500 and 1500, and 1500 or more). Ratings and derivatives of ratings in the equation (3.5) are calculated using the estimates of θ^L and ξ from the first stage. Demographic draws are taken from the CPS and are independent of those used in the first stage. Given the estimates of θ_{2m}^A and θ^C, I can back out θ_{1m}^A by equating the observed average revenue in each market with its predicted counterpart.

Next I discuss a variation in the data that identifies parameters θ^A and θ^C. The intuition for such identification is that estimating Equation 3.5 can be regarded as a linear regression in which θ_m^C is an intercept and θ_2^A is a coefficient of a variable that is a function of supplied quantity. In this case, the mean deviation of FOCs from zero in each market identifies the intercept θ_m^C. The slope

[3] Such an approach potentially ignores possible variance of the Ω^m estimator. The source of this variance might come from the finiteness of the CPS dataset and the distribution of Arbitron estimates.

parameter θ_2^A is identified by the size of the response of the firm to changes in quantity supplied by its competitors due to change in the market structure or demographics. Such a response, as mentioned in Section 1, is composed of listeners' demand feedback and the direct effect of quantity on CPP. Elasticity of listeners' demand, that determines the strength of the feedback, is consistently estimated in the first step. Therefore, one can subtract the difference out the feedback effect from the total response observed in the data. This allows to obtain the strength of the direct effect that directly identifies the slope of the CPP, θ_2^A. For example, if we look at the response of ad quantity reacting to the merger, the slope of listeners' demand alone predicts large increases in ad quantity. However in the data, we observe smaller increases or even decrease in the quantity supplied, depending on the market. Those differences are rationalized by a negative value of CPP slope, θ_2^A.

4 Results

This section presents estimates of the structural parameters. The next subsection discusses listeners' demand parameters. This is followed by results concerning advertisers' demand and marker power. The last subsection contains estimates of marginal cost and profit margin (before subtracting fixed cost).

4.1 Listeners' demand

Table 3 contains estimates of demand parameters for radio programming. The estimate of the mean effect of advertising on listeners' utility is negative and statistically significant. This is consistent with the belief that radio listeners have a disutility for advertising. When it comes to the mean effects of programming formats, Contemporary Hit Radio format gives the most utility, while the News/Talk format gives the least.

The second column of Table 3 contains variances of random effects for station formats. The higher a format's variance, the more persistent are the tastes of listeners for that format. For example, in response to an increased amount of advertising, if the variance of the random effect for that format is high, listeners tend to switch to a station of the same format. The estimates also suggest that tastes for the Alternative/Urban format are the most persistent.

Table 4 contains estimates of interactions between listener characteristics and format dummies. The majority of the parameters are consistent with intuition. For example, younger people are more willing to choose a CHR format while older people go for News/Talk. The negative coefficients on the interaction of Hispanic format with education and income suggests that less educated Hispanic people with lower income are more willing to listen to Hispanic stations. For blacks, I find a disutility for Country, Rock and Hispanic, and a high utility for Urban. This is consistent with the the fact that Urban radio stations play mostly rap, hip-hop and soul music performed by black artists.

4.2 Advertisers' demand

Tables 5 presents the weights for selected markets representing large, medium and small listener populations. They were computed using the 1999 edition of Radio Today publication and Common Population Survey aggregated from 1996 to 2006. It is interesting to compute a total impact coefficient that is the sum of all the columns of the table for each format. Not surprisingly, general interest formats like AC and News/Talk have the biggest impact on the price of advertising, while Spanish format has the smallest. The values on the diagonals of the matrices represent the formats' own effect of the quantity of advertising supplied on per-listener price. They are usually bigger than the off-diagonal values, that suggests that it is mostly the ad quantity in the same format that influences a per-listener price. In accord with an intuition, the formats with the most demographically homogenous listener pools, Urban/Alternative and Spanish, have the highest values of the own effects. On the other hand, general interest formats like CHR and Rock are charaterized by the smallest values of the own effect, measuring the fact that their target population of listeners is more dispersed across other formats. For cross effects, one notices that News/Talk is close to AC and Urban is close to CHR. This can be explained by, for example, the age of the listeners. In the former case the formats appeal to an older population while in the latter case to a younger one.

Estimates of the slope of inverse demand are presented in Table 6. In markets with less than 0.5m people radio stations have considerable control over the per-listener price. However, such control significantly drops in markets from 0.5m to 2m people, and it disappears completely in markets with more than 2m people, making radio stations essentially price takers. I suspect that this phenomenon can be explained by the fact that in larger markets there are more outside options for radio advertising. This can lead to tougher competition between media outlets, and make the inverse demand for advertising flatter. However, in small markets radio might be a primary advertising channel, because other media like the Internet or billboards are not as widespread. This gives radio stations more control over price.

4.3 Supply

The marginal costs of selling advertising minutes are presented in Table 7. The values of this cost range from \$356 per minute of advertising sold in Los Angeles, CA to \$11 in Ft. Myers, FL. 66% of the variation in marginal cost can be explained by variation in market population. A population increase of one thousand translates to about a 2 cent increase in marginal cost (with t-stat equal to 12). The high correlation between population and marginal costs can be explained by the fact that revenues per-minute of advertising are an increasing function of total market population. Suppose this surplus is split between radio station owners and advertisers' sales people according to the Nash Bargaining solution. In this case, the high correlation of revenue with population will translate into a high correlation of marginal cost with population.

From the revenues and marginal cost estimates, I can calculate variable profit margins. These are presented in the last last column of Table 7. The range is from 92% in Shreveport, LA to 15% in Honolulu, HI and Reno, NV. It is interesting that 38% of the profit margin variation can be explained by the variance in total ad quantity supplied and markets with high profit margins firms supply more advertising. The marginal effect of extra minute per day of broadcasted advertising translates into 0.6% of extra profit margin.

5 Counterfactual experiments

In this section I investigate the impact of consolidation on listener and advertiser welfare. First, I investigate the changes in the surplus of listeners and advertisers. In particular, I calculate how much market power was exercised on both of those groups. Second, I decompose market power into a variety component and extra market power that is manifested in changes in quantity supplied.

Before performing counterfactual calculations, consider descriptive relationships between concentration and prices. First, I regressed market Price Per Rating Point on a market's HHI, including market fixed effects. I find that higher concentration is correlated with higher prices in the advertising market, suggesting that radio station owners are exercising some amount of market power on advertisers. Second, I regressed total advertising supplied on the market's HHI with market dummies. Here I get a coefficient of 1.65(0.3). This is evidence of market power in the listener market. Because market power appears to be present in both market segments, I cannot definitely conclude who had more surplus extracted by radio station owners if I just use quantities and prices. In the next subsection I present the structural counterfactuals that answer this question.

5.1 Impact of mergers on consumer surplus

To isolate the impact of the Telecom Act on a surplus division between advertisers and listeners, I perform a counterfactual in which I recompute new equilibrium ad quantities under the old 1996 ownership structure and 1996 formats. This calculation is motivated by the fact that in 1996 many markets were at their ownership caps.

The total impact of consolidation on advertiser and listener welfare is presented in the last row of Table 8. It turns out that mergers decrased total ad quantity by roughtly 14 thousand minutes. That resulted in lowering average ad exposure by 4.8 persons-day-minutes (pdm), which is about 10% of the total ad load. The changes translated to about a 4.7 pdm increase in consumer welfare. Because we do not observe dollar prices in the listenership market we cannot compute the dollar value of this compensating variation. However, we can compute a rough estimate using the prices for the satellite radio. If we assume people buy satelite radio just to avoid advertising, we get a rough estimate of 1.5 cents per minute, or 730million dollars for each person-day-minute per year. The total effect would amount to $3.5b. This is of course a very loose upper bound on the overall welfare gain, however if make a conservative assumption that only 10% of the value of satellite radio is lack of advertising, we get $350m.

For advertisers, a decrease in quantity supplied leads to about 2.57% increase in per-listener prices, or a $300m decrease in advertiser surplus. I therefore conclude that the Telecom Act lead to a reallocation of surplus from advertisers to listenerss. Moreover, because the gain by listeners ($350m) is larger than the surplus lost by advertisers, I find that the Act created new surplus. This increase can be explained by the fact that listeners are more annoyed by ads than the value of an ad to the advertisers.

A deeper story can be told by looking seperately at small versus large markets. As mentioned in the previous section, radio stations have considerable control over prices in small markets, and no control in the large markets. Motivated by this fact, I present counterfactuals for markets with less than 0.5 population and more than 2m population. In smaller markets (see Table 9), stations contract advertising to exercise market power on advertisers. They supply more than 10,000 minutes less of advertising. That translates into a 7.3pdm decrease in ad exposure, which increases consumer surplus by 11.6pdm. However, prices rise by 6.4%, and cause a $230m loss in advertiser surplus. On the other hand in large markets (see Table 10) firms supply more than 2,000 extra minutes of advertising, which lowers consumer surplus by almost 2pdm. On balance, this does not affect advertiser surplus. I conclude that listeners gained form the Telecom Act only in small markets.

5.2 Effects of product variety and market power

Berry and Waldfogel (2001) suggest that the negative effects of ownership consolidation on listeners might be mitigated by format switching. They find that post-merger repositioning results in spatial competition leading to more variety, which they assume is beneficial for the listeners [4]. To quantify this effect, I compare surpluses computed imposing 1996 ownership and formats with surpluses computed imposing actual ownership and formats without ad quantity adjustments. That is, I fix ad quantities computed with 1996 ownership and formats. The results of this experiment are presented in the first row of Table 8. It turns out that if I do not account for quantity changes, the assertion of Berry and Waldfogel (2001) is true. In this case, listeners have a 1.3% larger surplus (about 6.6pdm) after consolidation and format switching. Listener surplus grows because of two factors: increased variety and decreased advertising exposure. The latter decreased even though I keep number of ad minutes fixed. However, in the real world, repositioning changes firms' incentives to set ad quantity, because it softens competition in the advertising market. The impact of quantity readjustments is presented in the middle row of Table 8. It turns out that both listeners and advertisers are worse off due to quantity adjustments. Listeners lose 1.9pdm and advetisers lose additional $150m in surplus.

[4] Similar results obtained using direct analysis of station playlists can be found in Sweeting (2008).

6 Robustness analysis

This section examines the robustness of my advertising model to different assumptions about competition among station formats. This step is motivated by the fact that the data concerning advertiser deals is incomplete. I deal with the incompleteness by proposing a stilyzed decision model for advertisers that uses publicly available data to predict substitution patterns between formats. These patterns directly detemine the market power of stations over advertsers, and can potentially alter the results of counterfactual experiments.

To investigate the robustness of the results, I reestimated the model under two alternative assumptions. The first scenario represents the extreme situation in which formats compete only between themselves. In particular, suppose that advertiser types get utility from only one particular format. In this case, equation (1.6) has $\omega_{ff} = 1$ and $\omega_{ff'} = 0$ if $f \neq f'$. The second scenario represents another extreme in which formats are perfect substitutes, i.e., there is only one type of advertiser who values all formats in the same way. Formally this means that $\omega_{ff'} = 1/8$, because there are 8 possible formats. The estimated model is in a sense in-between the these extreme alternatives, because it assumes that formats are imperfect substitutes.

Estimates of the inverse demand advertising slopes are presented in Table 11. The estimates show that the baseline model lies between the two extremes. When we assume oligopoly within a format, the estimated slope parameter θ_2^L is smaller than the one in the baseline model. On the other hand in the perfect substitutes model, the estimated slope tends to be higher. Despite the fact that there are statistical differences between the different models, the main qualitative assertion, that stations have more power in smaller markets, still holds. In order to assess the economic implication of those differences, I recomputed the estimated profit margin under different models. It turns out that the model with format oligopoly predicts on average a 2.4% higher profit margins than the baseline model. Conversely the model with perfect substitutes predicts 2.1% lower profit margin.

To draw final conclusions about the strength of the assumption about weights, I recomputed the main counterfactual using the alternative models. The results are presented in Table 12. The baseline again lies between the new counterfactuals. There is no qualitative change in the results. Moreover the percentage changes in consumer and advertiser surplus are almost the same.

Consequently, I conclude that the results of the paper are not sensitive to changes in the assumption about substitution between formats.

7 Conclusion

In this paper I analyze mergers in two-sided markets on the example of the 1996-2006 consolidation wave in U.S. radio industry. The goal of this study is to describe and quantify how mergers in the two-sided market differ from a differentiated product oligopoly setting. I make two main contributions. First, I recognize the fact two-sided markets consist of two types of consumers, who may be affected by the merger in different ways. For example, if extra market power causes the radio station to increase advertising, it will benefit consumers but hurt advertisers. Second, I disaggregate the impact of a merger on consumers into changes in the variety of available products and changes in supplied quantity of ads.

Radio is an important medium in the U.S., reaching about 94% of Americans twelve years old or older each week. Moreover, the average consumer listens to about 20h of radio per week and between 6am and 6pm more people use radio than TV or print media[5]. In 1996 the Telecommunication Act deregulated the industry by raising local ownership caps. This deregulation caused a massive merger wave, that reshaped the ownership structure, by moving from family based ownership into more corporate structures. I estimate that this consolidation raised consumer surplus by 1%, but lowered advertiser surplus by $300m. I find that the mergers created extra variety that increased listener welfare by $1.3%. On the other hand they softened competition and decreased advertiser welfare by $147m per year. Subsequent ad quantity adjustments led to a 0.3% decrease in listener welfare (with the variety effect it totals to the 1% increase) and an additional $153m decrease in advertiser welfare (with the variety effect it totals $300m).

References

Ackerberg, D. A., and M. Rysman (2005): "Unobserved Product Differentiation in Discrete-Choice Models: Estimating Price Elasticities and Welfare Effects," *RAND Journal of Economics*,

[5]Source: A.Richter (2006)

36(4), 771–788.

Argentesi, E., and L. Filistrucchi (2007): "Estimating market power in a two-sided market: The case of newspapers," *Journal of Applied Econometrics*, 22(7), 1247–1266.

A.Richter, W. (2006): *Radio: Complete Guide to the Industry*. Peter Lang Publishing.

Armstrong, M. (2006): "Competition in Two-Sided Markets," *The RAND Journal of Economics*, 37(3), 668–691.

Ashley Herweg, G. H. (1997): *Radios Niche Marketing Revolution FutureSell*. Focal Press.

Bain, J. (1968): *Industrial organization*. John Wiley & Sons.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–90.

Berry, S. T. (1994): "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25(2), 242–262.

Berry, S. T., and J. Waldfogel (2001): "Do Mergers Increase Product Variety? Evidence From Radio Broadcasting," *The Quarterly Journal of Economics*, 116(3), 1009–1025.

Brynjolfsson, E., Y. J. Hu, and M. D. Smith (2003): "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers," *Manage. Sci.*, 49(11), 1580–1596.

Bulow, J. I., J. D. Geanakoplos, and P. D. Klemperer (1985): "Multimarket Oligopoly: Strategic Substitutes and Complements," *The Journal of Political Economy*, 93(3), 488–511.

Chandra, A., and A. Collard-Wexler (2009): "Mergers in Two-Sided Markets: An Application to the Canadian Newspaper Industry," *Journal of Economics & Management Strategy*, 18, 1045–1070.

CRA International (2007): "Expost Merger Review: An Evaluation of Three Competition Bureau Merger Assessments," Discussion paper, CRA International.

Dube, J.-P. H., J. T. Fox, and C.-L. Su (2009): "Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation," NBER Working Papers 14991, National Bureau of Economic Research, Inc.

Dukes, A. (2004): "The Advertising Market in a Product Oligopoly," *Journal of Industrial Economics*, 52(3), 327–348.

Evans, D. S. (2002): "The Antitrust Economics of Two-Sided Markets," *SSRN eLibrary*.

Gal-Or, E., and A. Dukes (2006): "On the Profitability of Media Mergers," *Journal of Business*, 79(2), 489–526.

Gandhi, A. K., L. M. Froeb, S. T. Tschantz, and G. J. Werden (2005): "Post-Merger Product Repositioning," *SSRN eLibrary*.

Kaiser, U., and J. Wright (2006): "Price structure in two-sided markets: Evidence from the magazine industry," *International Journal of Industrial Organization*, 24(1), 1 – 28.

Kim, J., G. M. Allenby, and P. E. Rossi (2002): "Modeling Consumer Demand for Variety," *Marketing Science*, 21(3), 229–250.

L'Ecuyer, P., and T. H. Andres (1997): "A Random Number Generator Based on the Combination of Four LCGs," in *Mathematics and Computers in Simulation*, pp. 99–107.

Leeper, S. E. (1999): "Game of Radiopoly: An Antitrust Perspective of Consolidation in the Radio Industry," *Fed. Comm. L.J.*, 52(473).

Nevo, A. (2000): "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *RAND Journal of Economics*, 31(3), 395–421.

Rochet, J.-C., and J. Tirole (2006): "Two-Sided Markets: A Progress Report," *The RAND Journal of Economics*, 37(3), 645–667.

Rysman, M. (2004): "Competition between Networks: A Study of the Market for Yellow Pages," *The Review of Economic Studies*, 71(2), 483–512.

Sweeting, A. (2007): "Dynamic Product Repositioning in Differentiated Product Markets: The Case of Format Switching in the Commercial Radio Industry," NBER Working Papers 13522, National Bureau of Economic Research, Inc.

——— (2008): "The Effects of Horizontal Mergers on Product Positioning: Evidence from the Music Radio Industry," working paper, Duke University.

Appendices

A Tables

$r_{f|a}$

	Talk	Hits
Teens	1/5	4/5
Adults	3/5	2/5

$r_{a|f}$

	Teens	Adults
Talk	1/4	3/4
Hits	2/3	1/3

Ω

	Talk	Hits
Talk	0.56	0.44
Hits	0.28	0.72

Table 1: Simple example of advertising weights

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Number of stations	26.75	26.92	27.25	27.53	27.66	27.89	28.48	28.61	28.72	28.78	28.86
Number of owners	16.58	15.55	14.94	14.21	13.29	13.03	13.16	12.96	12.73	12.52	12.48
C3	0.77	0.83	0.88	0.91	0.97	0.95	0.93	0.93	0.93	0.93	0.90
Number of stations owned	4.43	5.10	5.66	5.94	6.58	6.32	6.31	6.34	6.42	6.38	6.28
Fraction of stations that changed ownership	0.12	0.12	0.10	0.11	0.12	0.03	0.04	0.03	0.03	0.03	NaN
Fraction of stations that changed format	0.11	0.11	0.13	0.12	0.12	0.13	0.10	0.11	0.11	0.11	NaN
Ad quantity	23.19	25.85	26.12	28.45	30.31	24.71	28.37	24.54	28.16	28.30	26.95
Price divided by price in 1996	1.00	0.96	1.08	1.10	1.26	1.51	1.42	1.51	1.39	1.37	1.43

Table 2: Panel data descriptive statistics

	Mean Effects (θ_1^L)	Random Effects (Σ_1)
Advertising	-1.106 (0.002)	0.030 (0.009)
AM/FM	0.861 (0.000)	-
AC, SmoothJazz, and New AC	-2.431 (0.008)	0.043 (0.004)
Rock	-1.559 (0.140)	0.004 (0.020)
CHR	-0.179 (0.025)	0.009 (0.006)
Alternative Urban	-2.339 (0.026)	0.348 (0.008)
News/Talk	-4.678 (0.010)	0.024 (0.002)
Country	-2.301 (0.006)	0.011 (0.003)
Spanish	-1.619 (0.004)	0.011 (0.001)
Other	-4.657 (0.004)	0.005 (0.002)
ρ	0.568 (0.091)	-

Table 3: Estimates of mean and random effects of demand for radio programming. Stars indicate parameter significance when testing with 0.1, 0.05 and 0.01 test sizes.

	Demographics characteristics (Π)					
	Age	Sex	Education	Income	Black	Spanish
AC, SmoothJazz, and New AC	−0.171 (0.001)	−0.341 (0.064)	0.602 (0.013)	−0.024 (0.003)	0.121 (0.012)	−1.014 (0.008)
Rock	−0.645 (0.072)	0.399 (0.031)	0.861 (0.006)	−0.147 (0.045)	−1.359 (0.007)	−1.643 (0.003)
CHR	−2.541 (0.015)	0.477 (0.080)	1.772 (0.006)	−0.291 (0.005)	1.946 (0.015)	0.463 (0.001)
Alternative Urban	−0.817 (0.008)	1.350 (0.018)	0.583 (0.025)	−0.141 (0.002)	3.152 (0.005)	0.267 (0.027)
News/Talk	0.329 (0.002)	1.228 (0.012)	0.237 (0.009)	0.093 (0.005)	−0.321 (0.001)	−1.649 (0.005)
Country	0.062 (0.004)	−0.149 (0.022)	0.133 (0.004)	−0.125 (0.003)	−1.548 (0.009)	−1.717 (0.002)
Spanish	−0.024 (0.013)	−0.908 (0.012)	−0.328 (0.018)	−1.140 (0.002)	−2.560 (0.004)	0.797 (0.003)
Other	0.263 (0.373)	0.624 (0.003)	0.338 (0.006)	−0.031 (0.063)	0.498 (0.001)	0.238 (0.002)

Table 4: Interaction terms between listeners' demographics and taste for radio programming.

Los Angeles, CA

	AC SmoothJazz New AC	Rock	CHR	Alternative Urban	News/Talk	Country	Spanish	Other
AC SmoothJazz New AC	**0.22**	0.10	0.11	0.09	0.17	0.14	0.00	0.17
Rock	0.15	**0.21**	0.12	0.09	0.16	0.13	0.01	0.12
CHR	0.18	0.12	**0.16**	0.16	0.10	0.13	0.03	0.13
Alternative Urban	0.11	0.05	0.17	**0.44**	0.06	0.05	0.00	0.12
News/Talk	0.17	0.10	0.05	0.05	**0.30**	0.13	0.00	0.21
Country	0.16	0.10	0.09	0.07	0.15	**0.22**	0.01	0.21
Spanish	0.03	0.04	0.11	0.02	0.01	0.03	**0.72**	0.04
Other	0.18	0.07	0.06	0.08	0.20	0.17	0.00	**0.23**
Total impact	1.20	0.79	0.87	0.99	1.15	1.00	0.77	1.23

Atlanta, GA

	AC SmoothJazz New AC	Rock	CHR	Alternative Urban	News/Talk	Country	Spanish	Other
AC SmoothJazz New AC	**0.20**	0.10	0.12	0.09	0.14	0.18	0.00	0.18
Rock	0.14	**0.21**	0.13	0.10	0.12	0.17	0.01	0.13
CHR	0.17	0.13	**0.17**	0.14	0.09	0.17	0.01	0.13
Alternative Urban	0.11	0.06	0.16	**0.40**	0.06	0.08	0.00	0.13
News/Talk	0.16	0.10	0.05	0.05	**0.25**	0.17	0.00	0.22
Country	0.15	0.09	0.08	0.06	0.13	**0.26**	0.01	0.22
Spanish	0.04	0.04	0.12	0.02	0.01	0.03	**0.71**	0.03
Other	0.16	0.07	0.06	0.07	0.16	0.23	0.01	**0.25**
Total impact	1.11	0.78	0.88	0.94	0.95	1.31	0.75	1.29

Knoxville, TN

	AC SmoothJazz New AC	Rock	CHR	Alternative Urban	News/Talk	Country	Spanish	Other
AC SmoothJazz New AC	**0.20**	0.11	0.16	0.11	0.10	0.16	0.01	0.16
Rock	0.13	**0.21**	0.14	0.11	0.10	0.18	0.01	0.12
CHR	0.16	0.12	**0.18**	0.14	0.08	0.17	0.02	0.13
Alternative Urban	0.12	0.06	0.16	**0.38**	0.06	0.08	0.00	0.13
News/Talk	0.16	0.13	0.10	0.09	**0.17**	0.16	0.01	0.18
Country	0.15	0.13	0.14	0.10	0.09	**0.22**	0.01	0.16
Spanish	0.05	0.05	0.11	0.02	0.02	0.04	**0.66**	0.05
Other	0.17	0.09	0.11	0.12	0.12	0.18	0.01	**0.21**
Total impact	1.12	0.90	1.11	1.05	0.74	1.21	0.72	1.14

Table 5: Product closeness matrices for chosen markets

Market population	less than .5m	between .5m and 1.5m	more than 1.5m
	1.34 (0.046)	0.35 (0.026)	0.00 (0.008)

Table 6: Slope of the inverse demand for ads θ_2^A, by market size

Market	Population (mil)	Marginal cost ($ per-miute)	Profit margin	Market	Population	Marginal cost	Profit margin
Los Angeles, CA	13,155	356.4 (5.15)	30%	Tulsa, OK	856	72.8 (2.13)	21%
Chicago, IL	9,341	180.0 (2.70)	34%	Knoxville, TN	785	54.3 (1.99)	27%
Dallas-Ft. Worth, TX	5,847	198.6 (5.60)	28%	Albuquerque, NM	740	27.4 (1.04)	36%
Houston-Galveston, TX	5,279	199.7 (4.20)	28%	Ft. Myers-Naples-Marco Island, FL	737	11.3 (0.94)	57%
Atlanta, GA	4,710	95.4 (3.37)	43%	Omaha-Council Bluffs, NE-IA	728	48.0 (0.91)	28%
Boston, MA	4,532	172.2 (3.68)	33%	Harrisburg-Lebanon-Carlisle, PA	649	29.7 (1.44)	42%
Miami-Ft, FL	4,174	134.3 (3.70)	28%	El Paso, TX	619	41.8 (4.12)	20%
Seattle-Tacoma, WA	3,776	128.7 (2.21)	29%	Quad Cities, IA-IL	618	51.3 (1.30)	23%
Phoenix, AZ	3,638	63.7 (1.84)	39%	Wichita, KS	598	38.9 (0.85)	25%
Minneapolis-St. Paul, MN	3,155	160.8 (4.66)	26%	Little Rock, AR	577	45.2 (1.64)	26%
St. Louis, MO	2,689	190.6 (5.38)	18%	Columbia, SC	577	60.0 (2.10)	23%
Tampa-St, FL	2,649	102.7 (2.09)	26%	Charleston, SC	569	59.6 (1.74)	19%
Denver-Boulder, CO	2,604	99.9 (1.40)	32%	Des Moines, IA	564	21.3 (0.92)	40%
Portland, OR	2,352	48.6 (1.35)	41%	Spokane, WA	540	24.5 (0.63)	28%
Cleveland, OH	2,134	170.6 (3.34)	24%	Madison, WI	520	93.6 (3.02)	22%
Charlotte, NC-SC	2,127	67.1 (1.96)	38%	Augusta, GA	510	30.9 (0.60)	24%
Sacramento, CA	2,100	47.6 (1.30)	42%	Ft. Wayne, IN	509	37.8 (1.35)	27%
Salt Lake City, UT	1,924	58.1 (1.19)	26%	Lexington-Fayette, KY	495	36.8 (1.59)	35%
San Antonio, TX	1,900	75.0 (2.27)	24%	Chattanooga, TN	471	41.5 (2.53)	29%
Kansas City, MO-KS	1,871	152.5 (2.87)	19%	Boise, ID	469	46.2 (3.73)	30%
Las Vegas, NV	1,752	47.7 (1.49)	32%	Jackson, MS	453	18.6 (2.03)	59%
Milwaukee-Racine, WI	1,713	74.6 (1.27)	25%	Eugene-Springfield, OR	439	27.4 (1.29)	31%
Orlando, FL	1,686	42.4 (1.77)	41%	Reno, NV	400	99.7 (1.64)	15%
Columbus, OH	1,685	70.2 (1.53)	30%	Shreveport, LA	359	19.8 (4.25)	92%
Indianapolis, IN	1,602	86.8 (2.32)	26%	Fayetteville, NC	337	38.1 (2.48)	46%
Norfolk, VA	1,583	196.8 (4.64)	17%	Springfield, MA	336	20.8 (0.87)	55%
Nashville, TN	1,342	40.5 (1.84)	38%	Macon, GA	276	34.4 (2.29)	26%
Greensboro-Winston, NC	1,329	53.5 (2.34)	32%	Binghamton, NY	255	37.5 (1.51)	27%
New Orleans, LA	1,294	91.2 (2.44)	24%	Lubbock, TX	248	57.7 (1.98)	18%
Memphis, TN	1,278	53.2 (1.82)	30%	Odessa-Midland, TX	231	21.4 (0.99)	27%
Jacksonville, FL	1,271	66.1 (1.64)	29%	Fargo-Moorhead, ND-MN	200	48.6 (2.42)	25%
Oklahoma City, OK	1,268	75.6 (1.35)	25%	Medford-Ashland, OR	184	27.7 (0.90)	28%
Buffalo-Niagara Falls, NY	1,150	141.5 (3.63)	19%	Duluth-Superior, MN-WI	159	43.3 (0.79)	20%
Louisville, KY	1,100	92.9 (2.36)	21%	Parkersburg-Marietta, WV-OH	157	31.7 (1.41)	21%
Richmond, VA	1,066	55.3 (1.47)	28%	Abilene, TX	149	23.0 (1.14)	26%
Birmingham, AL	1,030	85.8 (2.50)	24%	Eau Claire, WI	149	31.6 (2.77)	28%
Honolulu, HI	938	78.2 (2.39)	15%	Williamsport, PA	130	31.0 (1.13)	23%
Albany, NY	909	113.9 (3.18)	16%	Monroe, LA	124	14.2 (1.49)	64%
Grand Junction, CO	902	24.5 (0.67)	24%	Sioux City, IA	118	26.1 (0.96)	24%
Tucson, AZ	870	41.1 (0.93)	27%	San Angelo, TX	104	26.4 (1.36)	16%
Grand Rapids, MI	864	37.9 (0.79)	38%	Bismarck, ND	99	32.8 (1.65)	22%

Table 7: Estimated marginal cost (in dollars per minute of broadcasted advertising) and profit margins (before subtracting the fixed cost) for a chosen set of markets

	Consumer surplus	Average ad load	Advertiser surplus	Advertising minutes	Mean price index
Ownership change and format switching	11.7pdm +2.5%	-5.4pdm -17.3%	-118.1m -15.8%	-737min -1.0%	+1.34%
Ad adjustment	1.2pdm +0.3%	-2.2pdm -8.4%	-119.4m -19.0%	-8,216min -11.7%	+5.66%
Total impact	12.9pdm +2.8%	-7.5pdm -24.2%	-237.5m -31.8%	-8,953min -12.6%	+6.99%

Table 9: Counterfactuals for small markets (less than 500k people)

Market population	less than .5m	between .5m and 1.5m	more than 1.5m
Baseline model	1.34 (0.046)	0.35 (0.026)	0.00 (0.008)
Oligopoly within format	1.07 (0.036)	0.28 (0.061)	0.02 (0.009)
Perfect substitutes	1.44 (0.035)	0.32 (0.030)	0.01 (0.009)

Table 11: Slope of the inverse demand for ads θ_2^A, by market size

	Consumer surplus	Average ad load	Advertiser surplus	Advertising minutes	Mean price index
Ownership change and format switching	6.6pdm +1.3%	-6.4pdm -12.6%	-158.3m -16.3%	-2,491min -1.5%	+0.60%
Ad adjustment	-1.9pdm -0.4%	1.6pdm +3.6%	-146.1m -18.0%	-9,838min -5.9%	+2.09%
Total impact	4.7pdm +0.9%	-4.8pdm -9.5%	-304.4m -31.4%	-12,329min -7.3%	+2.67%

Table 8: Counterfactuals for all markets

	Consumer surplus	Average ad load	Advertiser surplus	Advertising minutes	Mean price index
Ownership change and format switching	2.6pdm +0.5%	-6.0pdm -11.0%	-1.0m -12.8%	-835min -2.0%	+0.01%
Ad adjustment	-4.4pdm -0.8%	4.6pdm +9.5%	0.7m +9.9%	3,081min +7.7%	-0.02%
Total impact	-1.8pdm -0.3%	-1.4pdm -2.5%	-0.3m -4.2%	2,245min +5.5%	-0.01%

Table 10: Counterfactuals for large markets (more than 2,000k people)

	Consumer surplus	Average ad load	Advertiser surplus	Advertising minutes	Mean price index
Baseline model	4.7pdm +0.9%	-4.8pdm -9.5%	-304.4m -31.4%	-12,329min -7.3%	+2.67%
Oligopoly within format	4.4pdm +0.8%	-4.5pdm -9.0%	-253.4m -31.3%	-9,056min -5.6%	+1.12%
Perfect substitutes	4.9pdm +0.9%	-5.3pdm -10.3%	-314.7m -32.7%	-16,648min -9.0%	+2.57%

Table 12: Robustness of counterfactuals

B Advertising demand: Micro foundations

In this section I present a model that rationalizes inverse demand for advertising (1.5)

Assume that there are $\mathcal{A}$ types of advertisers. Each type $a \in \mathcal{A}$ targets a certain demographic group(s) d_a. Let γ_2 be a total mass of advertisers and AS_a be a share of advertisers of type a in market m. Advertisers are also heterogeneous in their value of the ad slot in format f, and I assume that those values are distributed uniformly on the interval $[0, \gamma_{1f}]$. An advertiser of type a gets utility only if a listener of type d_a hears an ad. To compute the exact expected value of an advertising slot, advertisers need to know the demographic composition of each station in the market. Because advertisers are small, and such detailed data is not offered by Arbitron, it seems unlikely that they would be able to do that. Instead, I assume that they approximate those calculations using publicly available data contained in Arbitron's Radio Today publications. These publications provide nation-wide conditional probabilities $r_{f|a}$ of a consumer of type d_a choosing format f conditional on listening to the radio. Advertisers take these conditional probabilities as given and compute the market specific probabilities of obtaining correct listeners when advertising in each format. Such computations can be done by Bayes' Rule, i.e.

$$r_{a|f} = \frac{r_{f|a} LS_a}{r_f}$$

where $r_f = \sum_c r_{f|a} LS_a$ and LS_a is the population share of demographic group d_a, which is assumed to be known to the advertiser. Having listeners' distributions $r_{a|f}$ and station ratings r_j (available on Arbitron's website) at hand, advertisers compute the probability of successful targeting at station j to be $r_j r_{a|f}$, where f is a format of station j.

Radio stations quote costs-per-point CPP_{af} individually for each advertiser type and format. Advertisers decide if they want to purchase advertising after observing the CPPs and station ratings. Because advertisers are small and likely do not have much market power over radio station owners, I assume that they are price and rating takers[6]. Advertisers can purchase advertising from several stations at once; however, I assume away any potential complementarities.

In equilibrium, advertisers purchase advertising as long as their expected value is above price. Let q_a be the amount of advertising purchased by advertisers of type a. A marginal advertiser must be indifferent between purchasing advertising or not, so the clearing per-listener prices are given by

$$\text{CPP}_{af} = \gamma_{1f} r_{a|f} \left(1 - \frac{1}{\gamma_2 AS_a} q_a\right)$$

Given the clearing prices CPP_{af}, advertisers are indifferent when choosing between formats, so I assume that advertising is purchased proportionally to the target listeners' tastes i.e. $q_a = AS_a \sum_f r_{f|a} q_f$. If I make the simplifying assumption that $AS_a \approx LS_a$, then the arrival probability of an advertiser of type a at a station of format f would be equal to $r_{a|f}$. Therefore, expected per-listener price in

[6]This assumption is is motivated by the fact that about 75% is purchased by small local firms. Such firms' advertising decisions are unlikely to influence prices and station ratings in the short run.

format f is given by

$$\text{CPP}_f = \sum_a (r_{a|f})^2 \gamma_{1f} \left(1 - \frac{1}{\gamma_2} \sum_{f'} r_{f'|a} q_{f'}\right) =$$
$$= \gamma_{1f} \left(\sum_a (r_{a|f})^2\right) \left(1 - \frac{1}{\gamma_2} \sum_{f'} q_{f'} \left(\sum_a (r_{a|f})^2\right)^{-1} \sum_a (r_{a|f})^2 r_{f'|a}\right).$$

Finally, I obtain Equation (1.5)

$$p_j = \theta_{1f}^A r_j \left(1 - \theta_2^A \sum_{f' \in \mathbf{F}} \omega_{ff'}^m q_{f'}\right)$$

by setting $\omega_{jj'} = \left(\sum_a (r_{a|f})^2\right)^{-1} \sum_a (r_{a|f})^2 r_{f'|a}$, $\theta_2^A = \frac{1}{\gamma_2}$ and assuming that $\theta_1 = \gamma_{1f} \sum_a (r_{a|f})^2$ for all f. The last assumption basically means that niche formats (with listenership concentrated in one demographic bin) are less profitable for advertisers than general interest formats.

The presented model is only one of a number of ways to rationalize the weighted price equation (1.5) in which competition between formats is channeled though demographics. Other possibilities include: a local monopoly in which each advertiser type draws utility only from advertising on one particular station, and a format-monopoly in which each advertiser type targets only one format.

C Numerical considerations

To solve the optimization problem (3.3), I used a version of the Gauss-Newton method implemented in the commercial solver KNITRO. Using this state-of-the-art solver avoids certain convergence problems that are common to many non-linear estimators.

The iteration step of the KNITRO solver requires computing constraints, a Jacobian of the constraint, and an inverse of the inner product of this Jacobian (used to compute the approximate Hessian of the Lagrangian). The objective function and its Jacobian come essentially for free because of their simple nature.

To compute the constraints and their Jacobian, I employed a piece of highly optimized parallel C code. This allows the use a fairly large dataset (about 42,000 observations) and many draws (500 draws from Normal and CPS per date/market) when computing the constraints. When parallelizing the code, I was careful to maintain independence of the draws within and between threads. To achieve this, I implemented a version of a pseudo-random number generator (described in (L'Ecuyer and Andres 1997)). This generator enables us to create a desired number of independent pseudo-random feeds for each thread.

One iteration of the solver takes about two to three minutes on an 8-Core 3Ghz Intel Xeon processor and uses about 4GB of memory. About 90% of this computation is the inversion of a Hessian estimator within the KNITRO solver. This inversion cannot be parallelized because it is done inside the solver, without the user's control.